Cronfa
Swansea University's Research Repository

Identifying Host and Bacterial Biomarkers to Predict Sepsis

Matthew Lawrence Lewis

Submitted to Swansea University in fulfilment of the requirements for the Degree of Doctor of Philosophy

Swansea University

2023

Copyright: The Author, Matthew L. Lewis, 2024.

Summary

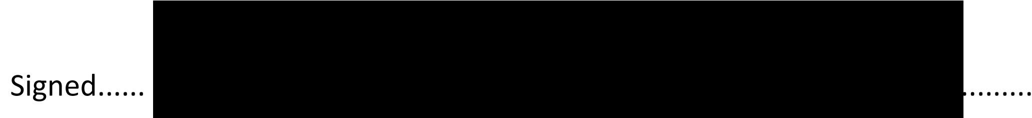
Sepsis is a serious concern for healthcare programmes worldwide and is associated with 11 million deaths annually. *E. coli* accounts for 20% of bloodstream infections worldwide and is responsible for 17% of sepsis related mortality. Early diagnosis and treatment are critical and delays in initiating antimicrobial therapy are linked to mortality. Identifying biomarkers to predict patients that might succumb to sepsis is vital to aid in early diagnosis. It was hypothesised that *E. coli* bacteraemia isolates and isolates from different sources of bacteraemia would elicit a distinctive host response and be genetically unique.

Blood culture positive isolates (n=165) were collected from the Hywel Dda University Health Board. Most of the isolates were assigned to the B2 and D phylogroups and belonged to either ST131 or ST73. Antimicrobial resistance in the collection was lower than national averages. Host models of infection were used to identify phenotypic responses of the bacterial collection. IL-8 and MIP3α were increased following stimulation by bacteraemia isolates compared to non-pathogenic strains. Greater IL-8 in whole blood was associated with a urinary and abdominal bacteraemia. Isolates that were resistant to human plasma elicited a higher IL-6, IL-8 and resistin response in whole blood compared to plasma sensitive isolates. Blood culture positive bacteraemia isolates had significantly more virulence factors than control isolates. Bacteraemia isolates expressed more P fimbriae genes. The S fimbrial adhesin genes were found to be significantly different between urinary and abdominal isolates. Abdominal isolates had significantly more *sfaC* (32%) while urinary isolates had more *sfaX* (22%). GWAS analysis revealed 6 potential gene targets based on bacterial phenotypes. These were *ynbC*, *yhgE*, *ybjE*, *yejF*, *tufB* and *yohF*.

The results contained within this thesis describe new targets for predicting bacteraemia and sepsis and underline the importance of using host and pathogen as sources for biomarkers.

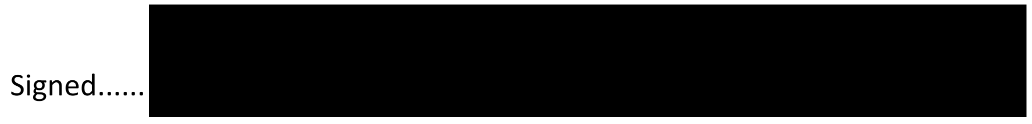
Declarations

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed......  ....

Date.....09/06/24..

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed...... 

Date........09/06/24...

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed...

Date.......09/06/24..

The University's ethical procedures have been followed and, where appropriate, that ethical approval has been granted.

Signed... ...

Date.........09/06/24..

Acknowledgements

First, I would like to thank my supervisor Professor Tom Wilkinson for putting up with me for 3+ years and his fantastic supervision during my PhD. This thesis would not have been possible without his tireless support in helping me plan experiments and discuss ideas, as well as the odd political discussion and the occasional wordle competition.

I would also like to thank Dr Llinos Harris, Dr Rowena Jenkins, Dr Lisa Williams, and Dr Aled Roberts for their help with lab work, poster preparations and project ideas.

Thank you to Dr Matthew Hutchins for sequencing the *E. coli* isolates and providing bioinformatics guidance. Thank you to Professor Kier Lewis and Dr Mike Simmons for kindly providing the isolates and a special thank you to Jane Regan for her fantastic efforts in organising the patient data.

I would also like to thank Stephen Leary for his great work as our laboratory technician his help with preparing various buffers and media was greatly appreciated.

Thank you to the members of the MID group both old and new, Heather, Ebony, Taylor, Sarah, Phil, Meg, Nerissa, Amy, and Dan for making my time at Swansea enjoyable.

Additionally, I would like to thank Professor Paul Morgan and Wioleta Zelek for kindly providing complement reagents.

Finally, I would like to thank my family for both financial and emotional support (mostly financial), during my time as a PhD student. Thanks Mum, Dad, Alex, Aunty Marie, and Uncle Andy!

Contents

Chapter 1. Introduction 11

1.1 Sepsis 11

1.1.1 Definition 11

1.1.2 Global impact 11

1.1.2 Sepsis diagnosis, treatment, and prevention 12

1.1.2.1 Diagnosis 12

1.1.2.2 Treatment 13

1.1.2.3 Prevention and the role of biomarkers 15

1.1.3 Risk factors for sepsis 17

1.1.4 Sepsis-associated pathogens 17

1.1.4 Origins of infection 20

1.2 *Escherichia coli* 20

1.2.1 Background 20

1.2.2 *E. coli* virulence factors 24

1.2.2.1 Adhesins 25

1.2.2.2 Protectins and invasins 25

1.2.2.3 Iron acquisition 26

1.2.2.4 Toxins 26

1.2.3 *E. coli* and antimicrobial resistance (AMR) 27

1.2.4 *E. coli* in sepsis 28

1.3 The immune system and the host response 30

1.3.1. Passive innate immunity and barriers 30

1.3.2 Induced innate immunity 30

1.3.2.1 Pathogen associated molecular patterns (PAMPS) 31

1.3.2.2 Damage associated molecular patterns (DAMPs) 32

1.3.2.3 Complement 33

1.3.2.4 Neutrophils 34

1.3.2.5 Macrophages and monocytes 35

1.3.2.6 Cytokines and the acute phase response 36

1.3.2.7 Coagulation 37

1.3.2.8 Cell death and immunity 38

1.3.3 Adaptive immunity 38
1.4 The immune system in sepsis 39
1.4.1 Summary 39
1.4.2 Cells and pathways affected by sepsis 39
1.4.2.1 LPS theory of sepsis 39
1.4.2.2 Barriers to infection in sepsis 40
1.4.2.3 Complement in sepsis 41
1.4.2.4 Neutrophils in sepsis 41
1.4.2.5 Monocytes and Macrophages in sepsis 42
1.4.2.6 Cytokines in sepsis 43
1.4.2.7 DAMPS in sepsis 44
1.4.2.7 Coagulation in sepsis 44
1.4.2.8 Adaptive immunity during sepsis 45
1.5. Biomarkers in sepsis 45
1.5.2 Current sepsis biomarkers 47
1.5.2.1 Procalcitonin (PCT) 47
1.5.2.2 C reactive protein 48
1.6 Current issues and gaps in the literature 51
1.7 Aims and objectives 52
Chapter 2: Materials and methods 54
2.1 Media and stock solutions 54
2.1.1 Media 54
2.1.1.1 LB Broth 54
2.1.1 Columbia blood agar plates 54
2.1.2 THP-1 Cell growth media 54
2.1.3 Experimental THP-1 cell media 54
2.1.4 Phosphate buffered saline 55
2.1.5 ELISA wash buffer 55
2.1.6 ELISA reagent diluent 55
2.1.7 ELISA stop solution 55
2.1.8 HEPES buffer 55
2.2 Reagents 55
2.2.1 ELISA colour solution 55
2.2.2 DNA extraction kit 56
2.2.3 Citrated human plasma 56
2.2.4 Healthy volunteer serum collection 56

2.3 Microbiology ... 56

2.3.1 *E. coli* strains from Hywel Dda University Health Board used in this study ... 56

2.3.1.2 K12 knockout mutants ... 56

2.3.1.3 Additional *E. coli* strains used in this study. ... 57

2.3.2 Culture of *E. coli* ... 57

2.4 Cell biology and cell culture ... 57

2.4.1 Whole blood model of infection ... 57

2.4.2 THP-1 cell model of infection ... 58

2.5 Assays ... 58

2.5.1 Complement depletion ... 58

2.5.2 Human cytokine ELISAs ... 58

2.5.3 Plasma resistance of *E. coli* ... 59

2.5.4 Serum bactericidal assay ... 59

2.6 Microbial genetics, sequencing and analysis ... 60

2.6.1 DNA isolation and extraction ... 60

2.6.2 Genome assembly ... 61

2.6.3 Genome annotation, pangenome generation and GWAS ... 61

2.6.4 MLST ... 61

2.6.5 Phylogroup Identification ... 62

2.6.6 Virulence factor and AMR gene identification ... 62

2.6.7 Virulence and antimicrobial resistance heat maps ... 62

2.7 *E. coli* associated patient data ... 62

2.8 Data analysis and statistics ... 62

Chapter 3: *E. coli* bacteraemia collection: their phenotypic and genetic traits and patient demographic data ... 64

3.1 Bacterial sepsis in Wales ... 64

3.2 Hywel Dda University Health Board ... 64

3.3 Properties of *E. coli* bacteraemia strains ... 66

3.3.1 Source of *E. coli* bacteraemia ... 66

3.3.2 UTI ... 66

3.3.3 GI and biliary tract ... 66

3.4 Antimicrobial resistance ... 67

3.6 Aims and objectives ... 68

3.7 Materials and methods ... 68

3.7.1 *E. coli* bacteraemia collection ... 68

3.7.2 Hospital acquired patient and isolate data ... 69
3.7.3 DNA extraction and genome processing ... 69
3.7.4 Phylotyping and sequence typing ... 70
3.8 Results ... 70
3.8.1.1 Antimicrobial resistance ... 70
3.8.2 Genetic traits of *E. coli* collection ... 71
3.8.2.1 Phylogroup assignment ... 71
3.8.2.2 Sequence typing ... 72
3.8.2.3 AMR genes of ST131 and ST73 isolates ... 73
3.8.3 Patient demographics and definitions ... 74
3.8.3.1 Patient demographics ... 74
3.8.4. Inter-relationships between data ... 77
3.8.4.1. Source and phylogroup ... 77
3.9 Discussion ... 79
Chapter 4: Host biomarkers induced with clinical *E. coli* strains from bacteraemia patients ... 83
4.1 Host biomarkers of infection ... 83
4.2 Early mediators of the immune response ... 83
4.2.1. IL-6 ... 83
4.2.2 IL-8 ... 85
4.2.3 MIP1α ... 85
4.2.4 MIP3α ... 86
4.2.5 Resistin ... 86
4.2.6 TNFα ... 87
4.3 Host and bacterial factors affecting outcomes in bacteraemia ... 87
4.3.1 Host factors affecting outcomes in bacteraemia ... 87
4.3.2 Bacterial factors affecting outcomes in bacteraemia-the importance of ExPEC strain in inducing divergent responses ... 88
4.4 Aims and objectives ... 90
4.5 Methods ... 90
4.5.1 *Ex-vivo* whole blood model of infection. ... 90
4.5.2 THP-1 infection model ... 90
4.5.3 Plasma resistance in *E. coli* ... 90
4.5.4 Serum bactericidal assay ... 91
4.6 Results ... 91
4.6.1 Whole blood response to *E. coli* isolates ... 91
4.6.2 Whole blood response to bacteraemia/non-pathogenic *E. coli* ... 92

4.6.3 Whole blood cytokine response to sepsis and non-sepsis *E. coli* isolates 95
4.6.4 Host response to source of infection 95
4.6.5 Host response to mortality 95
4.6.6 Cytokine responses to *E. coli* isolates in THP-1 cells 98
4.6.7 Plasma resistance of the *E. coli* collection. 102
4.6.8. Interrelationship between whole blood responses and plasma sensitivity 104
4.6.9 Serum bactericidal assay (SBA) assay 106
4.7 Discussion 107
Chapter 5 Descriptive and comparative genomics of virulence factors and AMR genes in *E. coli* bacteraemia in Southwest Wales 113
5.1 Introduction 113
5.1.1 *E. coli* as a commensal and pathogen 113
5.1.2 Genetic traits of ExPEC isolates 114
5.1.2.1 Virulence factors 114
5.1.2.2 Mechanisms of antimicrobial resistance in ExPEC 115
5.2 Aims and objectives. 116
5.2 Methods 117
5.2.1 Assembly 117
5.2.2 Phylogenetic tree 117
5.2.3 Annotation 117
5.2.4 Abricate 117
5.2.5 Comparing presence and absence of genes 118
5.4 Results 118
5.4.1 *E. coli* pangenome creation 118
5.2 Phylogenetic relationships of the *E. coli* collection 119
5.4.3 Total number of virulence factors and AMR genes 120
5.4.4 Top virulence factor and AMR genes 122
5.4.5 Incidence of Virulence factors and AMR genes in bacteraemia compared to ECOR isolates. 123
5.4.6 Incidence of established *E. coli* virulence factors in urinary and abdominal isolates 125
5.4.7 Incidence of established *E. coli* virulence factors in urinary and abdominal isolates in B2 phylogroup 127
5.5 Discussion 130
Chapter 6: Identification and functional confirmation of gene targets associated with clinical, and laboratory associated bacteraemia phenotypes. 134
6.1 Introduction 134
6.1.1 Bacterial whole genome wide association studies 134
6.1.2 Scoary 135

6.1.2 Genetic engineering and the *E. coli* Keio collection 136
6.1.2.1 Methods of genetic engineering 136
6.1.2.1 *E. coli* Keio collection 137
6.2 Aims and objectives 137
6.3 Methods 138
6.3.1 Bacterial GWAS 138
6.3.2 Whole blood infection model 138
6.3.3 THP-1 model of infection 138
6.4 Results 139
6.4.1 Genes associated with bacteraemia 139
6.4.2 Associations with origin of infection 140
6.4.3 Genes associated with Urinary bacteraemia isolates 142
6.4.4 Genes associate with sepsis isolates 142
6.4.5 Genes associated with patient and phenotypic data 143
6.4.6 Associations with *in vitro* laboratory phenotypes 143
6.5.7 Top gene targets identified through GWAS 146
6.4.8 *Ex vivo* whole blood and THP-1 cell cytokine response to *E. coli* K12 mutants 146
6.4.9 *E. coli* K12 mutant growth in human plasma 148
6.5 Discussion 148
Chapter 7: General discussion 152
7.1 Overarching thesis aims and hypothesis 152
7.2 Results chapter summaries 152
7.3 General discussion and future work 158
7.4 Limitations 160
Bibliography .... **Error! Bookmark not defined.**
Appendices 197
Appendix 2 197
Appendix 3 201
Appendix 4 207
Appendix 5 209
Appendix 6 222

Chapter 1. Introduction

1.1 Sepsis

1.1.1 Definition

Sepsis or blood poisoning is a serious condition arising from a dysfunctional host response to an infection. This exaggerated host immune response is frequently associated with bacteraemia (the presence of bacteria in the blood), resulting in systemic inflammatory changes which can cause damage to tissues and ultimately lead to multiple organ failure and death. Historic definitions of sepsis syndrome have been divided into three main types of systemic inflammatory response syndrome (SIRS) based on the severity and associated mortality; these were sepsis, severe sepsis, and septic shock [1]. However, definitions remain complex, a patient was said to suffer from SIRS if they presented with two or more of the following: temperature greater than 38°C or less than 36°C, heart rate greater than 90 bpm, respiratory rate greater than 20 breaths / minute, and a white blood cell count exceeding $12x10^9$ /L or below $4x10^9$ /L [1]. This system led to confusion and a lack of clear definitions for each syndrome, ultimately leading to a revised definition of sepsis, agreed by the third international sepsis taskforce in 2016. Sepsis was redefined as ***'life threatening organ dysfunction caused by a dysregulated host response to infection'*** and septic shock was redefined as ***'a subset of sepsis associated with greater mortality with profound circulatory, cellular and metabolic abnormalities.'*** The terms SIRS and severe sepsis were removed from these definitions [2].

1.1.2 Global impact

Sepsis represents a significant challenge to healthcare professionals worldwide and is recognised as a leading cause of mortality and morbidity but understanding its global scale is difficult. Low to middle income countries (LMIC) are often underrepresented in these estimates, primarily due to missing data on sepsis epidemiology. This discrepancy is highlighted in two reports. Firstly, a meta-analysis in 2016 assessing global incidence, using data from 27 studies across a range of high-income countries reports an estimated 19.4 million sepsis cases a year with an associated 5.3 million annual deaths [3]. In comparison a more recent report in the Lancet, using cause of death records to estimate the global

burden of sepsis including data from low-income countries, aimed to improve our understanding of global sepsis epidemiology. Using this method, the authors estimated 48.9 million cases of sepsis in 2017 linked to 11 million deaths (19.7% of all global deaths) [4]. This prediction is almost double previous estimates, confirming the previous bias towards studying high-income countries and underestimating the burden in low-income countries. It is clear that sepsis impacts public health on a global scale.

In the UK, there are an estimated 274,000 sepsis cases a year. Between 2001 and 2010 there were 226 547 deaths directly linked with sepsis (4.7% of all deaths), with an annual estimated cost to the NHS of £7.76 billion [5–7]. In the United States more than 750,000 people are diagnosed annually at an estimated cost of $16 billion [8]. Despite improvements in the supportive care of critically ill patients and improvements in early diagnosis and treatments, mortality rates remain high at between 17-26% [3]. In addition, sepsis patients who do survive often develop long term sequalae. These long-term adverse effects can result in readmission rates of 26% within 40 days and 48% within 180 days, with many patients having multiple readmissions [9], increasing the burden of sepsis related healthcare costs. The increasing global burden of sepsis has resulted in the World Health Organisation (WHO) confirming sepsis as a global health priority. This resolution aims to reduce global burden through improved 'prevention, diagnosis, and management of sepsis' [10,11].

1.1.2 Sepsis diagnosis, treatment, and prevention

1.1.2.1 Diagnosis

The diagnosis of an infection is guided by; i) clinical signs and symptoms (e.g., fever, white blood cell count, immune markers); ii) presence of signs of infection (e.g., dysuria, redness, swelling); and iii) microbiological confirmation (e.g., colony counts, PCR). Signs of infection combined with organ dysfunction are classified as sepsis [12]. Difficulties in this diagnostic method arise due to the heterogenous nature of sepsis infections. The signs of infection are not always present as they are often be masked by the treatment of pre-existing conditions. This is especially evident in nosocomial infections when patients may be treated for other diseases. In addition, the immunosuppressed, the elderly and critically ill patients (key risk factors) often lack these signs of infection. Furthermore, signs may also be masked in critically ill patients while receiving antimicrobial therapy resulting in negative blood culture

tests. The situation is also compacted by the fact that microorganisms may not need to pass into the blood to cause sepsis, due to the secretion of toxins and the local and systemic response to them, which would also result in a negative blood culture [13]. Multiple organisms can also be the cause of sepsis which can lead to complications assigning antimicrobial therapy [14,15]. Clinical scenarios like these provide physicians with poor evidence for specific diagnoses and unexplained organ dysfunction is often a sign of late-stage infections.

1.1.2.2 Treatment

Sepsis is further complicated by the limited treatment options available. Current treatment strategies revolve around eliminating the infection with antibiotics (if appropriate) and supportive therapy to maintain vital signs and organ function (Table 1.1).

Table 1.1: Recommended treatment options. Adapted from Gotts and Matthay [1]

Treatment	Function
Antibiotics	Treat bacterial infection (administration recommended within the first hour of diagnosis).
Fluids (Crystalloid and colloid)	Counteract fluid deficit caused by vasodilation and increased membrane permeability. Prevent blood pressure from dropping.
Vasopressors (Norepinephrine, epinephrine, and vasopressin)	Improve blood pressure, organ perfusion and arterial pressure.
Insulin therapy	Maintaining glucose levels during hypoglycaemia.
Lung protective ventilation	Reduce damage and inflammation in the lungs

While these treatments can save lives there are three clear limitations in their success:

Firstly, these treatment options do not contain strategies to target the systemic inflammatory responses closely linked to the sepsis syndrome. Over the last ten years,

numerous clinical trials have tested strategies that target some of the inflammatory pathways involved in sepsis (e.g LPS/TLR4, section 1.1.4), but have often resulted in poor efficacies compared to conventional treatments. More than 100 randomised clinical trials aimed at modulating the immune response to infection have been conducted and none have been successful. A classic example of a failed clinical immunomodulatory intervention trial in sepsis is highlighted by the recent failure of the TLR-4 antagonist Eritoran to reduce mortality in sepsis patients compared to controls [16]. Therefore, despite initial promise in animal models, there is poor translation to humans. Another example is the administration of activated protein C to patients, as although this was shown to reduce mortality in sepsis patients it was also associated with increased risk of serious bleeding [17].

Secondly, early diagnosis and treatment are critical for patient survival with multiple studies supporting earlier treatment associated with decreased mortality. The early administration of antibiotics can have a significant effect on disease outcome, with a 6-7% increase in mortality for each hour delay in administration [18,19]. Further evidence for the benefits of early treatment come from a 2014 retrospective study showing that patients receiving appropriate antibiotic treatment within the first hour of diagnosis had lower mortality than patients given antibiotics after 6 hours [20–22].

Thirdly, the increasing incidence of antimicrobial resistance is also a concern for the successful treatment of sepsis. Not only will antibiotic therapies be unsuccessful in treating sepsis caused by viral or fungal infection, but the incorrect use of antibiotics can have drastic consequences on patients. This is seen with the increase in beta-lactam producing *E. coli* in bacteraemia patients who received inappropriate antibiotics, as they a have a higher mortality rate than those receiving an appropriate initial treatment [23–26].

Attempts have been made to identify distinctive clinical phenotypes of sepsis patients to guide therapies and clinical trial designs. A retrospective analysis study using data from nearly 64,000 patients identified 4 phenotypes associated with different demographics. Patterns of organ dysfunction in these phenotypes were found to correlate with biomarkers and mortality. The phenotypes identified were α, β, γ and δ. The α phenotype was the most common and associated with lower vasopressor administration, the β phenotype comprised older patients with greater levels of chronic diseases, the γ phenotype was associated with greater levels of inflammation and pulmonary disfunction and patients in the δ group had

higher levels of liver dysfunction and septic shock [27]. In addition, patients in the γ and δ had greater levels of inflammatory markers (including IL-6) compared to the other two phenotypes. In addition, the δ phenotype also had higher levels of coagulation related biomarkers (such as thrombin-antithrombin complex), compared to the other phenotypes [27]. Studies like these are valuable as they provide data that will inform the design of clinical trials for sepsis therapies, improving the chances of success.

1.1.2.3 Prevention and the role of biomarkers

Patients with sepsis can experience rapid deterioration due to factors such as low blood pressure and release of endotoxin from causative organism, leading to the sudden occurrence of life threatening organ dysfunction [28,29]. Developments in the use of biomarkers aimed at identifying causative organism during sepsis infection may lead to improvements in both treatment options and preventative measures for at risk patient groups. The recent pandemic caused by Sars-CoV-2 has been shown to cause sepsis in a number of patients, causing cytokine storm and widespread immune dysfunction with increased risk of mortality [30].

Due to the widespread response by academics and healthcare professionals on a global scale our understanding of COVID pathogenesis has developed rapidly and has resulted in a number of different treatment options that are showing success in the clinic. For example, the monoclonal antibody Tocilizumab, which targets the IL-6 receptor, thus reducing inflammatory pathways, has been shown to reduce pathology and associated mortality of COVID19 in a dose and time dependant manner for patients with severe infection[29]. Additionally, Revens anti-sepsis drug RJX has recently completed a promising phase II trial with 9/12 patients with symptomatic high risk COVID, including three with hypoxemic respiratory failure showing rapid clinical recovery [31]. In addition to the different prophylactic treatment options now available for at risk Sars-Cov-2 patients, the vaccine programs across the globe have been shown to be a massive success, dramatically reducing the mortality rates and severe cases in hospitals, with an estimated 14.4 million deaths prevented in the first year of the vaccine programme alone [32].

Unfortunately, as sepsis can be caused by such a large range of pathogenic organisms [33–35], the antigenic differences between them makes vaccination against all causative agents

of sepsis unfeasible. Preventing sepsis currently revolves around increasing awareness, in both the public and healthcare professional communities, with the aim of reducing the time it takes to diagnose and treat. An example of a successful campaign is THINK SEPSIS from Health Education England which aimed to improve the diagnosis and management of sepsis within the NHS through education [36].

The identification of sepsis can also be improved using diagnostic tools. Scoring systems such as sequential organ failure assessment (SOFA) and the Acute Physiology and Chronic Health Evaluation II (APACHE II) system, can be useful to diagnose organ dysfunction in sepsis and advise treatment, although this is at a later stage of infection. The APACHE II system uses a variety of categories to evaluate patient condition including history of organ failure, age, heart rate, levels of sodium, potassium, and creatinine, as well as concentrations of white and red blood cells. The APACHE II scoring system was first introduced in 1981 and is still in use today [37]. More recently, the APACHE based systems have been updated including the APACHE III and APACHE IV, but they have not been fully tested and adopted to date [38].

The speed of onset and complexity of sepsis makes complete prevention difficult, therefore strategies to improve early detection methods remain viable options for improving outcomes. One method increasingly being used for many diseases, including cancer and sepsis, is the use of biological markers or biomarkers. Modern biomarkers are defined as ***a biological molecule found in blood, tissue or other bodily fluid which is indicative of either normal bodily processes or of a disease/medical condition***[39]. Biomarkers have been recently utilised by the science and healthcare industries as key tools to identify illnesses and help to guide treatments [40–42]. One of the most successful biomarkers identified is in the field of cancer research. PD-1/PD-1L is used as a biomarker to predict patients who are likely to respond to PD-1 blockage by the monoclonal antibody pembrolizumab [43].

To date there are several potential biomarkers to predict sepsis and associated infections and significant amounts of research are underway to improve the panel of biomarkers in use [44–48]. Given the multiple failures in finding an effective treatment, sepsis has proven itself an incredibly complicated disease to predict. A potential reason for this complexity is the sheer number of host cells and pathways that are affected during sepsis infections.

Increasing our understanding of the underlying molecular mechanisms of sepsis continually provides potential therapeutic targets and brings into light the complex heterogenous nature of sepsis. Current and potential biomarkers for sepsis will be discussed in detail later in the thesis (Chapter 4).

1.1.3 Risk factors for sepsis

There are several risk factors that can increase the risk of developing sepsis. The main risk factors and groups at risk according to the National Institute for Health and Care Excellence (NICE)[49], are shown below (Table 1.2).

Table 1.2: Risk factors associated with the development of Sepsis [49,50]

Risk Factor	Description
Age	Age over 75 or under 1 year (neonates)
Impaired immune system	Chemotherapy, impaired immune function due to illness e.g., diabetes, sickle cell disease, cancer, long term steroid use e.g., Rheumatoid arthritis
Recent surgery	Risk of infection at site of surgery, spreading to other parts of body/blood
Breach of skin integrity	Cuts, burns
Indwelling lines	Catheters etc. Can lead to bacterial growth on indwelling surface leading to infection.
IV drugs	Misuse of drugs leading to infection.
Hospital admission	Hospital admission within previous 28 days
Neutropenia	<500/µL at time of BSI
Antimicrobial therapy	Antimicrobial therapy within previous 28 days

1.1.4 Sepsis-associated pathogens

It is important to note that over the last 100 years there have been clear changing aetiologies in the groups of pathogenic bacteria causing sepsis, particularly in hospital

acquired sepsis (nosocomial). Before the advent of antibiotics Gram-positive bacteria were the most common cause of sepsis, particularly *Streptococcus pyogenes* and *Staphylococcus aureus* [51] Since antibiotic use became widespread (late 1940s and beyond) the amount of sepsis cases caused by Gram-positive bacteria began to fall between 1950-1980 and Gram-negative bacteria where the cause of sepsis in more than 50% of cases [51] . In contrast, for the years 1997-2000 Gram-negative and Gram-positive bacteria accounted for 30.5% and 39.7%, of total blood stream infections respectively. Recent reports (between 2013-2016) have indicated that Gram-negative bacteria account for 43.4% of total cases, while Gram-positive bacteria account for 33.1% of the top ten pathogens causing blood stream infections worldwide. This has led to Gram-negative bacteria having a greater association with sepsis and the development of the endotoxin theory of sepsis[52]. Indeed, animal models used to study sepsis infections often involve the injection of Lipopolysaccharide (LPS), a component of Gram-negative bacterial cell walls. Mice deficient in the receptor which recognises LPS, TLR-4, are hyporesponsive to LPS[53], which is known to cause severe inflammatory responses in mammalian hosts through the activation of TLR4 and subsequent activation of Nf-κβ [52,54].

The microorganisms that can cause sepsis are diverse at domain level, including both Gram-positive and Gram-negative bacteria, viruses, and fungi [55]. Sepsis cases are predominantly (≈80%) caused by bacterial infections [56]. In patients presenting as blood culture positive, Gram-negative, and Gram-positive bacteria represent 62.2% and 46.8% of cases respectively [57]. These numbers overlap due to the occurrence of polymicrobial sepsis, where more than one underlying organism is positive on blood culture. Although there is now increasing incidences of Gram-negative sepsis, Gram-positive bacteria are still a major cause. Indeed, many Gram-positive bacteria make the ‘Top 10’ lists of most common causes of bacteraemia. Among the top 5 bacteria found in blood infections in Wales are the Gram-positive organisms *Staphylococcus aureus* and *Streptococcus pneumoniae* (Table 1.3). Similar species have also been found to be the leading causes across the world with American and Chinese studies indicating a similar top 10 leading organisms [58,59].

In addition, Gram-positive bacteria are the dominant pathogens in sepsis where the initial source was from a bone/soft tissue infection (55.7%) or a cardiovascular infection (52.5%),

whereas Gram-negative bacteria are the leading cause of sepsis from a urinary tract infection (72%) and abdominal infection (38.4%) [60].

Although Gram-positive bacteria lack LPS they do have inflammatory PAMPS that can elicit damaging responses from the host immune system. The cell wall of Gram-positive bacteria comprises a thick layer of peptidoglycan, a structure comprised of glycan crosslinked by short peptides[61]. Peptidoglycan is recognised by multiple host receptors including, NOD receptors (NOD1 and NOD2), peptidoglycan recognition proteins (PGRP1) and toll like receptors (TLR2)[61,62]. Additionally, bacterial cell walls also contain lipoteichoic acid which can act as a major virulence factor and cause inflammation through the activation of TLR2. Both peptidoglycan and lipoteichoic acid have been shown to induce cytokines associated with sepsis (TNFα, IL-1β and IL-6)[63].

These studies reiterate the fact that sepsis as not just a problem associated with an uncontrolled host immune response to LPS.

Table 1.3: Top ten species causing bacteraemia in Wales for 2017. Data from Public Health Wales[64]

Organism	**Rate per 100,000 bed days**
Escherichia coli	81
Staphylococcus aureus	23
Klebsiella species	17
Enterococcus species	15
Streptococcus pneumoniae	14
Coagulase-negative *Staphylococcus*	13
Proteus species	7
Streptococcus viridians group	6
Pseudomonas aeruginosa	6
Streptococcus group	6

1.1.4 Origins of infection

Sepsis and bacteraemia are often secondary to an original site of infection. Sepsis can develop from many different sources of infection but the most common infections leading to sepsis are, respiratory infections (31.3%), abdominal infections (26.4%), urinary tract infections (18.4%), bone/soft tissue infections (13.3%) and cardiovascular infections (3.2%) [60]. Origin of infection is largely dependent on the causative organism and certain species are more likely to be associated with certain sources of infection. Species such as *S. pneumoniae* and Methicillin-sensitive *Staphylococcus aureus* (MSSA) are abundant in respiratory infections but are rarely found in the urinary tract, while *E. coli* is strongly associated with urinary tract and abdominal infections but is rarely found in respiratory infections [60].

1.2 *Escherichia coli*

1.2.1 Background

Escherichia coli (*E. coli*) is a highly versatile Gram-negative bacteria, capable of surviving a wide range of ecological niches [65]. *E. coli* is one of the first microorganisms to colonise the mammalian gut (within the first hours of life) where it establishes itself as the most abundant facultative anaerobe in a healthy microbiota [66]. As such, colonisation with *E. coli* does not usually cause adverse effects in the host, due to a symbiotic relationship with the host gut immune system; however, under circumstances which cause a breakdown in the healthy relationship between host and bacteria (e.g., in immunocompromised individuals), strains typically considered commensals can become opportunistic pathogens [67,68].

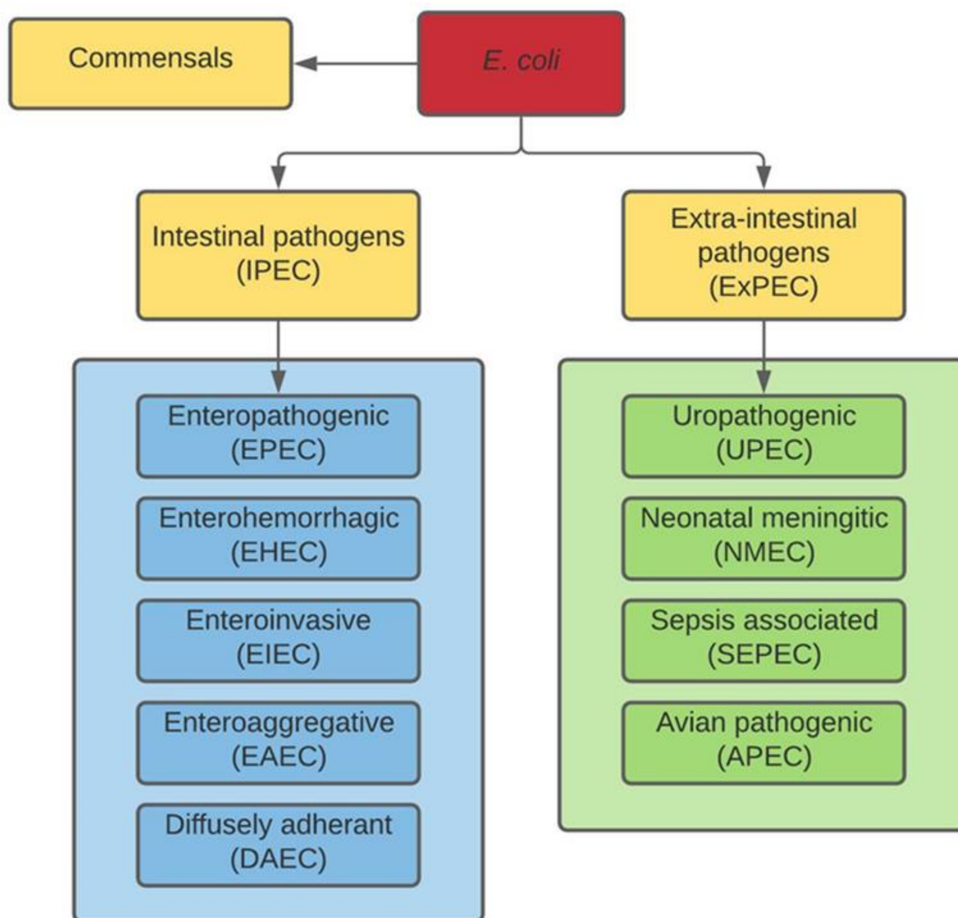


Figure 1.1: Pathotypes of *E. coli*. *E. coli* can be commensal of pathogenic, where pathogenic strains are defined by the site of infection as intestinal pathogenic *E. coli* (IPEC) and extraintestinal pathogenic *E. coli* (ExPEC). Adapted from [69]

E. coli strains can be considered commensal or pathogenic organisms (Figure 1.1). Human pathogenic strains of *E. coli* can be grouped based on the site of infection; these are defined as either intestinal pathogenic *E. coli* (IPEC) or extraintestinal pathogenic *E. coli* (ExPEC) [68]. The ExPEC strains are responsible for diseases such as sepsis and meningitis and can be further categorised into pathotypes based on the clinical features of disease, the most common of which are uropathogenic (UPEC), meningitis-associated *E. coli* (MNEC), septicaemia-associated *E. coli* (SEPEC) and avian pathogenic *E. coli* (APEC) (Figure 1.1) [68].

Genetics plays an essential role in the ability of *E. coli* to survive in a wide range of ecological niches. The genomes of *E. coli* contain a 'core genome' and an 'accessory genome'. The core genome contains genes involved in essential processes and ensure cellular survival (e.g., metabolic genes). The flexibility of the *E. coli* genome is evidenced by the fact that different *E. coli* isolates may only share 60-70% of their genes [70]. The accessory genome is a set of non-essential genes acquired by different bacterial strains through the processes of horizontal gene transfer and homologous recombination, that express virulence and survival factors often encoded on mobile genetic elements known as pathogenicity islands. The accessory genome of *E. coli* is known to contain a multitude of genes encoding virulence / survival factors and in resistance to antibiotics and evasion of the host immune system (Table 1.4). This genetic adaptability of *E. coli* also allows the acquisition of virulence and survival factors by commensal strains [71]. This is due to *E. coli* virulence genes often being located on transmissible genetic elements such as genomic islands and plasmids, which can be easily exchanged with other bacteria [72]. There is a plethora of *E. coli* virulence factors, which are often organised by their functional interaction with the host. These include Adhesins (cell binding), Protectins and Invasins (cell entry), Iron acquisition systems (resistance to iron), Toxins and 'Miscellaneous factors' (Table 1.4) [73].

The ECOR collection of *E. coli* was established in 1984 by Ochman and Selander to represent the genetic diversity of the species [74]. In this work they established 5 main phylogroups of *E. coli* namely, A, B1, B2, D and E. This remains true today, although, two additional phylogroups are now recognised (termed E and F) increasing the total number of phylogroups to 7 [75]. There is a broad relationship between the pathotype of *E. coli* and the phylogroup; most of the commensal strains of *E. coli* belong to the A phylogroup and the majority of the ExPEC strains belong to either B2 or D. In addition to organising *E. coli* isolates by pathotype they can also be distinguished based on the presence of certain antigens on their surface. These antigens are known as K, O and H antigens. *E. coli* has more than 80 K antigens which can combine in any way with the possible 187 O and 53 H antigens [76,77], highlighting the amount of variability present in the species.

Table 1.4: Virulence factors expressed by ExPEC *E. coli* strains ExPEC virulence factors can be grouped into categories according to their function. Adapted from [73]

Virulence factor associated with ExPEC	Gene(s)
ADHESINS	
Adhesion siderophore	*iha*
Dr binding adhesins	*afa/draBC*
E. coli common pilus	*ecpA*
F1C fimbriae	*foc* gene cluster
Heat-resistant haemagglutinin	*hra*
M fimbriae	*bmaE*
N-acetyl d-glucosamine-specific fimbriae	*gaf*
P fimbriae	*papACEFG*
S fimbriae	*sfa/sfaS*
Temperature sensitive haemagglutinin	*tsh*
Type 1 fimbriae	*fimH*
Adhesin and siderophore, bifunctional gene	*iha*
Invasion associated locus	*ial*
IRON ACQUISITION SYSTEMS	
Aerobactin receptor	*iutA*
Peri-plasmic iron binding protein	*sitA*
Salmochelin receptor	*iroN*
Siderophore receptor	*ireA*
Yersiniabactin receptor	*fyuA*
Yersiniabactin biosynthetic protein	*irp2*
Iron transporter	*feoB*
PROTECTINS AND INVASINS	
Colicin V	*cva*
Conjugal transfer surface exclusion protein	*traT*
Group 3 capsule	*kpsMT* II
Increased serum survival	*iss*
Invasion of brain endothelium	*ibeA*

K1/K2/K5 group 2 capsule variants	K1/K2/K5 genes
kpsM II group 2 capsule	*kpsM* II
Outer membrane protease T	*ompT*
TOXINS	
α-haemolysin	*hylD*
Cytolethal distending toxin	*cdtB*
Cytotoxic necrotising factor	*cnf1*
Enteroaggregative *E. coli* toxin	*astA*
Haemolysin A	*hylA*
Secreted autotransporter toxin	*sat*
Serine protease	*pic*
Vacuolating toxin	*vat*
enteroaggregative E. coli heat-stable enterotoxin (EAST1)	*astA*
Shiga toxins	*stx*
MISCELLANEOUS/OTHERS	
β-glucoronidase	*uidA*
Colibactin synthesis	*clb* & *clbB*
Uropathogenic-specific protein	*usp*
Flagellin variant	H7 *fliC*
Maltose and glucose-specific PTS transporter subunit IICB	*malX*
Pathogenicity island marker	*malX*
d-serine deaminase	*DsdA*
capsule synthesis	*kpsMTII/III*

1.2.2 *E. coli* virulence factors

E. coli utilise numerous mechanisms that aid the bacteria in subverting host responses. This can occur in a number of ways from avoidance of immune system through modification of PAMPs e.g. modification of LPS (thereby reducing recognition by PRRs), to secretion of proteases which can directly eliminate AMPs[78].

Additionally, virulence factors expresses by *E. coli* can also influence host responses by either promoting bacterial survival or in some cases manipulating host immune cells e.g. Toxin secretion. Common ExPEC virulence factor types are discussed below.

1.2.2.1 Adhesins

Adhesins are proteins found on the surface of pathogens which aid in surface attachment and colonisation. Surfaces can be either, biological (e.g., host cells), or non-biological (e.g., catheters). Adhesins facilitate the binding of pathogens to cellular surfaces during infection through specific interactions. Adhesins are located on fimbriae (pili), hair like appendages located on bacterial surfaces, or non-fimbrial adhesins. *E. coli* have numerous adhesins that aid in the attachment of bacteria to host surface structures, including the mucosal layer in the gut or other surfaces including the respiratory, intestinal, or urinary tract. Perhaps the most commonly identified adhesin is the Type 1 fimbrial adhesin encoded by the *fim* gene cluster. The type 1 fimbrial adhesin is expressed by 90% of uropathogenic *E. coli* isolates and research has shown it is crucial in establishing urinary tract infections in mice [79,80]. The type 1 fimbrial adhesin binds to mannose on host cells facilitating entry and attachment to epithelial cells [81].

1.2.2.2 Protectins and invasins

Protectins are proteins found on the surface of bacteria that aid in survival e.g., the K antigens on *E. coli,* which aid in the resistance to phagocytosis. The K antigens enable the bacteria to avoid the deposition of complement proteins C3b and C3d onto their cell surface and thus avoid recognition by phagocytic cells. Van Dijk et al 1979 found that heat inactivation of K antigens restored the ability of C3 proteins to bind to the bacterial cell surface [82].

Invasins are a class of proteins that are involved in the entry of pathogenic bacteria into host cells. Interestingly, the K antigen variant K1 has been shown to be essential for *E. coli* to cross the blood brain barrier and cause meningitis through the invasion of human brain microvascular endothelial cells (HBMECs). In addition, the adhesin protein *fimH* has been shown to facilitate invasion of ExPEC isolates across the intestinal epithelium of immunocompromised individuals [83].

1.2.2.3 Iron acquisition

Iron is essential for nearly all living organisms and iron dependent catabolism plays a role in a multitude of bacterial reactions including energy metabolism, respiration, lipid metabolism, amino acid synthesis and DNA metabolism [84]. Due to the essential nature of iron in bacterial metabolic reactions the withholding and/or removal of iron from the environment is an essential part of host nutritional immunity against pathogenic bacteria [85]. To combat the sequestration of iron by the host, bacteria have evolved a myriad of iron acquisition systems (Table 1.4). These include metallophores, siderophores, uptake pumps, and xenosiderophores [84]. Metallophores are secondary metabolites that enable bacteria to sequester and store metal ions from the extracellular environment, aiding in growth, survival and virulence [86]. An example is the use of the yersiniabactin receptor by UPEC *E. coli* isolates to bind copper and resist its toxic effects [87]. Siderophores are a subset of metallophores and the most widely studied. *E. coli* produce four types of siderophores; enterobactin, salmochelin, yersiniabactin and aerobactin, all of which have roles in sequestering iron away from host proteins ferritin, transferrin and lactoferrin [88].

1.2.2.4 Toxins

Bacterial toxins are potent molecules, usually secreted, that are significant factors in determining bacterial virulence. They have direct impact on host cells, which can often result in toxicity, and can manipulate the host response to infection, aiding the survival of the pathogen[89]. Common toxins that are important in ExPEC isolates include cytotoxic necrotizing factor (*cnf*), cytolethal distending toxin (*cdt*), vacuolating autotransporter toxin (*vat*) and secreted autotransporter toxin (*sat*). The toxin cnf is associated with ExPEC strains that cause meningitis as it contributes to the invasion of brain endothelial cells by *E. coli* isolates possessing the K1 capsule in neonatal meningitis models [90]. It also causes damage to the DNA of target cells. The toxin can bind to multiple cell types through cholesterol moieties and impair host defence in three main ways; i) inducing cell cycle arrest in epithelial cells, ii) induce the apoptosis of lymphocytes and iii) inducing a pro-inflammatory state in macrophages [91]. Vacuolating autotransporter toxin (*vat)* is a serine protease which has been identified in both APEC and UPEC *E. coli* strains. It induces the formation of vacuoles in target cells such as the bladder epithelial cells and causes cell death [92].

Secreted autotransporter toxin (*sat*) has been shown to induce cell death in kidney and bladder cells [93]. A more recent study has shown that sat is able to increase survival of ExPEC strains by cleaving three proteins of the complement cascade, thereby aiding in immune evasion [94].

Another toxin of relevance to ExPEC is the human genotoxin, colibactin which is encoded by the *pks* island. Colibactin is a hybrid polyketide-non-ribosomal peptide product encoded mainly on *E. coli* belonging to the B2 phylogroup, which has been linked to a DNA mutation in host cells leading to colorectal cancer [95].Colibactin is strongly associated with bacteraemia causing isolates [96].

1.2.3 *E. coli* and antimicrobial resistance (AMR)

Antimicrobial resistance (AMR) is a serious concern for global health [97–99]. Bacteria are becoming increasingly resistant to antibiotic therapies. A recent report published in the Lancet has highlighted the intense burden AMR has on healthcare systems worldwide. The authors conducted a rigorous study estimating the burden of AMR globally for 2019. They found that AMR resistance was associated with 4.95 million deaths and attributed to 1.27 million deaths. Authors also suggested that burdens were disproportionally distributed, with lower income countries bearing the largest burden [100]. The organism identified as being associated was the most deaths related to AMR was *E. coli,* followed by *Staphylococcus aureus, Klebsiella pnemoniae, Streptococcus pneumoniae, Acinetobacter baumannii* and *Pseudomonas aeruginisa* [100]. This is consistent with the leading causative agents of bacteraemia in South Wales (Table 1.1). Interestingly the pathogens causing the most AMR related deaths differed based on the income levels of the country. In the high-income countries group *E. coli* and *S. aureus* were the leading causative organisms with more than half of attributable deaths associated with these two pathogens. In contrast, in Sub-Saharan Africa the leading causative pathogens were *S. pneumoniae* and *K. pneumoniae* and had a smaller share of the burden combined [100].

Prospective studies have also highlighted the increasing incidence of AMR and reports indicate that by 2030 over half of *E. coli* isolated could become resistant to third generation cephalosporins (the broad-spectrum antibiotics used to treat Gram-negative bacterial infections) [101]. Public Health England have also published concerning data on the

increase of AMR in *E. coli* in 2018, where they found increasing resistance to several antibiotics, termed multidrug resistance (MDR). This included resistance to multiple classes of antibiotics including ceftazidime, cefotaxime, ciprofloxacin and gentamicin [102]. Given the propensity of *E. coli* to cause UTI's, it is also of concern that resistance to the first line antibiotics Nitrofurantoin or Trimethoprim is high, as shown in a study in Scotland of 7,845 patients with *E. coli* bacteraemia. Resistance rates to nitrofurantoin were the second highest of tested antibiotics with a 44.9% of isolates being resistance. Resistance rates to trimethoprim were lower at 9.8%, however, mortality rates were reported as being higher in patients with trimethoprim resistant infections[103].

The increase in the *E. coli* bacteraemia isolates expressing MDR[104,105], is worrying, as treatment for *E. coli* sepsis relies on antimicrobial therapies. A better understanding of disease pathology and host responses to infection are crucial in helping to guide treatment options and identifying potential markers that can guide treatment.

1.2.4 *E. coli* in sepsis

E. coli is a leading cause of bloodstream, infections worldwide accounting for 20% of cases and accounting for 17% of all sepsis related deaths in the United States[106]. In England, *E. coli* bacteraemia is the leading cause of sepsis with a mortality rate of 18%[107].

In South Wales it is the also the leading cause of bacteraemia. In part, this is due to an increased sepsis awareness programme in the area [108]. The *E. coli* phylotype groups B2 and D are the ones most associated with extraintestinal infections and multiple studies have highlighted their prevalence during sepsis infections compared to the other phylogroups. [109–111]. However, it is important to note that due to the plasticity of the *E. coli* genome any phylogroup of *E. coli* can cause extraintestinal disease. Indeed, in a recent study by Zhuge et al, the authors showed that an avian pathogenic *E. coli* (APEC) had similar virulence gene compositions compared to extraintestinal pathogenic *E. coli* causing disease in humans and was able to cause disease in both a mouse and rat model of sepsis infection [112]. Furthermore, as sepsis infections are associated with the host response, infection of an immunocompromised host can lead to sepsis infections in response to commensal bacteria, as a form of opportunistic infection [113]. This heterogenous nature of *E. coli* infections

combined with increasing antibiotic resistance of *E. coli* isolates found throughout the literature and through government surveillance studies can make treatment difficult [114].

Bloodstream infections associated with sepsis often arise from the ability of *E. coli* to cause infection at local sites [60]. The most common infections caused by *E. coli* infections are urinary tract infections [114]. These infections are often recurrent, and with the increasing incidence of antimicrobial resistance within *E. coli* remain a risk for sepsis development.

E. coli along with other *Enterobacteriaceae* are also associated with various acute intra-abdominal infections e.g., appendicitis, peritonitis, and cholecystitis [115,116]. Although fewer in number, compared to other pathogens, *E. coli* can also cause bone and soft tissue infections as well respiratory and cardiovascular infections, all of which can be the origin of a sepsis infection[60]. This diverse number of infection sites makes identifying common genes associated with sepsis difficult. Moreover, genes associated with a certain tissue site, may also affect virulence in another. For example, the gene *fimH* has been shown to be important in the attachment of *E. coli* UPEC isolates to urinary epithelial cells [117]. At the same time *fimH* has also been shown to play a role in translocation from the intestine to the kidneys, spleen, and lungs [118].

1.3 The immune system and the host response

The immune system is organised into three lines of defence; i) passive innate immunity and barriers (section 1.3.1); ii) induced innate immunity (section 1.3.2); and iii) adaptive immunity (section 1.3.3). In addition, there are cellular products, (e.g., cytokines), processes (e.g., cell death) and protein cascades (e.g., complement and coagulation) that contribute to immune mechanisms and maintain physiological homeostasis.

1.3.1. Passive innate immunity and barriers

Barriers are the first line of defence and include chemical, microbiological, and physical barriers. Perhaps the first and most basic line of defence against invading pathogens are physical barriers to infection. These barriers prevent invasion of pathogens into bodily sites. The skin and the mucosal membranes that line the gastrointestinal tract as well as epithelial layers of the respiratory system and urinary system comprise these barriers to entry. The skin epidermal surface provides protection though keratinized cells as well as Langerhans cells (tissue resident macrophages) which can recognise and kill pathogens and present antigens to cells of the adaptive immune system. Sweat secretions from the skin also act as an antimicrobial agent with its low pH and a washing action. Mucosal barriers including the gastrointestinal tract have a low pH unsuitable for many potential pathogens as well as containing a plethora of antimicrobial agents such as antimicrobial peptides e.g. defensins. Epithelial cells of the gut detect pathogens and signal to immune cells as well as preventing the translocation of pathogens to other tissue sites. They elicit a carefully controlled immune response, which is heavily regulated and inductive towards tolerance, is employed to elicit immune responses to potential invading pathogens, whilst maintaining homeostasis with the gut resident microbes, which themselves can also act as a barrier by preventing pathogen colonisation [119,120].

1.3.2 Induced innate immunity

The innate immune system is often described as the second line of defence against invading pathogens. The role of the innate immune system is to recognise and respond to foreign material and prevent its spread throughout the body by segregating it away for destruction. To do this the innate immune system relies on pattern recognition of microbes and ‘danger’ or damage associated alarm molecules (Figure 1.2), together with the phagocytic action of

monocytes and neutrophils. Therefore, pattern recognition leads to activation of inflammatory genes, leading to removal of the pathogen by phagocytosis or direct cellular killing.

1.3.2.1 Pathogen associated molecular patterns (PAMPS)

Pathogens express molecules that are conserved between organisms (e.g., LPS on Gram-negative bacteria), called pathogen associated molecular patterns (PAMPS) which can be recognised by innate immune cells through specialised receptors. Innate recognition of pathogens by immune cells relies on a set of germline-encoded pattern recognition receptor (PRRs) which recognise conserved PAMPs. There are several classes of PRRs including Toll-like receptors (TLRs), Nod-like receptors (NLRs), C-type lectin receptors (CLRs) and intracellular DNA sensors (e.g., cGAS). TLRs were the first to be discovered and are the most widely studied and understood.

In humans there are 10 members of the TLR family, which recognise different PAMPs. TLR proteins span membranes, either the outer membrane of the cell or internal membranes (e.g., lysosomes, phagosomes). Those found on internal membranes recognise DNA and RNA and are important for the recognition of intracellular pathogens. Each TLR is composed of three domains; the ectodomain which mediates PAMP recognition, the transmembrane domain, and a cytoplasmic Toll/IL-1 receptor (TIR) domain which is involved in signalling. Upon pathogen recognition by TLR and co-receptors (e.g., CD14), adaptor proteins which contain a TIR domain are recruited and initiate downstream signalling pathways, which ultimately lead to the activation of transcription factors, NF-κβ and IRFs which cause increased expression of inflammatory genes and type one interferons respectively[121]. (Figure 1.2).

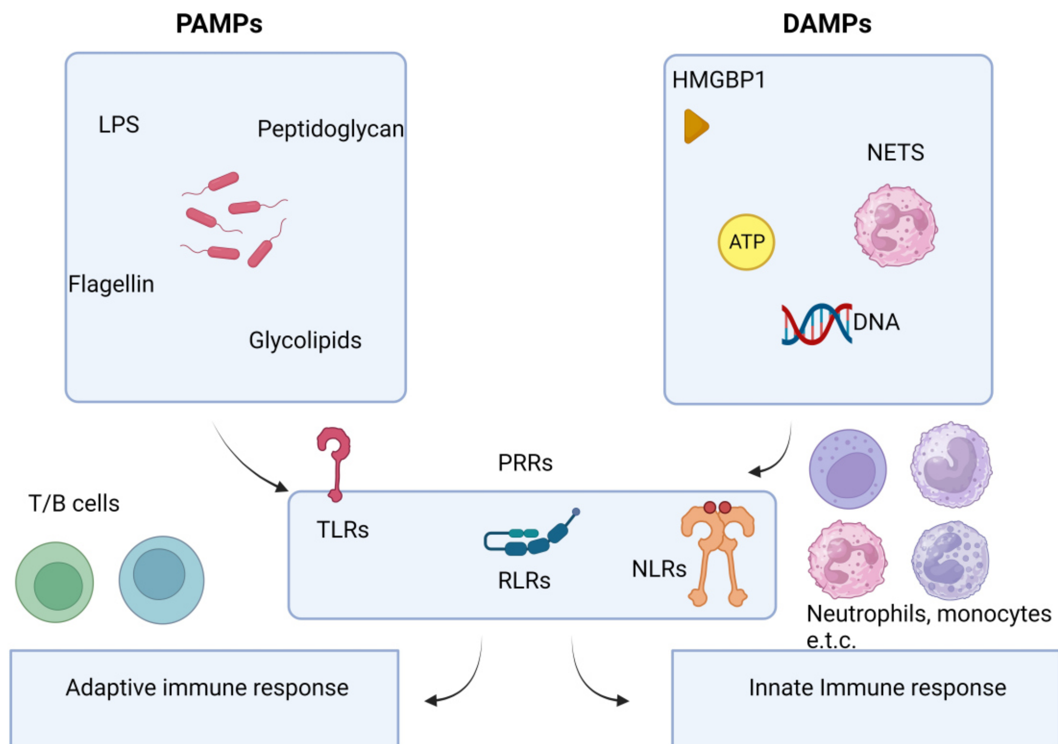


Figure 1.2: PAMPs and DAMPs. Created with BioRender.com and adapted from [122]

1.3.2.2 Damage associated molecular patterns (DAMPs)

Damage associated molecular patterns or DAMPs are host derived molecules able to trigger an immune response in the absence of infection. DAMPs are intracellular molecules, such as cytosolic, nuclear, or mitochondrial proteins, and thus are usually shielded from immune recognition. After tissue injury (e. g., infection) DAMPs are released and can be recognised by cells of the immune system [123]. High mobility group box protein 1 (HMGB1) is a protein which is highly conserved and expressed in all mammalian cells. It can act as a DAMP when it is released through cytoplasmic vesicles or from dying cells. HMGB1 activates the immune system by recruiting neutrophils to the site of injury [124], as well as binding to receptors and triggering the activation of macrophages and epithelial cells to produce inflammatory cytokines.

1.3.2.3 Complement

Complement is a collection of more than 30 proteins present in blood, either soluble or bound to membranes [125]. Complement plays a major role in the early response to infection and participates in the direct killing of pathogenic organisms and, along with antibodies, participates in the opsonisation of pathogens for phagocytosis by phagocytic cells[126,127]. There are three pathways activated in different ways but come together terminally to form the membrane attack complex (MAC). The pathways of complement are the classical pathway, the alternative pathway, and the lectin pathway (Figure 1.3). The classical pathway is activated by C1q binding either directly to a pathogen surface or binding to IgM and IgG antibodies on pathogen cell surfaces. The alternative pathway is activated by the deposition of C3 directly onto pathogen surfaces. The lectin pathway, as the name suggests, is initiated by proteins that recognise lectins on pathogen cell surfaces [128]. All the pathways converge at the activation of the protein C3 and the subsequent production of the anaphylatoxins C3a and C5a, which are chemoattractant to phagocytes, induce the expression of adhesion molecules, and increase the permeability of blood vessels [128].

The general effects of the cascade results in the activation of three functional effector pathways which aid in defence against pathogens, namely, inflammation through chemotaxis, exclusion by phagocytosis (opsonisation), and killing through the membrane attack complex (MAC). All pathways terminate with the formation of a membrane attack complex (MAC), a collection of bound complement proteins which have direct lytic action on pathogenic cells through the formation of pores in the cell membrane (Figure 1.3).

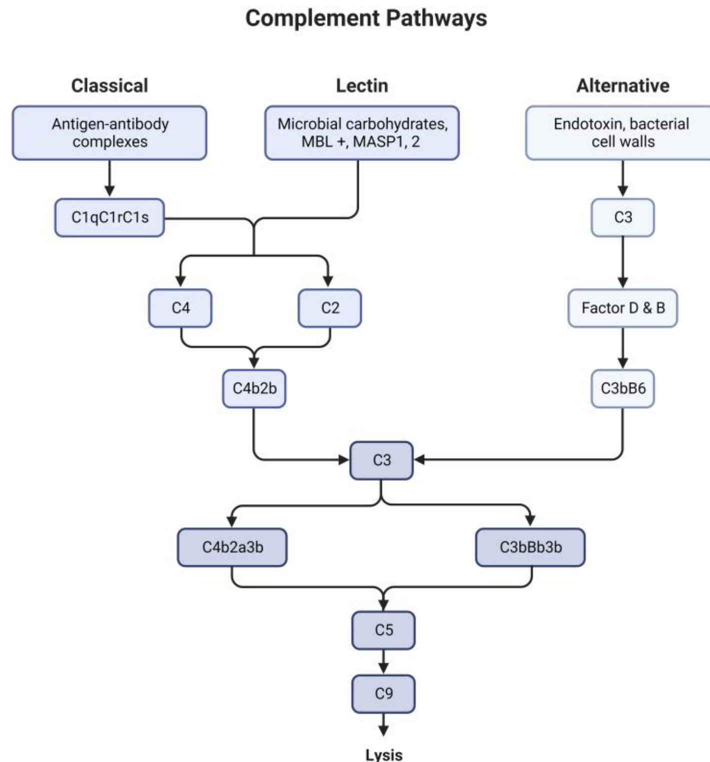


Figure 1.3: Complement activation

The three pathways of complement activation the classical pathway, the lectin pathway and the alternative pathway. All pathways converge with the attachment of C3 on pathogen cell surface. Created with BioRender.

The complement system has evolved to combat many different pathogens due to their role as soluble pattern recognition receptors[129]. Knock out studies for specific proteins of the cascade do not cause susceptibility to all pathogens, for example, it is well documented that deficiencies in the terminal complement proteins (those that form the MAC) only increase susceptibility to infections by *Neisseria* species [130–132].

1.3.2.4 Neutrophils

Neutrophils are the most abundant leukocytes in the body comprising 40-70% of all white blood cells [133]. They are granulocytic cells formed in the bone marrow from a myeloid precursor and circulate in the blood after exiting the bone marrow. Neutrophils are recruited to the site of injury/infection within minutes following chemotactic signals

released at the site of infection (e.g., IL-8 and C5a)[134]. Neutrophils migrate through blood vessels via interactions between cell surface selectins and integrins and enter the site of infection where they act to further increase the local immune response through the release of inflammatory cytokines as well as direct killing of microbes through; phagocytosis, degranulation, and the release of reactive oxygen species (ROS [135].Neutrophils are usually short-lived cells which undergo constitutive apoptosis after the clearing of infection. Dying neutrophils are recognised by monocytes and macrophages, stimulating phagocytosis. This process has an anti-inflammatory effect on phagocytes, suppressing pro-inflammatory cytokines and releasing immunosuppressive cytokines such as IL-10. This clearing of dead cells is a critical part of the resolution of infection [136,137]. Neutrophils also have the capacity to undergo a specialised form of cell death called NETosis in which they release chromatin and DNA into the extracellular matrix. This chromatin and DNA is first condensed and then mixed with antimicrobial peptides before being released into the extracellular matrix. The formation of these web-like structures into the extracellular matrix can then trap and kill pathogens [138].

1.3.2.5 Macrophages and monocytes

Monocytes circulate in the blood and through the spleen, and are specially adapted to ingest foreign particles, dead cells and debris and initiate inflammatory responses via the release of cytokines such as tumour necrosis factor alpha (TNFα). During infection monocytes are recruited to the infection site through CCL2/MCP-1 released from epithelial cells and neutrophils, here they differentiate into macrophages and undertake phagocytosis of foreign materials [139,140].

Tissue macrophages are acquired before birth and are long lived, self-renewing cells and their maintenance is not dependent on migration of monocytes from the bone marrow via haematopoiesis [141]. Tissue derived macrophages differentiate in tissues in response to cytokine stimuli from the surrounding environment, which generates a large tissue diversity in macrophage phenotypes. Broadly speaking macrophages can be categorised into three main types, M0, M1 and M2 macrophages. M0 macrophages are non-activated, while M1 macrophages are associated with inflammatory responses and the release of pro-inflammatory cytokines, M2 macrophages are characterised by their anti-inflammatory

nature [142]. The polarising conditions for M1 and M2 macrophages are IFN-γ, LPS or TNFα and IL-4, IL-13 respectively [139].

Phagocytosis is a cellular process of ingesting particles larger than 0.5µm in diameter and there are multiple mechanisms by which this can occur either through direct pattern recognition on pathogen cell surfaces via non-opsonic receptors (e.g., C-type lectin receptors) [143], or via the binding of opsonin's attached to a pathogen cell surfaces (e.g. complement or antibody via complement and Fc receptors respectively) [143]. All pathways lead to formation of a phagosome and eventual fusion with a lysosome, leading to the degradation of the pathogen into biomolecules e.g., peptides, lipids, and nucleic acids.

1.3.2.6 Cytokines and the acute phase response

Cytokines are small, secreted intercellular signalling molecules that contribute to inflammation. After pattern recognition by host immune cells, receptor signalling leads to the activation of transcription factors such as NF-κβ which stimulate the production of cytokines such as TNFα, IL-1β, IL-6 and IL-8. Cytokines can be released by a variety of different cell types including cells of the immune system e.g. T cells and macrophages as well as by other cell types such as epithelial cells, stromal cells and fibroblasts[144]. Cytokines are released from these cells and can act on cells of the immune system to stimulate or inhibit immunoactivity [145]. Cytokines can be divided into many categories including chemokines, interleukins, interferons, colony stimulating factors and transforming growth factors, which can have a wide range of effects. Cytokines can also be grouped according to function, those that act in a way to induce inflammation and those that suppress inflammation. Early acting pro-inflammatory cytokines such as TNFα and IL-1β increase early during infection, followed by later acting cytokines IL-6, IL-1Ra and IL-8. (Figure 1.3) and finally anti-inflammatory cytokines such as IL-10. Historically this concurrent release of pro-inflammatory cytokines was termed a cytokine storm [145] [137].

The acute phase response is part of the innate immune system's response to infection / trauma, which involves a systemic response of the body mediated by the cytokines TNFα and IL-6. Within a few hours of infection protein synthesis in the liver is upregulated in response to cytokine stimulation, due to production of proteins which are part of the acute phase response [146]. The acute phase comprises of approximately 200 proteins which are

mainly produced in the liver by hepatocytes, through the release of pro-inflammatory cytokines by monocytes, endothelial cells, and tissue resident macrophages, which are also able to produce some acute phase proteins [147].

Acute phase proteins can be either positive reactants or negative reactants, which increase and decrease in amount during the response respectively. Many acute phase proteins show a predictable increase during infection/injury and are measured as biomarkers of acute illness [148]. Examples of acute phase proteins include C-reactive protein, Serum amyloid A and complement proteins C3 and C4 [146].

1.3.2.7 Coagulation

Coagulation is a process that aims to stop bleeding through the formation of blood clots after tissue damage. It is comprised of three pathways, the intrinsic, extrinsic, and common pathways, that interact to form a stable blood clot. The intrinsic pathway involves the recognition of collagen, kallikrein and high molecular weight kininogen (HMWK) by Factor XII (Hageman Factor), which triggers downstream reactions and feeds into the common pathway through the activation of Factor IXa. The extrinsic pathway is initiated by damage to endothelial tissue, exposing tissue Factor (Factor III) to the blood. This then activates Factor X and feeds into the common pathway. The common pathway is activated by Factor X from either the intrinsic or extrinsic pathways and results in the cleavage of fibrinogen into fibrin, the activation of fibrin leads to crosslinks, which captures platelets and blood cells to form a clot. Coagulation has a role in the host immune response to infection through activating a variety of responses including direct bactericidal activity of blood clots through confining the bacteria inside the clot, restricting nutrient and water supply[149]. Blood clot retraction also cause pressure that can tear bacterial capsules[149]. In addition, compression of the blood clot can lead to the release of oxygen from erythrocytes, leading to killing of bacteria through oxidation[149].

Many pathogens have evolved mechanisms to modulate coagulation; the causative agent of plague, *Yersinia pestis* has mechanisms which disrupt clot formation and prevent trapping of the pathogen via fibrin/fibrinogen[150]. The bacteria possess a plasminogen activator which when absent from the bacteria leads to increased survival due to a more controlled

infection, through the greater infiltration of immune cells at site of infection and increased bacterial clearance preventing dissemination [150].

1.3.2.8 Cell death and immunity

Cellular death is a natural and critical process to maintaining homeostasis. Cell death was initially divided into three types: Type 1, (apoptosis), type II (autophagy) and type III (necrosis) [151]. There are now more than ten recognised mechanisms of cellular death: Autophagy, Entosis, methuosis, paraptosis, mitoptosis, parthanatos, ferroptosis, pyroptosis, NETosis, Necroptosis, Apoptosis, Anoikis and necrosis. These can be organised into three major categories; Programmed non-apoptotic cell death, programmed apoptotic cell death and non-programmed cell death [151]. These mechanisms function in different ways with programmed cell death initiated under the control of regulated signal transduction pathways, with apoptotic pathways involving the retention of membrane stability and caspase dependency and non-apoptotic pathways being associated with a rupture in cellular membranes and being caspase independent. Non-programmed cell death in contrast is stimulated by factors such as infection and injury [151]. Programmed cell death mechanisms contribute to a variety of immune processes that protect the body against infection. It is involved in immunity to viruses and other intracellular pathogens, inflammation, chemotaxis as well as roles in the adaptive immune response on the form of lymphocyte selection and immune tolerance mechanisms [152–154]. Critically, as individual pathogens are able to induce different cell death modalities it is clear that understanding death mechanisms may give insight into markers and therapies for disease [155].

1.3.3 Adaptive immunity

The third line of defence is the adaptive immune system and is comprised of three major cell types: T lymphocytes and B lymphocytes and dendritic cells. Thymus derived lymphocytes or T cells are critical components of the adaptive immune response. There are multiple types of T cells which all have distinctive roles in host immunity. These are, helper T cells (CD4), cytotoxic T cells (CD8), regulatory T cells (Tregs), Mucosal associated invariant T cells (MAIT), memory, Natural killer T cells (NKT cells), and gamma delta T cells (γδT cells). B cells provide humoral immune responses as they are producers of specific antibodies

against pathogens and aid host defence through; blocking entry at mucosal surfaces (IgA) to facilitating phagocytosis (IgM/IgG) through opsonisation and subsequent recognition of the Fc region of the antibody by Fc receptors present on phagocytic cells or NK cells.

Linking the innate immune system to the adaptive immune system are dendritic cells (DCs). Dendritic cells are phagocytic cells which function as antigen presenting cells (APCs) and are mainly located in the mucosal tissue such as the intestine. In the intestine dendritic cells reside in specialised tissue called gut associated lymphoid tissue (GALT), here they sample the intestinal environment for microorganisms via PAMPs. DCs can then travel to the lymph node and present antigens to T and B cells to generate tailor made immune responses [141].

1.4 The immune system in sepsis

1.4.1 Summary

The underlying pathology of sepsis was thought to be caused by an initial severe inflammatory response to an infectious agent followed by a complementary immunosuppressive response, ultimately resulting in compromised immunity and an inability to control the infection with a self-mediated inflammatory damage to tissues and organs[156]. However, it is now clear that immunosuppression starts earlier than previously thought and the dysregulation between contrasting cytokines leads to sepsis patients having profoundly different cytokine responses compared to healthy controls, with defects in both pro and anti-inflammatory cytokine production [157]. A multitude of host response pathways have been implicated in the developments of sepsis including the complement cascade, the coagulation pathway and cellular metabolism pathways. Numerous immune cell types have been implicated in disease progression including cells of both the innate and adaptive immune response [158].

1.4.2 Cells and pathways affected by sepsis

1.4.2.1 LPS theory of sepsis

Lipopolysaccharide (LPS), the major component of the outer membrane of Gram-negative bacterial cell walls has long been recognised as a major inducer of the immune system in mammals and has been historically associated with the classic cytokine storm seen in sepsis (Figure 1.4)[159]. Indeed, animal models of sepsis infection commonly use LPS to stimulate

the immune system into a sepsis like state [160]. LPS is comprised of 3 major components: the O antigen, Lipid A, and a core component of oligosaccharide. LPS binds strongly to the host immune system receptor TLR4 (Toll like receptor 4) in complex with CD14 and MD2 (myeloid differentiation factor 2), which is expressed on many innate immune cells such as monocytes and macrophages. The binding of this receptor complex to LPS induces a strong inflammatory response from the host cells through the signalling of the immune transcription factor NF-κB, triggering the transcription and release of pro-inflammatory cytokines such as TNF-α and IL-1.

Exposure to LPS has a profound effect on the host immune system as evidenced by the differential expression of 3714 genes in response to LPS[161]. However, Gram-positive bacteria lack LPS in their cell walls and instead have significantly greater amounts of peptidoglycan (a polymer of sugars and amino acids) in their cell walls. While still immunogenic, peptidoglycan does not produce as strong a response as LPS. Indeed, multiple studies have shown that gram negative bacteria elicit a greater immune response from the host compared to Gram-positive bacteria. Both IL-6 and C reactive protein were significantly greater with Gram-negative bacteria in a study of 515 culture positive blood samples from bacteraemia patients [162]. Similar studies have also found vastly increased cytokine profiles in Gram-negative bacteria compared with Gram-positive bacteria, with both pro and anti-inflammatory cytokines vastly increased in Gram-negative sepsis infections [163], which can influence disease as Gram-negative bacteria were found to be a major risk factor for sepsis prognosis [164].

1.4.2.2 Barriers to infection in sepsis

During sepsis infection breakdown and dysbiosis of the normal gut microbiome in response to pathogens is thought to drive multiple organ dysfunction. The normal functioning of the gut is perturbed in sepsis and drive systemic propagation of the disease[165]. Epithelial cell barriers become permeable during sepsis allowing the contents of the intestinal lumen (including pathogens and commensal organisms) to break away from their normal environment where they can cause injury either in the gut itself or in distal tissues[166]. Intestinal epithelial cell apoptosis is also greatly increased in sepsis further contributing to the breakdown of the intestinal barrier. Dysbiosis of the normal gut microbiome also occurs

in sepsis and shifts toward the uncontrolled proliferation of bacteria manifesting as a loss of diversity, dominance of pathogenic organisms-a shift referred to as the pathobiome [167,168].

1.4.2.3 Complement in sepsis

The complement cascade has been implicated in the destructive early host immune response during sepsis, particularly the anaphylatoxins C5a and C3a. C3a is increased in sepsis patients and correlates with a worse APACHE II score[169]. Another study investigating complement in critically ill patients found that both C5a and C3 was elevated in sepsis patients compared to trauma patients[170]. Complement proteins C5b-9 (MAC) have also been found to increase in patients with severe sepsis or septic shock[171]. Although essential for early host defence against pathogens, complement activation can have detrimental effects in the later stages of sepsis and is associated with enhanced organ dysfunction in a rat model of sepsis infection, with increased levels of complement proteins found in the heart, liver, and spleen [172]. Evidence for the use of complement inhibitors as therapeutics for sepsis pathways have shown promise in animal models, including primate models of sepsis as well as human blood infections. Antibodies against C5 in conjunction with antibodies against CD14 have been shown to reduce inflammatory responses and respiratory burst in human whole blood infected with *E. coli* [173,174]. The anti-C5a and anti-CD14 has also been shown to be effective in a porcine model of sepsis, attenuating the inflammatory response and increasing survival[175,176]. Mouse models also show the role of C5a driving inflammation in sepsis. Mice lacking the complement receptor C5aR1 have increased survival, with increased pathogen clearance and liver function compared to control mice[177]. In contrast mice with deficiencies in C3 and C5 alone have increased bacterial burdens and are unable to clear infection, highlighting the complex nature of sepsis and the delicate balance needed between too much or too little complement activation to control infections. [178].

1.4.2.4 Neutrophils in sepsis

Several apoptotic signalling pathways appear to be impaired in neutrophils from sepsis patients; sepsis patient neutrophils have increased levels of the anti-apoptotic proteins Mcl-

1 and Bcl-xL and decreased levels of caspase-8, resulting in the prolonged survival of neutrophils in sepsis patients, mediated by C5a and LPS[179]. In addition to delayed apoptosis, neutrophils from sepsis patients have impairments in migration, rolling, adhesion and transmigration as well as antimicrobial activity [180–182]. NETs have been shown to induce tissue damage during sepsis[183], and increased numbers of NETs has been associated with increased acute respiratory distress syndrome (ARDS) severity in patients with microbial pneumonia [184].

1.4.2.5 Monocytes and Macrophages in sepsis

Monocytes can be classified according to their expression of CD14 a co-receptor which binds LPS in the presence of TLR4 [185] , and CD16 (FcγRIII) a Fc receptor which recognises IgG. Classical monocytes are $CD14^{++}$ and $CD16^{-}$, while non classical monocytes are $CD14^{+}$ and $CD16^{++}$ and intermediate monocytes which are $CD14^{+}$ and $CD16^{+}$ [186]. Normally classical monocytes make up more than 90% and are by far the most common monocytes circulating throughout the body. CD16 expression is increased during sepsis infections which is indicative of increased phagocytosis. However, HLA-DR expression has been found to be lower in monocytes of sepsis patients indicating a decrease in the presentation of antigens on monocyte cell surfaces [187]. Other studies have noted similar results in the alteration of monocyte function during sepsis. Shalova et al 2015 noted that monocytes from sepsis patients showed a pro-inflammatory gene expression pattern during infection but that these cells also showed a reduction in inflammatory cytokine and chemokine production and impaired antigen presentation when challenged with LPS indicating functional reprogramming of the cells during infection [188]. The authors also show evidence to suggest that HIF1α is a key mediator of this switch from an inflammatory to a suppressive state, through regulation of multiple different genes involved in inflammation, regulation, tissue remodelling and anti-microbial action[188].

Macrophages are heavily involved with the response to bacterial pathogens and are a major source of both pro-inflammatory and anti-inflammatory cytokines during sepsis [189]. Macrophages share a similar fate to monocytes during sepsis infections and studies in mice have shown a dysregulated function of both pro and anti-inflammatory actions of macrophages. Post-sepsis mice had impaired wound healing as well as a decrease in the

expression of inflammatory cytokines IL-1β, IL-12 and IL-23 due to epigenetic modifications in the bone marrow which are passed on to peripheral macrophages [190].

1.4.2.6 Cytokines in sepsis

The pattern of cytokine production during sepsis is time dependent (Figure 1.4). Early pro-inflammatory cytokine release leads to a cytokine storm resulting in damaging inflammation, this is then followed by an increase in anti-inflammatory cytokines, which suppress the immune system. The high levels of both pro and anti-inflammatory cytokines lead to a dysregulated immune response. Multiple studies have shown a dysregulated cytokine response in sepsis patients. Studies have shown increased levels of the anti-inflammatory cytokines IL-10 and lower levels of the pro-inflammatory cytokine TNFα, which correlate with mortality [191]. Other studies have highlighted the dysregulation associated with sepsis, even in surviving patients. Both whole blood and cells from sepsis patients have drastically reduced levels of TNFα, IL-6 and IL-1β in response to endotoxin. [192,193]. It is likely that these patients either have rapid production of both pro and anti-inflammatory cytokines, a predominance of anti-inflammatory cytokines, or a general depression in cytokine responses [157]. This mixed response of patients can make finding single biomarkers for sepsis incredibly difficult due to the heterogeneous nature of the disease. Indeed, patients have been shown to have varied cytokine responses (both pro-inflammatory and anti-inflammatory) when multiple cohort studies were compared. Patients which fit into either the α or β sepsis phenotypes have lower levels of IL-6 and IL-10 when compared with patients who fit into the γ or δ sepsis phenotypes (using data from four cohort studies) [27].

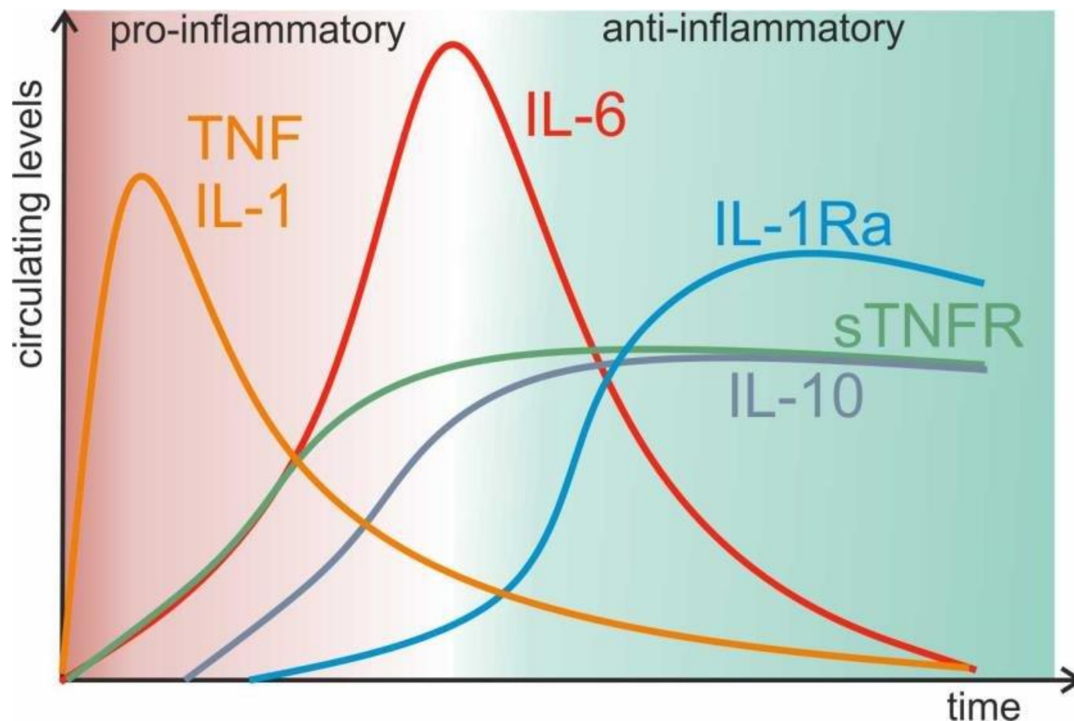


Figure 1.4 Pattern and levels of cytokine release in sepsis over time. Figure from [194], figure URL https://www.mdpi.com/1422-0067/19/5/1442

1.4.2.7 DAMPS in sepsis

Sepsis severity has been shown to correlate with DAMPs. High mobility group box 1 (HMBGB1), extracellular cold-increasing RNA-binding protein (eCIRP) and H3 have all been found to be at increased levels in sepsis, leading to increased tissue inflammation and injury [195–197].

During sepsis HMGB1 is elevated [198] , and studies have shown that targeting and inhibiting HMGB1 improves outcomes in sepsis [199]. Other DAMPs such as eCIRP, Histones, cell free DNA and even ATP have also been linked with sepsis [200], contributing to the release of cytokines throughout sepsis infections and acting in concert with PAMPs to drive dysregulation of the immune system.

1.4.2.7 Coagulation in sepsis

Abnormalities in the coagulation cascade are incredibly common in patients suffering from sepsis, although effects may range in severity. Reports indicate that 70% of sepsis patients will develop thrombocytopenia and haemostatic changes [201]. Thrombocytopenia is

defined as a deficiency of platelets in the blood and platelets are significantly reduced in several ways during sepsis. There is a markedly decreased production of platelets in septic patients along with increased destruction as well as increased consumption [202]. In addition, pathogenic bacteraemia causing *E. coli* induce apoptosis in platelets via the intrinsic apoptotic pathway and action of α-hemolysin [203].

1.4.2.8 Adaptive immunity during sepsis

Sepsis infections have been associated with a significant reduction in the number of CD4 T cells, CD8 T cells and B cells [204]. This depletion occurs through apoptosis via both the intrinsic (mitochondrial dependent) and the extrinsic (FAS/FASL associated) pathways [205]. This rapid depletion of effector immune cells results in an inability to control infection. Effector T and B cells that remain in sepsis patients become exhausted, evidenced by T cells harvested from the spleen of patients who have died of sepsis showing a decreased capacity to produce pro-inflammatory cytokines such as IFNγ, TNFα and IL-6 [206]. This depletion of effector immune cells can cause long term problems in patients who do recover from sepsis. This is evidenced by the fact that 60% of sepsis survivors are readmitted to a hospital after one year often due to an infection[207]. This is in conjunction with other morbidities brought about by sepsis including post sepsis syndrome (physical and psychological long-term effects with symptoms ranging from extreme fatigue to insomnia) and increased risk of cardiovascular disease[207].

1.5. Biomarkers in sepsis

A possible alternative diagnostic tool for identifying early bacteraemia is by the use of biomarkers. A good broad definition of a biomarkers *is "any substance, structure or process that can be measured in the body or its products and influence or predict the outcome of disease"* [208,209]. However, a more specific definition from the perspective of this thesis would be "*a biological molecule found in blood, tissue or other bodily fluid which is indicative of either normal bodily processes or of a disease/medical condition"*[39].

The most investigated biomarkers in sepsis research involve components of early host immune pathways, as these are usually relatively easy and quick to measure. To date

however there is no clear consensus on a quintessential biomarker for sepsis infection[210]. The complexity and overarching changes to host cellular pathways in addition to the multitude of organisms capable of causing sepsis infections does not lead itself to a one size fits all marker of disease.

1.5.1 Current methods and limitations of pathogen detection

Identifying sepsis before the onset of severe organ damage is extremely beneficial to patients. To cause widespread disease, bacteria must be able to cross into the bloodstream and infect other organs. Identification of bacteria in the blood is the main way in which bacteraemia is diagnosed. This is done by directly culturing the pathogenic organism/s from a patient blood sample. Blood culture is a good way of being able to identify the pathogenic organism down to the species level but there are several drawbacks to this method of diagnosis. The main limitation is the time it takes to culture the microorganisms, with blood cultures taking at least a day, if not longer for slower growing organisms and more time still to confirm AMR profiles. Other issues include the relatively low number of circulating organisms 1-10CFU/ml and the fact that some organisms are difficult to culture, as well as false negative results [55,211].

While CFU counts from blood cultures can be low this does not reflect the true immunological burden of bloodstream infections, as blood is an incredibly hostile environment for bacteria there will inevitably be a lot of dead bacteria causing immunological responses from the host as well as bacteria from cells clumped together and from inside circulating phagocytic cells [55]. An alternative measure which considers dead, clumped and bacteria in phagocytes is genome copies (GC). Genome copies can be identified using PCR based approaches to blood stream infections. Using this measure, it has been estimated that the average genome copies during BSI were between $1x10^3$ and $1x10^4$ [211].

Although PCR methods have gained popularity, due to their rapidity, there are still problems with their implementation and as such are currently used alongside traditional blood culture methods. Problems facing direct bacterial diagnosis from blood samples centre mostly around disruption of the PCR reaction by factors in the blood, including DNA from Leucocytes, DNA binding IgG, iron, and Heparin[212].

Infection models aimed at identifying biomarkers derived from host response to infection can be helpful tools for researchers aiming to identify markers of infection. Basic infection models involve the use of immortalised cell lines such as THP-1cells. Immortalised cell lines allow for the study of cellular mechanisms and function. The THP-1 cell line is a monocytic cell line extensively used to study the functions of monocytes/macrophages. Using this model allows for the study of potential biomarkers of interest during infection [213].

Other models such as whole blood models of infection can provide additional information, due to the larger variety of cell types present in blood e.g. neutrophils and monocytes. Recent research from our lab has shown that whole blood infection models can be used to bacterial infections at the strain level [214].

1.5.2 Current sepsis biomarkers

The diagnosis of sepsis remains challenging, but some conventional biomarkers remain useful. Over 100 biomarkers have been put forward as potentially useful in sepsis infections but to date few are routinely used. The most common biomarkers in use are procalcitonin and C reactive protein however, there are some promising biomarkers that could prove useful in the future (Table 1.5). Serum amyloid A (SAA) and acute phase protein which has shown promise as a biomarker for neonatal sepsis[215]. IL-6 and IL-8 in conjunction with PCT have been recognised as promising sepsis biomarker candidates[216,217] CD64 (FcγR1) an Fc receptor whose expression upregulates in response to cytokines also shows promise as a sepsis biomarker for disease prognosis as well as diagnosis [218,219]. Finally, over production of lactate has long been associated with increase inflammatory states and there is renewed interest in investigating its potential as a sepsis biomarker[220,221].

1.5.2.1 Procalcitonin (PCT)

PCT is the peptide precursor of the hormone calcitonin which is involved with calcium homeostasis. Levels of PCT in healthy people are below the limit of detection of most assays, but increase dramatically during infections, particularly infections of bacterial origin. PCT increases significantly in response to many bacterial infections including abdominal infections, meningitis, pneumonia, urinary tract infections and sepsis and is a promising marker to help guide antimicrobial therapy [222]. Levels of PCT increase substantially within 6-12 hours post infection and remain high until infection is controlled by antibiotics or the

host immune system. Blood levels reduce by half daily, making it a valuable resource in guiding antimicrobial treatment [222]. This trait of PCT as a marker of antibiotic effectiveness is a valuable tool in aiding the unnecessary prescribing of antibiotics, which is known to be associated with increases in antibiotic resistance. However, early PCT levels are often influenced by the initial cause of sepsis and its severity and not necessarily the severity of sepsis itself. In addition, there have been reported problems with assay sensitivity when detecting PCT with minor changes as well as elevated PCT levels in non-infectious diseases such as chronic kidney disease (CKD)[223,224].

1.5.2.2 C reactive protein

CRP is a protein involved in the acute phase response. It is synthesised in the liver in response to IL-6 released during infection. CRP can directly bind to pathogen surfaces where it activates the complement cascade and can lead to direct killing of the pathogen or opsonise them for direct uptake by phagocytes, along with several other pro-inflammatory effects. CRP can increase up to 1000-fold at sites of infection [225,226]. It can be used to discriminate between different pathologies and is elevated to different levels in different diseases. Levels of 10mg/dL indicate marked elevation and point to either acute bacterial infections, viral infections or trauma and levels over 50mg/dL indicate acute bacterial infections in 90% of cases[225].

1.5.2.3 Biomarker panels

It has become increasingly clear that a single biomarker for the prediction and diagnosis of sepsis with high degrees of sensitivity and specificity will not be found and so many researchers have turned their attention to a combination approach. Some panels of biomarkers show promise to help with the diagnosis of sepsis. A combination of C-reactive protein, leucine-rich alpha glycoprotein-1 and serum amyloid A in urine has been shown to be elevated in sepsis patients compared to patients with SIRS and healthy controls. As this study identified the biomarkers in urine this may be a potentially beneficial non-invasive method to diagnose sepsis [227]. Another study has found that a combination panel of biomarkers identified through machine learning is more effective at predicting 30-day

mortality than the gold standard procalcitonin and C- reactive protein [228]. This newly emerging research area shows promise for the future of sepsis biomarker development. Combined with advances in technologies such as microfluidics, electrochemistry, and artificial intelligence, panels of biomarkers are showing promise as a point of care detection method for sepsis infection [229].

Table 1.5 Current and promising sepsis biomarkers for detection of sepsis. Table from [229]

Sepsis biomarker	Category	Source	Normal concentration for healthy individuals	Severe sepsis concentrations	Peak onset time after stimulus	Sample type
CRP	Acute-phase protein	Liver	<3 µg/mL	>50 µg/mL	4–6 h	Whole blood, serum and urine
PCT	Acute-phase protein	Thyroid gland	<0.05 ng/mL	>2 ng/mL	12–24 h	Whole blood and serum
SAA	Acute-phase protein	Liver	<10 µg/mL	>1 mg/mL	8–24 h	Whole blood and serum
IL-6	Pro or anti-inflammatory cytokine	Monocytes, endothelial cells, and adipose tissue	<25 pg/mL	>1000 pg/mL	6 h	Whole blood, serum and cerebrospinal fluid
IL-8	Pro or anti-inflammatory cytokine	Macrophages	<10 pg/mL	>234 pg/mL	1–3 h	Whole blood, serum and cerebrospinal fluid
CD64	Cell marker	Monocytes	<8 mcL	>800 mcL	24 h	Whole blood and serum
Lactate	Other	Myocyte tissue	<2 nmol/L	>3.9 mmol/L	24 h	Whole blood, serum and urine

1.6 Current issues and gaps in the literature

Although there are sepsis biomarkers that are used in the clinic setting there is a desperate need for improved specificity as conflicting reports in the literature can lead to confusion in their accuracy. Single molecule biomarker studies have repeatedly shown inconsistency to successfully diagnose and predict the outcome of sepsis infections. There are currently no sepsis biomarkers that can correctly diagnose and predict sepsis outcome with 100% accuracy[210]. Furthermore, while there has been work done *in vivo* animal models to study differential host response to bacterial pathogens, animal genetic backgrounds (e.g., mice) can produce vastly different host responses. To date, there have been no studies investigating differential host response to collections of *E. coli* strains from bloodstream infections in human models of infection. Moreover, although studies have associated the presence of certain genetic traits of *E. coli* to aid sepsis diagnosis there have been no studies that have investigated the effect of these genetic traits on the host immune response. Indeed, genetic studies to identify at risk patients have mostly focused on host genetic traits during sepsis infections (e.g., mutations in host immune factors, such as TLRs). There has been little research on the potential of using bacterial virulence factors as biomarkers to aid in the prediction / diagnosis of sepsis.

1.7 Aims and objectives

This thesis aims to take a combined approach to identifying potential *E. coli* sepsis biomarkers by considering both *E. coli* and host response genes following infection. It is predicted that *E. coli* isolates from bacteraemia infection will not only be genetically distinct but also elicit a unique host immune response compared to control isolates. It is also predicted that *E. coli* from different clinically defined sources of infection will be genetically distinct and elicit a unique, identifiable host immune response. This study will:

Chapter 3 Establishing a collection of *E. coli* and linked patient and phenotypic data and basic genomic data

- Acquire *E. coli* collection from bacteraemia patients from the Hywel Dda University Health Board.
- Acquire patient demographic data from the blood culture positive isolates
    - Sepsis/non sepsis
    - Origin of infection
    - Coinfections etc
- Identify AMR resistance of the *E. coli* collection from the patient data.
- Perform basic genetic organisation of the *E. coli* bacteraemia isolates
    - Phylogroup identification
    - Sequence type identification

Chapter 4 Generating phenotypes: In vitro and ex vivo responses induced by *E. coli* collection

- Investigate panel of potential cytokine biomarkers using *ex vivo* whole blood infection models of bacteraemia. This panel will be used to characterise the host response to *E. coli* isolates from varying origins of infection.
- Investigate host response to *E. coli* collection using THP-1 model of infection.
- Characterise *E. coli* collection survival and growth in human plasma
- Assess serum bactericidal capacity of human serum on select *E. coli* isolates.

Chapter 5 Descriptive and comparative genomics of VFs and AMR genes associated with the *E. coli* collection

E. coli isolates will be whole genome sequences and then analysed for:

    - Generating a pangenome
    - Constructing a whole genome phylogenetic tree
    - Establishing the core/accessory genome
    - Identifying and counting the presence and absence of virulence factors
    - Identifying and counting the presence and absence of AMR genes
- Comparison of established virulence factors in bacteraemia whole genomes grouped by:
    - Origin of infection (Urinary vs Abdominal)
    - Bacteraemia vs ECOR
    - B2 phylogroup of urinary vs abdominal isolates

Chapter 6: Identification of genetic targets associated with bacteraemia and infection and functional confirmation.

- Comparison of whole genomes (GWAS method) grouped by phenotypic data collected from other chapters:
    - Association with bacteraemia
    - Association with Origin of infection
    - Association with urinary vs other bacteraemia
    - Association with sepsis / non sepsis (Chapter 3)
    - Association with mortality (Chapter 3)
    - Association with in vitro/ex vivo phenotypes (Chapter 4)
- Top gene targets identified through GWAS will be investigated further using knockout *E. coli* K12 isolates in
    - infection models and
    - survival/growth in plasma/serum

Chapter 2: Materials and methods

2.1 Media and stock solutions

2.1.1 Media

Media was prepared using 200-1000mL of distilled water (H_2O) and sterilised by autoclaving.

2.1.1.1 LB Broth

Luria Bertani Broth (LB) tablets were purchased from Sigma Aldrich (Gillingham UK), 1 tablet was dissolved per 50mL of H_2O. LB was made up in 200mL or 1000mL distilled water.

10g/L Tryptone

5g/L Yeast extract

5g/L NaCl

2.2g/L inert binding agents

For studies involving mutant *E. coli* K12 isolates, LB was supplemented with 25µg/mL of Kanamycin sulphate (Thermo Fisher, Loughborough UK).

2.1.1 Columbia blood agar plates

Columbia blood agar plates were purchased from Oxoid (Basingstoke UK).

2.1.2 THP-1 Cell growth media

500mL RPMI 1640 (Gibco, Montana USA) was supplemented with 50mL of heat inactivated FBS (Thermo Fisher Scientific, Loughborough UK) and 5mL of Penicillin and Streptomycin (Pen/Strep), (Thermo Fisher Scientific Loughborough UK) along with 5mL of L-glutamine (Thermo Fisher Scientific Loughborough UK). Was used for THP-1 cell line maintenance.

2.1.3 Experimental THP-1 cell media

Experimental THP-1 media used for bacterial infections made using 500mL of 1640 RPMI (Gibco, Montana USA), 50mL of Human serum, plasma derived (TSC Biosciences, Buckingham UK) and 5mL of L-Glutamine (Thermo Fisher Scientific, Loughborough UK). Plasma derived human serum was not tested for inflammatory responses prior to use however an uninfected control was used for each experiment. The same batch was used for all experiments.

2.1.4 Phosphate buffered saline

Phosphate buffered saline was purchased in both tablet and powder form from VWR life sciences (Leicestershire UK) and 1 tablet was dissolved in 100ml dH_2O. A 1L stock solution was prepared for microbiological and ELISA work. For powdered solutions 10x PBS was made in 100mL dH_2O and this was then added to 900mL dH_2O as required.

Per tablet (powder) when dissolved:

137nM Sodium Chloride

2.7 nM Potassium chloride

10 nM phosphate buffer

2.1.5 ELISA wash buffer

Tween 20 0.05% (Sigma Aldrich, Gillingham UK) was added to 1000mL PBS. Pipette tips were ejected into the PBS and mixed thoroughly due to viscosity of Tween.

2.1.6 ELISA reagent diluent

Bovine serum albumin (BSA) Protease free powder (Thermo Fisher, Loughborough UK) was diluted 1% w/v in 50mL dH_2O per ELISA plate.

2.1.7 ELISA stop solution

Hydrochloric acid (HCL) 37% (Sigma Aldrich, Gillingham UK) was diluted to a 1 M solution by adding 41.3mL HCl to 458.7 mL dH_2O.

2.1.8 HEPES buffer

HEPES buffer was prepared as previously described previously [230] 0.01 M HEPES, 0.15 M NaCl, 135 nM CaCl2, 1 mM MgCl2, at final pH 7.4.

2.2 Reagents

2.2.1 ELISA colour solution

KPL SureBlue TMB Microwell Peroxidase substrate (Sera Care, Massachusetts USA) and used as the ELISA colour solution.

2.2.2 DNA extraction kit

DNA mini kit Qiagen (Manchester UK) used for DNA extractions for sequencing used as per manufacturer's instructions.

2.2.3 Citrated human plasma

Citrated human plasma for use in growth assays was purchased from TCS Biosciences (Buckingham UK).

2.2.4 Healthy volunteer serum collection

Whole blood was collected from healthy volunteers (Ethics:13/WA/0190) in gold top blood collection tubes. Blood was then centrifuged at 2000g and serum collected and stored at -20°C. Some serum samples for use in bactericidal assay were obtained as a gift from Paul Morgan and Wioleta Zelek (Cardiff University).

2.3 Microbiology

2.3.1 *E. coli* strains from Hywel Dda University Health Board used in this study

List of *E. coli* strains obtain from the HDUHB can be found in Appendix 2.1

2.3.1.2 K12 knockout mutants

The *E. coli* strain K12 BW25113 and single gene knockout mutants were purchased from Horizon ™

Table 2.1 Single gene knockout mutants used in this study purchased from Horizon ™

Single gene knockout mutants
ybjE
yhgE
tufB
yohF
ynbC
yejF

2.3.1.3 Additional *E. coli* strains used in this study.

Table 2.2 List of additional *E. coli* strains used in this study

Strain	Comments	Source
ATCC 35218	Canine isolate, USA	ATCC
ATCC 25922	Control strain for AMR testing	ATCC
NCTC 1093	Isolated from human faeces	NCTC
NCTC 9001	Isolated from human urine	NCTC
K12	Laboratory control strain	Laboratory stocks

2.3.2 Culture of *E. coli*

E. coli isolates were received from the Hywel Dda University Health Board on agar slopes. Single colonies were picked and used to inoculate 5mL of LB broth. Frozen stocks were made with Microbank beads. For routine culture of *E. coli, isolates* from bead stocks were streaked onto Columbia blood agar plates and incubated overnight at 37˚C. Single colonies were picked from agar plates and used to inoculate 5mL of LB broth. For mutant studies, media was supplemented with 50µg/mL Kanamycin (Thermo Fisher, Loughborough UK).

2.4 Cell biology and cell culture

2.4.1 Whole blood model of infection

Whole blood was obtained from healthy volunteers (Ethics:13/WA/0190) in green top heparin blood collection tubes and 1/1.5ml of blood was added to Eppendorfs. Bacterial solutions (100µL or 150µL), or RPMI for control was added to the whole blood before samples were incubated on a rotator at 10rpm in a 37°C incubator for 2-6 hours. Ultrapure LPS InvivoGen (Loughborough UK), was also included in the whole blood solutions at a concentration of 1µg/mL. LPS was sonicated in water bath for 5 minutes before use in assays.

Overnight cultures of each isolate were set up by inoculating 5ml of LB with bacteria (from frozen glycerol stocks and cultured overnight at 37°C. OD readings were taken for each isolate and corrected to an OD 600nm of 0.1 in LB. The appropriate volume of broth sample

was added to LB to make up 1ml and centrifuged for 5 minutes at 16000g so that a pellet formed at the bottom of the tube. Media was then removed from the Eppendorfs, without disturbing the pellet and the pellet was resuspended in 1ml RPMI.

For frozen sample preparation, blood samples for each of the isolates were centrifuged at 7000rpm-9000rmp for 5 minutes. The resulting supernatant was then removed without disturbing the red blood cells blood and placed into fresh Eppendorfs. Tubes were the frozen at -20°C and defrosted before use in future experiments.

2.4.2 THP-1 cell model of infection

THP-1 cells were obtained from laboratory stocks and grown in RPMI (Gibco, Montana USA) supplemented with 10% FCS and 1% pen strep solution in T-75 flasks. For infection model, *E. coli* from an overnight culture were corrected to an OD 600nm of 0.1 in RPMI. THP-1 cells were seeded at 500,000 cells per mL in 24 well plates in antibiotic free RPMI with 10% plasma derived serum (TSC biosciences, Buckingham UK). Then, 100µL of bacteria (~$8x10^6$ CFU per mL) was added to cells and plate was incubated for 4 hours at 37˚C. After incubation, media was transferred from each well to a 1.5mL Eppendorf tube and centrifuged for 5 minutes at 9050g for 5 minutes. Supernatant was removed and transferred to a new 1.5mL Eppendorf and stored at -20˚C.

2.5 Assays

2.5.1 Complement depletion

Complement was depleted in the human plasma by heating at 56°C for 30 minutes in a heat block before the addition of *E. coli* isolates.

2.5.2 Human cytokine ELISAs

To measure cytokine levels in whole blood infected with bacterial isolates,
Human Duoset® Human ELISAs (R&D Systems, Minnesota USA) were performed as per manufactures instructions for IL-6, IL-8, MIP1α, MIP3α, Resistin and TNFα. Briefly 96- well half area plates (Greiner Bio, Kremsmünster Austria) were coated with the Capture Antibody, diluted to working concentration. Plates were then incubated overnight at room temperature. Wells were washed with Tween 0.05% in PBS and blocked with Reagent Diluent (1% BSA in PBS) for a minimum of 1 hour. Plates were washed again 3 times, and

50µl of blood supernatant samples or standards were added to the wells, plates were then sealed with an adhesive strip and incubated for 1.5-2hours. After washing an additional 3 times with 0.05% Tween 20 in PBS, 50µl of capture antibody (diluted to working concentration) was added to each well before an additional 1.5-2 hour incubation. Plates were washed 3 times then Streptavidin-HRP was added to each well and incubated for 20 mins, avoiding direct sunlight. Plates were then washed 3 times again, and the substrate solution was added to each well, incubated for 20 mins again avoiding direct sunlight and then the stop solution (HCL) was added to each well. OD was then measured at 450nm using a BMG plate reader. Standard curves were calculated using Mars software (BMG, London UK) and 4-parameter fit based on blank corrected data was used to ascertain the concentration of cytokine present in each sample.

2.5.3 Plasma resistance of *E. coli*

E. coli isolates from glycerol stocks were placed in 5ml of LB and incubated overnight at 37°C 200RPM. Bacteria were corrected to and OD600nm of 0.1 and diluted 1/10 in citrated human plasma (TSC biosciences, Buckingham UK) plasma and bacterial solutions were then transferred into a 96 well plate and incubated overnight at 37°C. OD600nm readings were taken every hour for 24 hours. Controls of both the LB and Plasma were included, and plates were blanked to the 0hrs time point. For confirmation of serum resistance patterns, a subset of isolates were grown in serum derived from gold top blood collection tubes. For collection of serum, blood was incubated at room temperature in gold top tubes for 30 minutes before being centrifuged at 2000g. Serum was then extracted and stored at -20˚C.

2.5.4 Serum bactericidal assay

To determine serum bactericidal activity in selected blood isolates a modified version of the serum bactericidal assay was used, combining a classical serum bactericidal assay with a luminescence bacterial cell viability assay Bac-Titre Glo™(Promega, Wisconsin USA). A modified Hepes buffer with supplemented cations was also used to keep complement enzymes active. Briefly OD readings of *E. coli* overnight cultures were measured at 600nm and corrected to 0.2 OD Bacteria was then centrifuged at 16000g for 5 minutes, supernatant was removed, and pellet was resuspended in HEPES buffer. Then, 20µl of this was combined

with 20µl human serum (gifted by Paul Morgan and Wioleta Zelek, Cardiff University) and supplemented with Hepes buffer up to 100ul. Bacterial and serum suspensions were then incubated for 3 hours at 37°C. 50µl of bacterial and serum cultures was transferred to a white 96 well plate (Grienier, Kremsmünster Austria) and 50µl of Bac-Titre Glo™ was added to each well. Luminescence was then measured on a BMG plate reader. For colony counting, 20µl of each bacterium and serum well was transferred to a new 96 well plate and serially diluted in PBS before being plated on Horse blood agar plates and counted after 24 hours incubation at 37°C.

2.6 Microbial genetics, sequencing and analysis

Table 2.3 List of software/programmes used in this study

Programme	Manufacturer
Prokka	GitHub - tseemann/prokka: :zap: Rapid prokaryotic genome annotation
Roary	Roary: the pan genome pipeline (sanger-pathogens.github.io)
Abricate	GitHub - tseemann/abricate: :mag_right: Mass screening of contigs for antimicrobial and virulence genes
MLST	GitHub - tseemann/mlst: :id: Scan contig files against PubMLST typing schemes
Sickle	GitHub - najoshi/sickle: Windowed Adaptive Trimming for fastq files using quality
Sythe	GitHub - vsbuffalo/scythe: A 3'-end adapter contaminant trimmer
QUAST	GitHub - ablab/quast: Genome assembly evaluation tool
seqtk	GitHub - lh3/seqtk: Toolkit for processing sequences in FASTA/Q formats
Clermont Phylotyper	ClermonTyping - Index (iame-research.center)
Morpheus	Morpheus (broadinstitute.org)
iTOL	iTOL: Interactive Tree Of Life (embl.de)
Gubbins	GitHub - nickjcroucher/gubbins: Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins
Prism	GraphPad Software Inc (San Diego, USA)
Lucid Chart	https://www.lucidchart.com/
Biorender	Scientific Image and Illustration Software | BioRender
Growthcurver	growthcurver package - RDocumentation

2.6.1 DNA isolation and extraction

All DNA extractions were carried out using the QIAamp DNA mini kit (Qiagen, Manchester UK) according to manufacturer's instructions. An overnight culture of the

strains selected for sequencing was grown in 5 ml LB broth as described in Section 2.3 and then 1 ml of culture was pipetted into a 1.5 ml microcentrifuge tube and centrifuged for 5 minutes at 6300g in an Eppendorf Centrifuge 5415 R. The supernatant was discarded, and a further 1 ml of culture added and centrifuged as described previously. The DNA extraction protocol was then followed as per the manufacturer's instructions and quantified (ng/ml) using a Nanodrop spectrophotometer MD1000 (Labtech, East Sussex, UK). Samples were frozen at -20°C before being sequenced at the Swansea Genome Centre.

2.6.2 Genome assembly

All genetic analysis was conducted on the Swansea University Linux virtual machine unless otherwise stated. Raw read fastq files were trimmed using sickle [231], and sythe [232], before genome assembly with spades [233]. Resulting assemblies were quality controlled with QUAST[234], before assemblies with low contig number were removed. Contigs <1000bp were removed from all isolates using seqtk [235], before annotation.

2.6.3 Genome annotation, pangenome generation and GWAS

E. coli whole genome sequence assemblies from spades were annotated using the prokaryotic genome annotation tool Prokka[236]. A pangenome was created using the stand alone pan genome pipeline roary[237]. ECOR genomes were included in the pangenome creation, genomes downloaded from [238]. After running Roary the gene presence/absence file was combined with a binary trait file and submitted to Scoary[239], for identification of genes associated with traits e.g., source of infection.

2.6.4 MLST

To identify sequence types in the *E. coli* collection isolates were run through the software MLST (Seeman T, unpublished)[240]. Whole genome sequences uploaded onto linux virtual machine were ran against PubMLST typing schemes for ST identification [241].

2.6.5 Phylogroup Identification

E. coli phylotypes were identified using the online *in silico* Clermont Phylotyper [75]. Whole genome sequences from bacteraemia, water and asymptomatic bacteriuria sources and the ECOR collection, were uploaded and phylotype was determined and recorded.

2.6.6 Virulence factor and AMR gene identification

E. coli virulence factors and AMR genes were identified using Abricate software[242], using the databases: Card[243], Resfinder[244], and VFDB[245].

2.6.7 Virulence and antimicrobial resistance heat maps

Abricate output was uploaded onto online heatmap generator Morpheus [246].

2.7 *E. coli* associated patient data

Ethics applications were approved to collect patient demographic data associated with the *E. coli* collection from the blood culture positive patients from HDUHB. The study details were i) title-Host and bacterial biomarkers to predict sepsis; ii) REC reference-20/WA/0127; iii) Protocol number-RIO-024-19; iv) Research ethics committee-Wales REC 7; and v) an IRAS project ID-262334. The extracted parameters included age, sex, sepsis development, nosocomial or community infection, source/origin of infection, co-morbidities, and mortality (Table 3.3). In addition, laboratory tests related to antimicrobial sensitivity were also included. All relevant paperwork is attached as separate files (Appendix 2.2, 2.3, 2.4 and 2.5) for reference please see (Appendix table 2.1).

2.8 Data analysis and statistics

Bacterial growth in plasma

The R package Growth curver was used to calculate AUC for bacterial growth in plasma.

ELISA

ELISA data was analysed with either one-way ANOVA with Kruskal Wallis test or for comparisons between two phenotypes with Mann Whitney U test, in Graphpad Prism. Standard deviations for grouped cytokine data were calculated in Microsoft Excel. Graphs show mean +/- SEM. *=p<0.05, **=p<0.01, ***=p<0.001.

Genetic data

Common *E. coli* virulence factor data was analysed with Chi Squared tests using Excel. In built statistics from scoary were used for GWAS studies.

Chapter 3: *E. coli* bacteraemia collection: their phenotypic and genetic traits and patient demographic data

3.1 Bacterial sepsis in Wales

Bacteraemia and the subsequent development of sepsis is a serious healthcare concern on a global scale, which also places a large burden on local healthcare systems. Despite the decreasing incidence and mortality (age standardised) of sepsis, it remains a major healthcare issue representing nearly 20% of global deaths [4]. This equates to more deaths than bowel, breast and prostate cancer combined [247].

In the UK, sepsis accounts for between 5% and 7% of all deaths [248]. It is particularly concerning in the intensive care unit, as 27% of all ICU admissions in England and Wales were for severe sepsis with a mortality rate of almost 50% [248]. In Wales the estimated number of deaths is 1800 annually, associated with a cost of £125 million [249,250]. However, this is expected to be higher due to underestimates in non-critical care reporting in Wales, where studies have found sepsis to be prevalent in 2% of patients on the general ward [251].

Bacteraemia and sepsis are much more likely to occur in older patients. Incidence rates in 2018 in England, Wales, and Northern Ireland for patients over the ages of 75 were between 400 (females) and 500 (males) per 100,000, which is more than double the incidence rate of the next youngest age group 65-74 (around 100-150 per 100,000) [102]. In addition, the annual financial cost to the NHS for *E. coli* bacteraemia was estimated at £14,346,400 in part due to increasing AMR resistant isolates being found in the clinic driving up the costs of treatment [252].

3.2 Hywel Dda University Health Board

The Hywel Dda University Health Board (HDUHB) is the NHS health board for the west of Wales which covers the areas of Carmarthenshire, Ceredigion and Pembrokeshire (Figure 3.1A). The board provides healthcare services in this area for a population of around 390,000 people. There are four main hospitals managed by the board: Bronglais General in Aberystwyth, Glangwili General in Carmarthen, Prince Philip in Llanelli and Withybush General in Haverfordwest along with five community hospitals and 48 general practices [253]. HDUHB has an aging population with 3.2% over the age of 85, and the number of people aged over 65 is expected to increase to 45% by 2033 [253]. HDUHB has the highest

incidence of bacteraemia in Wales with a rate of 117.54 per 100,000 in 2018 (Figure 3.1B). Unfortunately, more recent reports have not included updated numbers from the Health Board most likely because of the focus on COVID. Incidents of *E. coli* bacteraemia were seen to rise in Wales during 2002-2016. This has however been attributed to improved sepsis surveillance as well as successful use of sepsis bundles for rapid diagnosis and management. This is evidenced by the fact that although positivity results increased throughout the period, this increase was attributable to an increase in the number of blood cultures taken[254].

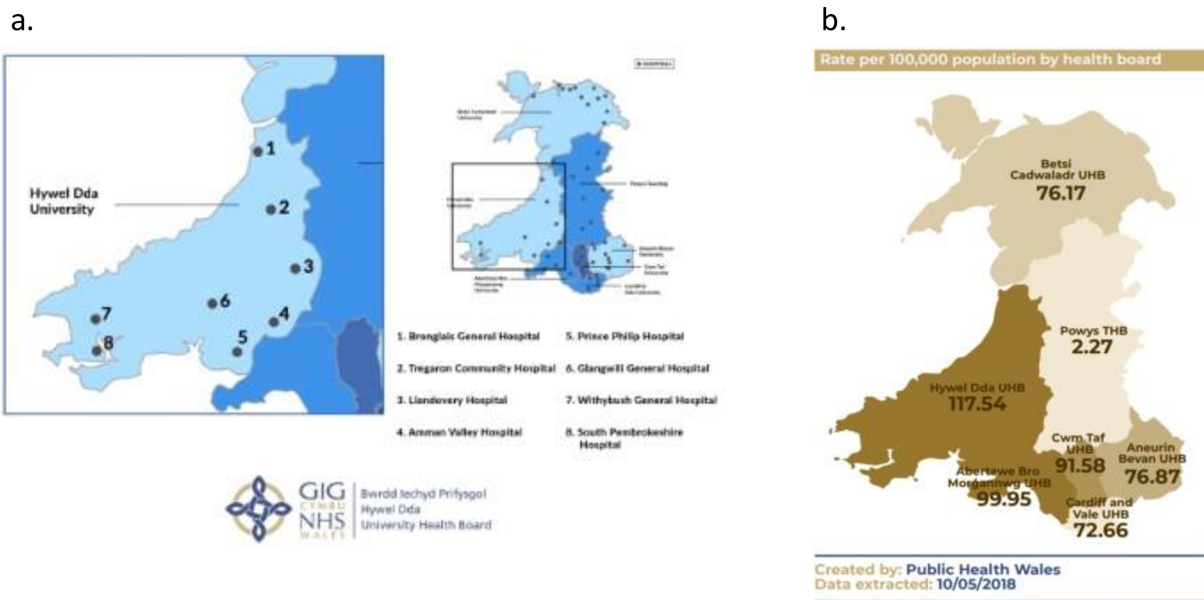


Figure 3.1: The Hywel Dda University Health Board area

a. Map of Hywel Dda University Health Board showing hospital locations), b. Bacteraemia rates in Wales. Figure adapted from HDUHB site and HDUHB annual report 2018 [255]

3.3 Properties of *E. coli* bacteraemia strains

3.3.1 Source of *E. coli* bacteraemia

E. coli bacteraemia can develop from a variety of different anatomical sources. Therefore, bacteraemia can occur through the spread of the bacteria from a primary source of infection into the blood stream either naturally during infection or through interventions such as surgery.

E. coli is a highly versatile and adaptable pathogenic organism and as such can cause disease in a multitude of host sites[68]. The capacity to translocate from the intestine to other sites is a characteristic of ExPEC infections. Extraintestinal infectious foci, lead to the development of bacteraemia. In *E. coli,* bacteraemia primarily occurs due to an infection at another site with the most common infection site being the urinary tract [256].

3.3.2 UTI

Most *E. coli* blood infections have a urinary tract origin (50% of cases) [257]. In addition, UTIs are also the most frequent bacterial infection in elderly patients who themselves are the most effected demographic by sepsis. Urinary tract infection can present from benign and asymptomatic to septic shock [258]. Bacteraemia commonly results from urinary tract infections and severity of initial local infections appears to correlate with systemic infection severity. Approximately 20-30% of patients admitted with more serious forms of urinary tract infection (e.g., complicated UTI, kidney inflammation, fever [259–261]), have bacterial positive blood cultures with *E. coli* accounting for 70% of them [262,263]. However, elderly patients who have UTI and bacteraemia can also present with asymptomatic urinary tract infection [264,265]. This vastly increases the difficulty of diagnosis and emphasises the need for biomarkers of local early infection as guiding early treatment is critical in reducing mortality in these patients [264].

3.3.3 GI and biliary tract

Infections originating in the abdomen can be variable and complex and include infections of the GI and biliary tract.

The GI tract can also be a source of *E. coli* bacteraemia and the abdomen is often reported as the source of bacteraemia in patients who have undergone surgery [266,267]. The

abdomen is the second most common source of bacteraemia and over 66% of surgical patients with sepsis have an intra-abdominal source of infection[268].

E. coli bacteraemia from an abdominal origin has also been associated with increased risk of disease severity [269], potentially due to the comorbidities associated with patients undergoing surgery, as these patients are generally unwell, they are more likely to have increased risk of developing sepsis and increased risk of mortality if sepsis does occur due to the inability to establish an effective response to infection.

E. coli bacteraemia infections which originate from the biliary tract tend to occur in patients with underlying structural abnormality such as choledocholithiasis [266,267]. Indeed, structural abnormalities of pre-existing conditions and newly diagnosed have been reported in 69% of cases of biliary bacteraemia[270]. Interestingly, time to antibiotic administration in these biliary infections did not significantly reduce mortality (as has been reported for sepsis/bacteraemia in general) [271].

3.4 Antimicrobial resistance

The increasing trend of AMR is a true threat to healthcare systems worldwide. For instance, in 2014 it was estimated that by 2050, 10 million people a year will die due to AMR related infections or treatment complications [272]. Although, this report has been reported as sensationalist, and likely over-estimates natural resistant and sensitive phenotypes and their ability to develop resistance [273]. Nevertheless, AMR rates are predicted to rise, causing unprecedented problems for healthcare systems. For instance, a recent comprehensive report in the Lancet highlights the global AMR resistance of 23 human pathogens across 204 countries and territories. In 2019 there were a predicted 4.95 million deaths associated with bacterial antimicrobial resistance and *E. coli* was the leading cause of death [100].

AMR of *E. coli* bacteraemia isolates has been reported to be associated with in-hospital death, as well as being strongly associated with increased hospital stay (in third generation cephalosporin and piperacillin/tazobactam resistance, 1.58 days and 1.23 days increased respectively [274]. AMR resistance in ExPEC is strongly linked to phylogroup and sequence type and these isolates can often resist β-lactam antibiotics through the production of β-lactamases [275]. EXPEC belonging to the sequence type 131 have become strongly

associated with the production of β-lactamases and thus resistance to β-lactam antibiotics such as the Penicillins, Cephalosporins, Carbapenems, Monobactams and β-lactamase inhibitors [276]. *E. coli* bloodstream infections caused by β-lactamase producing *E. coli* strains have 10% higher mortality rate compared to those that do not [277].

3.5 Aims and objectives

The high incidence of bacteraemia in HDUHB suggested to us that its local hospitals would be a good source of ExPEC bacteria for the study of biomarker identification and virulence mechanisms.

This chapter aims to

- Acquire an *E. coli* collection from bacteraemia patients from the HDUHB.
- Acquire linked patient demographic data from the blood culture positive *E. coli* isolates
    - Sepsis/non sepsis
    - Origin of infection
    - Coinfections etc
- Identify and characterise the AMR resistance profiles of the *E. coli* collection.
- Perform basic genetic organisation of the *E. coli* bacteraemia isolates
    - Phylogroup identification
    - Sequence type identification

3.6 Materials and methods

3.6.1 *E. coli* bacteraemia collection

In total 179 *E. coli* isolates were received from Public Health Wales on agar slopes. All isolates were obtained from the HDUHB between 2018 and 2020 [254]. A loop was touched gently on the bacterial smear of each agar slopes and used to inoculate blood agar plates, which were incubated at 37°C for 24 hours. Of the isolates received, 14 were non-culturable, 7 of the isolates were from an environmental water source and 6 were from asymptomatic bacteriuria patients while the remaining isolates were from blood culture positive bacteraemia (n=152), (Figure 3.2).

3.6.2 Hospital acquired patient and isolate data

Ethics applications were approved to collect patient demographic data associated with the *E. coli* collection from the blood culture positive patients from HDUHB. The study details were i) title-Host and bacterial biomarkers to predict sepsis; ii) REC reference-20/WA/0127; iii) Protocol number-RIO-024-19; iv) Research ethics committee-Wales REC 7; and v) an IRAS project ID-262334. The extracted parameters included age, sex, sepsis development, nosocomial or community infection, source/origin of infection, co-morbidities, and mortality (Table 3.3). In addition, laboratory tests related to antimicrobial sensitivity were also included.

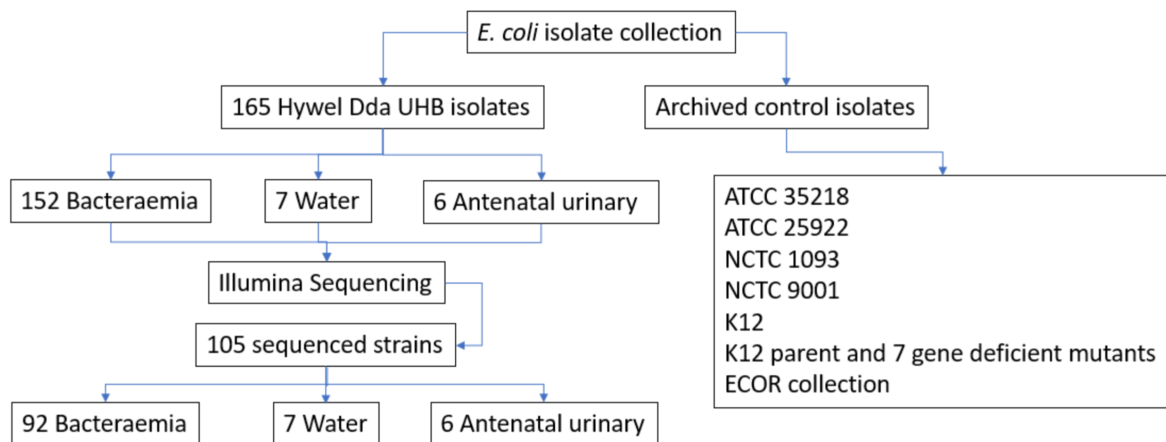


Figure 3.2 Summary of *E. coli* isolates from HDUHB Two major sources of isolates included i) isolates from the Hywel Dda University Health Board (HDdUHB); ii) Archived isolates from the Microbiology and Infectious Disease group. New isolates from the HDdUHB were submitted for sequencing, annotated and subjected to gene-by-gene analysis

3.6.3 DNA extraction and genome processing

In a subset of 105 isolates (Figure 3.1) DNA was extracted using Qiagen kits according to manufacturer's instructions (section 2.4.4) ng/uL was recorded using a nanodrop. All DNA sequencing was performed by Dr Matthew Hitchings at the Swansea genome centre (section 2. 4.4).

Raw read files were processed as per section 2.6.1 for genome assembly and annotated with Prokka (section 2.6.2) using default parameters The ECOR collection were included in

the pangenome creation[238]. Pan genome assembly was conducted with Roary [237] using K12MG1655 as a reference genome (Genbank U00096.3).

3.6.4 Phylotyping and sequence typing

E. coli phylotypes were identified using the online *in silico* Clermont Phylotyper[278]. Whole genome sequences from bacteraemia, water and asymptomatic bacteriuria sources, were uploaded and phylotype was determined and recorded. For sequence type confirmation, the software MLST was used[279,280]. Whole genome sequences uploaded onto linux virtual machine were ran against PubMLST typing schemes for ST identification[281].

3.7 Results

Having collected blood culture positive *E. coli* isolates from hospitals within HDUHB, experiments focused on characterising their properties and extracting data associated with the patients who suffered the infections. Therefore, three sets of interdependent data will be presented in this chapter and include i) antimicrobial resistance traits of the *E. coli* collection (3.8.1); ii) basic genetic traits of the *E. coli* collection (3.8.2); and patient demographic data associated with the *E. coli* collection (3.8.3).

3.7.1.1 Antimicrobial resistance

Antimicrobial sensitivity was determined for each *E. coli* isolate by in-hospital testing and recorded along with patient demographic data. Of the antibiotics tested the isolates showed most resistance against ampicillin and amoxicillin with 64% and 62% of isolates showing resistance respectively (Table 3.1). High resistance rates (>30%) were also observed against cotrimoxazole (31%), and moderate resistance rates (>10%) were observed against co-amoxiclav/augmentin (19%), piperacillin/tazobactam (15.1%), ceftazidime (13.2%), and cefotaxime (11.8%). No resistance was observed against imipenem or meropenem (Table 3.1).

Table 3.1: Antimicrobial resistance of *E. coli* collection (first panel antibiotics).

Antibiotic	Abbreviation	Total (percentage)
Ampicillin	AMP	97 (64)
Ceftazidime	CAZ	20 (13.2)
Ciprofloxacin	CIP	15 (9.9)
Cefotaxime	CTX	18 (11.8)
Ertapenem	ERT	1 (0.7)
Gentamicin	GENT	13 (8.6)
Imipenem	IMP	0
Meropenem	MEM	0
Piperacillin/Tazobactam	PIPT	23 (15.1)
Amikacin	AMI	4 (2.6)
Amoxicillin	AMX	94 (62)
Co-amoxiclav/Augmentin	AUG	29 (19)
Cotrimoxazole	SXT	47 (31)

3.7.2 Genetic traits of *E. coli* collection

E. coli isolates were sequenced in house at Swansea University. A total of 92 *E. coli* bacteraemia isolates 7 environmental water isolates and 6 isolates from asymptomatic bacteriuria were sequenced (Figure 3.1). Multiple online tools (section 2.6) were used to determine phylogroup, sequence type and serotype.

3.7.2.1 Phylogroup assignment

To determine the phylogroup that each of the isolates belonged to, whole genomes were uploaded onto the online ez Clermont phylotyping tool (section 2.6.4) and phylogroups were recorded. The majority of the bacteraemia isolates (>70%) belonged to the B2 phylogroup (Table 3.2). Similarly, five of the six asymptomatic bacteriuria isolates were also B2 phylogroup (the remaining isolate being D). In contrast, the environmental water isolates were more diverse; x3 cryptic, x2 B1, x1 A and x1 D group.

Table 3.2: Phylogenetic grouping of *E. coli* bacteraemia isolates using Clermont online phylotyping.

Phylogroup	Total (percentage)
A	6 (6.5)
B1	3 (3.3)
B2	71 (77)
C	1 (1)
D	3 (3.3)
E	1 (1)
F	1 (1)
U/Cryptic	6 (6.5)

3.7.2.2 Sequence typing

E. coli contigs were run through mlst (Version 2, June 1991)[282], on a Linux virtual machine to identify sequence types. A total of 37 different sequence types were identified by mlst (Figure 3.3). Three sequence types were associated with over half of the isolates. The most common identified sequence types were ST131 (22%), ST73 (20%) and ST12 (10%). Most of the remaining 34 sequence types were only represented by one or two isolates indicating the diversity of *E. coli* isolates able to cause bacteraemia.

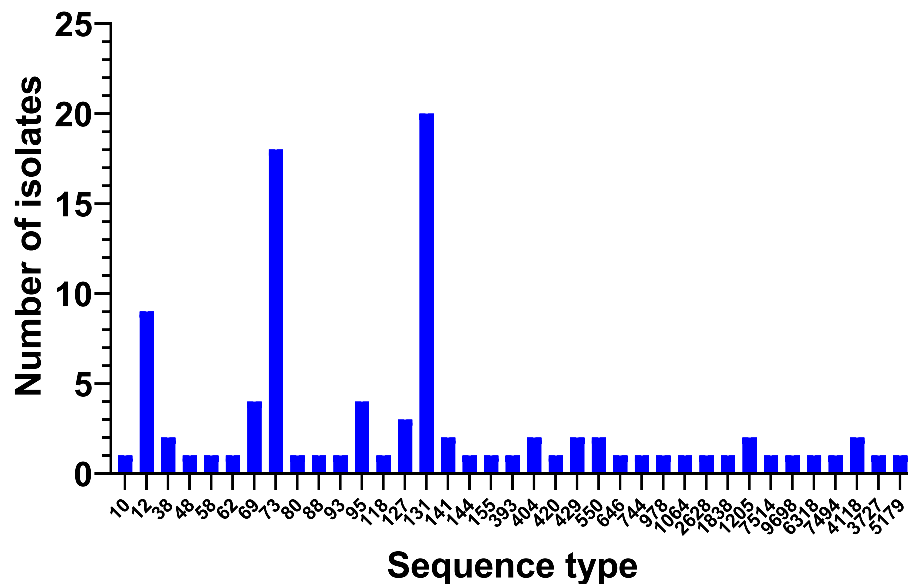


Figure 3.3: Sequence type of *E. coli* bacteraemia isolates.

Whole genome sequence contigs were analysed by software mlst, on a Linux virtual machine using the *E. coli* PubMLST database

3.7.2.3 AMR genes of ST131 and ST73 isolates

Given that *E. coli* ST131 and ST73 were the major sequence tyoes identified in our collection and that the same sequence types are also of major global importance, AMR to therapeutic agents was assessed (Figure 3.4 a and b).

ST131 isolates showed increased resistance to antibiotics and resistance to more antibiotic classes than the ST73 isolates, 22/29 ST131 isolates were resistant to Ampicillin whilst 7/17 ST73 isolates were resistant. 5/29 ST131 isolates were resistant to Gentamicin, compared to 0 from ST73, 21/29 of the ST131 isolates were resistant to Amoxiciccin compared to 7/17 of the ST73 isolates and, 20/29 of the ST131 isolates were resistant to Augmentin (Amoxicillin/Clavulanate) compared to 5/17 of the ST73 isolates (Figure 3.4). All of the isolates form ST131 and ST73 were sensitive to Imipenem and Meropenem (Figure 3.4).

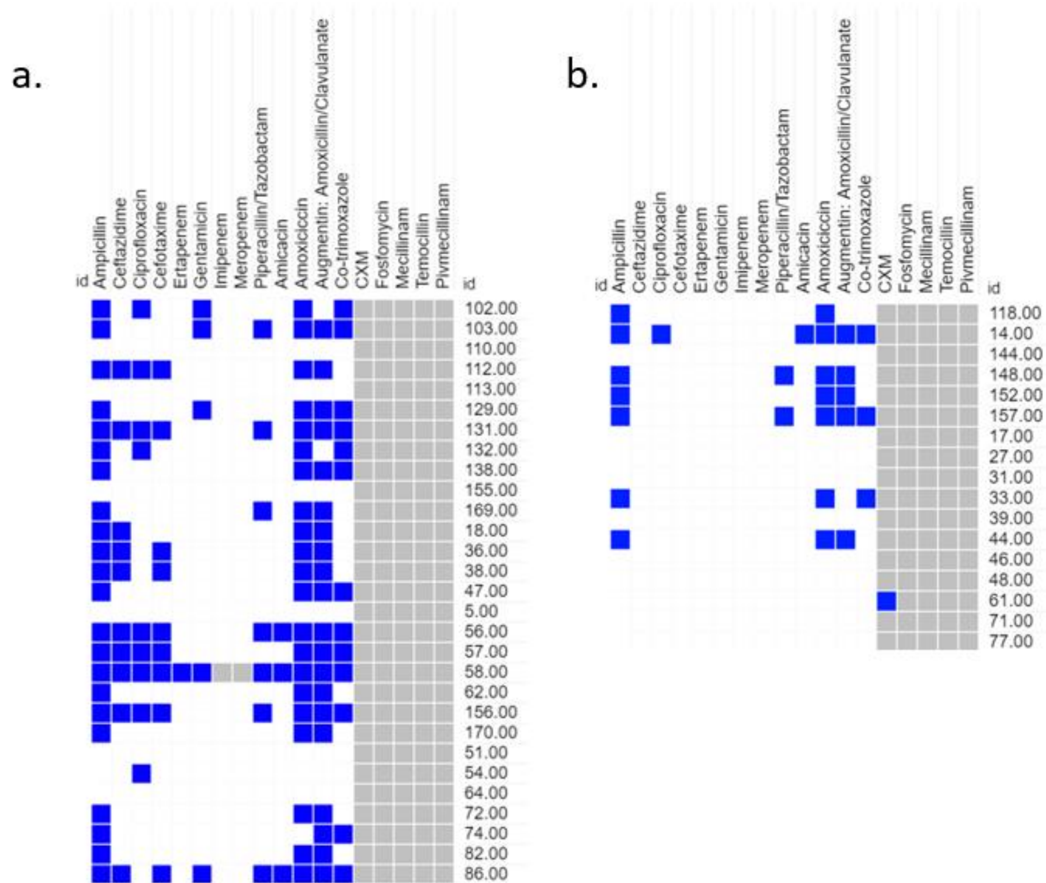


Figure 3.4: Presence and absence of antimicrobial genes in ST131 and ST73 bacteraemia isolates

Heatmaps showing antimicrobial resistance profiles of two most common sequence types 131(a.), and 73 (b.). Heat maps produced using Morpheus online software. Blue indicates presence of resistance and white indicates no resistance. Grey indicates antibiotics not tested.

3.7.3 Patient demographics and definitions

3.7.3.1 Patient demographics

Blood culture positive *E. coli* isolated between 2016-2019 were collected by the Hywel Dda University Health Board and transferred to Swansea University. After correcting for duplicates and isolates that did not culture, a total of 152 bacteraemia isolates (Figure 3.2) with associated patient data were collected. Patient data that was collected included, age,

data admitted to hospital, co-morbidities, source of infection, antimicrobial resistance profiles (Table 3.3). In addition, 7 *E. coli* from environmental water sources were also received from the hospital and 6 *E. coli* isolates from asymptomatic bacteriuria during pregnancy.

The mean and median age of the patient population was 73.84 and 75 respectively (Table 3.3), although there was also a large range in the age of the patients, from a neonate, 118 days to a patient 103 years old. In this collection of patients males and females were represented equally with 49% male and 51% female patients (Table 3.3). The majority of the isolates collected in this study came from a urinary source of infection (48%) with other/unknown being the second most common (24.3%), whilst an abdominal source of infection accounted for 21 % (Table 3.3).

there was a high incidence of co-morbidity associated with the patients. The most common co-morbidity was Urinary tract infection (15%), followed by cancer (11.8%), co-infection (8.6%) and diabetes (6.6%) (Table 3.3). Nearly half of patients went onto develop sepsis (46.75%) and 44% of bacteraemia cases were hospital acquired. Overall mortality among the bacteraemia patients was 13.2% (Table 3.3)

Table 3.3: Patient demographic data.

E. coli bacteraemia cases from 152 patients from the years 2016-2019 were collected along with patient demographic data. (n=152) Intra-abdominal sepsis isolates and abdominal isolates were grouped together as per hospital recommendation.

Patient Characteristics	**Total (percentage) n=152**
Age (years)	
Mean	73.84
Median	75
Range	118 (days)- 103 years
Gender	M 77 (49), F 80 (51)
Source of infection	
Urinary	73 (48)
Abdominal	32 (21)
Biliary	10 (6.6)
Other/Unknown	37 (24.3)
Sepsis	71 (46.7)
Nosocomial	67 (44)
Mortality	20 (13.2)
Comorbidities	
Cancer	18 (11.8)
Diabetes	10 (6.6)
Urinary tract infection	23 (15)
Catheter	11 (7.2)
Cardiovascular disease	5 (3.3)
Kidney disease	6 (4)
Cholecystitis	4 (2.6)
Pancreatitis	2 (1.3)
Co-infection	13 (8.6)
COPD	4 (2.6)
Dementia	5 (3.2)
Hospital acquired pneumonia	2 (1.3)

3.7.4. Inter-relationships between data

3.7.4.1. Source and phylogroup

Investigations into the source of infection from the 4 most common phylogroups (Figure 3.5) confirmed the most common phylogroup was B2, with 71 isolates. Of these, 33 of the isolates were from a urinary source of infection, while the second most common were from another/unknown source (Figure 3.5 B2). Phylogroups A, B1 and D had very few isolates so a relationship to origin of infection was not possible to assess (Figure 3.5 A,B1,D). The ANC urine isolates were predominantly from the B2 phylogroup (Figure 3.5 B2) while the water isolates were spread between the phylogroups D, A and B1 or the unknown/cryptic group (Figure 3.5 A, B1 D).

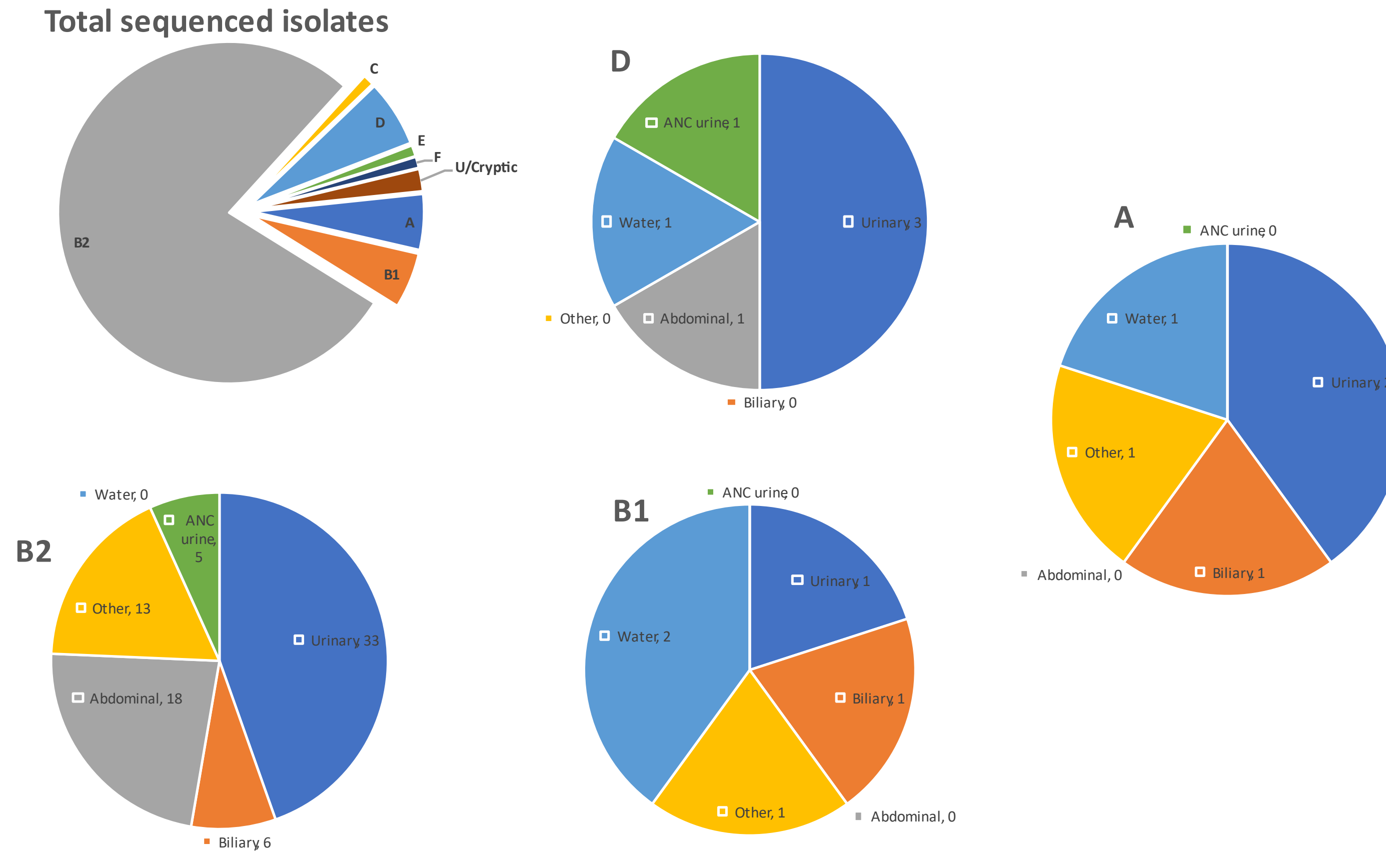


Figure 3.5 Inter-relationship between four common *E. coli* phylogroups and source of infection.

The figur shows five pie charts. Top left= total isolates phylogoups, top middle. phylogroup D, middle bottom. phylogoup B1 , bottom left phylogoup B2, right phylogroup A. *E. coli* isolates included were sequenced bacteramia isolates, environmental water isolates and Asymptomtpmatic bacteriurea isolates from antenatal checkups (ANC urine). Three bacteraemia isolates were excluded from analysis due to unclear infection source.

3.8 Discussion

This chapter aimed to acquire a collection of bacteraemia isolates from the HDUHB and characterise phenotypic and basic genetic traits. In addition, this chapter also investigated the patient records associated with these bacteraemia infections to confirm interdependencies [28,29].

AMR sensitivity data was acquired from the hospital along with patient demographic data. Due to this the specific minimum inhibitory concentrations of antibiotics needed for those isolates that were sensitive are not known, which is the case for standard EUCAST broth methods of AMR sensitivity testing[283]. However, resistance levels in these *E. coli* isolates appear to be below the average reported in England. For instance, Ciprofloxacin resistance levels were lower in this study than they are in the English AMR surveillance report. Average levels of ciprofloxacin in England were 17%, whereas in this study a lower rate of 9.9% was observed (Table 3.1). Furthermore, there were also significantly lower levels of resistance to Co-amoxiclav. In this study 19% of isolate were resistant however, general resistance levels in England were more than twice as high at 41.2% [284]. This general trend may be explained by the time difference between the collection date of the isolates in the current study and when the most recent report was published.

In terms of the characteristics of the *E. coli* found in this study there was again similarities between previously reported BSI and sepsis-associated infections. By far the most common phylogroup associated with ExPEC BSIs is the B2 phylogroup as has been reported in multiple studies [285,286]. The same is true in this study with phylogroup B2 making up 74% of *E. coli* bacteraemia cases (Table 3.2). The second most common phylogroup, group D, which is also consistent with other studies, although this phylogroup often has a higher incidence than found in this study [285].

Interestingly, the majority of the isolates from asymptomatic bacteriuria group were also from the B2 phylogroup and the remainder of these isolates were from the D phylogroup. As these are the two most common phylogroups in bacteraemia and sepsis it is logical to conclude that this may be a potential source of *E. coli* neonatal infections. *E. coli* is a major neonatal pathogen, and a leading cause of early onset sepsis and *E. coli* is passed from the mother during birth which in some cases can lead to the development of neonatal sepsis and meningitis[287,288]. The most prevalent phylogroups in *E. coli* neonatal sepsis are the

B2 and D phylogroups [289]. Further investigation into this potential source of *E. coli* associated neonatal sepsis/meningitis is warranted.

The most prevalent sequence types identified in this study were ST131 and ST73 (Figure 3.3), this is in line with previous reports in the literature [290,291]. This study identified that *E. coli* from sequence type 131 have higher levels of AMR compared to isolates from the phylogroup 73 (Figure 3.3). This conforms with literature reports of isolates from sequence type 131 being a major global health concern regarding antimicrobial resistance [292–296]

Underlying source of infection was identified at the hospital and provided along with patient demographic data. In this study the majority of the isolates (48 %) came from a urinary source of infection (Table 3.2). Given that 15% of patients were suffering from UTIs (Table 3.2) and that *E. coli* is the leading cause of urinary tract infections this is not surprising. Indeed, other studies have also indicated a predominant urinary source of bacteraemia [269,297,298]. This study had a large percentage of isolates (24.35%) that came from an unknown source of infection. As this data was acquired from the hospital it is unknown the reason why source of infection could not be identified. It is possible that an infection source in this case was hard to determine due to the nature of sepsis spreading throughout the body and infecting other organ sites, making it difficult to determine the origin. This is particularly true during later stage infections where inflammation is systemic.

The incidence of each source of infection in this study is consistent with those found in the literature. A report published in the Journal of Hospital Infection found similar levels of all sources of infection. In this study authors compared the rates of different sources of infection from the Imperial College Healthcare Trust to national data. A urinary source of infection was the most common in both the Imperial College and national data (~35% and ~55% respectively) [299]. The rate of urinary source of infection in this study was 48% which more closely matches national data.

It also appears that source of infection can vary based by geographic region. While nationally, gastrointestinal sources were rare (<10%), data from the Imperial College Healthcare Trust showed a significantly more prevalent amount with 20% of bacteraemia's from the gastrointestinal tract. This study had very similar amounts from an abdominal source of infection with 21% (Table 3.2), indicating a closer match to the Imperial College Trust local infections compared to national levels for gastrointestinal sources [299].

Respiratory infections were rare in this study with only one isolates identified from a respiratory source. While in general ICU respiratory infections are responsible for a high percentage of infections, those attributable to *E. coli* are particularly rare. Indeed, literature reports suggest that respiratory *E. coli* infections account for only 0.7% of cases which matches what was found with this collection of isolates [300].

Bloodstream infections are extremely common in older populations and are often fatal. Studies have shown that septic shock can occur in up to 39% of patients and older patients have a higher risk of developing septic shock compared to their younger counterparts [301]. Additionally, the most common source of bloodstream infections in older populations appears to be the urinary tract [301]. This is in line with the data presented in this study as the mean age of the population in this study was 73.84 and the majority of the isolates came from a urinary source of infection (Table 3.3).

One striking difference between this study and others is the fact that urinary catheters are commonly associated with BSIs from a urinary source but only a small minority of patients in this study had indwelling urinary catheters (7.2%) (Table 3.3). This is in stark contrast to a previous study investigating community acquired BSIs which showed that nearly 44% of patients over 65 had an indwelling catheter and that these patients had significantly higher risk of a urinary source of infection [302]. The differences observed between these two studies could be due to missing data from this study. Full patient records were not given as only brief notes on each of the patients was provided by the hospital.

Other common comorbidities found in this study also closely mirror reported risks for BSIs. Chronic comorbidities are present in 54-65% of all sepsis patients [303,304], and certain co-morbidities have been associated with an increased risk of sepsis infection, such as cancer, diabetes, HIV and chronic liver disease [304]. This study found that UTI was the most common source of co-morbidity which correlates with previous literature reports [264,305]. The number of nosocomial infections identified in this study were higher than average, collated studies on literature have reported hospital associated infections to be between 9-33% of total infections caused by *E. coli* [306], in this study nosocomial infections accounted for 44% of total infections. Mortality was found to be lower than the mortality rate for England at 13.2% compared to 18.2% [307]. However, it is withing the expected range of mortality rates found in other studies [308,309].

There were some limitations associated with the *E. coli* collection described in his chapter. Due to available funding only a subset of the *E. coli* could be sequenced. In total, 92 bacteraemia isolates out of 152 were sequenced and included in the genetic analysis. This represented the population used in this study with a spread of sources of infection with the most isolates being from a urinary source. This may have removed some of the power in the genetic analysis section of future results chapters (Chapter 6).

Patient data was not collected by the research team and relied on the expertise of our clinical colleagues to access numerous databases. In some cases, diagnosis was not final, and this made producing a final table difficult.   Indeed, determination of co-morbidities was especially difficult in this regard. Future studies include more time for data extraction from clinical records.

However, this collection appears to be a relatively standard collection of *E. coli* bacteraemia isolates with strong correlations with other reported *E. coli* bacteraemia collections. The patient population in this study are generally older which matches the typical bacteraemia patient profile. The antimicrobial resistance rates are lower than average, and isolates belonging to the sequence type 131 are more resistance to antibiotics than the second most common sequence type identified in this study 73. The source of infection was consistent with the high incidence of urinary origin of infection in other National studies.

Chapter 4: Host biomarkers induced with clinical *E. coli* strains from bacteraemia patients

4.1 Host biomarkers of infection

Biomarkers using proteins involved with the host innate immune response have great potential to aid clinicians in the early diagnosis of many diseases. Although a wide variety of sepsis biomarkers have been studied, many have failed due to lack of specificity and sensitivity [310,311]. A review in 2020 identified 258 biomarkers that had previously been investigated in relation to bacterial sepsis, highlighting the complex nature of the disease and the many different immune pathways involved in the host's immune response to the disease [310,311]. Despite this, there are biomarkers that are currently in use in the clinic, however, there is no biomarker that can successfully; diagnose, predict, and track the treatment of sepsis [312].

Three main biomarkers are used for bacterial sepsis and are considered the current gold standard: C-reactive protein (CRP), procalcitonin (PCT), and Interleukin 6 (IL-6)[313–315]. Indeed, PCT and CRP often define the benchmark for comparing new potential sepsis biomarkers due to their relatively good predictive ability [310]. However, there are limitations on the usefulness of these biomarkers largely due to their inducibility in non-infectious states [316]. New biomarkers that can differentiate between infections and non-infectious inflammatory states are crucial for guiding or even removal of appropriate antimicrobial therapies, in bacterial sepsis. Furthermore, the rise in antimicrobial resistance makes it more important than ever to use antibiotics both appropriately and sparingly [317–320].

4.2 Early mediators of the immune response

Early mediators of immunity have great potential to predict the presence of bacterial infections and cytokines show promise as biomarkers of infection due to extensive studies done on their relationships and role with disease [321].

4.2.1. IL-6

IL-6 is an acute phase pleiotropic cytokine produced early during the host response to infection. It is produced by macrophages, endothelial cells and epithelial cells after

recognition of PAMPs (e.g LPS) and has roles in the cytokine networks driving inflammation and in the resolution of inflammation, together with effects on T cell differentiation (e.g., Th17 cells). IL-6 is released locally at the site of infection and in endothelial cells where it is translocated via the bloodstream, to the liver where it acts on hepatocytes to induce the production of acute phase proteins such as CRP, serum amyloid A (SSA), haptoglobin (Hp) fibrinogen and albumin [322]. These proteins have prominent systemic effects on the host and lead to several physiological changes such as fever, anorexia, and catabolism of muscle cells [323]. IL-6 has several other pro-inflammatory effects including the production of neutrophils in the bone marrow, Th17 cell differentiation, downregulation of Treg cells and increased antibody production in B cells [322].

IL-6 is an important cytokine required for controlling *E. coli* infections, including the induction of fever and recruitment of immune cells to local tissues[324,325]. IL-6 has been shown to drive inflammation in systemic infection models, contributing to tissue and organ damage. Indeed, IL-6 deficient mice have reduced survival in intraperitoneal *E. coli* infections compared to wild type mice, corresponding with lower CFU in the liver and higher numbers of neutrophils during infection[326]. The importance of IL-6 mediated neutrophil recruitment has also been shown in a pneumoniae model of infection whereby IL-6 deficient mice had increased lung CFU counts and decreased neutrophil infiltration to the lungs during infection[327].

During *E. coli* infection, IL-6 has been shown to control the sequestering of iron by promoting the production of siderophore binding molecules on the surface of macrophages, thus restricting iron sources [328].

It appears that IL-6 induced phosphorylation of the signal transducer and activator of transcription (STAT1 and STAT3) is the contributing factor to enhancing neutrophil recruitment and decreasing bacterial burdens in both the lungs and the urinary tract [327,329].

Research has shown that Gram-negative and Gram-positive bacterial infections can be identified based on levels of IL-6 and IL-8 enabling discrimination between Gram-negative and Gram positive infections. In a study on bloodstream infections in Iran, patients with *E. coli* or *Klebsiella pneumonia* infections had significantly higher levels of IL-6 and IL-10 when compared with Gram-positive bacterial infections [330]. The authors also show promising

data on the ability of IL-6 to discriminate between species patients with *Klebsiella pneumoniae* infections eliciting a significantly higher IL-6 response compared to patients with *E. coli* infections. Although limited by group size, more research is warranted in species level discrimination of bacterial BSIs.

Thus, IL-6 is a potentially important cytokine in discriminating between different species in bacterial infections, which may help to aid physicians in the rapid diagnosis and appropriate treatment of bacterial sepsis.

4.2.2 IL-8

IL-8 or CXCL8 is a chemokine that is produced by many different cell types including macrophages, monocytes, neutrophils, and epithelial cells. It is a chemoattractant to neutrophils, basophils, and T- cells and is involved in the activation of neutrophils during inflammation. IL-8 appears to be an important chemokine in driving sepsis pathology. Functional polymorphisms in the promoter regions of the IL-8 gene (IL-8 -251 A/T) have been shown to be associated with decreased risk of sepsis in a Chinese population[331]. Despite conflicting evidence of associations between risk of sepsis and gene polymorphisms, IL-8 plays a vital role in both protecting against and exacerbating infections. Increased IL-8 levels in patients with sepsis has been associated with septic shock and increased mortality[332]. More recent studies have corroborated this association, as well as showing IL-8 levels are significantly increased in patients with sepsis induced acute renal failure (ARF) and acute respiratory distress syndrome [333]. IL-8 has been shown to be a promising biomarker in neonatal sepsis,[334] as well as in predicting sepsis development in burns patients (a group of patients particularly at risk of sepsis due to the loss of the protective barrier of the skin).

4.2.3 MIP1α

MIP1α or CCL3 is a CC chemokine family member derived from a multitude of immune cells including monocytes, neutrophils, and lymphocytes. MIP1α has a primary role in chemotaxis, whereby immune cells are recruited to the site of infection, resulting in increased local inflammation. MIP1α is chemotactic to monocytes, macrophages, T cells,

neutrophils, eosinophils, basophils, DCs, NK cells and coronary endothelial cells [335]. Other effects of MIP1α include actin polymerisation, trans-endothelial migration, histamine release and granzyme release [335]. Gene knockout studies in mice have revealed the importance of MIP1α in the response to viral infections, with knockout mice having delayed clearance of Influenza virus[336]. In addition, MIP1α has been shown to be a potent M-tropic-HIV-1 replication inhibitor, through binding CCR5 (a viral entry point for HIV), and higher levels of MIP1α and MIP1β are associated with asymptomatic HIV-1 infection [337,338]. In a mouse caecal ligation and puncture (CLP) septic model, mice showed increased levels of MIP1α 8 hours after challenge [339]. In a cohort of 38 human patients 45% of patients showed increased levels of MIP1α but this did not correlate with disease progression and outcome [340]. In agreement with this, in humans, Knapp et al also showed that increased levels of MIP1α and MIP1β were found in sepsis patients with both Gram-negative and Gram-positive infections and there was no association with MIP1α and disease outcome [341].

4.2.4 MIP3α

Macrophage inflammatory protein 3 alpha (MIP3α), also known as CCL20 is a chemokine belonging to the CC family and is a strong chemoattractant of lymphocytes. The primary source of MIP3α secretion appears to be epithelial cells and it has been linked with the formation of mucosal-associated lymphoid tissue (MALT) via recruitment of lymphocytes, neutrophils, and dendritic cells (DCs) to skin and gut barriers. For instance, MIP3α promotes the recruitment of Langerhans cells (skin tissue resident macrophages) in the skin. Like MIP1α, MIP3α has been implicated in the inhibition of HIV infection in the female reproductive tract [342]. In addition, there is some evidence to suggest that MIP3α has direct antibacterial effects on a range of bacterial organisms including *E. coli* [343]. MIP1α and MIP3α both showed potential to distinguish between bacterial infections of different bacterial species in a previous study in our lab [344].

4.2.5 Resistin

Resistin is a peptide hormone belonging to the secretory protein family resistin-like molecules (RELMs), which was first discovered in rodents and found to induce insulin

resistance[345]. There are four members of the RELMs family in mice (RELMα, RELMβ, RELMγ and resistin) and two in humans (RELMβ and resistin)[345]. In mice resistin is expressed almost solely in adipose tissue, whereas in humans, expression in adipose tissue is much lower and it is expressed in other tissues including the bone marrow and the lung [346]. In contrast to mice, human resistin is predominantly expressed by monocytes and macrophages indicating a role for resistin in inflammation. Studies have shown that pro-inflammatory cytokines such as TNFα and IL-6 as well as LPS profoundly increase expression of resistin in PBMCs[347], and that resistin can induce the expression of TNFα and IL-12 in PBMCs, an effect that is mediated by NF-κB [348,349]. Resistin levels correlate with disease severity and shock in sepsis infections [350]. Increased levels of resistin have also been found in intra-abdominal sepsis [351], indicating the potential for resistin to be used as a biomarker, to discriminate between source of infection and disease severity.

4.2.6 TNFα

Tumour necrosis factor alpha (TNFα) is an inflammatory cytokine produced my macrophages and monocytes early in the inflammatory process. It was discovered in the 1970s and was called tumour necrosis factor due to its cytotoxic activity against cancer cells. [352]. TNFα belongs to the TNF superfamily, a family of type two transmembrane proteins which are released from the membrane via proteolytic cleavage. The active cytokine form can then bind to TNF receptors (TNF-R family). There are also decoy receptors which act in sequestering TNFα, to control cellular toxicity [353]. TNFα has a role in various diseases, from arthritis to tuberculosis, but has shown promise as a potential biomarker for early onset neonatal sepsis and late onset neonatal sepsis with sensitivity and specificity for early onset of 0.66, 0.76 and late onset of 0.68 and 0.87 respectively [354].

4.3 Host and bacterial factors affecting outcomes in bacteraemia

4.3.1 Host factors affecting outcomes in bacteraemia

A multitude of factors are at play in the response to *E. coli* bacteraemia and sepsis including both host and bacterial factors[13,311]. With respect to the host, Burdet et al, found that ‘bacteria from a non-urinary source’ and ‘age under 3 months’ were the primary risk factors of severity in children with bacteraemia [355]. Similarly, Lefort et al found that age,

immunocompromised status, portal of entry and previous hospitalisation before bacteraemia were the major predictors of mortality [356]. While Matrinez et al found that an immunocompromised condition, pneumonia and >7 hours to bacterial detection were the best predictors of mortality [357]. This is consistent with the risk factors of sepsis highlighted in the introduction (Table 1.2). Given the large number of conditions affecting the immune system which are considered as risk factors for sepsis it is not surprising that immune status is consistently found to be a main determinant of bacteraemia outcome.

4.3.2 Bacterial factors affecting outcomes in bacteraemia-the importance of ExPEC strain in inducing divergent responses

ExPEC can cause disease in a wide range of body sites, distal to the original infectious focus, which can lead to differential responses in host immune response. ExPEC isolates are a genetically diverse group of pathogens, each bacterial typically possess ~5000 genes but these genes can vary by 20-30% between individual strains [358]. These genes can impact on bacterial survival, fitness and virulence and animal studies have noted the importance of genes involved in; adhesion, iron exploitation, metabolism, transport, toxins, LPS and capsule proteins are of importance influence in extra intestinal infections [358–360][361].

With respect to the bacteria factors influencing bacteraemia, Gram-negative bacteria elicit greater CRP and IL-6 responses from the host compared to Gram-positive bacteria [362]. In addition, ExPEC isolates induce higher inflammation and are more likely to result in mortality compared to Gram-positive bacteria and other Gram-negative bacteria [363]. Additionally, studies have also linked specific genes to disease outcomes. For instance, patients infected with isolates carrying the *cnf* and bla_{TEM} genes had increased risk of severe illness by 6.75 and 2.59 times respectively [364].

Barber et al used zebrafish to assess the divergent host responses to two similarly lethal ExPEC strains, CFT073 and F11, which both cause death in zebrafish embryos by 24 hours. The isolates encode the K2 capsular antigen and O6 surface antigen but possess different flagellar serotypes. The F11 infections were associated with the development of more pronounced pathology including pericardial edema, ulceration, and erosion of the tail fin, compared to the CFT073 strain where these phenotypes were less common. F11 infection was also associated with tissue perfusions in 50% of the fish but none were observed in fish

infected with CFT073. Transcriptional analysis showed that infection with F11 affected the expression of a third more genes than CFT073 including some transcripts that were suppressed in CFT073 infected fish [363]. The authors show that the two different isolates can elicit divergent immune responses from the host; for instance, INF-γ2 is induced by F11 but not CFT073. In contrast, higher levels of IL-10 and IL-8 are induced in response to CFT073 than F11 [363]. These results indicate a large degree of variability in the host response to bacteria dependent on the background of the bacteria itself.

ExPEC strains also produce differential responses in adaptive immunity. O'Brian et al 2018 have shown that two urosepsis strains which are common model organisms for UTI infections in mice have drastically different kinetics of bacterial clearance as well as susceptibility to reinfection. Mice infected with the strain UTI189 could be infected indefinitely, being unable to clear infection. However, mice infected with CFT073 cleared the infection within 8 weeks. In addition, reinfection of mice treated with antibiotics with strain UTI189 resulted in recolonisation of the urinary tract whereas, mice were protected against UTI re-infection with CFT073. Interestingly T cells were found to be the determining factor in mouse immunity to recurrent infections. Depletion of CD4+ and CD8+ T cells from mice abolished the protection against re-infection with CFT073. Such data confirms different responses of the adaptive immune system to individual *E. coli* strains [365].

Finally, it is becoming increasingly clear that even commensal *E. coli* can promote inflammation that may compromise gut integrity and enhance translocation. In a mouse model of *E. coli* infection Kittana et al found that commensal *E. coli* of different sequence types elicited distinctive immune responses after inflammatory stimulus[365]. Isolates belonging to the sequence type ST129 and ST374 were associated with the production of inflammatory cytokines GM-CSF, IL-6 and IFN-γ. in contrast, isolates belonging to the sequence types ST150 and ST468 were associated with mild inflammation and low levels of pro-inflammatory cytokines [356]. The authors illustrated that IL-6 was the main driver of disease severity.

4.4 Aims and objectives

The significant evidence for differential responses observed to *E. coli* and in particular ExPEC strains in animal models raises the question about the diversity of responses that might be detected in response to the bacteraemia collection described in Chapter 3 in human models. The aims of this chapter are to investigate the cytokine responses generated by the collection of *E. coli* isolates identified in Chapter 3 and whether these responses could be used to discriminate certain bacterial phenotypes. Specific objectives are to:

- Investigate panel of potential cytokine biomarkers using *ex vivo* whole blood infection models of bacteraemia. This panel will be used to characterise the host response to *E. coli* isolates from varying origins of infection.
- Investigate host response to *E. coli* collection using THP-1 model of infection.
- Characterise *E. coli* collection survival and growth in human plasma
- Assess serum bactericidal capacity of human serum on select *E. coli* isolates.

4.5 Methods

4.5.1 *Ex-vivo* whole blood model of infection.

Whole blood from healthy volunteers was infected with *E. coli* isolates for four hours before supernatants were removed and stored for cytokine analysis (Chapter 2.4.1).

4.5.2 THP-1 infection model

THP-1 cells in an undifferentiated state were used for infection studies and 500000 cells per well were infected with *E. coli* as per methods (Chapter 2.4.2). Cytokine responses of IL-8 and TNFα were assessed using ELISAs (R&D Systems) as per manufacturer's instructions (Chapter 2.4.3).

4.5.3 Plasma resistance in *E. coli*

Overnight *E. coli* cultures were corrected to an OD600nm of 0.1 in LB. Bacteria were then diluted 1/10 in citrated human plasma (TCS biosciences) in a 96 well plate. OD 600 was determined every hour for 24 hours (Chapter 2.4.6). For appendix growth in serum analysis

the same protocol was followed but citrated human plasma was replaced with serum from gold top blood collection tubes.

4.5.4 Serum bactericidal assay

The importance of complement in immunity to *E. coli* was tested in serum bactericidal assays in which complement C5a was specifically depleted. C5a was depleted with 10B6 antibody (a gift from Prof Paul Morgan, Cardiff University) and compared to C5a depleted serum (also a gift from Paul Morgan, Cardiff University).

4.6 Results

Host responses of the *E. coli* bacteraemia strains and controls were determined and organised as; i) whole blood responses to *E. coli* (4.6.2-4.6.6), ii) THP-1 responses to *E. coli* (4.6.7 – 4.6.8) and iii) plasma resistance, killing responses to *E. coli* (4.6.1 and 4.6.8, 4.6.9)

4.6.1 Whole blood response to *E. coli* isolates

The *ex vivo* whole blood infection model was used to investigate the host response to *E. coli.* A panel of cytokines identified in a previous PhD were determined; IL-6, IL-8, MIP-1α, MIP-3α and resistin (Figure 4.1)[344]. These cytokines were measured 4 hours post infection with *E. coli* isolates from bacteraemia (n=84), water sources(n=7), references isolates (n=4) and K12 isolate controls (n=2). Most isolates elicited a significantly increased response ($p<0.0001$) in all the cytokines tested 4 hours post infection with the exception of Resistin, compared to control untreated blood (Figure 4.1). All cytokines except for resistin had levels approaching limit of detection for ELISA assays (IL-8, MIP1α, resistin 31.2pg/mL, MIP3α 7.2 pg/mL and IL-6 3.1 pg/mL) in pre-infected (T=0) and control (RPMI) treated blood. When cytokine responses were grouped from all the tested isolates (Figure 4.1), the IL-6 responses had the highest concentration (50-100,000pg/ml). Each biomarker had a distinctive distribution with IL-6, MIP-1 and resistin having wide ranges (as confirmed by standard deviations of 13682.89, 10996.32 and 11332.28 pg/mL respectively). In contrast, IL-8 and

MIP-3 had more discrete ranges (as confirmed by standard deviations of 4781.87 and 4556.73pg/mL respectively) (Figure 4.1).

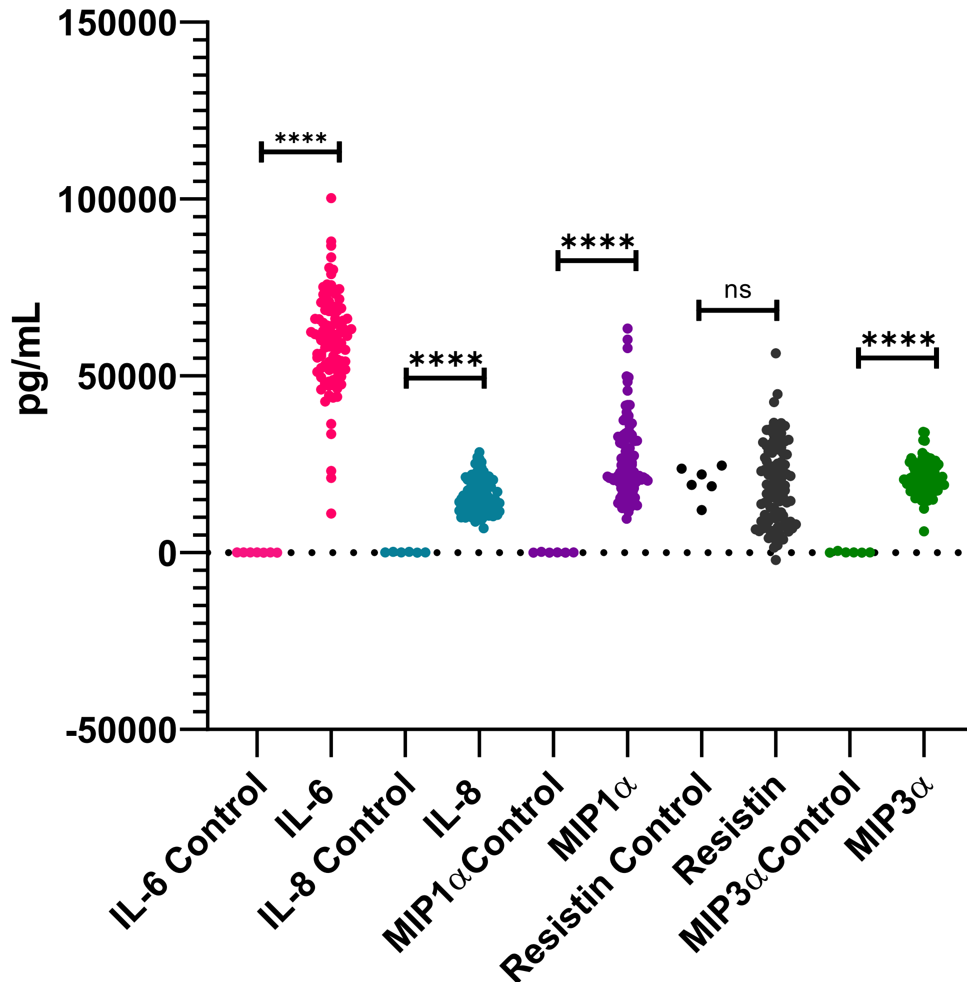


Figure 4.1: Cytokine responses to *E. coli* isolates in an *ex vivo* whole blood infection model.
Healthy volunteer whole blood was infected with *E. coli* isolates for 4 hours before the platelet poor plasma was isolated for ELISA determination. Isolates included in the blood infection model were 85 clinical bacteraemia isolates, 7 water isolates, lab K12, Keio parent K12 and four reference isolates. Dots represent average of at least 2 separate donors for each isolate. *E. coli* induced cytokine responses compared to control untreated cytokine levels with Mann-Whitney test ****p ≤0.0001.

4.6.2 Whole blood response to bacteraemia/non-pathogenic *E. coli*

E. coli isolates were grouped as bacteraemia and non-pathogenic, where non-pathogenic included the laboratory strain K12 as well as *E. coli* from environmental water sources and reference isolates; ATCC 35218, ATCC 25922, NCTC 1093 and NCTC 9001. Then, cytokine

responses in whole blood were reassessed (Figure 4.2a-e). Host IL-6, MIP-1α and Resistin responses in whole blood were not significantly different between *E. coli* from patients with bacteraemia or non-pathogenic isolates (Figures 4.2a, 4.2c and 4.2e respectively). In contrast, significant differences were observed between bacteraemia and non-pathogenic groups in the amount of both IL-8 and MIP3α produced (Figures 4.2b and 4.2d respectively).

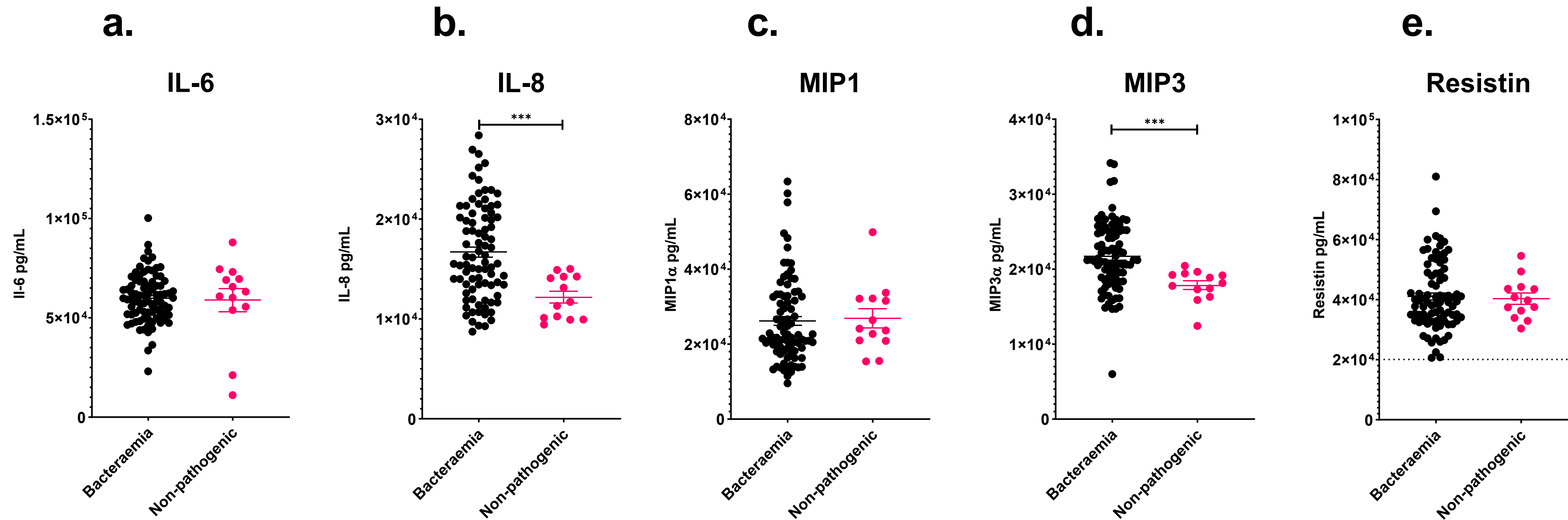


Figure 4.2 Host cytokine response in an *ex vivo* whole blood infection-effect of pathogenicity.

Whole blood from healthy volunteers was infected with *E. coli* isolates from either bacteraemia or non-pathogenic source for 4 hours and the platelet poor plasma isolated prior to host cytokine determination by ELISA. (***p ≤0.0005, Mann-Whitney test) a. IL-6, b. IL-6, c. MIP1α, d. MIP3α, and e. Resistin. Each point on the graph represents the average of at least two different blood donors. Error bars show mean ± SEM. Resistin background control levels indicated as line on graph.

4.6.3 Whole blood cytokine response to sepsis and non-sepsis *E. coli* isolates

Using patient hospital data from Chapter 3, cytokine data was reassessed, and isolates were grouped by the diagnosis of sepsis or non-sepsis (Figure 4.3). No significant differences were observed in any of the cytokine responses when grouped in this manner (Figure 4.3).

4.6.4 Host response to source of infection

In order to assess the ability of the cytokine panel to discriminate between *E. coli* from different sources of infection the patient hospital data was again used to group data by clinically defined source of infection (Figure 4.4a-e). No differences could be detected in IL-6, MIP-1α and resistin responses when grouped by source of infection and *E. coli* controls (Figure 4.4a, c and d). IL-8 responses grouped by source of infection and *E. coli* controls showed a significant increase between the urinary or abdominal isolate responses and those from water isolates (Figure 4.4b). Consistent with this result MIP-3α responses from abdominal isolates were significantly increased compared to responses from water isolates (Figure 4.4e).

4.6.5 Host response to mortality

No significant differences were observed for any of the cytokines tested when isolates were grouped by mortality (Appendix 4.1).

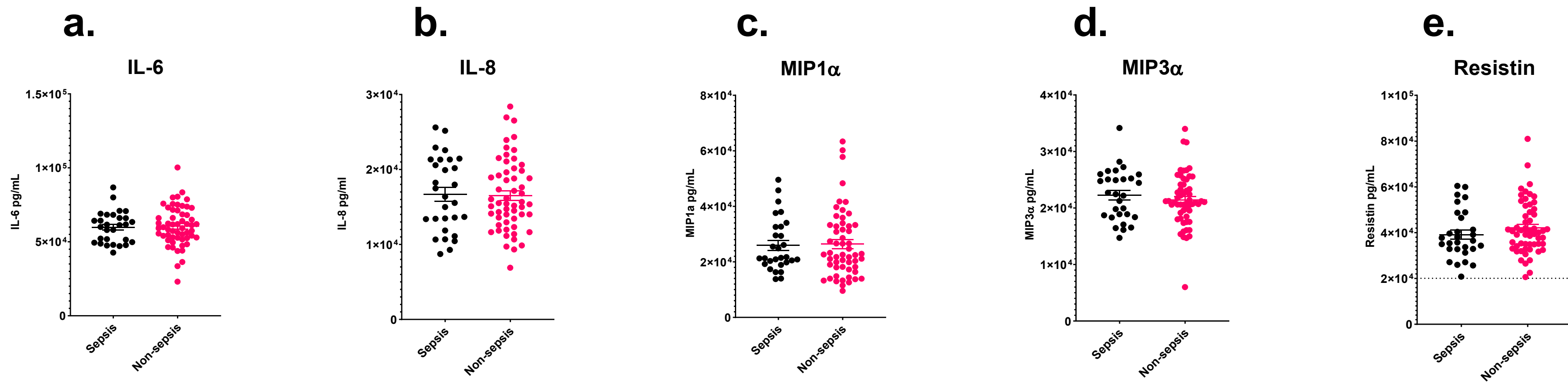


Figure 4.3: Host cytokine response in an *ex vivo* whole blood model-role of sepsis.

Whole blood from healthy volunteers was infected with *E. coli* isolates from a bacteraemia source for 4 hours and the platelet poor plasma isolated prior to host cytokine determination by ELISA. Data was then grouped using patient data from chapter 3 as 'sepsis' or 'non-sepsis' and reanalysed. a. IL-6, b. IL-8, c. MIP1α, d. Resistin, and e MIP3α. Each point on the graph represents the average of three independent measurements. Data shows the mean ± SEM. No significant difference was detected between groups Mann Whitney test.

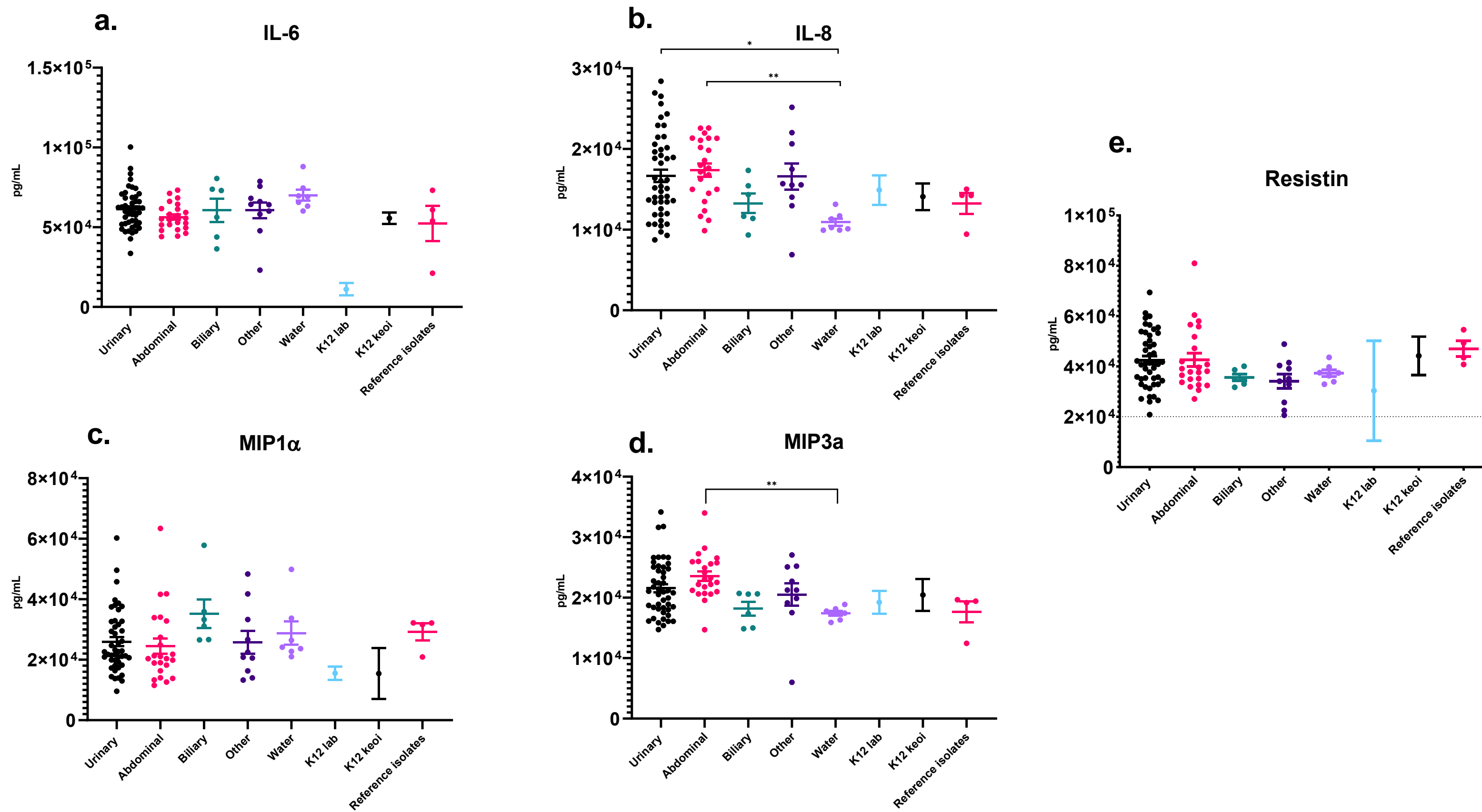


Figure 4.4: Host cytokine response to an *ex vivo* whole blood model- role of source of infection.

Whole blood from healthy volunteers was infected with *E. coli* isolates from either bacteraemia or non-pathogenic sources for 4 hours and the platelet poor plasma isolated prior to host cytokine determination by ELISA. Data was then grouped using patient data from chapter 3 by source of infection and by *E. coli* control isolates and reanalysed. a. IL-6, b. IL-8, c. MIP1α, d. MIP3α, and e. Resistin. Each point on the graph represents the average of three independent measurements. Data shows the mean ± SEM (One-way ANOVA with Kruskal-Wallis test* p ≤0.05, **p ≤0.01). Average background levels of Resistin indicated on graph by dotted line (20060pg/mL).

4.6.6 Cytokine responses to *E. coli* isolates in THP-1 cells

The Covid pandemic restricted the availability of healthy volunteer blood for the *ex vivo* infection model. Therefore an *in vitro* corollary was developed by infecting THP-1 monocytes with *E. coli* and the early cytokines IL-8 and TNFα determined. On average, *E. coli* treatment elicited significantly higher IL-8 and TNFα compared to untreated controls (Figure 4.5). IL-8 and TNFα showed unique response distributions, (Figure 4.5). IL-8 had a significantly greater range than TNFα as confirmed by their standard deviations (1168 vs 328). In addition, there was a group of isolates identified which elicited low levels of IL-8 (<500pg/mL), and TNFα (approaching limit of detection), in the THP-1 cells (Figure 4.5). To investigate the reason for this low IL-8 production, Trypan Blue cell viability assays were carried out on a subgroup of high and low cytokine producers (Figure 4.6). This analysis confirmed that isolates eliciting a low cytokine response in THP-1 cells were the most cytotoxic to the cells, indicating that the cells were dying before an appropriate immune response could occur. In contrast isolates that elicited a high cytokine response had lower amounts of cytotoxicity (Figure 4.6). This was also the case for the K12 positive control that elicited a cytokine response but was not cytotoxic (Figure 4.6). No significant differences were detected in IL-8 and TNFα production when isolates were grouped by source of infection (Figure 4.7) despite a wide range of responses being observed for IL-8 production (Figure 4.7a.)

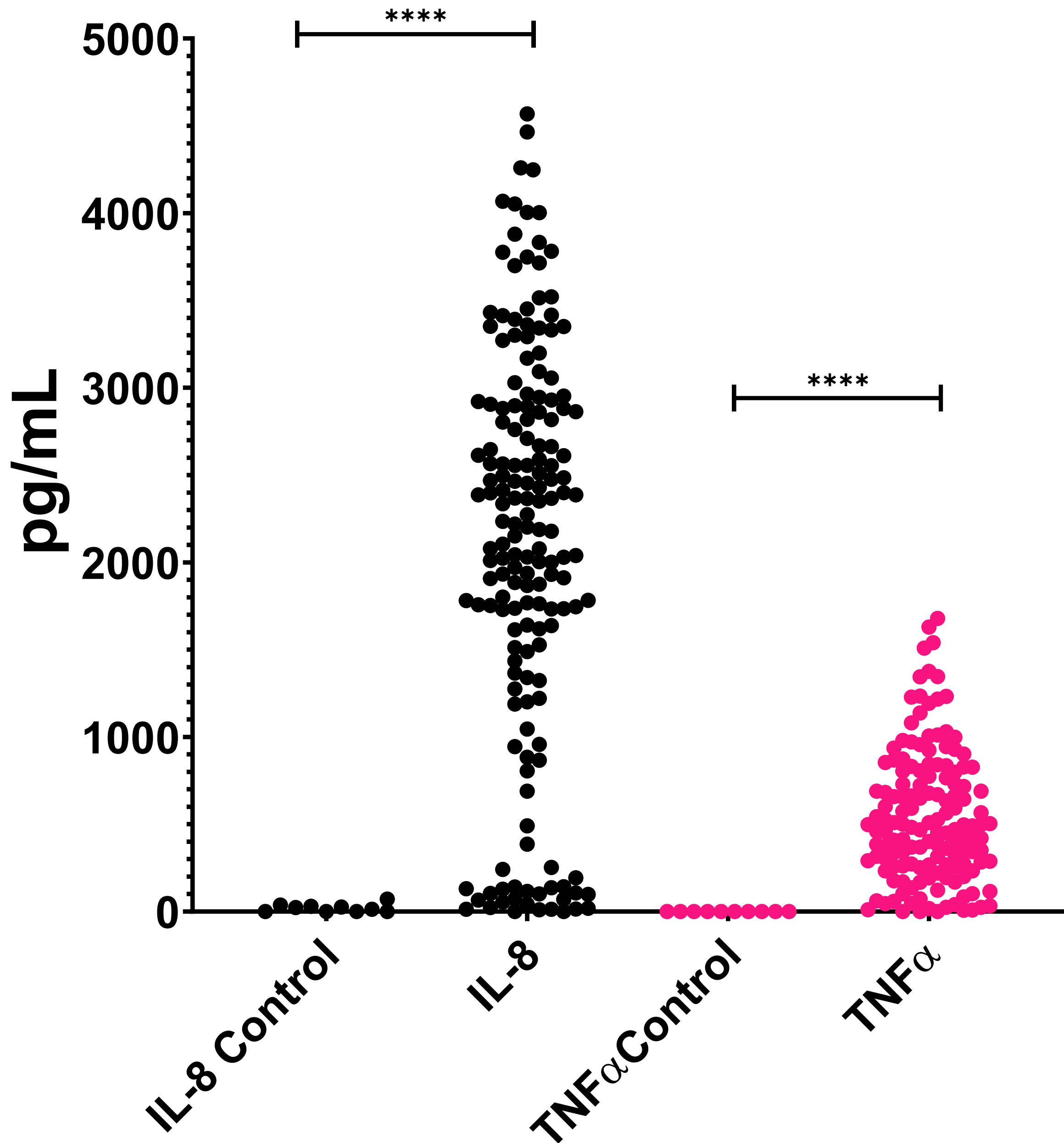


Figure 4.5 Distribution of *E. coli* induced IL-8 and TNF responses in THP-1 cells

THP-1 cells were infected with *E. coli* isolates for 4 hours before supernatant was removed for ELISA determination. Isolates included in the THP-1 infection model were 165 bacteraemia isolates, 7 water isolates, Keio parent K12 and four reference isolates and untreated controls. Dots represent average of three independent experiments. Control untreated cytokine responses compared to *E. coli* treated responses with Mann-Whitney test ****p ≤0.0001.

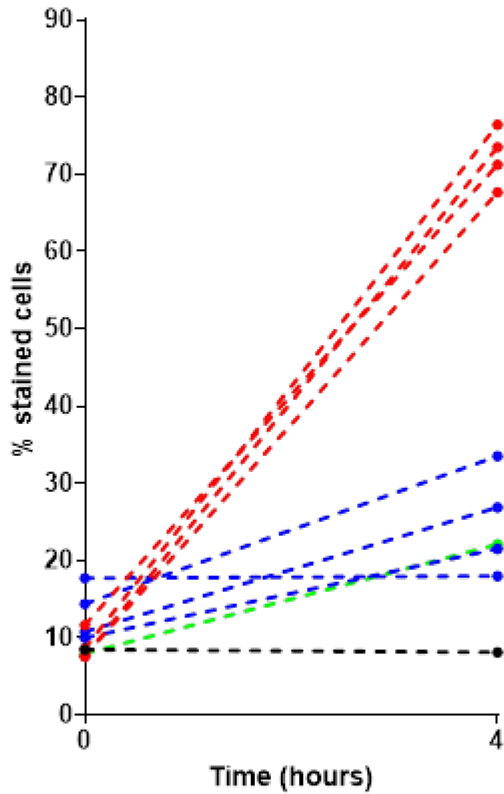


Figure 4.6 THP-1 viability in high and low cytokine producing bacteraemia strains
THP-1 cells (blue). THP-1 cells were subject to infection model as per methods (section 2.4.2) Four isolates with low cytokine responses (red), four isolates with high cytokine responses (blue), a K12 control (green) and untreated control (black) were included. Cells were stained with Trypan blue after 4 hours infection and data presented as % stained (dead) cells.

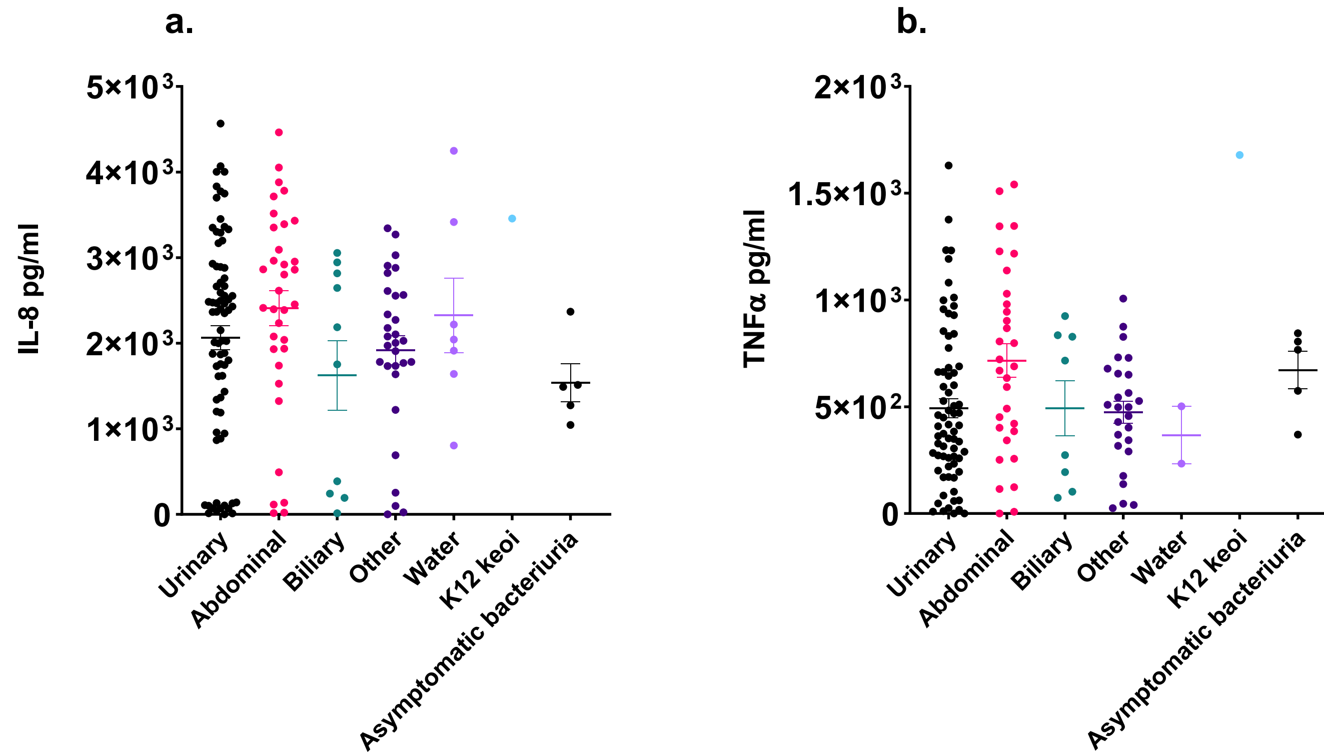


Figure 4.7: *E. coli* induced cytokine responses from THP-1 cells organised by source of infection

THP-1 cells were infected with *E. coli* for 4 hours and the supernatant was removed for cytokine determination by ELISA Data was then grouped using patient data from chapter 3 by source of infection and by *E. coli* control isolates and reanalysed. a. IL-8 and b. TNFα. Each point represents the average three independent experiments. Data shows the mean ± SEM and are corrected after subtracting values from control untreated wells. No significant difference detected $p>0.05$ (One-way ANOVA with Kruskal-Wallis test).

4.6.7 Plasma resistance of the *E. coli* collection.

Serum resistance is a defining trait in ExPEC bacteraemia and therefore resistance to human plasma was investigated (plasma was used as a proxy for serum due to commercial availability at the time) in the collection. In total 179 *E. coli* isolates were grown in human plasma (made using sodium citrate), but one K12 isolate, and three asymptomatic bacteriuria isolates were not included in this analysis. *E. coli* isolates were grown in 90% plasma for up to 24 hours and OD 600nm readings were taken every hour (example curves, Figure 4.8). There was a large spread in the distribution among the isolates growing in plasma with some isolates being able to grow early (within 2 hours) while others only grew after a prolonged lag phase (Figure 4.8) or were unable to grow in plasma (Figure 4.8). To account for this variability growth curves were analysed using the Analysis of area under curve (AUC) by assigning *E. coli* isolates and arbitrary number 1-168 and using the slope function in excel on each of the OD readings during the 24-hour time points. AUC analysis (Figure 4.9) confirmed the variability amongst isolates. Growth of the isolates in plasma ranged from those unable to grow and hence low/no gradients (thus lower AUC scores) and those that can grow with higher gradients and those in-between. This pattern was consistent with experiments where *E. coli* isolates were grown in human serum (gold top blood tubes) (Appendix 4.2). Using the growth in plasma data isolates were grouped as resistant (n=112), and sensitive (n=57). Isolates were determined resistant if any increase in OD was observed in 2 out of 3 replicates in growth curve experiment

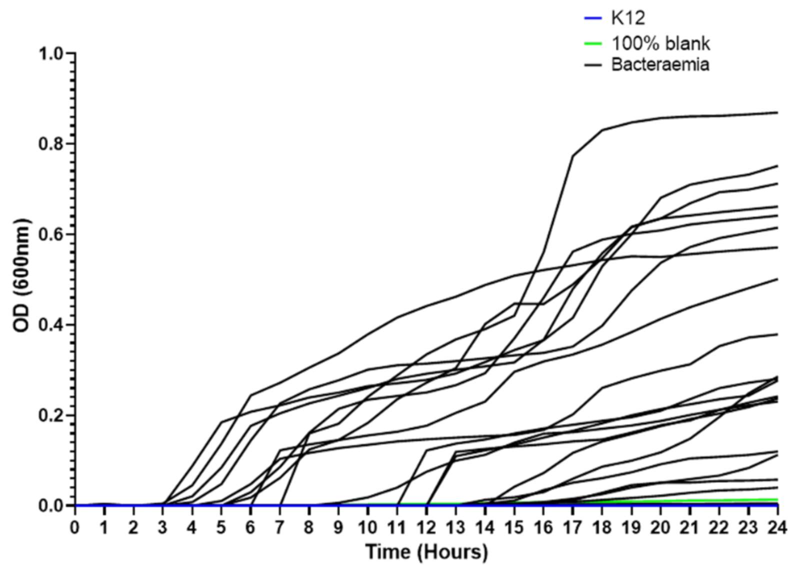


Figure 4.8: Selection of *E. coli* growth curves

Representative selection of *E. coli* isolates grown in 90% citrated human plasma for 24 hours. 100% blank indicates 100% plasma, untreated with *E. coli.* Each line represents the mean of three independent experiment (n=30)

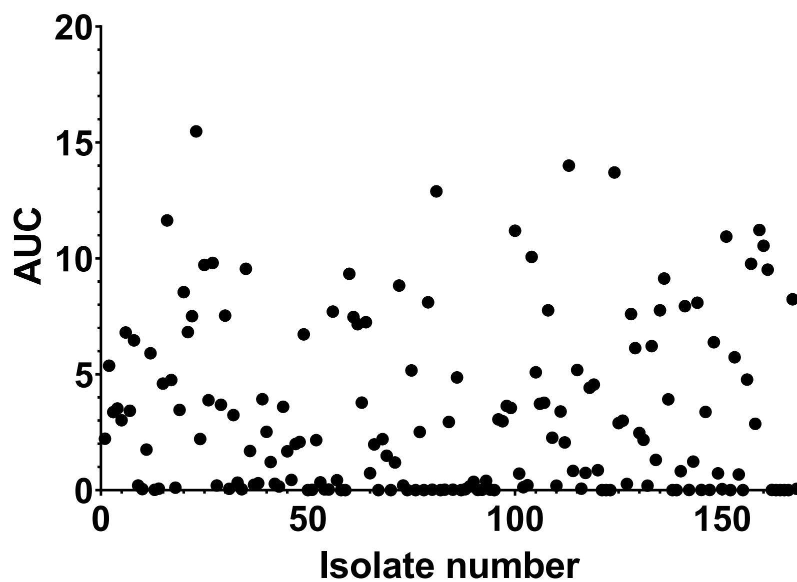


Figure 4.9: Logistic Area under curve (AUC) analysis of *E. coli* growth curves.

Logistic AUC calculated using Growth curver R package. Isolates were assigned arbitrary number 1-168 for X axis.

4.6.8. Interrelationship between whole blood responses and plasma sensitivity

The cytokine data was now reassessed in isolates grouped by their capacity to grow in citrated human plasma (Figure 4.10). Significantly higher levels of IL-6, IL-8 and resistin were found in response to isolates that were serum resistant, compared to serum sensitive isolates (Figures 4.10a, b, and e respectively). No significant differences were observed in the levels of MIP1α and MIP3α in response to the serum resistant isolates, however there was a trend towards increasing levels of these cytokines in response to infection with these isolates.

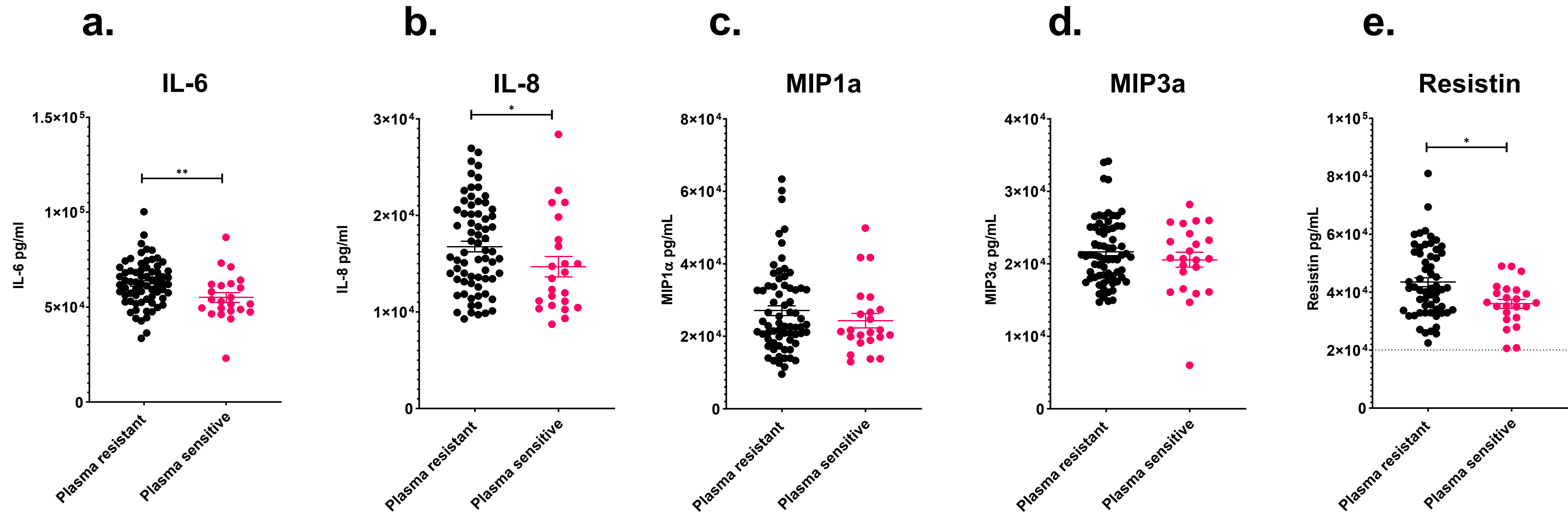


Figure 4.10: Host cytokine response in *ex vivo* whole blood infection model to plasma resistant/sensitive *E. coli* isolates.

Whole blood from healthy volunteers was infected with *E. coli* isolates from either bacteraemia or non-pathogenic sources for 4 hours and the platelet poor plasma isolated prior to host cytokine determination by ELISA. Data was then grouped using phenotype data from chapter 4 as plasma sensitive and plasma resistant and reanalysed. a. IL-6, b. IL-8, c. MIP1α, d. MIP3α, and e resistin. Each point on the graph represents the average of three independent measurements. Data shows the mean ± SEM. (Mann-Whitney test, ***$p \leq 0.0005$). Resistin background levels are indicated on graph with line.

4.6.9 Serum bactericidal assay (SBA) assay

The previous experiments determining serum sensitivity and resistance of the *E. coli* collection stimulated two new research questions which were beyond the scope of the thesis but were initiated towards the end of the research work. They were. 1) Is there a high throughput bactericidal assay for serum killing? 2) What is the role of complement in serum killing of the *E. coli* collection? Here preliminary data based on a selection of isolates, is presented on a new serum bactericidal assay and the use of complement specific reagents to investigate mechanism.

Three serum resistant (23, 30 and 110, blue) and sensitive isolates (K12, 9 and 131, orange) were used in this investigation (Figure 4.11). *E. coli* isolates from the serum resistant group were able to survive in untreated human serum and were unaffected by either C5 specific antibody or being grown in C5 depleted serum (except for strain 110 which had slightly improved survival). In contrast the sensitive isolates tested were unable to survive to significant levels in the untreated serum. Their growth was rescued by culture with a C5 antibody or with C5 deficient serum. All *E. coli* grew poorly in the assay buffer HEPES.

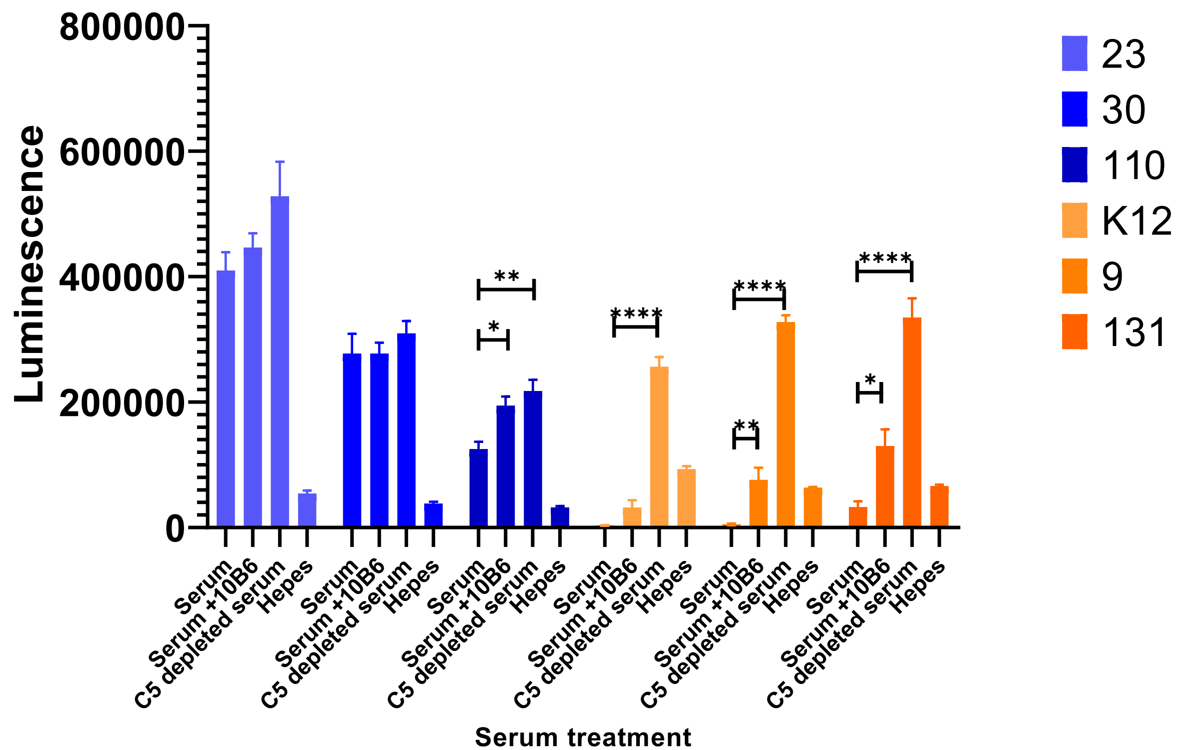


Figure 4.11: Serum bactericidal assay.

Three serum resistant isolate (23, 30, 110) and three serum sensitive isolates (K12,9,131) were corrected to an OD of 0.2 before being resuspended in HEPES buffer. Isolates were then exposed to serum with various treatments (1:1 with HEPES), untreated, 10B6 aB treated and C5 depleted or HEPES alone for 2 hours before being mixed with BacTitre-Glo. Data represent the average luminescence of three independent experiments ± SEM (one-way ANOVA with Tukey test, *p= ≤0.05, **p= ≤0.01, ****p= ≤0.0001).

4.7 Discussion

Sepsis represents a major challenge to healthcare professionals worldwide due to the complex nature of disease pathology. The interplay between each *E. coli* strain and each unique host results in complex infection dynamics. This chapter aimed to determine; i) the cytokine responses to the *E. coli* collection in whole blood; ii) cytokine responses to the *E. coli* collection in THP-1 cells and iii) the plasma resistance / sensitivity profile of the *E. coli* collection. A subset of the isolates was included in the *ex vivo* whole blood infection before an *in vitro* THP-1 model was developed due to COVID restrictions.

Host associated biomarkers to help distinguish between *E. coli* from different sources of infection would be greatly beneficial in guiding correct antibiotic therapies. In addition,

many patients suffering from *E. coli* sepsis will already be hospitalised or have previous medical histories associated with local *E. coli* infections (e.g., UTI). Indeed, UTI infections and subsequent development of sepsis account for approximately 25% of all cases [366]. Accurately identifying links between, the cytokine response, the type of *E. coli,* and the different sources of infections would be highly advantageous in guiding antibiotic therapies. To that end this chapter aimed to determine the cytokine profile induced by the *E. coli* bacteraemia isolates from patients in the Hywell Dda University Health Board.

All cytokines tested in the whole blood model were significantly increased compared to untreated controls (excluding resistin), indicating that they are good markers of *E. coli* infection in this model. This was expected due to the inherent inflammatory nature of adding bacteria to human blood. Isolates that caused bacteraemia elicited a higher IL-8 and MIP3α cytokine response in healthy volunteer blood, compared to non-pathogenic control organisms (Figure 4.4 b. and d.). Higher IL-8 responses have been observed in isolates that are able to traverse epithelial barriers and cause infection in other sites, with increased IL-8 production in two epithelial cell lines Caco-2 and HT-29 cells as well as THP-1 monocytic cells [367]. Since crossing intestinal barriers (the natural environment for *E. coli)* is essential for the development of sepsis this could be a useful marker for differentiating between strains of *E. coli*.

IL-8 may have uses in syndromes associated with sepsis. Indeed, many sepsis patients have pre-existing conditions that can result in the so-called leaky gut syndrome[368]. This is where gut barrier integrity is compromised, allowing normal commensal bacteria to spread into the bloodstream and to other organs and tissues. It is interesting to speculate that commensal *E. coli* escaping in this manner may elicit a lower IL-8 response in the blood. IL-8 could therefore be a useful marker in distinguishing from pathogenic and usually non-pathogenic *E. coli* isolates. Indeed, combining information from the host (low IL-8) with information about the bacteria may also prove fruitful. For instance, non-pathogenic isolates tend to belong to the phylogroups A and B1 and isolates that go on to cause sepsis in immuno-compromised individuals, usually have a higher number of antimicrobial resistance genes (e.g., increased resistance to ciprofloxacin, and higher β-lactamase production) [369]. This could be a potential avenue of guiding antimicrobial therapies to counteract these resistant isolates in immunocompromised individuals. However, other

studies have indicated high degrees of antimicrobial resistance in isolates belonging to the B2 and D phylogroups [370], so care should be taken, especially with the increasing incidence of antimicrobial resistance amongst all *E. coli* strains. It is also not necessarily the case that all isolates from immunocompromised patients are less virulent than those found in patients with lower comorbidities. For instance, there were no significant differences in cytokine production when isolates belonging to phylogroups B2 and D were compared to B1 and A. (Appendix 4.3). This is potentially due to the low number of isolates that came from these phylogroups from the isolates used in this study. These results confirm the complex nature of the sepsis syndrome.

Given the predominant amount of urinary source isolates within the *E. coli* collection tested, this indicated that isolates from the urinary tract may elicit more IL-8 and MIP3α compared to control isolates. This was true for IL-8 (Figure 4.4), which was found to be significantly increased in response to the isolates from a urinary source of infection. However, the isolates with the greatest significant differences in IL-8 production compared to controls were the abdominal isolates. In addition, MIP3α levels for abdominal isolates were significantly higher than the environmental water isolates but this difference was not identified for the urinary source. This could potentially be due to the limited number of isolates from both an abdominal source and water isolates and hence, less variation in the whole blood model. Further investigation into the mechanisms of immunity against *E. coli* from different sources would allow a greater understanding of this observed effect.

Although IL-8 has been investigated substantially as a biomarker for sepsis, less research has been done on MIP3α. Although levels were shown recently to be elevated in non-surviving sepsis patients compared to survivors [371]. Interestingly, there was no relationship between disease outcome and cytokine responses to *E. coli* isolates found in our study (Appendix 4.1). However, there are also some indications that MIP3α may be useful in other disease states such as Biliary Atresia, a disease of the liver and bile ducts that occurs in infants [372]. More research into MIP3α and IL-8 as potential biomarkers to discriminate between *E. coli* isolates from different sources of infection is warranted.

No significant differences were found between isolates in the THP-1 model of infection (Figure 4.8). COVID and laboratory access restricted research to two cytokines being tested in this model. This data did not corroborate the significantly different IL-8 response seen in

the blood infection model (Figure 4.6). This may be due to the increased complexity and number of cell types and other inflammatory factors involved in the immune response in blood. For example, neutrophils are known to produce IL-8 early on in infection, although studies have shown that monocytes secrete more IL-8 (per cell) than neutrophils [373]. However, despite differences in cell output there is a higher abundance of neutrophils in the blood than monocytes-the relative contribution of each cell is difficult to assess.

Isolates that were resistant to human plasma showed higher cytokine responses than isolates that were sensitive to plasma. Significantly higher levels of IL-6, IL-8 and Resistin responses were found against isolates that were resistant to plasma (Figure 4.4). This is interesting as it suggests a greater immune response is mounted to resistant isolates in comparison to sensitive isolates. This may be due to the resistant isolates surviving longer in the blood and therefore having a higher immunological burden than isolates that are sensitive that get lysed by complement or phagocytosed before a large immune response can occur. Indeed, cytokine levels in the kidneys have been found to correlate directly with *E. coli* CFU count in acute pyelonephritis[374].

Whole blood models have significant advantages and limitations. Whole blood models of infection allow the measurement of human cytokine responses to live bacterial infections to be elucidated without requiring the use of animal models. The models require very little processing, for molecular analysis by ELISA and cellular analysis by flow cytometry of surface markers. Therefore, whole blood models of sepsis could help reduce the use of animal models of sepsis. There is not a truly clinically relevant animal model of sepsis infection and hundreds of different models have been tested [375–377]. Mice are the most used models of sepsis, and the genetic similarity of current mice models allows for homogenous research, and available genetically modified lines allow for testing the contributions of individual genes to disease. They are also relatively inexpensive and easy to source. However, mice are significantly less sensitive to LPS than humans. The common dosage used in mouse LPS models of sepsis is between 1-25mg/Kg [378,379]. The LD50 dose in mice is 1000-10000 times greater than the dose required to induce severe illness in humans [380].

The clinical relevance of whole blood models may therefore be more useful than animal models in determining biomarkers in sepsis. With the addition of immunomodulatory agents (e.g., cytokines or TLR agonists etc), researchers can also mimic the

immunosuppressive state present in many patients before the onset of sepsis. Using an *ex vivo* whole blood model also removes the need for human endotoxemia models of sepsis infections, which can be beneficial for studying the early stages of sepsis [381,382], do not model sepsis to the full extent and have limitations such as the maximum dose of LPS that can be used due to toxicity concerns[383].

However, there are some potential disadvantages of using ex-*vivo* whole blood models in biomarker research, the most problematic being the use of anticoagulants. These are necessary to prevent clotting however, they can interfere with signalling processes, particularly those than involve the complement system. Heparin, a common anticoagulant inhibits membrane fixation of complement component C3 at high concentrations and may even enhance complement activity at lower concentrations [384]. This can affect cytokine responses as well as interfere with the coagulation cascade which is important in sepsis infections. The use of anticoagulants is, however, unavoidable due to the short shelf life of blood outside the body. There are, however, potential different options for using anticoagulants that are more specific and thus have lesser effects on the complement and coagulation cascades, e.g., hirudin is a specific inhibitor of thrombin. Hirudin is more expensive than traditional anticoagulants such as sodium citrate and heparin and is often not available in large quantities making it difficult to use in large scale studies [385].

Whole blood models also require healthy volunteers for blood donations which can result in biomarker variation. For this study there were no significant differences in the control samples between donors for most of the markers tested. However, for resistin there was some variability in baseline levels and so the average baseline was shown on the graph. Another disadvantage of using whole blood is that although it is a good model for systemic infection it does not offer any information on local inflammatory events that precede sepsis. Given timing difficulties faced because of COVID-19 it was not possible to look at levels of inflammation in cell types relevant to the origin of infection (e.g gut and urinary epithelial cells). This may be an encouraging direction for a future PhD studentship.

The importance of complement proteins in the host response to *E. coli* was investigated using a serum bactericidal assay. This assay tested the growth of plasma sensitive/resistance bacteria in complement complete and two forms of complement inactive serum (C5 depleted and mAb treated) (Figure 4.9). All sensitive strains had increased numbers of

bacteria in the serum depleted treatments with both the 10B6 mAb treated and the C5 depleted serum having more bacteria than the active complement treated wells (Figure 4.9). This highlights the importance of complement in confining bacterial growth in blood during early infection stages. Given that this restriction may not occur in whole blood models of infection that use anticoagulants (which interfere with complement), it may be that serum sensitive isolates have cytokine levels similar to resistance isolates solely based on bacterial dose

Data from this chapter indicated that cytokines have the potential to discriminate between *E. coli* from different sources of infection and based on bacterial traits such as resistance to serum/plasma. Using cytokine data to identify such information during disease has significant time advantages over traditional methods of acquiring this information, such as those relying on bacterial culture. Advances in cytokine detection technology such as AlphaLISA allows detection of cytokines within 15-20 minutes[386].. A marked improvement in the hours/days needed for culture based methods and a critical consideration in time sensitive diseases such as sepsis.

This chapter investigated host responses and potential biomarkers from the host in response to *E. coli* bacteraemia, the next chapter will focus on genome associated biomarkers and virulence factors from *E. coli*.

Chapter 5 Descriptive and comparative genomics of virulence factors and AMR genes in *E. coli* bacteraemia in Southwest Wales

5.1 Introduction

5.1.1 *E. coli* as a commensal and pathogen

E. coli is a highly diverse Gram-negative bacteria belonging to the family *Enterobacteriaceae* and the genus *Escherichia*. The Escherichia family consists of three closely related members *E. albertii*, *E. fergusonii* and *E. coli.* In addition, five clades are recognised, which are phenotypically undistinguishable from *E. coli* named Clade I-V [387]. *E. coli* can be grouped into seven main phylogroups namely A, B1, B2, C, D, E and F as well as the cryptic clade 1 which appear phenotypically indistinguishable [388]. Commensal *E. coli* are often associated with clade A and B1. Intestinal pathogenic *E. coli* (IPEC) often associated with clades A, B1 or D, whereas extraintestinal pathogenic *E. coli* (ExPEC) are often associated with clades B2 and D. To date, *E. coli* is still the most studied bacteria in microbiology and infectious disease[389], and a large amount of information is available about its genetic make-up. Indeed, over 13000 genes have been identified across thousands of strains of *E. coli* and has stimulated research into new concepts in comparing and analysing genomes.

The concept of a pangenome was introduced in 2005 and comprises all known genes from all strains within a clade, thereby representing the potential whole gene repertoire of the bacteria under study[390]. The pangenome can be divided into genes that are shared between each strain and are usually required for cellular survival (the core genome) and genes that vary between each individual strain either being present or absent (accessory or variable genome)[391]. The *E. coli* pangenome is highly diverse, in any given strain the number of genes shared with all other *E. coli* strains is approximately 1000, while non-shared/unique genes number in the tens of thousands [392]. Interestingly, with the number of *E. coli* genomes ever increasing, the number of core genes appear to be decreasing with time, whilst the number of accessory genes is increasing, a trend expected to continue as more and more genomes are sequenced[392–396]. The large *E. coli* pangenome reservoir (increasing accessory genome) allows for the large range of diversity seen in the species and highlights the implications this diversity has on the pathogenesis of the species [397–399].

5.1.2 Genetic traits of ExPEC isolates

5.1.2.1 Virulence factors

E. coli is a highly adaptable organism which can cause disease in a wide range of bodily sites. Critical to this is the ability to evade and subvert a diverse host immune response at various sites. This is achieved through the expression of multiple genes which play a role in the virulence of the bacteria (virulence factors).

ExPEC isolates express a wide range of virulence factors which increase their ability to survive and evade the immune system and cause disease. The most common functions of ExPEC virulence factors are adhesins, protectins and invasins, toxins, iron acquisition systems, and the capsule. Critically, these virulence factors are often shared between different pathotypes of *E. coli* (Figure 5.1) and may aid in discrimination and identification.

Multiple studies have been conducted investigating the genetic traits of *E. coli* bacteraemia infections. Adhesins such as the P-fimbriae (*pap)*, have been found to be more prevalent in bacteraemia strains compared to commensal control strains [400]. Protectins such as *traT, kps,* and *iss*, invasins such as *ibe,* toxins such as *cnf* and *cdt,* and iron acquisition systems such as *iroN* and *chuA* have also been found to be associated with isolates that can cause sepsis[73]. Some of these virulence factors also play a role in the translocation of *E. coli* from the gastrointestinal tract into the blood with the genes *papC, sfaD/E, cnf1, hlyA* and *iutA* being predicted as genes enabling translocation in patients with haematological malignancies [401]. The capsule of *E. coli* has also been implicated in sepsis, particularly the K1 capsule which has strong links with the development of meningitis and sepsis in the neonate[402–404].

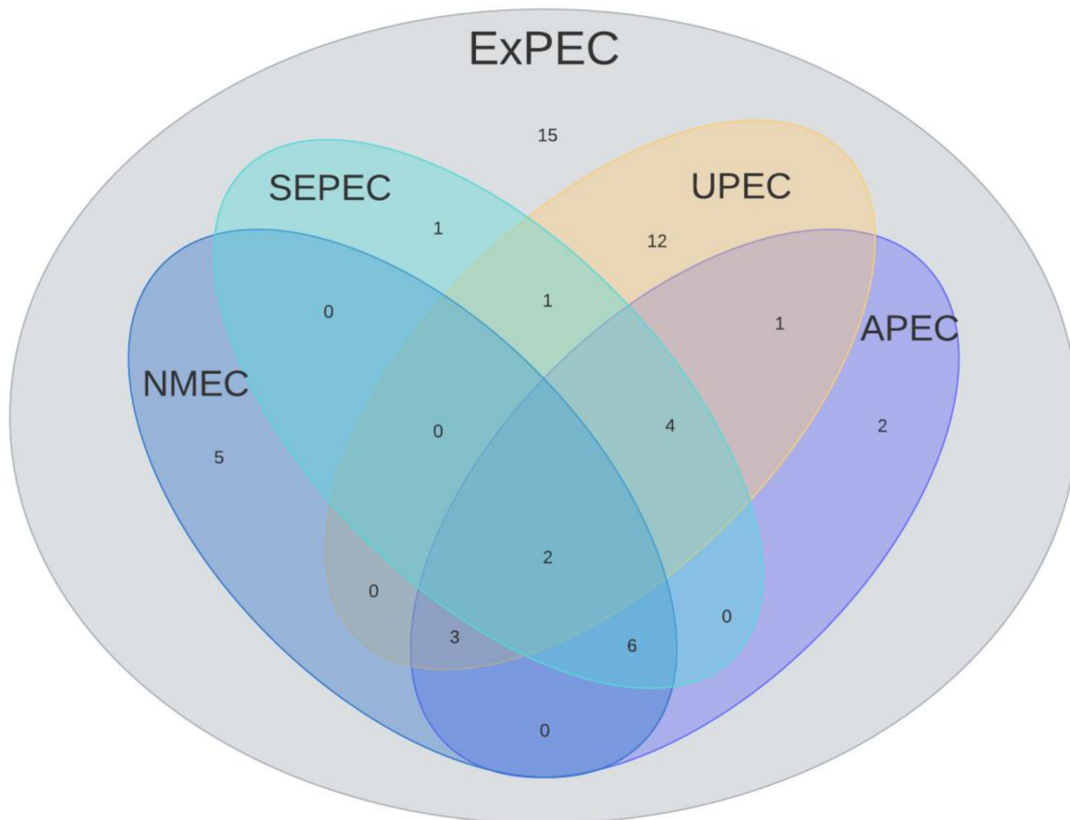


Figure 5.1 Venn diagram of shared virulence factors in ExPEC

The table of common virulence factors (Appendix 5.1) were submitted to a literature search for associations with different ExPEC pathotypes. The shared virulence factors are indicated by overlapping ovals. ExPEC- extra-intestinal, UPEC- uropathogenic, APEC- avian pathogenic, NMEC-Neonatal meningitic, SEPEC-sepsis associated. Created using Lucid chart.

5.1.2.2 Mechanisms of antimicrobial resistance in ExPEC

ExPEC have several genes providing functional resistance to antibiotic therapies, and they do this by four mechanisms; i) limiting drug uptake; ii) modification of antimicrobial drug target; iii) inactivation of the drug itself and iv) active efflux of the drug from the intracellular space [405]. The most common mechanism of limiting drug uptake employed by bacteria is the formation of biofilms. Bacterial biofilms are a complex community of bacteria which attach to and colonise surfaces (including tissues) [406]. *E. coli* biofilms are associated with urinary tract infections as well as indwelling medical device infections (e.g.,

catheters). There is evidence that traditional antimicrobial therapies are less effective against *E. coli* biofilms [407].

ExPEC can increase resistance through the modification of the antibiotic target, RNA polymerase, which has been shown to lower the efficacy of rifampicin. This is manifested by mutations in the RNA polymerase gene, *rpoB*, where mutations to its β subunit reduce binding affinity and decrease antibiotic efficacy [408].

Direct antibiotic modification by *E. coli* is achieved through β-lactamases (*bla* genes), which hydrolyse the β-lactam ring of penicillin antibiotics rendering them inactive [409] [410]. *E. coli* also express extended spectrum β-lactamases (ESBLs) encoded by *tem-1* and *shv-1* which are becoming an increasing concern for public health. Between 2013 and 2014 ESBL producing *E. coli* were found in 11% of all isolates from human faeces and caused more than 5000 bacteraemia cases annually in the UK [411]. Of particular importance is the emergence of *E. coli* expressing the β-lactam CTX-M-15 which has become globally disseminated through isolates of the sequence type 131. The acquisition of the CTX-M-15 β-lactam has firmly cemented this sequence type as a human pathogenic strain over the course of two decades [412].

Removal of antibiotics from the intracellular space in *E. coli* is performed by efflux pumps. These membrane transporters actively remove active antibiotics from the cytoplasm. Common *E. coli* efflux pumps include AcrB-TolC and EmrAB-TolC which are associated with resistance to multiple antibiotics [413],[414,415].

5.2 Aims and objectives.

This chapter aims to create a *E. coli* pangenome using whole genome sequencing as well as publicly available whole genome sequences and identify virulence factors and AMR genes present in the bacteraemia collection described in Chapter 3. Virulence factors and AMR genes will be compared to the historical standard ECOR collection and within sources of infection. Specifically:

E. coli isolates will be whole genome sequenced and then analysed and characterised by:

- Generating a pangenome

- Constructing a whole genome phylogenetic tree
- Establishing the core/accessory genome
- Identifying and counting the presence and absence of virulence factors
- Identifying and counting the presence and absence of AMR genes

- Comparison of established virulence factors in bacteraemia whole genomes grouped by:
    - Origin of infection (Urinary vs Abdominal)
    - Bacteraemia vs ECOR
    - B2 phylogroup of urinary vs abdominal isolates

5.2 Methods

5.2.1 Assembly

Raw Illumina genome reads were uploaded onto a Linux virtual machine (Swansea Genome Centre) and trimmed using sickle and scythe[232]. before assembly using Spades [233], Genome assemblies were quality controlled with Quast before assemblies with low contig (Contigs <1000bp) were removed from all isolates before annotation using seqtk[235].

5.2.2 Phylogenetic tree

The newick file generated from creating the roary pangenome was uploaded onto the interactive tree of life website (iTOL version 6.8) and phylogenetic tree was generated rooting to *E. coli* K12 strain MG1655.

5.2.3 Annotation

E. coli isolates were annotated as per section 2.6.2 using prokka [236].

5.2.4 Abricate

To identify virulence factors and antimicrobial resistance genes present in the *E. coli* bacteraemia isolates abricate was used as a mass screening tool to detect virulence and

antimicrobial resistance genes. For virulence factors the database used was VFDB[245], and for AMR genes the database CARD[243], was used.

5.2.5 Comparing presence and absence of genes

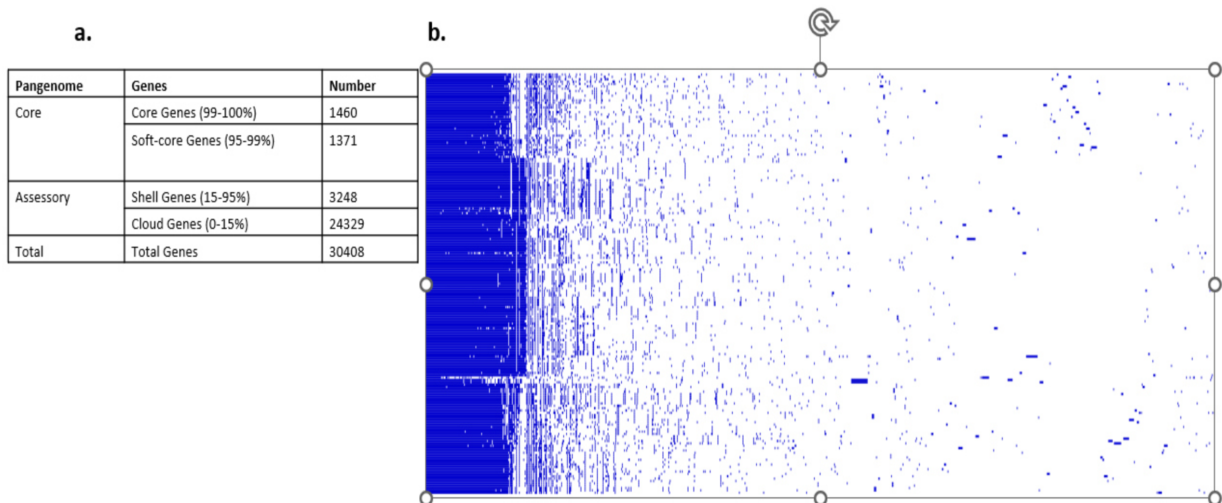
For common ExPEC virulence factors abricate outputs were manually searched for genes of interest. Tables were constructed and analysed with the Chi-squared test using Excel. Heatmaps were generated using the online heatmap making tool Morpheus [416].

5.4 Results

5.4.1 *E. coli* pangenome creation

A pangenome was created for all sequenced *E. coli* isolates using the software roary [237]. The ECOR collection (71 isolates) of *E. coli*, the Keio collection parent K12 strain, two reference *E. coli* isolates (ATCC 35218, ATCC 25922), 92 bacteraemia isolates, 7 water isolates and 6 ANC urine isolates (179 total) The K12 isolate MG1655 was used as a reference.

In total 30408 genes were identified in the collection. Of these 1460 were core genes (present in >99% of isolates), 1371 were soft core genes (found in 95-99% of isolates). The majority of genes identified were accessory genes with 3248 shell genes (present in 15-95% of isolates) and 24329 cloud genes identified under 15% of isolates (Figure 5.2).



Pangenome	Genes	Number
Core	Core Genes (99-100%)	1460
	Soft-core Genes (95-99%)	1371
Assessory	Shell Genes (15-95%)	3248
	Cloud Genes (0-15%)	24329
Total	Total Genes	30408

Figure 5.2 ***E. coli*** **pangenome** a. core and accessory genes of the *E. coli* collection b. roary gene presence/absence matrix. Each row in the matrix represents a genome while each individual blue/white dot represent the presence (blue) or absence (white) of a gene.

5.2 Phylogenetic relationships of the *E. coli* collection

To confirm the phylogenetic relationships between the isolates, the roary output newick file was uploaded onto iTOL and coloured according to the phylogroup identified using the Clermont method (Section 2.6.4). Isolates clustered well into the phylogroups identified using the Clermont method although some inaccuracies were observed (Figure 5.2), which is to be expected based on 95% accuracy of the Clermont method[417].

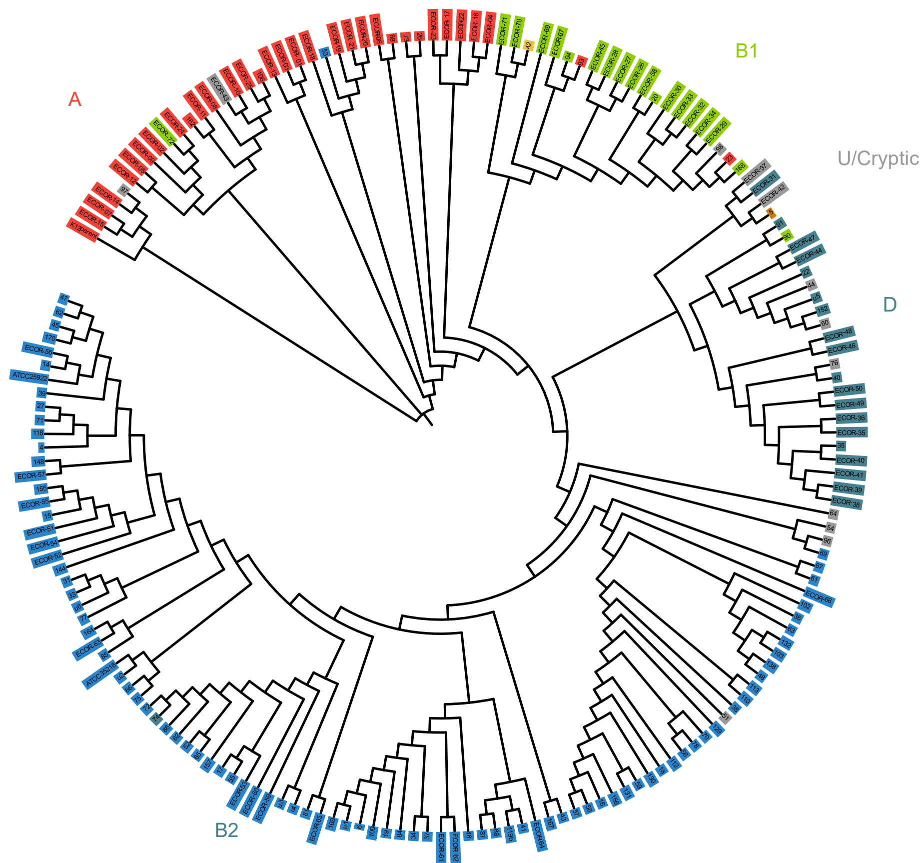


Figure 5.2 Phylogenetic tree of *E. coli* bacteraemia isolates

Ninety-two bacteraemia isolates, 7 water isolates, 2 reference isolates and 71 isolates from the ECOR collection were included in the genetic analysis. Isolates were coloured according to their phylogroup as identified in section 2.6.4. Tree was rooted to K12 reference organism MG1655.

5.4.3 Total number of virulence factors and AMR genes

To identify the number of virulence factors and AMR genes in the *E. coli* collection the genomes of all sequences isolates were analysed using the abricate software which incorporates multiple databases. Blood culture positive bacteraemia isolates had significantly more virulence factors than both the water isolates and the ECOR isolates in both the *E. coli* VF database and vfdb (Figure 5.3a and b respectively). Surprisingly the isolates from asymptomatic bacteriuria had significantly more virulence factors than the bacteraemia isolates in both databases (Figure 5.3a, b).

Similar amounts of AMR genes were identified in bacteraemia, water, asymptomatic bacteriuria, reference and ECOR isolates using the CARD database (Figure 5.4b). Using the Resfinder database, much lower overall AMR genes were identified but significantly more were found in the bacteraemia isolates compared to the water and ECOR isolates (Figure 5.4a).

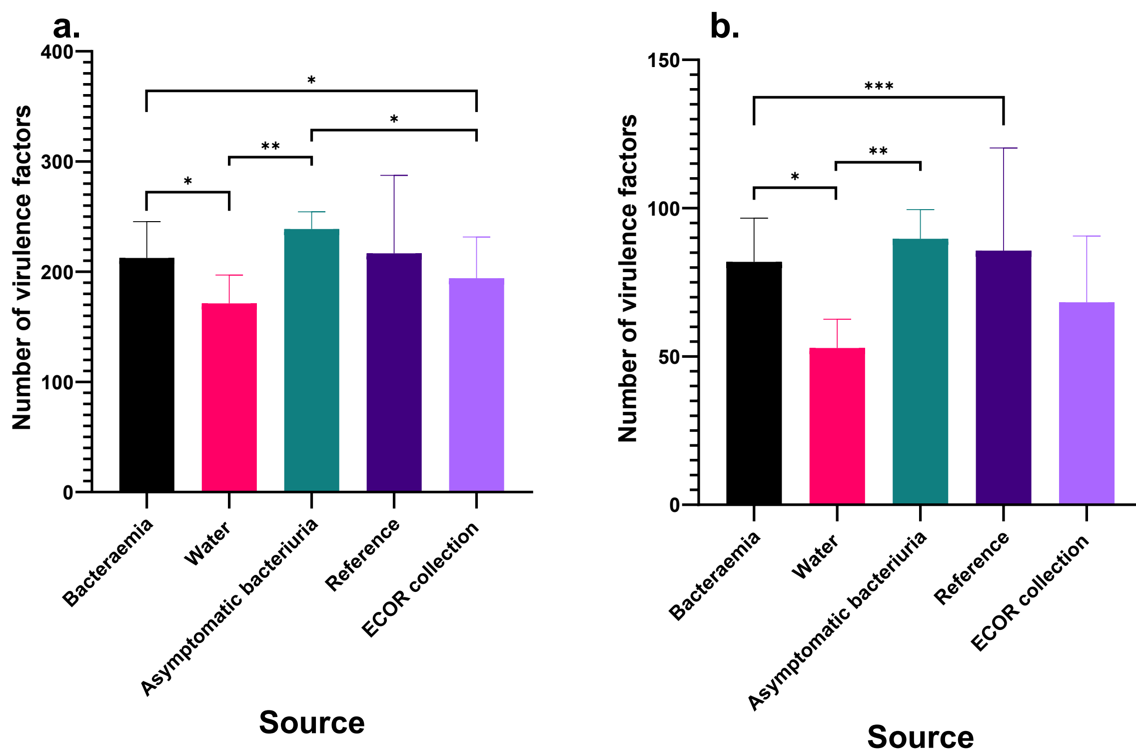


Figure 5.3: Number of virulence factors present in *E. coli* isolates and ECOR collection. Abricate was used on the isolate programs screening for virulence factors and antimicrobial resistance genes form various databases a. E.colivf database, b. virulence factor database (vfdb). Data represents mean number of genes per isolate ± SEM. Differences between groups were calculated by Kruskal-Wallis test with P * = ≤0.05, **= ≤0.001. ***= 0.0001

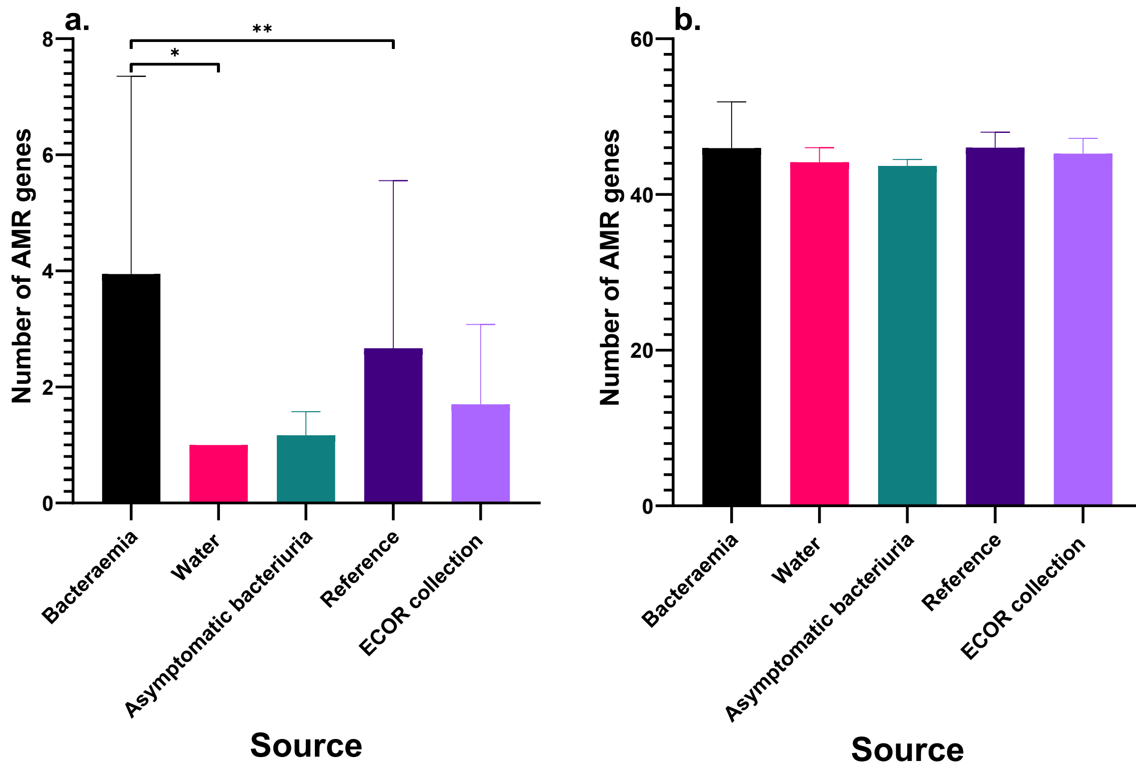


Figure 5.4: Number of AMR genes present in *E. coli* isolates used and isolates ECOR collection.

Abricate was used on the isolate programs screening for virulence factors and antimicrobial resistance genes form various databases a. card, b. resfinder. Data represents mean number of genes per isolate ± SEM. Differences between groups were calculated by Kruskal-Wallis test with P * = ≤0.05, **= ≤0.001. *** ≤ 0.0001

5.4.4 Top virulence factor and AMR genes

The top 50 virulence factors identified from vfdb include several key *E. coli* virulence factors (Appendix 5.2). The genes *csgF, entD, fepD, fepG* and *ompA* were identified in all 178 of the *E. coli* genomes. The *csgF* gene is a curli assembly component, *entD* is involved in the synthesis of the iron transport molecule enterobactin. The *fep* genes are involved in that transport of ferric enterobactin. Other top genes identified amongst the sequences tested include the fimbrial adhesin encoding genes *fim* and the siderophore synthesis genes *ybt* (Appendix 5.2). In the bacteraemia isolates (n=92) the most common virulence factors were also *csgF, entD, fepD, fepG* and *ompA,* which were identified in all the bacteraemia isolates (Appendix 5.3).

Abricate software was also used to identify the AMR genes present in the collection. Using the card database, the top identified AMR genes were identified (Appendix 5.4). Ten genes were found in all 178 isolated namely, *acrB*, *acrD, bacA*, *baeS*, CRP, *emrA*, *eptA*, *ampH*, H-NS, *mdtA* and *tolC.* Among the most commonly identified AMR genes were efflux pumps and other transporters associated with antimicrobial resistance.

The presence and absence of virulence factors and AMR genes was analysed and presented visually as heat maps for abricate (Appendix 5.5), vfdb (Appendix 5.6) and CARD (Appendix 5.7). The data in these heat maps was used to compare groups of *E. coli* isolates.

5.4.5 Incidence of Virulence factors and AMR genes in bacteraemia compared to ECOR isolates.

The ECOR collection is a collection of known genetic sequences from *E. coli* isolates that represent the genetic diversity of the species. Bacteraemia isolates expressed more of the P fimbriae genes (*pap) papC* and *papG* (Table 5.1) than the ECOR isolates. In addition, many of the genes encoding the S fimbriae (*sfa*) were also greater in the bacteraemia isolates (Table 5.1). Other genes that are more frequent in bacteraemia isolates include; the group 2 capsule synthesis gene *kspM*, the hemin receptor molecule gene *chuA*, the ferric aerobactin receptor *iutA*, intracellular pathogen resistance genes *irp1* and *irp2*, the aerobactin synthesis genes *iucABCD*, the yersinabactin uptake receptor *fyuA*, the salmochelin genes *iroBCDEF*, the heamolysin A gene *hlyA*, the streptothricin acetyltransferase gene *sat*, the vacuolating autotransporter gene *vat*, the cytotoxic necrotizing factor gene *cfn1*, and the brain microvascular endothelial cell invasion gene *ibeA* (Table 5.1). In contrast, relatively similar levels of the F1C fimbriae genes (*foc*) and the type 1 pili genes *fimA* and *fimH* were detected between the ECOR isolates and the bacteraemia isolates (Table 5.1).

Abricate was unable to locate several genes commonly identified in *E. coli* isolates. These genes were screened for but could not be identified namely*: focB, focl, focl-2, focx, focy, afaB, afaC, matA, matC, matF, csgA, tsh, tia, traT, neuC, cvaA, Iss2, sitA, sitB, sitC, sitD, ireA, ompT, cvaC, aatA,* and *usp*.

Table 5.1 Differences in *E. coli* virulence factors between bacteraemia isolates and the ECOR collection. Data was collected using abricate, Genes not found were removed from the table (section 5.4.3). Statistical analysis was performed using excel Chi-squared function, *=p-value <0.05

Gene name	Bacteraemia n (%)	ECOR n (%)	P value (Chi^2)
papA	7 (7)	5 (7)	0.8949
papC	48 (52)	18 (25)	0.0076*
papG	34 (37)	10 (14)	0.0053*
sfaA	0 (0)	1 (1)	0.2550
sfaB	21 (23)	11 (15)	0.2948
sfaC	26 (28)	9 (13)	0.0332*
sfaD	24 (26)	11 (15)	0.1478
sfaE	11 (12)	5 (7)	0.3207
sfaF	11 (12)	5 (7)	0.3207
sfaG	14 (15)	4 (6)	0.0679
sfaH	3 (3)	2 (3)	0.8725
sfaS	1 (1)	1 (1)	0.8542
sfaX	19 (21)	10 (14)	0.3243
focA	13 (14)	8 (11)	0.6136
focC	13 (14)	6 (8)	0.2923
focD	14 (15)	6 (8)	0.2214
focF	12 (13)	7 (10)	0.5549
focG	16 (17)	9 (13)	0.4460
focH	21 (23)	8 (11)	0.0828
afaA	5 (5)	3 (4)	0.7297
afaC-I	5 (5)	2 (3)	0.4239
afaD	5 (5)	2 (3)	0.4239
draA	2 (2)	0	0.2141
draD	3 (3)	1 (1)	0.4541
draP	7 (8)	3 (4)	0.3872
fimA	86 (93)	65 (92)	0.8990
fimH	90 (98)	67 (94)	0.8234
kpsM	75 (82)	27 (38)	0.0005*
kpsT	14 (15)	10 (14)	0.8517
chuA	70 (76)	17 (24)	0.00001*
iutA	53 (58)	21 (30)	0.0084
iucA	54 (59)	21 (30)	0.0066
iucB	53 (58)	21 (30)	0.0084
iucC	50 (54)	21 (30)	0.0175*
iucD	53 (58)	21 (30)	0.0084*
iroB	36 (39)	14 (20)	0.0265*

iroC	34 (37)	14 (20)	0.0443*
iroD	36 (39)	16 (23)	0.0629
iroE	34 (37)	16 (23)	0.0993
iroN	33 (36)	14 (20)	0.0569
irp1	76 (83)	33 (46)	0.0052
irp2	78 (85)	37 (52)	0.0138*
fyuA	82 (89)	37 (52)	0.0061*
ompA	92 (100)	71 (100)	1.0000
hlyA	31 (34)	6 (10)	0.0008*
pic	20 (22)	13 (18)	0.6295
sat	37 (40)	13 (18)	0.0123
vat	47 (51)	14 (20)	0.0012*
cnf1	32 (35)	6 (8)	0.0006*
cdtA	0 (0)	0 (0)	
cdtB	0 (0)	0 (0)	
astA	9 (10)	10 (14)	0.4251
ibeA	12 (13)	2 (3)	0.0272
Total isolates	**92**	**71**	

5.4.6 Incidence of established *E. coli* virulence factors in urinary and abdominal isolates

Isolates from a urinary or abdominal source were the origins of infection with the highest incidence in this study (Chapter 3). To investigate differences between these two sources of infection the heat maps of virulence factors and AMR genes were used (Appendix 5.5-5.7).

Isolates from a urinary source had a higher incidence of genes encoding P fimbriae *papA* (9% vs 0%), *papC* (64% vs 36%) and *papG* (47% vs 23%) than those from a urinary source (Table 5.2).

Urinary isolates also had more of the aerobactin biosynthesis genes *iut* compared to abdominal isolates (29% vs 9%) as well as the secreted autotransporter toxin, *sat* (42% vs 18%) (Table 5.2). Genes associated with iron regulation (*iro)* were found in similar percentage in both the urinary and abdominal group (Table 5.2).

The S fimbrial adhesin genes *sfaC* and *sfaX* S were also found to be significantly more different between urinary and abdominal isolates. Abdominal isolates had significantly more *sfaC* (32%) compared to urinary, while urinary isolates had more *sfaX* (22%) compared to abdominal isolates.

Table 5.2: Differences in *E. coli* virulence factors between bacteraemia isolates from urinary and abdominal origins of infection. Presence of commonly identified Virulence factors in *E. coli* isolates from the two major sources of infection. Genes not found were removed from the table (section 5.4.3). Statistical analysis was performed using excel Chi-squared function, *=p-value <0.05

Gene	Urinary n (%)	Abdominal n (%)	P value (Chi squared)
papA	4 (9)	0 (0)	0.1620
papC	29 (64)	8 (36)	0.1464
papG	21 (47)	5 (23)	0.1396
sfaA	0 (0)	0 (0)	0
sfaB	12 (27)	4 (18)	0.5045
sfaC	0 (0)	7 (32)	0.0002*
sfaD	13 (29)	5 (23)	0.6477
sfaE	7 (16)	2 (9)	0.4978
sfaF	7 (16)	2 (9)	0.4978
sfaG	8 (18)	4 (18)	0.9707
sfaH	1 (2)	2 (9)	0.2121
sfaS	1 (2)	0 (0)	0.4844
sfaX	10 (22)	0 (0)	0.0270*
focA	6 (13)	3 (14)	0.9746
focC	6 (13)	3 (14)	0.9746
focD	7 (16)	3 (14)	0.8486
focF	6 (13)	2 (9)	0.6370
focG	8 (18)	3 (14)	0.6944
focH	12 (27)	4 (18)	0.5045
afaA	3 (7)	0 (0)	0.2259
afaC-I	3 (7)	0 (0)	0.2259
afaD	2 (4)	1 (4)	0.9854
draA	1 (2)	1 (4)	0.6052
draD	3 (7)	0 (0)	0.2259
draP	4 (9)	1 (4)	0.5411
fimA	42 (93)	22 (100)	0.7932
fimH	43 (96)	22 (100)	0.8623
kpsM	39 (87)	18 (82)	0.8399
kpsT	6 (13)	6 (27)	0.2055
chuA	33 (73)	17 (77)	0.8609
iutA	29 (64)	9 (41)	0.2296
iucA	28 (62)	8 (36)	0.1751
iucB	28 (62)	10 (45)	0.3921
iucC	27 (60)	7 (32)	0.1283
iucD	28 (62)	10 (45)	0.3921

iroB	19 (42)	9 (41)	0.9378
iroC	18 (40)	8 (36)	0.8225
iroD	20 (44)	9 (41)	0.8364
iroE	18 (40)	8 (36)	0.8225
iroN	18 (40)	8 (36)	0.8225
irp1	37 (82)	18 (82)	0.9863
irp2	38 (84)	18 (82)	0.9121
fyuA	40 (89)	20 (91)	0.9346
ompA	45 (100)	22 (100)	1.0000
hlyA	16 (36)	6 (27)	0.5785
pic	9 (20)	6 (27)	0.5546
sat	19 (42)	4 (18)	0.1147
vat	25 (56)	12 (55)	0.9583
cnf1	16 (36)	7 (32)	0.8063
cdtA	0 (0)	0 (0)	0
cdtB	0 (0)	0 (0)	0
astA	7 (16)	0 (0)	0.0643
ibeA	5 (11)	6 (27)	0.1252
Total	**45**	**22**	

5.4.7 Incidence of established *E. coli* virulence factors in urinary and abdominal isolates in B2 phylogroup

Isolates from an B2 phylogroup were the most commonly identified in this study (Chapter 3). To establish differences between a urinary source and an abdominal source for the B2 phylogroup the isolates were screened for common virulence factors using the output from abricate (Appendix 5.5-5.7) and the list of common virulence factors (Table 1.4).

Thirty-three B2 isolates from a urinary source were compared to 18 isolates from an abdominal source of infection. Two genes were identified as being significantly different (Chi squared test). The genes *kpsT* (ATP binding cassette) and *ibeA* (invasin) were found to be more prevalent in the abdominal group (despite the lower total number of isolates) p= 0.0193 and 0.0082 respectively (Table 5.3) compared to the urinary group.

The genes *fimA, fimH* and *ompA* were found in all of the urinary isolates and all of the abdominal isolates (Table 5.3).

Table 5.3 Differences in *E. coli* virulence factors between B2 phylogroup bacteraemia isolates from a urinary and abdominal source. Presence of commonly identified Virulence factors in B2 *E. coli* isolates from a urinary and abdominal source of infection. Genes not found were removed from the table (section 5.4.3). Statistical analysis was performed using excel Chi squared function, *=p-value <0.05

Gene	Urinary n (%)	Abdominal n (%)	P value (Chi Squared)
papA	4 (12)	0 (0)	0.1944
papC	23 (70)	8 (44)	0.4579
papG	16 (48)	5 (28)	0.5087
sfaA	0 (0)	0 (0)	0
sfaB	12 (36)	4 (22)	0.5872
sfaC	12 (36)	7 (39)	0.2005
sfaD	13 (39)	5 (28)	0.5544
sfaE	7 (21)	2 (11)	0.6248
sfaF	7 (21)	2 (11)	0.6248
sfaG	8 (24)	4 (22)	0.4440
sfaH	1 (3)	2 (11)	0.0843
sfaS	1 (3)	0 (0)	0.5165
sfaX	9 (27)	5 (28)	0.3133
focA	6 (18)	3 (17)	0.5074
focC	6 (18)	3 (17)	0.5074
focD	7 (21)	3 (17)	0.6047
focF	6 (18)	2 (11)	0.7010
focG	8 (24)	3 (17)	0.6514
focH	12 (36)	4 (22)	0.5871
afaA	3 (9)	0 (0)	0.2611
afaC-I	3 (9)	0 (0)	0.2611
afaD	1 (3)	0 (0)	0.5165
draA	0 (0)	0 (0)	0
draD	2 (6)	0 (0)	0.3589
draP	3 (9)	0 (0)	0.2611
fimA	33 (100)	18 (100)	0.0627
fimH	33 (100)	18 (100)	0.0627
kpsM	30 (91)	16 (89)	0.0899
kpsT	5 (15)	6 (33)	0.0193*
chuA	31 (94)	17 (94)	0.0682
iutA	19 (58)	8 (44)	0.4098
iucA	19 (58)	7 (39)	0.4922
iucB	19 (58)	8 (44)	0.4098
iucC	18 (55)	6 (33)	0.5060
iucD	18 (55)	8 (44)	0.3715
iroB	14 (42)	8 (44)	0.1838
iroC	14 (42)	8 (44)	0.1838
iroD	15 (45)	8 (44)	0.2305

iroE	14 (42)	7 (39)	0.3112
iroN	13 (39)	8 (44)	0.1411
irp1	28 (85)	15 (83)	0.0983
irp2	29 (88)	15 (83)	0.1175
fyuA	31 (94)	17 (94)	0.0682
ompA	33 (100)	18 (100)	0.0627
hlyA	15 (45)	6 (33)	0.5028
pic	9 (27)	6 (33)	0.1588
sat	15 (45)	4 (22)	0.4303
vat	23 (70)	12 (67)	0.1558
cnf1	15 (45)	7 (39)	0.3657
cdtA	0 (0)	0 (0)	0
cdtB	0 (0)	0 (0)	0
astA	5 (15)	0 (0)	0.1469
ibeA	4 (12)	6 (33)	0.0082*
Total	**33**	**18**	

5.5 Discussion

This chapter aimed to characterise the virulence factors and AMR genes of the *E. coli* bacteraemia collection and compare to the standard reference ECOR collection.

Several virulence factors were identified in all the isolates included in the genetic analysis (including the ECOR collection). Genes identified in all the *E. coli* isolates include several genes involved in iron metabolism (Table 5.2). More specifically, genes involved in the enterobactin synthesis (*ent*), and transport system were present in most isolates. These genes encode for enterobactin which is a siderophore involved in the uptake of bacterial iron but may also have additional roles. Enterobactin has also been shown to be involved in bacterial survival in toxic copper environments [418]. It has also been noted that commensal bacteria also carry *ent* genes and that in the intestinal environment these genes act on host ATP synthases to promote mitochondrial iron uptake in a 'tug of war' over iron availability between bacteria and host cells [419]. It is perhaps not surprising then that members of the ECOR collection possess these genes as they were all originally isolates from mammalian hosts[74].

Antimicrobial resistance genes were also present in all but one of the isolates used in the analysis. The genes involved in the inner membrane transporter *acrB* and *acrD* were found in all but one of the isolates sequenced (Table 5.3). These genes, along with *tolC* (an outer membrane channel), were also found in 178 isolates and form efflux pumps. Such, efflux pumps are a common bacterial defensive system which allow bacteria to regulate their intercellular environment by removing toxic substances, including antibiotics. Interestingly, both *acr* genes (through the action of *TolC*) and enterobactin genes may play roles in the protection of *E. coli* against antimicrobial peptides. For example, the peptide WRWYCR, a potent antimicrobial, had increased effectiveness against *E. coli* Keio collection mutants lacking genes from the enterobactin synthesis pathway and *TolC* [420]. The importance of these genes in this defence system warrants further research and may explain why both are present in nearly all isolates used for genetic analysis in this study.

AcrAB-TolC is the major antibiotic efflux pump of *E. coli* [421]. Overactivity of this efflux pump is attributed with resistance to fluoroquinolones and tetracycline [421–423]. In addition to actively pumping out antibiotics from inside the bacterial cells the efflux pump AcrAB-TolC also has a role in resistance to other host defence mechanisms as *acrAB* mutants

are hypersensitive to bile salts and fatty acids [424] The AcrAD-TolC has been shown to participate in the efflux of aminoglycosides as deletion of the *acrD* gene reduced resistance to amikacin, gentamicin, neomycin, kanamycin and tobramycin [425].

Interestingly there were no significant differences in the number of virulence factors and AMR genes in the ECOR collection compared to the asymptomatic bacteriuria, the water, or the reference isolates tested (Figure 5.3 and 5.4). The similarity between bacteraemia and asymptomatic bacteriuria isolates appears surprising. However, the majority of the asymptomatic bacteriuria isolates are from the B2 phylogroup, which has been shown to possess more virulence factors than the other phylogroups, which could explain why more were not seen in the bacteraemia group[416]. In addition, there were only 6 of the asymptomatic bacteriuria isolates and if more were included in the study there may have been more further observable differences between them and the bacteraemia isolates.

This chapter also aimed to compare the numbers of common *E. coli* virulence factors associated with bacteraemia between bacteraemia isolates and the ECOR collection and the two most common sources of infection (urinary and abdominal). In total, 20 out of 53 common ExPEC virulence factors were significantly higher in bacteraemia isolates compared to ECOR isolates (Table 5.2). Of these, *chuA, kpsM, cnf1,* and *hlyA* were the most statistically significant (Table 5.2).

The gene *chuA* codes for an outer membrane hemin receptor which had been associated with colonisation with the urinary tract as well as ExPEC isolates [426]. Interestingly, in this paper, the authors link the isolates carrying the genes *chuA, vat, fyuA and yfcV* with effective colonisation of the urinary tract in a mouse model of urinary tract infection and growth in human urine [426]. In the current thesis, three of these genes were identified as being associated with bacteraemia isolates compared to the ECOR collection. These genes may be worthwhile investigating and may prove useful in future for predicting urinary tract infections more likely to progress to bacteraemia.

Another interesting link between this research and previously published data is the association of the genes *fyuA* and *irp2* with bacteraemia (Table 5.2). The gene *fyuA* is also involved in iron uptake and encodes the yersinabactin receptor and is required for biofilm formation in urinary tract infections[427], *irp2* is an iron regulated siderophore under the control of the transcriptional regulator *Fur*[428]. Previously these genes have been shown to

be present on the same pathogenicity island and are associated with pathogenicity in APEC isolates as gene knockouts had impaired capacity to bind to chicken fibroblasts [429]. Not only does this indicate a potential reservoir for pathogenic genes present in *E. coli* populations outside human ExPEC isolates but also highlights the plasticity of *E. coli* in general. Furthermore, the genes *iucA* and *iucD,* which are involved in aerobactin synthesis have been shown to be important contributing factors to pathogenicity in chickens as mutants show reduced colonisation and persistence [430]. Utilising genetic targets like *irp2*, *fyuA, iucA* and *iucD* as markers of potential *E. coli* pathogenicity may be a promising source of investigation. Particularly with concerns growing about potential food producing animal reservoirs of *E. coli* infections [431].

In addition to comparing bacteraemia isolates to the ECOR collection the isolates from a urinary source of infection were also compared to those from an abdominal source of infection (Table 5.3). In total, two genes were identified as being significantly different. Interestingly, one of each of the genes were found to be related to each source of infection. The genes identified were *sfaC* (abdominal) and *sfaX* (urinary). The gene *sfaA* encodes for a S-fimbriae major subunit while *sfaX* encodes a regulatory protein. The gene *sfaX* is associated with newborn meningitis and has a role in motility of these isolates through the regulation of motility related genes [432]. Interestingly, the infection of neonates with *E. coli* occurs through the genital tract during birth. Given that this gene has been associated with urinary isolates in this study it may prove a worthwhile gene target predicter for isolates which may cause neonatal sepsis infections.

The B2 phylogroup contained the most *E. coli* isolates in this study (Chapter 3). Therefore, B2 isolates from a urinary source of infection were compared to isolates from an abdominal source (Table 5.3). Two genes were identified as being significantly different between the groups. The gene *kpsT* encodes an ATP binding cassette, and in conjunction with the gene *kpsM* encodes the information required for transport of the capsular polysaccharide K1 (the most common K type for neonatal meningitis) to the cell surface[433]. K capsular antigens have been identified more frequently in bacteraemia strains compared to faecal isolates[434], indicating another potential source of neonatal sepsis infections. The second gene that was more prevalent in abdominal isolates was the gene *ibeA* which encodes a 50-kDa invasin. Interestingly, this gene has linked neonatal sepsis strains to invasion of the

blood brain barrier by *E. coli* K1 strains in neonatal meningitis [435]. Additionally, this gene has also been linked to the development of colorectal cancer and is being investigated as a potential biomarker for this disease [436]. Further investigation of these genes in abdominal isolates is warranted, particularly concerning neonatal sepsis.

This chapter has used a simple approach to compare the presence and absence of ExPEC virulence factors in groups of isolates and identified significant differences in pathogenic genes. The next chapter (Chapter 6) will use more advanced strategies to compare groups of isolates making use of bacterial GWAS methods to screen the pangenome for associations to phenotypic traits.

Chapter 6: Identification and functional confirmation of gene targets associated with clinical, and laboratory associated bacteraemia phenotypes.

6.1 Introduction

6.1.1 Bacterial whole genome wide association studies

Genome-wide association studies (GWAS) interrogate tens, hundreds or even thousands of genomes for variations in genetic content when grouped by a certain trait or phenotype. GWAS confirms statistical associations between variants and a trait or phenotype[437,438]. Studies involving human GWAS have progressed faster than bacterial GWAS due to the limited inherent nature of variation in the human genome[439]. Genomic data from human and bacterial GWAS studies come from different sources. Human genomic data comes from SNP genotyping chips while bacterial genetic information usually comes from genetic sequencing[439]. There are powerful examples of successful human GWAS studies demonstrating associations between FTO and obesity, PTPN22 and autoimmune disease and the IL-12/23 pathway in crohn's disease[440–443].

The primary source of human genomic variation is bi-allelic SNP variation [444]. In contrast, variation in the bacterial genome occurs through three major mechanisms:

1) Single nucleotide polymorphisms- point mutations, small insertions or deletions which occur during transcription [445]
2) Gene presence/absence- loss or gain of entire genes via several mechanisms including, horizontal and lateral gene transfers of mobile genetic elements, phage infections and plasmid acquisition [445]
3) Copy number variations and sequence inversions- acquisition of additional gene copies from mobile genetic elements or large-scale deletions/duplications [445]

Considering these mechanisms of variation, deciphering genetic associations in bacterial populations is more complicated and requires different tools to that of human GWAS studies [439].

To date there have been several bacterial GWAS studies investigating genetic traits and their association with disease. For example, genetic analysis of the pangenome of *Staphylococcus epidermidis* has shown correlation between genetic elements and pathogenic traits of the bacteria such as biofilm formation and IL-8 production[446]. Additionally, GWAS analysis of *Helicobacter pylori* has shown a link between multiple genes

and the development of gastric cancer including the presence of the gene *babA*, which encodes for an adhesin[447]. GWAS analysis can also be utilised to investigate and identify bacterial genotypes with resistance to therapy, such as with *Mycobacterium tuberculosis* where resistance was associated with 13 non-canonical loci which did not contain genes encoding for known resistance mechanisms[448]. GWAS studies can also be utilised to investigate associations with diseases affecting livestock, a potential reservoir for future pandemics. Studies investigating *E. coli* APEC (a pathotype which shares many features to ExPEC), infections in chickens have identified 143 genes with diverse functions associated with colibacillosis [449]. In addition, the presence of *E. coli* carrying the *pks* pathogenicity island which synthesises colibactin, has been associated with the development of mutations linked to colorectal cancer[450].

With the advent of genomic technologies and the lowering cost of sequencing, bacterial GWASs are a developing field of research aiming to understand how variations in microbial genomes affect host and pathogen phenotypes e.g., drug resistance and disease prognosis. Combined with classical gene knockout studies they can be a powerful tool which give greater understanding to disease progression and may serve as potential markers of infection in addition, identifying genes and mechanisms associated with infection may lead to a greater diversity of therapeutic options for many infectious diseases[451]. Several different tools have become available for identifying genetic associations between bacterial traits all of which have advantages and disadvantages[439,452].

6.1.2 Scoary

Scoary is an anagram of ‘scoring’ and ‘roary’ (the pangenome pipeline), which is a new tool used in linking genetic associations to phenotypic traits [239]. Scoary attributes a score to the pangenome of a bacterial population to find associations between genes and phenotypic traits. There are several advantages to using the Scoary software for running GWAS studies, firstly it is rapid, allowing users to screen genome data in a few minutes rather than the weeks needed for standard SNP analysis [239]. Secondly, it is simple to use, for users with limited bioinformatic knowledge and makes minimal assumptions about evolutionary processes allowing users to skip the need for ill-informed mutation rate parameters [239]. Scoary assigns ‘genetic association’ to binary categorical phenotypic data meaning that the user must organise traits into these binary categories. This is not without

flaw, especially when working with phenotypes that are ‘continuous variables’ rather than binary in nature. In addition, it has been reported that while some fine scale genetic differences and phylogenetic clustering may be detected there can be a tendency to discard large volumes of useful data [445,453].

6.1.2 Genetic engineering and the *E. coli* Keio collection

6.1.2.1 Methods of genetic engineering

Genetic engineering of microorganisms has been carried out for centuries with the selection of specific traits for beneficial process (such as producing fermented foods) in a process known as “forward genetics” i.e., selecting for traits of interest[454,455]. In recent years there has been an advent on the opposite of this approach “reverse genetics”. This involves modifying the genomes of microorganisms, either through mutating genes of interest, inserting genes of interest or, knocking out genes of interest, the impact of which can be assessed with functional analysis[454,456].

These genetic engineering methods have enabled association of many *E. coli* genes with specific phenotypes of interest such as serum resistance. For instance, the K2 capsule has been found to be important in both serum survival and for colonisation of the urinary tract by using mutant K2 strains of the urosepsis isolate CFT073 [457]. Other genes implicated in serum resistance in *E. coli*, through mutant studies include, the outer membrane protein *OmpA,* as well as cold shock proteins *CspC* and *CspE* [458–460].

There are several methods that are used to knock out genes of interest; early methods involved integrating a selectable plasmid into the target gene locus via homologous recombination [461]. The plasmid contained the cloned gene fragment with a temperature-sensitive replication origin. At low temperatures there was replication but at higher temperatures replication was reduced. After transformation, it is then possible to select for the integration of the plasmids at high temp 44°C, with a second growth step at 30°C resulting in a second recombination event. Depending on where the second recombination event occurred the bacterial chromosome would have either undergone gene replacement or retained the original copy of the gene of interest [461]. A modernised version of this is now used; the λ-red system. This system is also dependent on recombination but uses the λ-Red recombinase which is transformed into bacterial cells along with linear DNAs containing

homologous DNA sequences, two flippase recognition target (FRT) sequences and an antibiotic selection marker, which can be removed after recombination with a helper plasmid expressing the FRT recombinase [462]. In addition to directly altering or removing a gene of interest alternative approaches using antisense RNAs may also be utilised to repress RNA generated from translation, thereby silencing the gene [454].

6.1.2.1 *E. coli* Keio collection

The λ- Red system of gene knockout has been utilised to generate in-frame single gene knockouts of the widely studied *E. coli* strain K12 (BW25113) [463] In this study, authors generated single gene knockouts for all non-essential genes for the *E. coli* K12 strain used, totalling 3985 gene knockouts [463]. As these single gene mutants are commercially available, this is a valuable tool for conducting functional genomic studies. The Keio collection has been utilised by researchers in several different areas, screening the collection for gene functionality. Antimicrobial sensitivity profiles have been determined with the collection aiming to generate a 'barcode' and aid in the design of new antimicrobials [464]. In addition, researchers have also screened the collection for genes that may influence harmful algal blooms[465], as well as metabolic flux analysis[466]. The wide-ranging applications of this collection make this a useful functional genomic analysis tool.

6.2 Aims and objectives

This chapter aims to compare *E. coli* whole genomes (GWAS) using the pangenome created by roary in chapter 3, with the laboratory phenotypes generated in chapter 4 and the clinical phenotypes generated in chapter 5.

Therefore, Scoary will be utilised to identify:

- *E. coli* genes associated with bacteraemia
- *E. coli* genes associated with *E. coli* origin of infection
- *E. coli* genes associated with urinary vs other bacteraemia isolates
- *E. coli* genes associated with and non/sepsis causing isolates(Chapter 3)
- *E. coli* genes associated with mortality (Chapter 3)

- *E. coli* genes associated with *in vitro/ex vivo* phenotypes (Chapter 4)

Furthermore, the top gene targets identified:

- Will be screened for their presence within the Keio collection
- Will have inducible cytokine responses to wild type and mutant strains compared in
    - Whole blood infection model
    - THP-1 infection model

6.3 Methods

6.3.1 Bacterial GWAS

The *E. coli* pangenome generated in chapter 5 was used as the input for scoary for GWAS analysis (section 2.6.3). Trait files were generated from either patient data (Chapter 3) or laboratory phenotype data (Chapters 3 and 4).

6.3.2 Whole blood infection model

The *ex vivo* whole blood infection model (section 2.4.1) was used to assess cytokine responses to *E. coli* single gene knock out mutants (Table 2.1). Cytokine responses were measured using ELISA following manufacturer's instructions (section 2.5.2).

6.3.3 THP-1 model of infection

E. coli mutants (Table 2.1) were also used in the THP-1 infection model (section 2.4.2). Cytokine responses were measured using ELISA following manufacturer's instructions (section 2.5.2).

6.4 Results

6.4.1 Genes associated with bacteraemia

To identify genes associated with blood culture positive *E. coli* bacteraemia isolates, GWAS was performed using the software scoary[239]. For this analysis 92 strains from the bacteraemia collection were compared to 87 non-bacteraemia strains from the ECOR collection and reference isolates (including K12, water isolates and asymptomatic bacteriuria). The results of the scoary GWAS (Table 6.1) confirmed the genes most associated with bacteraemia were the β-lactamase gene *bla*, the 34kDa membrane antigen *tpd,* the ferric aerobactin receptor *iutA,* elongation factor Tu *tufB, php_1* a putative hydrolase and *ybjE* a putative transporter (Table 6.1). The elongation factor Tu gene *tufB* had the best combined sensitivity and specificity with 91.3 and 65.6 respectively (Table 6.1).

Table 6.1: Genes identified by Scoary analysis as being associated with bacteraemia isolates. Hypothetical proteins were filtered out and genes were selected based on sensitivity and specificity >40%. Top genes targets highlighted in yellow to take forward for functional analysis see table 6.6)

Gene	Annotation	Sensitivity	Specificity	Naive_p value	Bonferroni_p value
bla	Beta-lactamase TEM	42.3913	94.2529	6.97E-09	0.0002
tpd	34 kDa membrane antigen	42.3913	90.8046	4.20E-07	0.0125
iutA	Ferric aerobactin receptor	47.8260	86.2069	8.88E-07	0.0265
tufB	elongation factor Tu	91.3043	65.517	4.63E-16	1.38E-11
php_1	putative hydrolase	82.6087	50.5747	3.26E-06	0.0972
ybjE	putative transporter	95.6522	49.4253	1.03E-12	3.09E-08

6.4.2 Associations with origin of infection

To investigate genes associated with source of infection, bacteraemia isolates from a urinary (n=45) or abdominal source (n=22) were compared to all other isolate genomes (including the ECOR collection, (n=71). Two genes were identified as significantly associated with an abdominal source of infection, *yfaA* a putative protein and *glcA* a glycolate/lactate H+ symporter (Table 6.2). Genes associated with the urinary tract included, *tufB* which encodes elongation factor Tu, *rusA_1* a Holliday junction resolvase, *fhuC_2* an iron hydroxamate ABC transporter subunit, *znuB_2* a zinc transporter subunit, *ybjE* a putative transporter, *yohF* a putative oxidoreductase, *ydeP* an acid resistance protein and *ydeQ* a putative fimbrial like adhesin protein.

Table 6.2 Genes identified by Scoary analysis associated with source of infection.

Isolates from either a urinary or abdominal source of infection were compared to all other genomes in the data set. Highest sensitivity/specificity combinations were selected. Top genes highlighted in yellow to take forward for functional analysis see table 6.6 and the full list shown in appendix 6.2 and 6.3.

Gene	Association	Annotation	Sensitivity	Specificity	Naive_p	Bonferroni_p
yfaA	Abdominal	putative protein	100	18.3007	0.02701	1
glcA	Abdominal	glycolate / lactate:H+ symporter	90.90	30.0654	0.04213	1
tufB	Urinary	elongation factor Tu	91.11	47.2868	1.81E-06	0.0538
rusA_1	Urinary	endodeoxyribonuclease RUS (Holliday junction resolvase)	75.56	46.5116	0.0132	1
fhuC_2	Urinary	iron (III) hydroxamate ABC transporter - ATP binding subunit	88.88	39.5349	0.0004	1
znuB_2	Urinary	Zn2+ ABC transporter - membrane subunit	93.33	37.9845	4.57E-05	1
ybjE	Urinary	putative transporter	95.55	34.8837	2.57E-05	0.763
yohF	Urinary	putative oxidoreductase with NAD(P)-binding Rossmann-fold domain protein	86.67	34.1085	0.0076	1
ydeP	Urinary	acid resistance protein	88.89	30.2326	0.01007	1
ydeQ	Urinary	putative fimbrial-like adhesin protein	88.88	26.3566	0.0385	1

6.4.3 Genes associated with Urinary bacteraemia isolates

To investigate associations between bacteraemia isolates from a urinary source and bacteraemia isolates from all other sources 45 strains from a urinary source and 42 strains from other bacteraemia sources were used. Genes associated with urinary origin of bacteraemia included, *iutA* the ferric aerobactin receptor, *agaC_1* a galactosamine PTS permease, *yqiG_2* a putative membrane protein and *xapB* a xanthosine H+ symporter. Full list of genes found associated with urinary bacteraemia isolates can be found in Appendix 6.3.

Table 6.3 Genes associated with a urinary source of bacteraemia.

Isolates from a urinary source of bacteraemia were compared to bacteraemia isolates from all other sources. Genes with highest sensitivity and specificity were selected as potential targets. Full list is available (Appendix 6.3).

Gene	Annotation	Sensitivity	Specificity	Naive_p	Bonferroni_p
iutA	Ferric aerobactin receptor	22.22222	100	0.001165	1
agaC_1	galactosamine PTS permease - cryptic	15.55556	100	0.012383	1
yqiG_2	putative membrane protein	55.55556	71.42857	0.016643	1
xapB	xanthosine:H+ symporter XapB	75.55556	7.142857	0.040197	1

6.4.4 Genes associate with sepsis isolates

Associations between bacteraemia isolates that did or did not cause sepsis were investigated by comparing 42 isolates where the keyword ‘sepsis’ was listed among patient data notes to 49 isolates missing the keyword. The genes *yjjQ*¸a putative DNA-binding transcriptional regulator, *yehY* an ABC transporter and *yicI* an alpha-xylodisase were found to be associated with isolates that caused sepsis (Table 6.4). Although, entries with non-unique gene names for *yjjQ* and *yicI* were found (Appendix 6.10), indicating either multiple copies of the gene or an error with the annotation software as both were found to be identical with a BLAST search.

Table 6.4 Genes associated with sepsis/non sepsis bacteraemia isolates which went to cause sepsis as identified from patient data (Chapter 3). Genes with highest sensitivity and specificity were selected. Full list (Appendix 6.10).

Gene	Annotation	Sensitivity	Specificity	Naive_p	Bonferroni_p
yjjQ	putative DNA-binding transcriptional regulator	90.69767	32.65306	0.01028	1
yehY	YehW/YehX/YehY/YehZ ABC transporter	100	12.2449	0.02816	1
yicI	alpha-xylosidase	44.18605	77.55102	0.043964	1

6.4.5 Genes associated with patient and phenotypic data

Genes associated with both patient data parameters such as nosocomial infections and mortality were tested with scoary. In this analysis, for nosocomial infections 23 isolates with a nosocomial source were compared to 57 isolates from a community acquired source. Isolates were excluded if data was missing from patient data. For mortality associations, the 17 isolates resulting in patient death, were compared to 72 isolates, where the patient was discharged from hospital.

Genes associated with mortality included *yhgE* a putative transport protein, *arfB* a peptidyl-tRNA hydrolase, *gmm* a GDP-mannosyl hydrolase and *cpsB* a mannose guanylyltransferase (Table 6.5). Although sensitivity was high for all these genes the sensitivity ranged from between 29 and 39 (Table 6.5). Genes associated with nosocomial infections were *yhdJ_2* and *allD*. For full lists of genes found associated with patient phenotypes please see Appendix 6.5 and 6.6.

6.4.6 Associations with *in vitro* laboratory phenotypes

To determine genetic associations with *in vitro* phenotypes, isolates were compared according to their ability to; i) survive in human plasma; ii) induce cytokine responses in whole blood and iii) induce cytokines in THP-1 cells. For plasma/serum resistance 58 resistant isolates (from bacteraemia, water, asymptomatic bacteriuria, and reference isolates) were compared to 43 sensitive isolates. Survival in human plasma was associated with the genes *yejF, ampG_2* and *msbA* which encode for an ABC transported ATP binding

subunit, a muropeptide H+ symporter and a putative multidrug export ATP binding cassette, respectively (Table 6.5).

No significant associations were confirmed with cytokine levels in whole blood. For cytokine levels in THP-1 cells, IL-8 responses of ‘less than 1000pg/mL’ and ‘less than 500 pg/mL’ were tested. For less than ‘500pg/mL’ 17 isolates with low cytokine levels were compared to 84 isolates with responses greater than 500pg/ml. For less than 1000pg/mL 25 isolates with low cytokine levels were compared to 76 isolates with responses greater than 1000pg/mL. Genes that were associated with ‘less than 1000pg/mL’ observed in THP-1 cells (TNFα and IL-8) were *flgI* and *eutJ* a flagellar ring protein and a putative chaperonin respectively (Table 6.5). The gene *ynbC* a putative hydrolase, was associated with low IL-8 levels in THP-1 cells (less than 500pg/mL) (Table 6.5). Full list of genes found to be associated with low cytokine levels in THP-1 cells can be found in Appendix 6.8 and 6.9.

Table 6.5 Genes associated with patient data and with laboratory phenotypes.

Isolates associated with patient data and plasma resistance were compared to other bacteraemia isolates that were not associated with the trait. Genes with highest sensitivity and specificity are shown. Top genes highlighted in yellow to take forward for functional analysis see table 6.6 and full gene lists can be found in Appendix 6.5-6.9)

Gene	Annotation	Sensitivity	Specificity	Naïve p	Bonferroni_p	Association
yhgE	putative transport protein	100	38.89	0.000975	1	Mortality
arfB	peptidyl-tRNA hydrolase, ribosome rescue factor	100	34.72	0.002318	1	Mortality
gmm	GDP-mannose mannosyl hydrolase	94.11	31.94	0.03379756	1	Mortality
cpsB	mannose-1-phosphate guanylyltransferase	100	29.16	0.009304	1	Mortality
yhdJ_2	DNA adenine methyltransferase	13.04	59.65	0.019549098	1	Nosocomial
allD	ureidoglycolate dehydrogenase	100	19.30	0.028352288	1	Nosocomial
yejF	peptide ABC transporter - ATP binding subunit	97.92	19.51	0.010405	1	Plasma resistance
group_7879 (yejF)	peptide ABC transporter - ATP binding subunit	2.08	80.49	0.010405	1	Plasma resistance
ampG_2	muropeptide:H+ symporter	97.92	14.63	0.045272	1	Plasma resistance
group_9771 (msbA)	Putative multidrug export ATP-binding/permease protein	97.92	14.63	0.045272	1	Plasma resistance
flgI	flagellar P-ring protein FlgI	100	26.83	0.001643	1	Low cytokine THP-1 (<1000pg/mL)
group_8081(flgI)	flagellar P-ring protein FlgI	0	73.17	0.001643	1	Low cytokine THP-1 (<1000pg/mL)
eutJ	putative chaperonin, ethanolamine utilization protein	96	21.95	0.040671	1	Low cytokine THP-1 (<1000pg/mL)
ynbC	Putative hydrolase	100	55.95	6.24E-06	0.139319184	Low cytokine THP-1 (<500pg/mL)

6.5.7 Top gene targets identified through GWAS

The functional role of the gene targets in whole blood and THP-1 infection models was investigated. Firstly, a list of the top gene targets with the highest sensitivities and specificities was compiled so that mutant strains could be acquired using the *E. coli* Keio collection (Table 6.6 and yellow annotation in previous tables). A total of 6 mutants and one wild-type strain were tested and included *tufB, ybjE, yhgE, yohF, yejF* and *ynbC.*

Table 6.6 Selected *E. coli* gene targets.

Five important phenotype-gene associations were selected from a Scoary analysis. The Scoary microbial pan-GWAS was used to compare and analyse gene association using gene presence / absence file generated in Roary, and a traits file detailing the presence/absence of selected traits.

Gene name	Annotation	Association	Naïve p value	Sensitivity	Specificity	K12 knock out
ynbC	Putative hydrolase	Low IL-8 (<500pg/mL)	6.24E-06	100	55.85	Yes
yhgE	Putative transport protein	Mortality	0.001	100	38.89	Yes
ybjE	Putative transporter	Bacteraemia	1.03E-12	95.65	49.43	Yes
yejF	Peptide ABC transporter-ATP binding subunit	Plasma resistance	0.0105	97.92	19.51	Yes
tufB	Elongation factor Tu	Urinary/bacteraemia	1.81E-06	91.11	47.29	Yes
yohF	Putative oxidoreductase	Urinary	0.008	86.67	34.11	Yes

6.4.8 *Ex vivo* whole blood and THP-1 cell cytokine response to *E. coli* K12 mutants

The functional effect of removing genes of interest (Table 6.6) was investigated using *ex vivo* whole blood and THP-1 cell infection models. In addition, the mutant *fimA* was included in mutant studies due to its role during the initial stages of urinary tract infection as well as in- the abdominal tract and neonatal sepsis [467–469].In whole blood infections, the cytokines IL-8, MIP3α and TNFα were investigated. Greater levels of cytokine responses were induced by LPS in whole blood than in the THP-1 cells (Figure 6.1 versus 6.2). While there was more variation in the blood model than the THP-1 model, there were no significant differences between the wild-type strain and the mutants for any of the cytokines tested (Figure 6.1).

In the THP-1 model, the cytokines IL-8 and TNFα were tested. All the bacteria tested induced significantly more IL-8 and TNFα than LPS controls (Figure 6.2), although there were no significant differences between any of the mutant strains and the wild type (Figure 6.2).

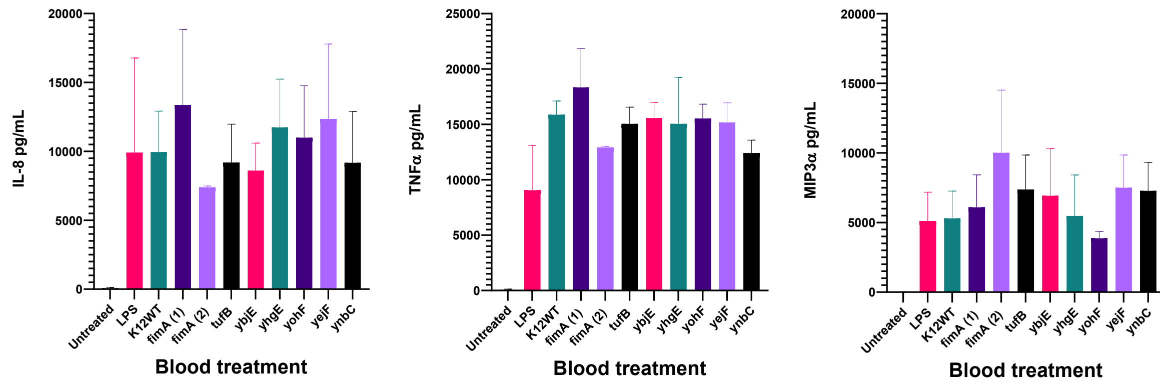


Figure 6.1 Whole blood cytokine response to *E. coli* keoi collection mutants and wild type strains. *Ex vivo* whole blood was infected with *E. coli* K12 mutants or wild type strain or LPS. After 4 hours platelet poor plasma was collected for cytokine quantification by ELISA. No statistically significance found. Kruskal-Wallis test.

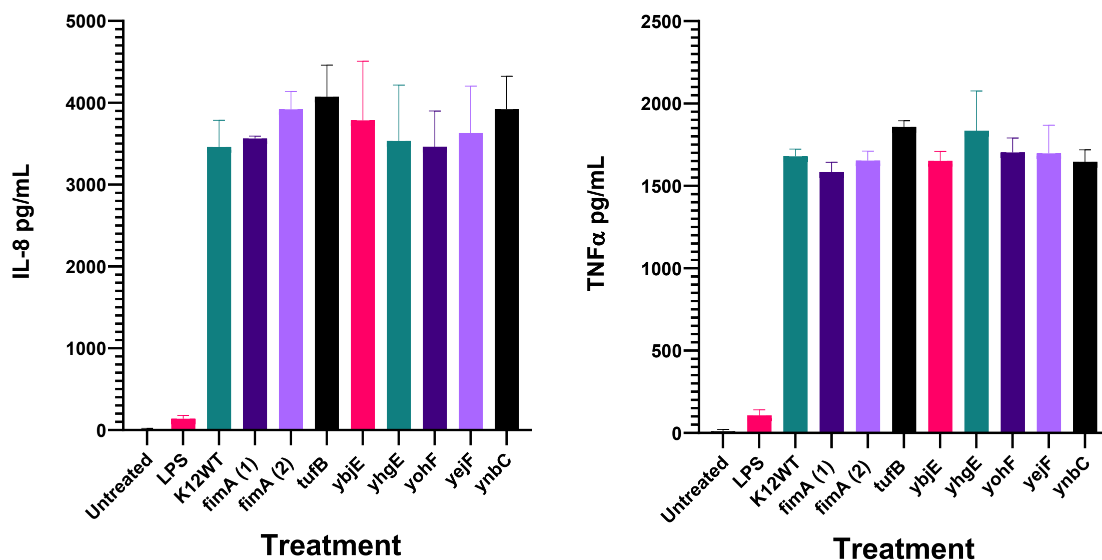


Figure 6.2 THP-1 cytokine response to *E. coli* mutants.
THP-1 cells were infected with *E. coli* Keio mutants of wild type strain for 4 hours. Supernatants were collected for cytokine quantification by ELISA. No statistically significant differences found. Kruskal-Wallis test.

6.4.9 *E. coli* K12 mutant growth in human plasma

The previous analysis showed that *tufB* was one of our lead targets for a role in bacteraemia (tables 6.1, 6. 2 and 6. 6). Therefore, to determine whether *tufB* affected growth in plasma, wild type K12 and a *tufB* mutant were grown in human plasma for 16 hours (Figure 6.3). No significant differences were observed between the strains. The mutant strain did show slightly earlier log growth (9 versus 11 hours) in 11.25% plasma compared to the wild type although this was not significant (Figure 6.3).

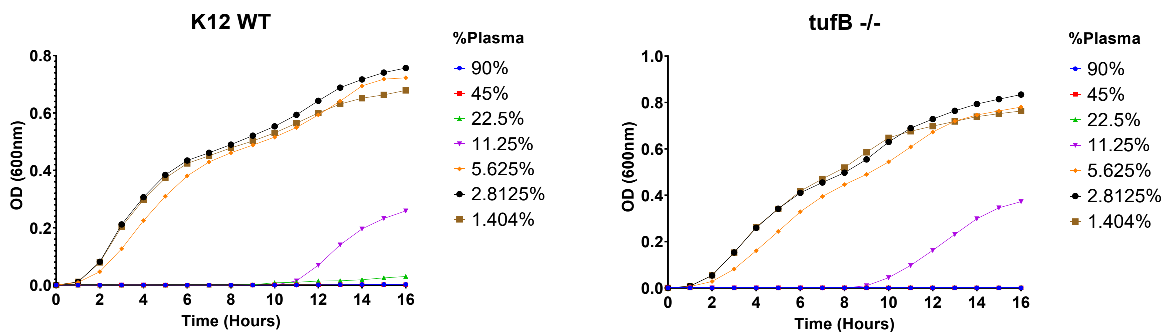


Figure 6.3: *E. coli* keoi collection wild type and tufB knockout strain growth in human plasma. *E. coli* mutant *tufB*-/- and wild type was grown in varying concentrations of human plasma (0-90%) for 16 hours. OD 600nm readings were taken every hour.

6.5 Discussion

This chapter aimed to use the scoary GWAS analysis pipeline on genome sequences from the *E. coli* bacteraemia collection to identify associated genes of interest with bacterial phenotypes generated earlier in the thesis from clinical or laboratory data. Furthermore, potential gene targets were tested with *E. coli* gene deficient mutants. Scoary resulted in significant numbers of gene targets (Tables 6.1-6.5). One advantage of the Scoary method is the generation of an output incorporating specificity and sensitivity, which allowed the potential for ranking gene targets by their strength of association. Another advantage of using Scoary is the ability to use relatively few isolates in comparisons with as few as 20 samples in some cases being sufficient to infer a causal link[239].

Gene targets

The gene targets chosen for mutant studies were chosen based on several qualifying factors including, having the highest significance in each of the phenotype associations made and importantly the availability of a K12 mutant (Table 6.5).

Three genes identified in the final gene list were protein transporters. The gene *ybjE* was found to be associated with bacteraemia isolates. This gene encodes a lysine exporter which controls export of lysine from the cell as well as mediating resistance to the toxic lysine antimetabolite L-thialysine [470,471]. Mortality was found to be associated with the *yhgE* gene, which is the only member of the putative transporter *yhgE* family and its function remains unknown[472]. Under anaerobiosis, its expression is activated by *arcA*, and given this link the gene may play a role in anaerobic metabolism[473,474]. The third protein transporter in the final gene list was the ABC transporter *yejF.* This gene has been found to be important for resistance to antimicrobial peptides in both *Brucella meltensis* and *Salmonella enterica* [475,476]. Apart from direct resistance to host defence (*yejF)* it is unclear how these genes contribute to pathogenicity and virulence in *E. coli* isolates, and further investigations into the role these genes play is required.

The final two genes included in the final table were the transcription factors *yohF* and *tufB.* There is scant literature on the *yohF* gene, and it is predicted to be a oxidoreductase. In stark contrast there is a vast amount of literature on *tufB* which encodes Elongation factor Tu (EF-Tu). EF-Tu is a G protein catalysing the binding of aminoacyl-tRNA to the ribosome in living cells [477]. It is one of the most abundant proteins found in bacteria accounting for 6% of the total protein expressed in *E. coli* [477]. Interestingly, EF-Tu is the target of the class of antibiotics known as elfamycins. Due to their poor pharmacokinetics, they are not approved for human use, however as the commercial antimicrobial pipeline slows on the global background of increasing AMR, they are a consideration for the future [478,479] In addition, *tufB* may be a promising vaccine candidate as mice immunised with EF-Tu were significantly protected against lethal challenge with *Streptococcus pneumoniae* [480]. EF-Tu has also been shown to have a diverse number of roles in multiple species of bacteria, including the ability to moonlight on the bacterial cell surface where it may have several functions[481–483]. It has also been shown to be present in outer membrane vesicles (OMVs) secreted by Shiga toxin producing *E. coli* [484]. Given the association with *E. coli* bacteraemia strains in

this study, *tufB* could be a potential avenue for biomarker investigation or even a potential therapeutic vaccine target. Indeed, recent research has linked EF-Tu with ExPEC isolates and the ability to bind and acquire iron from holo-transferrin[485].

Ef-Tu can also be recognised by cells of the immune system. EF-Tu is recognised by plant cells in the *Aribidopsis* species where an elongation factor receptor (EFR), binds Ef-Tu and stimulates a defence responses against the plant pathogenic bacteria *Agrobacterium*[486]. In colon epithelial cells HT29, recombinant Ef-Tu is recognised (in the presence of sCD14) and stimulates an inflammatory IL-8 response[487]. In addition, the stomach pathogen *Helicobacter pylori* may utilise secreted Ef-Tu to bind to THP-1 cells which may play a role in pathogenesis[488].

Intriguingly, EF-Tu has been linked with serum and complement resistance in multiple species of bacteria through binding to Factor H of the alternative complement pathway and has also been shown to inhibit C3b deposition by binding plasminogen[481]. To investigate whether *tufB* had a serum/plasma resistance effect in the current isolates *E. coli* wild type and mutants were grown in varying concentrations of human plasma. No significant differences were found between the two, indicating that *tufB* is not associated with serum resistance in *E. coli* (Figure 6.3). However, this study used citrated human plasma as a substitute for human serum, (limited during COVID). It has been shown that complement proteins are inhibited by a lack of calcium (the mechanisms of anti-coagulation for sodium citrate) [489–491]. This could explain why there was no significant difference between the mutant and the wild type, more investigation is needed into the role of EF-Tu in serum killing.

Of note, the genes not included in the mutant studies, but which demonstrated strong significant associations were *bla, tpd, iutA*, and *php_1* (Appendix 6.1) did not have an available K12 mutant. The gene *iutA* which encodes the ferric aerobactin receptor was also found to be associated with bacteraemia using the Chi-squared test (Chapter 5, Table 5.2). This highlights the importance that iron metabolism plays during infection[84,492–494].

It was of concern that GWAS analysis also identified resistance genes associated with bacteraemia; β-lactamase gene *bla*. As discussed previously β-lactams are a serious concern for healthcare systems worldwide as they confer resistance to commonly prescribed antibiotics. As incorrect antimicrobial therapies are associated with increased mortality and

length of time in hospital[495] it is a concern that β-lactamase genes are becoming associated with bacteraemia[496–498]. Screening antimicrobial resistance genes panels could be adapted to add further β-lactamase variants to stratify patients at risk of developing sepsis. This may help alleviate mortality based on incorrect antimicrobial therapy use in nosocomial acquired infections as these patients are more likely to be administered incorrect antibiotics [495], but would not be helpful for community acquired infections.

Another gene associated with the urinary source isolates in this study was *fhuC (*Appendix 6.2). This forms part of the Fhu system which is comprised of the genes *fhuABCDE* [494]. This system can transport a variety of ferric complexes, including aerobactin and has previously been associated with UPEC infections as well as being upregulated in a murine UTI model[494,499,500]. Given the association with urinary bacteraemia isolates in this study it is worth investigating what effect this system has on bacteraemic potential of urinary *E. coli* isolates.

Urinary isolates were also compared to other bacteraemia isolates in this study, and one gene of interest was the aerobactin gene *iucA.* This was also found to be associated with bacteraemia in Chapter 5 indicating that this may be due to the skew towards a urinary source in this study. Although, as previously mentioned this skew is not unusual as *E. coli* is the leading cause of UTIs worldwide [501].

Limitations

This study had a few limitations; firstly, phylogeny was not controlled for in this study which may impact the strength of association in the gene targets found. It may prove necessary to adjust Scoary commands to infer phylogeny in the future. Secondly, there were also several genes found in this study that were duplicated in the scoary outputs and further investigation of gene names would be required. Blast investigation into these genes showed that they were the same gene (data not shown). More stringent parameters when running scoary analysis may help to eliminate this in future studies. Alternative programmes to scoary, such as PySEER and Phenotype Seeker [445], are also available and could be a potentially useful tool to corroborate findings and strengthen genetic and phenotypic links.

Finally, an important limitation of this chapter is the use of K12 as a model organism. *E. coli* K12 is a non-pathogenic organism and by its nature is less virulent than bacteraemia isolates[502]. It is also serum sensitive and lacks many of the virulence factors of pathogenic strains. This limited the number of genes that could be tested in the *in vitro* models purely due to the lack of an available mutant (Appendix 6.1). Future studies where mutants are generated in pathogenic isolates (e.g CFT073) will give a fuller understanding of the pathogenic genes in these ExPEC isolates[503].

This chapter has identified associations with *E. coli* genes and clinical and laboratory phenotypes, which will prove useful as lead targets for biomarker identification in the treatment of bacteraemia.

Chapter 7: General discussion

7.1 Overarching thesis aims and hypothesis

This research aimed to take a combined approach to identifying *E. coli* sepsis biomarkers by considering both the variability in *E. coli* genetic makeup and the variability in induced host response to infection. It was hypothesized that *E. coli* strains causing bacteraemia would be genetically distinct and elicit a distinctive and unique host immune response. It was also predicted that isolates from defined clinical e.g., sepsis/non sepsis, source of infection(Appendix 3.4) and laboratory (serum resistance, cytokine production) defined groups would be genetically distinct and would elicit a distinct immune response.

7.2 Results chapter summaries

Chapter 3

Chapter 3 aimed to collect, archive, and characterise a collection of *E. coli* isolates taken from bacteraemia patients in the HDUHB along with associated patient demographic data (e.g co-morbidities, origin of infection, co-infections etc). This chapter also aimed to identify AMR resistance of the *E. coli* collection from hospital performed AMR testing (collected with patient data) and to perform basic genetic organisation including sequence type and phylogroup.

In total, 152 bacteraemia isolates were collected from the Hywel Dda University Health Board and used in this study, (along with associated patient data), 7 water isolates and 6 ANC urine control strains (section 3.7.1). Most of the bacteraemia isolates belonged to the phylogroup B2 (Table 3.2) and two sequence types ST131 and ST73 predominated (Figure 3.3). This is in line with reports of other bacteraemia collections as B2 is the most identified phylogroup[504–506], and ST131 and ST73 being the most common sequence types in bacteraemia isolates [507,508].

The *E. coli* isolates from this study came from three main sources of infection, the urinary tract, the abdominal tract, and the biliary tract as well as others which were classified as either unknown or other. The urinary tract was the most common source of infection compared to the other groups. This is not surprising given the prevalence of *E. coli* in infections of the urinary tract[509]. Indeed, 15 % of patients in this study were suffering from UTI infections (Table 3.2) and other studies have also reported a predominant urinary source of bacteraemia[297,298,510].

Antimicrobial resistance data showed that, isolates in this study had lower levels of resistance to multiple antibiotics compared to the English surveillance program (Table 3.1)[284]. However, it is likely that the bacteria were collected earlier for this study (2018-2020) than those in the surveillance programme report (2021-2022). The increasing trend in AMR resistance levels in all bacterial species including *E. coli* may account for these differences [274,284,511], Indeed, resistance levels of *E. coli* have been seen to increase by approximately 15% since 2017[511]. Further collections and AMR profiling of *E. coli* bacteraemia isolates from the HDUHB would allow resistance profiles of these isolates over time to be determined.

<u>Chapter 4</u>

Chapter 4 aimed to determine the diversity of host cytokine responses using two models (whole blood and THP-1) of infection to *E. coli* isolates collected in chapter 3 and determine if cytokine responses could be used to discriminate between certain bacterial phenotypes. It also aimed to determine the plasma resistance of the *E. coli* collection and investigate the role of complement proteins in bacterial killing.

Significantly more IL-8 and MIP3α was found in response to bacteraemia strains compared to non-pathogenic strains used in this study (Figure 4.2). IL-8 has previously been shown to be of predictive value for severe infection during burn injury [512], and, in combination with PCT and IL-10, has been shown to be required for the early diagnosis of neonatal sepsis [513]. This may indicate its promise as a potential biomarker. Less research has been done on the value of MIP3α as a marker for bacteraemia and sepsis, although recent research suggests it is elevated in lethal sepsis infection[371]. Further research on MIP3α viability as a biomarker of bacteraemia and sepsis is warranted.

This study also showed variation in the host response induced by isolates from different sources of infection. Both urinary and abdominal bacteraemia isolates elicited significantly more IL-8 than isolates from a water source (Figure 4.4) and abdominal isolates induced significantly more MIP3α than water isolates (Figure 4.4). This may partly be explained by the fact that the majority of the water isolates did not belong to either of the main bacteraemia phylogroups (B2 and D) as isolates from these two phylogroups have been shown to express more virulence factors and be associated with biofilm formation[510,514–516]. It may also be explained by the limited number of water isolates used in this study (7).

Isolates that were plasma resistant showed increased levels of IL-6, IL-8 and resistin compared to plasma sensitive isolate (Figure 4.9). Overall, this is not surprising and may be attributed to the greater PAMP dose stimulating immune cells due to increased numbers of bacteria surviving throughout the assay. This is, however, worth investigating further as resistance to plasma/serum can be manipulated by multiple bacterial factors [517]. Indeed, *E. coli* has been shown to modulate host immune responses, including through directing the host response to produce inhibitory antibodies, preventing complement activation [517,518]. Individual genes from *E. coli* have also been linked to the modulation of the host immune response including the secreted autotransporter (*sat)*, which is able to cleave complement proteins (C2, C3, C4, C5, C6 and C8)[519]. Additionally, the gene *prc* has also been shown to modulate survival of bacteria in serum albeit through a different mechanism. Mutant strains of *prc* have been shown to have modulated outer membrane properties

making them more susceptible to complement mediated lysis, indicating a role in outer membrane regulation[520].

The current gold standard for sepsis biomarkers are PCT and CRP[314,521]. While it is unlikely that any of the potential biomarkers used in this study will replace the use of PCT and CRP in the clinic, there is very good potential for panel biomarker approaches to the identification of sepsis[45–47,227,228,522,523]. This thesis highlights three potential markers (IL-6, IL-8 and MIP3α) for use in identifying sepsis and bacteraemia and discriminating between bacterial phenotypes. Indeed, one application for the future may be to apply machine learning to these potential markers of infection. Studies have shown that 30-day mortality can be predicted more accurately with a combination panel of biomarkers identified through machine learning compared to PCT and CRP [228].

Chapter 5

Chapter 5 aimed to generate an *E. coli* pangenome using a selection of the isolates acquired from HDUHB, 92 bacteraemia isolates, 7 water isolates, and 6 ANC urine isolates. The ECOR collection and reference ATCC isolates (Table 3.2) were also included in the pangenome. The pangenome was then used to identify virulence factors and AMR genes which were compared between bacteraemia and ECOR as well as sources of infection.

This thesis identified that 19/53 common ExPEC virulence factors were significantly increased in the bacteraemia isolates compared to the ECOR collection (Table 5.2). Of these 12 were associated with iron uptake. This underlines the importance of these virulence factors in ExPEC pathogenesis and is consistent with previous studies. [494,524,525]. Two genes encoding components of the S fimbriae system (*sfaC* and sfaX) were identified in this study as being associated with different sources of infection. This is a particularly exciting result as very few genes have been associated with a source of infection. Both genes encode regulatory proteins and the gene *sfaX* has previously been associated with the development of neonatal meningitis through its influence on motility of these isolates[432]. The association of components of the same virulence system with different sources of infection in this study (Table 5.3) is an interesting area for future research. It might be interesting to

speculate that the S fimbriae system is responsible for 'deciding' the initial adhesive steps taken in each local infection site.

The virulence factors *csgF, entD, fepD, fepG* and *ompA* were identified in all isolates of the *E. coli* collection. Interestingly, the gene *ompA*, has been associated with resistance to serum [526,527]. Given that there was variability in the capacity of *E. coli* isolates to grow in both human plasma and serum (Figure 4.7 and appendix 4.2) it may be that the presence of genes such as *ompA* alone, is not sufficient to predict traits (e.g., serum resistance).

This also raises the question about whether individual genes can be used to identify ExPEC bacteria as there is known overlap between pathotypes (Figure 5.1). Given the plasticity of the *E. coli* genome this is not surprising[68]. Indeed, studies on commensal *E. coli* from animal and human sources have shown that a variety of *E. coli* from different sequence types harbour a large degree of virulence associated genes highlighting their potential for ExPEC pathogenesis[528]. *E. coli* virulence factors are often clustered together on pathogenicity islands which are acquired through horizontal gene transfer[529–531]. It is therefore easy to envision the transfer of genetic material from a pathogenic strain to a usually commensal strain. Presence/absence of gene/genes alone may not be sufficient to identify pathogenic variants of *E. coli.* Gene expression studies may be an alternative approach to evaluating genes of interest in *E. coli* isolates under various conditions.
A recent single cell multiplexed sequencing study using *E. coli* have highlighted heterogeneity within populations of the same *E. coli* strain[532]. The authors suggest a bet-hedging model by which subpopulations of *E. coli* can pre-emptively activate the acid tolerance genes *gadA* and *gadB* for protection in an acidic environment e.g., the stomach[532]. This invites the question as to whether other genes follow a similar pattern and is worth further investigation, especially for virulence genes associated with pathogenic phenotypes i.e., serum resistance. Indeed, some virulence factors have been shown to be regulated by the same process known as phase variation. The *pap* operon has been shown to be epigenetically regulated by the deoxyadenosine methylase *dam* which can lead to *pap* on and *pap* off variants within a population of bacteria.
This is a potential avenue for further research and could be the best way not only to identify potential biomarkers but also to determine mechanisms of action for many virulence genes.

This in turn opens the possibility of future target for treatment options such as small molecule and monoclonal antibody therapies.

Chapter 6

Chapter 6 aimed to compare *E. coli* whole genomes with clinical and laboratory phenotypes using a GWAS approach. This involved combining the pangenome with clinical phenotypes created in chapter 3, and the laboratory phenotypes generated in chapter 4. It also aimed to investigate the effect of single gene knockouts of *E. coli* genes of interest on host response.

After analysis with scoary a top gene list was created using genes with the highest sensitivity and specificity for traits tested as well as availability in the K12 knockout model. The genes identified were, *ynbC, yhgE, ybjE, yejF, tufB* and *yohF* (Table 6.5). There is limited research on these gene targets in the literature although their functions as transporters and transcription factors point to a more subtle effect on virulence. The gene which showed the most promise was the elongation factor Tu encoding gene, *tufB.* This is an ancient and conserved protein which has recently been linked to a multitude of roles in pathogenesis via moonlighting on the cell surface, in addition to its main role catalysing the binding of tRNA to ribosomes[477,481,483]. Research regarding its role in *E. coli* pathogenesis is still limited although there are recent reports implicating a role in binding and acquiring iron from holo-transferrin[485]. Given the association with bacteraemia and a urinary source of infection in this study further research into its role in *E. coli* pathogenesis is needed. Understanding the mechanism by which *tufB* increases virulence may not only lead to a deeper understanding but also to potential avenues for treatment elongation factor Tu is the target of the antimicrobial drug class of elfamycins, this could spark new interest in improving the efficiency and tolerance of these unused antibiotics[479]. Additionally, it may be worth trialling the detection of *tufB* detection in bacteraemia patient blood samples to investigate its use as a biomarker for *E. coli* bacteraemia.
Although there was no phenotype identified in the infection models with single gene knockout K12 strains there is potential to use a model pathogenic organism in place of K12 as during this study a transposon library for the urosepsis strain CFT073 was developed [503].

7.3 General discussion and future work

This research brings numerous benefits to the scientific community, diagnostic laboratories, and patients. The bacterial and host biomarkers identified in this study could have direct benefits to the NHS in diagnosis and guiding treatment of bacteraemia and sepsis infections. The whole blood model of infection used in this study could be a potentially useful tool in investigating differences in phenotypes of not only different bacterial species but also within species [46,47]. Unfortunately, due to timing constraints and the need to develop a new model to replace blood (COVID 19) there was not enough time to investigate the immunosuppressed blood model of infection previously used by our group[533]. Future work on this model of infection for bacteraemia would be an obvious starting point for future studies as it would allow mimicry of the immunosuppressed state of many bacteraemia patients, further validating biomarkers identified in this study.

To validate host immune biomarkers in a clinical setting, further studies could involve the direct detection of cytokines in patient blood samples. Thus, future ethics applications could study involve isolating serum in patients with bacteraemia in HDUHB. Unfortunately, the current study focused on bacteria analysis only, but the extra patient blood sample is an area where future research should focus. Studies have highlighted the changing cytokine profiles with different stages of sepsis [534,535], and given this, it would be beneficial to measure cytokine levels over the course of infection using a small pilot population, to differentiate cytokine responses based on bacterial phenotypes such as origin of infection.

Detection of genes associated with bacterial virulence may lead to implementation of screening at-risk patients. Indeed, virulence factor detection has been successful for several organisms including, *Pseudomonas aeruginosa* (Quorom sensing signalling molecules), *Staphylococcus aureus*(*nuc*) (and *Helicobacter pylori* (*rdxA, pbp1)* [536,537]. Methods involved in such an implementation may involve PCR, cell free DNA detection, proteolytic signatures, mass spectrometry, next generation sequencing and advanced modelling strategies. Indeed, bacterial genetic markers of infection may provide a way forward in the early diagnosis of bacteraemia and sepsis. Advances in genetics-based pathogen detection methods are showing early signs of promise as a recent study using MinION nanopore sequencing has been shown to identify *Klebsiella pneumoniae* within 2 hours as well as

provide enough information to predict virulence genes and AMR resistance[538]. Similar studies involving *E. coli* may prove challenging due to the number of virulence factors and genome plasticity of the species, however detection and AMR resistant profiles have been detected in blood samples spiked with *E. coli*[539]. Alternative assays such as DNA based microarrays are also showing promise, allowing the detection of 44 different pathogens, 360 virulence factors and 409 AMR genes simultaneously [540]. While the presence of a virulence factor may not necessarily correlate with pathogenicity, especially in bacteria with high genome plasticity such as *E. coli,* a list of lead gene targets is an exciting and simple advance for future trial.

It has not escaped our attention that the work contained within this thesis developed from a successful surveillance program to detect *E. coli* bacteraemia's in HDUHB between 2002 and 2016[254]. In this study, our clinical colleagues concluded that a major strength was the management of sepsis, but a major weakness was reducing *E. coli* bacteraemia's. Therefore, can this thesis address this weakness and generate societal impact in the health board? The study had a clear aim associated with biomarker and target identification, and this was achieved with lead targets associated with host (IL-6, IL-8, resistin and MIP3α) and pathogen (*ynbC, yhgE, ybjE, yejF, tufB* and *yohF*). Translating these targets into a pilot study in bacteraemia patients, where host biomarkers may be detected by ELISA and the virulence factors detected by PCR is a clear goal going forward.

In the future, there is a clear goal in expanding key aspects of the current study. The most important issue would be to increase the size of the dataset. Firstly, by increasing the number of isolates and secondly the number of geographic regions (health boards) where they are isolated. Thirdly, there is need to even up group sizes and thus including more abdominal isolates. These adjustments would allow for more advance modelling and multivariate analysis of the data enabling better phenotype-genotype links to be generated. Recent work has suggested that discrimination down to the strain level is possible using host immune response data[214].

This study used roary to generate the pangenome prior to analysis of the *E. coli* collection, but there are other options available, including Panaroo, PPanGGoLin and PIRATE[541–543].

Given the fact that roary is no longer actively updated, it could prove beneficial to adopt alternatives for future studies [544]. Both PIRATE and PPanGGolin generate a gene presence/absence file in their output files which can be used for pangenome association studies. In addition, software such as Panaroo claims to be able to account for some common errors attributed to alternative approaches such as the classification of homologous genes [543]. While there are advantages to using roary, as the input directly feeds into downstream analysis software such as scoary, there are now methods for converting outputs to a roary format allowing the output of PIRATE to be used as the input for scoary[542].

Scoary was the GWAS tool used in this study, and it remains the most cited microbial GWAS software[545]. There are alternative programmes available reviewed here [445] and some make use of advances in machine learning based methods such as PySEER, Kover and PhenotypeSeeker[546–548]. It is also expected that advances in machine learning technologies and improved user training will lead to the influx of tools which support machine learning [445]. A detailed study comparing different bioinformatics pipelines could be useful in determining any differences in the genes found to be associated to bacterial traits and will likely solidify associations made further enhancing the results of this study.

7.4 Limitations

This study was not without limitations, an important one being the attempt to study the functional effects of *E. coli* gene target deletion in whole blood and THP-1 models. For this, mutant studies were conducted with the K12 Keio collection. Firstly, the lack of available mutants in the Keio collection did not allow for the testing of many of the genes associated with bacterial phenotypes to be tested (Chapter 6) and the final gene candidate list was limited to six. Therefore, no difference in the host response to *E. coli* mutants was observed in either of the host infection models (Figure 6.1 and 6.2). Using a clinical bacteraemia *E. coli* as the genetic background would not only allow more genes of interest to be investigated in host models of infection but would also provide significantly more insight into the effect of bacterial gene deletion on host response. Indeed, a transposon mutant library was created in CFT073 recently that could fill this role [503].

The patient associated data collected during this work relied on ethics applications and our clinical colleagues. Data on some of the patients is incomplete (COVID pandemic) and there are clear opportunities here for further investigation of patient records. This aspect of the research is time consuming and requires trained professionals to investigate notes and databases of patient records. This is surely an investment for the future.

This thesis originally intended to make use of the whole blood model alone, but access to whole blood was restricted during COVID. This led to developing an alternative model and the use of the THP-1 model which, did provide interesting results (particularly the bacteria mediated killing of THP-1 cells) but is not as robust as the whole blood model. Thus, only a subset of clinical isolates was subjected to whole blood infection. Future work should run the whole blood infection on remaining isolates.

The whole blood model of infection is an efficient and ethical way of measuring cytokine responses to systemic infections with advantages over animal and human models of infection. However, there are limitations, firstly this study was limited to using healthy volunteer donors which occasionally did not show up for appointments despite renumeration and restrictions placed on the number of blood donations for individuals in a given time period. This should be accounted for in future studies as there was a large degree of variability in cytokine responses between donors. Indeed, this variability in the cytokine responses between donors highlights the complex nature of sepsis from the perspective of the host response to infection.

In addition, as discussed in chapter 4 use of anticoagulants in blood is unavoidable due to the limited shelf life of blood in their absence, but they may have effects on the complement cascade. An alternative anticoagulant which selectively inhibits thrombin and does not inhibit complement is hirudin which may be of use for future studies, however this is more expensive and generally harder to produce in large batches [385,549].

Further work including direct analysis of patient blood samples may help account for this shortcoming. Alternatively, co-culture models of infection including tissue specific cultures such as organoids may also be of use for identifying tissue specific immune responses associated with bacterial phenotype.

References

[1] Gotts JE, Matthay MA. Sepsis: pathophysiology and clinical management. BMJ 2016;353:i1585. https://doi.org/10.1136/bmj.i1585.

[2] Singer M, CS D, Seymour C, al et. The third international consensus definitions for sepsis and septic shock (sepsis-3). JAMA 2016;315:801–10.

[3] Fleischmann C, Scherag A, Adhikari NKJ, Hartog CS, Tsaganos T, Schlattmann P, et al. Assessment of Global Incidence and Mortality of Hospital-treated Sepsis. Current Estimates and Limitations. Am J Respir Crit Care Med 2016;193:259–72. https://doi.org/10.1164/rccm.201504-0781OC.

[4] Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. The Lancet 2020;395:200–11. https://doi.org/10.1016/S0140-6736(19)32989-7.

[5] York Health Economics Consotrium. The cost of sepsis care in the UK. 2017.

[6] The UK Sepsis Trust. The Sepsis Manual. 2019.

[7] McPherson D, Griffiths C, Williams M, Baker A, Klodawski E, Jacobson B, et al. Sepsis-associated mortality in England: an analysis of multiple cause of death data from 2001 to 2010. BMJ Open 2013;3:e002586. https://doi.org/10.1136/BMJOPEN-2013-002586.

[8] Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. Crit Care Med 2001;29:1303–10.

[9] Dolin HH, Papadimos TJ, Stepkowski S, Chen X, Pan ZK. A Novel Combination of Biomarkers to Herald the Onset of Sepsis Prior to the Manifestation of Symptoms. SHOCK 2018;49:364–70. https://doi.org/10.1097/SHK.0000000000001010.

[10] Reinhart K, Daniels R, Kissoon N, Machado FR, Schachter RD, Finfer S. Recognizing Sepsis as a Global Health Priority — A WHO Resolution. New England Journal of Medicine 2017;377:414–7. https://doi.org/10.1056/NEJMp1707170.

[11] World Health Organisation. Seventieth World Health Assembly-Improving the prevention,diagnosis and clinical management of sepsis. 2017.

[12] Vincent JL. The Clinical Challenge of Sepsis Identification and Monitoring. PLoS Med 2016;13. https://doi.org/10.1371/journal.pmed.1002022.

[13] Sriskandan S, Cohen J. The pathogenesis of septic shock. Journal of Infection 1995;30:201–6. https://doi.org/10.1016/S0163-4453(95)90647-9.

[14] Bishara J, Leibovici L, Huminer D, Drucker M, Samra Z, Konisberger H, et al. Five-year prospective study of bacteraemic urinary tract infection in a single institution. European Journal of Clinical Microbiology and Infectious Diseases 1997;16:563–7. https://doi.org/10.1007/BF02447917.

[15] Mackowiak PA, Browne RH, Southern PM, Smith JW. Polymicrobial sepsis: An analysis of 184 cases using log linear models. American Journal of the Medical Sciences 1980;280:73–80. https://doi.org/10.1097/00000441-198009000-00002.

[16] Opal SM, Laterre P-F, Francois B, LaRosa SP, Angus DC, Mira J-P, et al. Effect of Eritoran, an Antagonist of MD2-TLR4, on Mortality in Patients With Severe Sepsis. JAMA 2013;309:1154. https://doi.org/10.1001/jama.2013.2194.

[17] Bernard GR, Vincent J-L, Laterre P-F, LaRosa SP, Dhainaut J-F, Lopez-Rodriguez A, et al. Efficacy and safety of recombinant human activated protein C for severe sepsis. N Engl J Med 2001;344:699–709. https://doi.org/10.1056/NEJM200103083441001.

[18] Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. Crit Care Med 2013;41:580–637. https://doi.org/10.1097/CCM.0b013e31827e83af.

[19] Ferrer R, Martin-Loeches I, Phillips G, Osborn TM, Townsend S, Dellinger RP, et al. Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program. Crit Care Med 2014;42:1749–55. https://doi.org/10.1097/CCM.0000000000000330.

[20] Ferrer R, Martin-Loeches I, Phillips G, Osborn TM, Townsend S, Dellinger RP, et al. Empiric Antibiotic Treatment Reduces Mortality in Severe Sepsis and Septic Shock From the First Hour. Crit Care Med 2014;42:1749–55. https://doi.org/10.1097/CCM.0000000000000330.

[21] Liu VX, Fielding-Singh V, Greene JD, Baker JM, Iwashyna TJ, Bhattacharya J, et al. The Timing of Early Antibiotics and Hospital Mortality in Sepsis. Am J Respir Crit Care Med 2017;196:856–63. https://doi.org/10.1164/RCCM.201609-1848OC.

[22] Sherwin R, Winters ME, Vilke GM, Wardi G. Does Early and Appropriate Antibiotic Administration Improve Mortality in Emergency Department Patients with Severe Sepsis or Septic Shock? Journal of Emergency Medicine 2017;53:588–95. https://doi.org/10.1016/j.jemermed.2016.12.009.

[23] Kang CI, Kim SH, Wan BP, Lee KD, Kim H Bin, Kim EC, et al. Bloodstream infections caused by antibiotic-resistant gram-negative bacilli: Risk factors for mortality and impact of inappropriate initial antimicrobial therapy on outcome. Antimicrob Agents Chemother 2005;49:760–6. https://doi.org/10.1128/AAC.49.2.760-766.2005.

[24] Garnacho-Montero J, Ortiz-Leyba C, Herrera-Melero I, Aldabó-Pallás T, Cayuela-Dominguez A, Marquez-Vacaro JA, et al. Mortality and morbidity attributable to inadequate empirical antimicrobial therapy in patients admitted to the ICU with sepsis: a matched cohort study. J Antimicrob Chemother 2008;61:436–41. https://doi.org/10.1093/JAC/DKM460.

[25] Garnacho-Montero J, Garcia-Garmendia JL, Barrero-Almodovar A, Jimenez-Jimenez FJ, Perez-Paredes C, Ortiz-Leyba C. Impact of adequate empirical antibiotic therapy on the outcome of patients admitted to the intensive care unit with sepsis. Crit Care Med 2003;31:2742–51. https://doi.org/10.1097/01.CCM.0000098031.24329.10.

[26] Schwaber MJ, Carmeli Y. Mortality and delay in effective therapy associated with extended-spectrum beta-lactamase production in Enterobacteriaceae bacteraemia: a systematic review and meta-analysis. J Antimicrob Chemother 2007;60:913–20. https://doi.org/10.1093/JAC/DKM318.

[27] Seymour CW, Kennedy JN, Wang S, Chang CCH, Elliott CF, Xu Z, et al. Derivation, Validation, and Potential Treatment Implications of Novel Clinical Phenotypes for Sepsis. JAMA 2019;321:2003–17. https://doi.org/10.1001/JAMA.2019.5791.

[28] Dhoble A, Chung W. Rapid Hemodynamic Deterioration and Death due to Acute Severe Refractory Septic Shock. J Clin Med Res 2009;1:109. https://doi.org/10.4021/JOCMR2009.04.1238.

[29] Durán-Méndez A, Aguilar-Arroyo AD, Vivanco-Gómez E, Nieto-Ortega E, Pérez-Ortega D, Jiménez-Pérez C, et al. Tocilizumab reduces COVID-19 mortality and pathology in a dose and timing-dependent fashion: a multi-centric study. Scientific Reports 2021 11:1 2021;11:1–8. https://doi.org/10.1038/s41598-021-99291-z.

[30] Koçak Tufan Z, Kayaaslan B, Mer M. COVID-19 and Sepsis. Turk J Med Sci 2021;51:3301–11. https://doi.org/10.3906/SAG-2108-239.

[31] Uckun FM, Orhan C, Powell J, Sahin E, Ozercan IH, Volk M, et al. Non-clinical safety profile and pharmacodynamics of two formulations of the anti-sepsis drug candidate Rejuveinix (RJX). Biomed Pharmacother 2021;141. https://doi.org/10.1016/J.BIOPHA.2021.111823.

[32] Watson OJ, Barnsley G, Toor J, Hogan AB, Winskill P, Ghani AC. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. Lancet Infect Dis 2022;22:1293–302. https://doi.org/10.1016/S1473-3099(22)00320-6.

[33] Opota O, Croxatto A, Prod'hom G, Greub G. Blood culture-based diagnosis of bacteraemia: state of the art. Clinical Microbiology and Infection 2015;21:313–22. https://doi.org/10.1016/J.CMI.2015.01.003.

[34] O'Dwyer MJ, Starczewska MH, Schrenzel J, Zacharowski K, Ecker DJ, Sampath R, et al. The detection of microbial DNA but not cultured bacteria is associated with increased mortality in patients with suspected sepsis-a prospective multi-centre European observational study. Clin Microbiol Infect 2017;23:208.e1-208.e6. https://doi.org/10.1016/J.CMI.2016.11.010.

[35] O'Dwyer MJ, Starczewska MH, Schrenzel J, Zacharowski K, Ecker DJ, Sampath R, et al. The detection of microbial DNA but not cultured bacteria is associated with increased mortality in patients with suspected sepsis-a prospective multi-centre European observational study. Clin Microbiol Infect 2017;23:208.e1-208.e6. https://doi.org/10.1016/J.CMI.2016.11.010.

[36] NHS. THINK sepsis. Https://Www.e-LfhOrgUk/Programmes/Sepsis/ n.d.

[37] KNAUS WA, ZIMMERMAN JE, WAGNER DP, DRAPER EA, LAWRENCE DE. APACHE—acute physiology and chronic health evaluation: a physiologically based classification system. Crit Care Med 1981;9.

[38] Ko M, Shim M, Lee SM, Kim Y, Yoon S. Performance of APACHE IV in Medical Intensive Care Unit Patients: Comparisons with APACHE II, SAPS 3, and MPM0 III. Acute and Critical Care 2018;33:216–21. https://doi.org/10.4266/ACC.2018.00178.

[39] Henry NL, Hayes DF. Cancer biomarkers. Mol Oncol 2012;6:140. https://doi.org/10.1016/J.MOLONC.2012.01.010.

[40] Póvoa P, Coelho L, Dal-Pizzol F, Ferrer R, Huttner A, Conway Morris A, et al. How to use biomarkers of infection or sepsis at the bedside: guide to clinicians. Intensive Care Med 2023;49:142–53. https://doi.org/10.1007/S00134-022-06956-Y.

[41] Dias S, Pheiffer C, Abrahams Y, Rheeder P, Adam S. Molecular Biomarkers for Gestational Diabetes Mellitus. Int J Mol Sci 2018;19. https://doi.org/10.3390/IJMS19102926.

[42] Sarhadi VK, Armengol G. Molecular Biomarkers in Cancer. Biomolecules 2022;12. https://doi.org/10.3390/BIOM12081021.

[43] Friedlaender A, Bauml J, Banna GL, Addeo A. Identifying successful biomarkers for patients with non-small-cell lung cancer. Lung Cancer Manag 2019;8:LMT17. https://doi.org/10.2217/LMT-2019-0009.

[44] Komorowski M, Green A, Tatham KC, Seymour C, Antcliffe D. Sepsis biomarkers and diagnostic tools with a focus on machine learning. EBioMedicine 2022;86:104394. https://doi.org/10.1016/j.ebiom.2022.104394.

[45] Conway SR, Wong HR. Biomarker Panels in Critical Care. Crit Care Clin 2020;36:89–104. https://doi.org/10.1016/J.CCC.2019.08.007.

[46] Tang Y, Ling N, Li S, Huang J, Zhang W, Zhang A, et al. A panel of urine-derived biomarkers to identify sepsis and distinguish it from systemic inflammatory response syndrome. Scientific Reports 2021 11:1 2021;11:1–11. https://doi.org/10.1038/s41598-021-99595-0.

[47] Casserly B, Read R, Levy MM. Multimarker panels in sepsis. Crit Care Clin 2011;27:391–405. https://doi.org/10.1016/J.CCC.2010.12.011.

[48] Barichello T, Generoso JS, Singer M, Dal-Pizzol F. Biomarkers for sepsis: more than just fever and leukocytosis-a narrative review. Crit Care 2022;26. https://doi.org/10.1186/S13054-021-03862-5.

[49] NICE. Sepsis: recognition, diagnosis and early management. 2016.

[50] Mitchell E, Pearce MS, Roberts A. Gram-negative bloodstream infections and sepsis: risk factors, screening tools and surveillance. Br Med Bull 2019;132:5–15. https://doi.org/10.1093/BMB/LDZ033.

[51] Polat G, Ugan RA, Cadirci E, Halici Z. Sepsis and Septic Shock: Current Treatment Strategies and New Approaches/Sepsis ve Septik Sok: Mevcut Tedavi Stratejileri ve Yeni Yaklasimlar. Eurasian J Med 2017;49:53–9.

[52] Śmiechowicz J. The Rationale and Current Status of Endotoxin Adsorption in the Treatment of Septic Shock. J Clin Med 2022;11. https://doi.org/10.3390/JCM11030619.

[53] Hoshino K, Takeuchi O, Kawai T, Sanjo H, Ogawa T, Takeda Y, et al. Cutting Edge: Toll-Like Receptor 4 (TLR4)-Deficient Mice Are Hyporesponsive to Lipopolysaccharide: Evidence for TLR4 as the Lps Gene Product. The Journal of Immunology 1999;162:3749–52. https://doi.org/10.4049/JIMMUNOL.162.7.3749.

[54] Sharif O, Bolshakov VN, Raines S, Newham P, Perkins ND. Transcriptional profiling of the LPS induced NF-κB response in macrophages. BMC Immunol 2007;8:1–17. https://doi.org/10.1186/1471-2172-8-1/TABLES/4.

[55] Opota O, Jaton K, Greub G. Microbial diagnosis of bloodstream infection: Towards molecular diagnosis directly from blood. Clinical Microbiology and Infection 2015;21:323–31. https://doi.org/10.1016/j.cmi.2015.02.005.

[56] Dolin HH, Papadimos TJ, Chen X, Pan ZK. Characterization of Pathogenic Sepsis Etiologies and Patient Profiles: A Novel Approach to Triage and Treatment. Microbiol Insights 2019;12:117863611882508. https://doi.org/10.1177/1178636118825081.

[57] Dolin HH, Papadimos TJ, Chen X, Pan ZK. Characterization of Pathogenic Sepsis Etiologies and Patient Profiles: A Novel Approach to Triage and Treatment. Microbiol Insights 2019;12:117863611882508. https://doi.org/10.1177/1178636118825081.

[58] Feezor RJ, Oberholzer C, Baker H V, Novick D, Rubinstein M, Moldawer LL, et al. Molecular characterization of the acute inflammatory response to infections with gram-negative versus gram-positive bacteria. Infect Immun 2003;71:5803–13. https://doi.org/10.1128/IAI.71.10.5803-5813.2003.

[59] Yan ST, Sun LC, Jia HB, Gao W, Yang JP, Zhang GQ. Procalcitonin levels in bloodstream infections caused by different sources and species of bacteria. Am J Emerg Med 2017;35:579–83. https://doi.org/10.1016/j.ajem.2016.12.017.

[60] Umemura Y, Ogura H, Takuma K, Fujishima S, Abe T, Kushimoto S, et al. Current spectrum of causative pathogens in sepsis: A prospective nationwide cohort study in Japan. Int J Infect Dis 2021;103:343–51. https://doi.org/10.1016/J.IJID.2020.11.168.

[61] Wolf AJ, Underhill DM. Peptidoglycan recognition by the innate immune system. Nature Reviews Immunology 2017 18:4 2018;18:243–54. https://doi.org/10.1038/nri.2017.136.

[62] Wang JE, Dahle MK, McDonald M, Foster SJ, Aasen AO, Thiemermann C. Peptidoglycan and lipoteichoic acid in gram-positive bacterial sepsis: receptors, signal transduction, biological effects, and synergism. Shock 2003;20:402–14. https://doi.org/10.1097/01.SHK.0000092268.01859.0D.

[63] Kang SS, Sim JR, Yun CH, Han SH. Lipoteichoic acids as a major virulence factor causing inflammatory responses via Toll-like receptor 2. Arch Pharm Res 2016;39:1519–29. https://doi.org/10.1007/S12272-016-0804-Y.

[64] Public Health Wales. Antibacterial Usage in Primary Care in Wales 2013-2017. 2018.

[65] García A, Fox JG. A One Health Perspective for Defining and Deciphering Escherichia coli Pathogenic Potential in Multiple Hosts. Comp Med 2021;71:3–45. https://doi.org/10.30802/AALAS-CM-20-000054.

[66] Bittinger K, Zhao C, Li Y, Ford E, Friedman ES, Ni J, et al. Bacterial colonization reprograms the neonatal gut metabolome. Nat Microbiol 2020;5:838. https://doi.org/10.1038/S41564-020-0694-0.

[67] Kaper JB, Nataro JP, Mobley HLT. Pathogenic Escherichia coli. Nat Rev Microbiol 2004;2:123–40. https://doi.org/10.1038/nrmicro818.

[68] Leimbach A, Hacker J, Dobrindt U. E. coli as an All-Rounder: The Thin Line Between Commensalism and Pathogenicity. Curr Top Microbiol Immunol, vol. 358, 2013, p. 3–32. https://doi.org/10.1007/82_2012_303.

[69] Geurtsen J, de Been M, Weerdenburg E, Zomer A, McNally A, Poolman J. Genomics and pathotypes of the many faces of Escherichia coli. FEMS Microbiol Rev 2022;46. https://doi.org/10.1093/FEMSRE/FUAC031.

[70] Dobrindt U, Chowdary MG, Krumbholz G, Hacker J. Genome dynamics and its impact on evolution of Escherichia coli. Med Microbiol Immunol 2010;199:145–54. https://doi.org/10.1007/S00430-010-0161-2/FIGURES/5.

[71] Proença JT, Barral DC, Gordo I. Commensal-to-pathogen transition: One-single transposon insertion results in two pathoadaptive traits in Escherichia coli -macrophage interaction. Scientific Reports 2017 7:1 2017;7:1–12. https://doi.org/10.1038/s41598-017-04081-1.

[72] Hacker J, Hentschel U, Dobrindt U. Prokaryotic Chromosomes and Disease. Science (1979) 2003;301:790–3. https://doi.org/10.1126/SCIENCE.1086802.

[73] Sarowska J, Futoma-Koloch B, Jama-Kmiecik A, Frej-Madrzak M, Ksiazczyk M, Bugla-Ploskonska G, et al. Virulence factors, prevalence and potential transmission of extraintestinal pathogenic Escherichia coli isolated from different sources: recent reports. Gut Pathogens 2019 11:1 2019;11:1–16. https://doi.org/10.1186/S13099-019-0290-0.

[74] Ochman H, Selander RK. Standard reference strains of Escherichia coli from natural populations. J Bacteriol 1984;157:690–3.

[75] Beghain J, Bridier-Nahmias A, Nagard H Le, Denamur E, Clermont O. ClermonTyping: An easy-to-use and accurate in silico method for Escherichia genus strain phylotyping. Microb Genom 2018;4. https://doi.org/10.1099/mgen.0.000192.

[76] Kunduru BR, Nair SA, Rathinavelan T. EK3D: an E. coli K antigen 3-dimensional structure database. Nucleic Acids Res 2016;44:D675. https://doi.org/10.1093/NAR/GKV1313.

[77] Fratamico PM, DebRoy C, Liu Y, Needleman DS, Baranzoni GM, Feng P. Advances in molecular serotyping and subtyping of Escherichia coli. Front Microbiol 2016;7:644. https://doi.org/10.3389/FMICB.2016.00644/BIBTEX.

[78] Reddick LE, Alto NM. Bacteria Fighting Back – How Pathogens Target and Subvert the Host Innate Immune System. Mol Cell 2014;54:321. https://doi.org/10.1016/J.MOLCEL.2014.03.010.

[79] Johnson JR, Stamm WE. Diagnosis and treatment of acute urinary tract infections. Infect Dis Clin North Am 1987;1:773–91. https://doi.org/10.1016/s0891-5520(20)30149-5.

[80] Tchesnokova V, Aprikian P, Kisiela D, Gowey S, Korotkova N, Thomas W, et al. Type 1 Fimbrial Adhesin FimH Elicits an Immune Response That Enhances Cell Adhesion of Escherichia coli. Infect Immun 2011;79:3895. https://doi.org/10.1128/IAI.05169-11.

[81] Tchesnokova V, Aprikian P, Kisiela D, Gowey S, Korotkova N, Thomas W, et al. Type 1 fimbrial adhesin FimH elicits an immune response that enhances cell adhesion of Escherichia coli. Infect Immun 2011;79:3895–904. https://doi.org/10.1128/IAI.05169-11.

[82] Van Dijk WC, Verbrugh HA, Van Der Tol ME, Peters R, Verhoef J. Role of Escherichia coli K capsular antigens during complement activation, C3 fixation, and opsonization. Infect Immun 1979;25:603. https://doi.org/10.1128/IAI.25.2.603-609.1979.

[83] Poole NM, Green SI, Rajan A, Vela LE, Zeng XL, Estes MK, et al. Role for FimH in extraintestinal pathogenic Escherichia coli invasion and translocation through the intestinal epithelium. Infect Immun 2017;85. https://doi.org/10.1128/IAI.00581-17/SUPPL_FILE/ZII999092195S1.PDF.

[84] Zhang Y, Sen S, Giedroc DP. Iron Acquisition by Bacterial Pathogens: Beyond Tris-Catecholate Complexes. ChemBioChem 2020;21:1955–67. https://doi.org/10.1002/CBIC.201900778.

[85] Frawley ER, Fang FC. The Ins and Outs of Bacterial Iron Metabolism. Mol Microbiol 2014;93:609. https://doi.org/10.1111/MMI.12709.

[86] Ghssein G, Ezzeddine Z. The Key Element Role of Metallophores in the Pathogenicity and Virulence of Staphylococcus aureus: A Review. Biology (Basel) 2022;11. https://doi.org/10.3390/BIOLOGY11101525.

[87] Koh EI, Robinson AE, Bandara N, Rogers BE, Henderson JP. Copper import in Escherichia coli by the yersiniabactin metallophore system. Nat Chem Biol 2017;13:1016. https://doi.org/10.1038/NCHEMBIO.2441.

[88] Cavas L, Kirkiz I. Characterization of siderophores from Escherichia coli strains through genome mining tools: an antiSMASH study. AMB Express 2022;12:1–13. https://doi.org/10.1186/S13568-022-01421-X/TABLES/5.

[89] Ghazaei C. Advances in the Study of Bacterial Toxins, Their Roles and Mechanisms in Pathogenesis. Malays J Med Sci 2022;29:4. https://doi.org/10.21315/MJMS2022.29.1.2.

[90] Polin RA, Abman SH, Rowitch DH, Benitz WE, Fox WW. Fetal and Neonatal Physiology, 2-Volume Set. Elsevier; 2017. https://doi.org/10.1016/B978-0-323-35214-7.00175-X.

[91] Scuron MD, Boesze-Battaglia K, Dlakic M, Shenker BJ. The cytolethal distending toxin contributes to microbial virulence and disease pathogenesis by acting as a tri-perditious toxin. Front Cell Infect Microbiol 2016;6:168. https://doi.org/10.3389/FCIMB.2016.00168/BIBTEX.

[92] Díaz JM, Dozois CM, Avelar-González FJ, Hernández-Cuellar E, Pokharel P, de Santiago AS, et al. The Vacuolating Autotransporter Toxin (Vat) of Escherichia coli Causes Cell Cytoskeleton Changes and Produces Non-lysosomal Vacuole Formation in Bladder Epithelial Cells. Front Cell Infect Microbiol 2020;10:299. https://doi.org/10.3389/FCIMB.2020.00299/BIBTEX.

[93] Vieira PCG, Espinoza-Culupú AO, Nepomuceno R, Alves MR, Lebrun I, Elias WP, et al. Secreted autotransporter toxin (Sat) induces cell damage during enteroaggregative Escherichia coli infection. PLoS One 2020;15. https://doi.org/10.1371/JOURNAL.PONE.0228959.

[94] Freire CA, Silva RM, Ruiz RC, Pimenta DC, Bryant JA, Henderson IR, et al. Secreted Autotransporter Toxin (Sat) Mediates Innate Immune System Evasion. Front Immunol 2022;13. https://doi.org/10.3389/FIMMU.2022.844878/FULL.

[95] Mousa WK. The microbiome-product colibactin hits unique cellular targets mediating host–microbe interaction. Front Pharmacol 2022;13. https://doi.org/10.3389/FPHAR.2022.958012.

[96] Johnson JR, Johnston B, Kuskowski MA, Nougayrede JP, Oswald E. Molecular Epidemiology and Phylogenetic Distribution of the Escherichia coli pks Genomic Island. J Clin Microbiol 2008;46:3906–11. https://doi.org/10.1128/JCM.00949-08.

[97] WHO. Antimicrobial resistance. 2021.

[98] Toner E, Adalja A, Gronvall GK, Cicero A, Inglesby T V. Antimicrobial Resistance Is a Global Health Emergency. Health Secur 2015;13:153. https://doi.org/10.1089/HS.2014.0088.

[99] Prestinaci F, Pezzotti P, Pantosti A. Antimicrobial resistance: a global multifaceted phenomenon. Pathog Glob Health 2015;109:309. https://doi.org/10.1179/2047773215Y.0000000030.

[100] Murray CJ, Ikuta KS, Sharara F, Swetschinski L, Robles Aguilar G, Gray A, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet 2022;399:629–55. https://doi.org/10.1016/S0140-6736(21)02724-0/ATTACHMENT/B227DEB3-FF04-497F-82AC-637D8AB7F679/MMC1.PDF.

[101] Alvarez-Uria G, Gandra S, Mandal S, Laxminarayan R. Global forecast of antimicrobial resistance in invasive isolates of Escherichia coli and Klebsiella pneumoniae. International Journal of Infectious Diseases 2018;68:50–3. https://doi.org/10.1016/J.IJID.2018.01.011.

[102] PHE. Laboratory surveillance of Escherichia coli bacteraemia in England, Wales and Northern Ireland: 2018. 2019.

[103] McCowan C, Bakhshi A, McConnachie A, Malcolm W, Sje B, Santiago VH, et al. E. coli bacteraemia and antimicrobial resistance following antimicrobial prescribing for urinary tract infection in the community. BMC Infect Dis 2022;22:1–10. https://doi.org/10.1186/S12879-022-07768-7/TABLES/7.

[104] Alhashash F, Wang X, Paszkiewicz K, Diggle M, Zong Z, McNally A. Increase in bacteraemia cases in the East Midlands region of the UK due to MDR Escherichia coli ST73: high levels of genomic and plasmid diversity in causative isolates. J Antimicrob Chemother 2016;71:339–43. https://doi.org/10.1093/JAC/DKV365.

[105] Palma N, Pons MJ, Gomes C, Mateu J, Riveros M, García W, et al. Resistance to quinolones, cephalosporins and macrolides in Escherichia coli causing bacteraemia in Peruvian children. J Glob Antimicrob Resist 2017;11:28–33. https://doi.org/10.1016/J.JGAR.2017.06.011.

[106] Russo TA, Johnson JR. Medical and economic impact of extraintestinal infections due to Escherichia coli: Focus on an increasingly important endemic problem. Microbes Infect 2003;5:449–56. https://doi.org/10.1016/S1286-4579(03)00049-2.

[107] Thornton H V., Hammond A, Hay AD. Urosepsis: a growing and preventable problem? The British Journal of General Practice 2018;68:493. https://doi.org/10.3399/BJGP18X699317.

[108] Simmons MD, Daniel S, Temple M. Sepsis programme successes are responsible for the increased detection of bacteraemia. J Hosp Infect 2019;101:93–9. https://doi.org/10.1016/j.jhin.2018.04.009.

[109] Jauréguy F, Carbonnelle E, Bonacorsi S, Clec'h C, Casassus P, Bingen E, et al. Host and bacterial determinants of initial severity and outcome of Escherichia coli sepsis. Clin Microbiol Infect 2007;13:854–62. https://doi.org/10.1111/j.1469-0691.2007.01775.x.

[110] Royer G, Darty MM, Clermont O, Condamine B, Laouenan C, Decousser JW, et al. Phylogroup stability contrasts with high within sequence type complex dynamics of Escherichia coli bloodstream infection isolates over a 12-year period. Genome Med 2021;13:1–17. https://doi.org/10.1186/S13073-021-00892-0/FIGURES/7.

[111] Bukh AS, Schønheyder HC, Emmersen JMG, Søgaard M, Bastholm S, Roslev P. Escherichia coli phylogenetic groups are associated with site of infection and level of antibiotic resistance in community-acquired bacteraemia: a 10 year population-based study in Denmark. J Antimicrob Chemother 2009;64:163–8. https://doi.org/10.1093/JAC/DKP156.

[112] Zhuge X, Zhou Z, Jiang M, Wang Z, Sun Y, Tang F, et al. Chicken-source Escherichia coli within phylogroup F shares virulence genotypes and is closely related to extraintestinal pathogenic E. coli causing human infections. Transbound Emerg Dis 2021;68:880–95. https://doi.org/10.1111/TBED.13755.

[113] Jauréguy F, Carbonnelle E, Bonacorsi S, Clec'h C, Casassus P, Bingen E, et al. Host and bacterial determinants of initial severity and outcome of Escherichia coli sepsis. Clin Microbiol Infect 2007;13:854–62. https://doi.org/10.1111/J.1469-0691.2007.01775.X.

[114] Public Health England. Annual epidemiological commentary: Gram-negative bacteraemia, MRSA bacteraemia, MSSA bacteraemia and C. difficile infections, up to and including financial year April 2018 to March 2019. 2019.

[115] Zhu W, Winter MG, Byndloss MX, Spiga L, Duerkop BA, Hughes ER, et al. Precision editing of the gut microbiota ameliorates colitis. Nature 2018;553:208–11. https://doi.org/10.1038/NATURE25172.

[116] Chen CY, Chen YC, Pu HN, Tsai CH, Chen WT, Lin CH. Bacteriology of acute appendicitis and its implication for the use of prophylactic antibiotics. Surg Infect (Larchmt) 2012;13:383–90. https://doi.org/10.1089/SUR.2011.135.

[117] Martinez JJ, Mulvey MA, Schilling JD, Pinkner JS, Hultgren SJ. Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. EMBO J 2000;19:2803. https://doi.org/10.1093/EMBOJ/19.12.2803.

[118] Poole NM, Green SI, Rajan A, Vela LE, Zeng XL, Estes MK, et al. Role for FimH in Extraintestinal Pathogenic Escherichia coli Invasion and Translocation through the Intestinal Epithelium. Infect Immun 2017;85. https://doi.org/10.1128/IAI.00581-17.

[119] Barbara G, Barbaro MR, Fuschi D, Palombo M, Falangone F, Cremon C, et al. Inflammatory and Microbiota-Related Regulation of the Intestinal Epithelial Barrier. Front Nutr 2021;8:623. https://doi.org/10.3389/FNUT.2021.718356/BIBTEX.

[120] Constant DA, Nice TJ, Rauch I. Innate immune sensing by epithelial barriers. Curr Opin Immunol 2021;73:1–8. https://doi.org/10.1016/J.COI.2021.07.014.

[121] Kawasaki T, Kawai T. Toll-like receptor signaling pathways. Front Immunol 2014;5:461. https://doi.org/10.3389/fimmu.2014.00461.

[122] Tang D, Kang R, Coyne CB, Zeh HJ, Lotze MT. PAMPs and DAMPs: signal 0s that spur autophagy and immunity. Immunol Rev 2012;249:158–75. https://doi.org/10.1111/J.1600-065X.2012.01146.X.

[123] Denning NL, Aziz M, Gurien SD, Wang P. Damps and nets in sepsis. Front Immunol 2019;10:2536. https://doi.org/10.3389/FIMMU.2019.02536/BIBTEX.

[124] Abraham E, Arcaroli J, Carmody A, Wang H, Tracey KJ. Cutting Edge: HMG-1 as a Mediator of Acute Lung Inflammation. The Journal of Immunology 2000;165:2950–4. https://doi.org/10.4049/JIMMUNOL.165.6.2950.

[125] Bayly-Jones C, Bubeck D, Dunstone MA. The mystery behind membrane insertion: a review of the complement membrane attack complex. Philosophical Transactions of the Royal Society B: Biological Sciences 2017;372. https://doi.org/10.1098/RSTB.2016.0221.

[126] Dunkelberger JR, Song WC. Complement and its role in innate and adaptive immune responses. Cell Research 2010 20:1 2009;20:34–50. https://doi.org/10.1038/cr.2009.139.

[127] Charles A Janeway J, Travers P, Walport M, Shlomchik MJ. The complement system and innate immunity 2001.

[128] Sarma JV, Ward PA. The complement system. Cell Tissue Res 2011;343:227–35. https://doi.org/10.1007/s00441-010-1034-0.

[129] Murugaiah V, Varghese PM, Beirag N, Decordova S, Sim RB, Kishore U. Complement Proteins as Soluble Pattern Recognition Receptors for Pathogenic Viruses. Viruses 2021;13. https://doi.org/10.3390/V13050824.

[130] Figueroa JE, Densen P. Infectious diseases associated with complement deficiencies. Clin Microbiol Rev 1991;4:359–95. https://doi.org/10.1128/cmr.4.3.359.

[131] Figueroa J, Andreoni J, Densen P. Complement deficiency states and meningococcal disease. Immunol Res 1993;12:295–311.

[132] Lewis LA, Ram S. Meningococcal disease and the complement system. Virulence 2014;5:98–126. https://doi.org/10.4161/viru.26515.

[133] Kumar S, Dikshit M. Metabolic Insight of Neutrophils in Health and Disease. Front Immunol 2019;10:2099. https://doi.org/10.3389/FIMMU.2019.02099.

[134] Schreiber A, Xiao H, Jennette JC, Schneider W, Luft FC, Kettritz R. C5a Receptor Mediates Neutrophil Activation and ANCA-Induced Glomerulonephritis. J Am Soc Nephrol 2009;20:289. https://doi.org/10.1681/ASN.2008050497.

[135] Iwaniuk A, Jablonska E. Neutrophils in Health and Disease: From Receptor Sensing to Inflammasome Activation. Int J Mol Sci 2023;24. https://doi.org/10.3390/IJMS24076340.

[136] Savill J, Dransfield I, Gregory C, Haslett C. A blast from the past: Clearance of apoptotic cells regulates immune responses. Nat Rev Immunol 2002;2:965–75. https://doi.org/10.1038/nri957.

[137] Henson PM, Bratton DL. Antiinflammatory effects of apoptotic cells. Journal of Clinical Investigation 2013;123:2773–4. https://doi.org/10.1172/JCI69344.

[138] Denning NL, Aziz M, Gurien SD, Wang P. Damps and nets in sepsis. Front Immunol 2019;10:2536. https://doi.org/10.3389/FIMMU.2019.02536/BIBTEX.

[139] Yao Y, Xu XH, Jin L. Macrophage Polarization in Physiological and Pathological Pregnancy. Front Immunol 2019;10:792. https://doi.org/10.3389/FIMMU.2019.00792.

[140] Austermann J, Roth J, Barczyk-Kahlert K. The Good and the Bad: Monocytes’ and Macrophages’ Diverse Functions in Inflammation. Cells 2022;11. https://doi.org/10.3390/CELLS11121979.

[141] Varol C, Mildner A, Jung S. Macrophages: Development and tissue specialization. Annu Rev Immunol 2015;33:643–75. https://doi.org/10.1146/annurev-immunol-032414-112220.

[142] Mantovani A, Sica A, Sozzani S, Allavena P, Vecchi A, Locati M. The chemokine system in diverse forms of macrophage activation and polarization. Trends Immunol 2004;25:677–86. https://doi.org/10.1016/J.IT.2004.09.015.

[143] Lafuente EM, Niedergang F, Rosales C. Editorial: Phagocytosis: Molecular Mechanisms and Physiological Implications. Front Immunol 2020;11:586918. https://doi.org/10.3389/FIMMU.2020.586918.

[144] Schilling JD, Martin SM, Hung CS, Lorenz RG, Hultgren SJ. Toll-like receptor 4 on stromal and hematopoietic cells mediates innate resistance to uropathogenic Escherichia coli. Proc Natl Acad Sci U S A 2003;100:4203–8. https://doi.org/10.1073/PNAS.0736473100.

[145] Kany S, Vollrath JT, Relja B. Cytokines in Inflammatory Disease. Int J Mol Sci 2019;20. https://doi.org/10.3390/IJMS20236008.

[146] Gruys E, Toussaint MJM, Niewold TA, Koopmans SJ. Review: Acute phase reaction and acute phase proteins. J Zhejiang Univ Sci B 2005;6:1045. https://doi.org/10.1631/JZUS.2005.B1045.

[147] Mantovani A, Garlanda C. Humoral Innate Immunity and Acute-Phase Proteins. N Engl J Med 2023;388:439. https://doi.org/10.1056/NEJMRA2206346.

[148] Jain S, Gautam V, Naseem S. Acute-phase proteins: As diagnostic tool. J Pharm Bioallied Sci 2011;3:118. https://doi.org/10.4103/0975-7406.76489.

[149] Minasyan H, Flachsbart F. Blood coagulation: a powerful bactericidal mechanism of human innate immunity. Https://DoiOrg/101080/0883018520181533009 2019;38:3–17. https://doi.org/10.1080/08830185.2018.1533009.

[150] Esmon CT, Xu J, Lupu F. Innate Immunity and Coagulation. J Thromb Haemost 2011;9:182. https://doi.org/10.1111/J.1538-7836.2011.04323.X.

[151] Yan G, Elbadawi M, Efferth T. Multiple cell death modalities and their key features (Review). World Acad Sci J 2020;2:39–48. https://doi.org/10.3892/WASJ.2020.40/HTML.

[152] Zhou CB, Fang JY. The role of pyroptosis in gastrointestinal cancer and immune responses to intestinal microbial infection. Biochim Biophys Acta Rev Cancer 2019;1872:1–10. https://doi.org/10.1016/J.BBCAN.2019.05.001.

[153] Orvedahl A, MacPherson S, Sumpter R, Tallóczy Z, Zou Z, Levine B. Autophagy protects against Sindbis virus infection of the central nervous system. Cell Host Microbe 2010;7:115–27. https://doi.org/10.1016/J.CHOM.2010.01.007.

[154] Miao EA, Rajan J V., Aderem A. Caspase-1-induced pyroptotic cell death. Immunol Rev 2011;243:206–14. https://doi.org/10.1111/J.1600-065X.2011.01044.X.

[155] Riera Romo M. Cell death as part of innate immunity: Cause or consequence? Immunology 2021;163:399. https://doi.org/10.1111/IMM.13325.

[156] Gyawali B, Ramakrishna K, Dhamoon AS. Sepsis: The evolution in definition, pathophysiology, and management. SAGE Open Med 2019;7. https://doi.org/10.1177/2050312119835043.

[157] Hotchkiss RS, Monneret G, Payen D. Immunosuppression in sepsis: a novel understanding of the disorder and a new therapeutic approach. Lancet Infect Dis 2013;13:260. https://doi.org/10.1016/S1473-3099(13)70001-X.

[158] van der Poll T, van de Veerdonk FL, Scicluna BP, Netea MG. The immunopathology of sepsis and potential therapeutic targets. Nat Rev Immunol 2017;17:407–20. https://doi.org/10.1038/nri.2017.36.

[159] Van Der Poll T, Van De Veerdonk FL, Scicluna BP, Netea MG. The immunopathology of sepsis and potential therapeutic targets. Nat Rev Immunol 2017;17:407–20. https://doi.org/10.1038/nri.2017.36.

[160] Buras JA, Holzmann B, Sitkovsky M. Animal Models of sepsis: setting the stage. Nature Reviews Drug Discovery 2005 4:10 2005;4:854–65. https://doi.org/10.1038/nrd1854.

[161] Calvano SE, Xiao W, Richards DR, Felciano RM, Baker H V., Cho RJ, et al. A network-based analysis of systemic inflammation in humans. Nature 2005;437:1032–7. https://doi.org/10.1038/nature03985.

[162] Abe R, Oda S, Sadahiro T, Nakamura M, Hirayama Y, Tateishi Y, et al. Gram-negative bacteremia induces greater magnitude of inflammatory response than Gram-positive bacteremia. Crit Care 2010;14. https://doi.org/10.1186/CC8898.

[163] Surbatovic M, Popovic N, Vojvodic D, Milosevic I, Acimovic G, Stojicic M, et al. Cytokine profile in severe gram-positive and gram-negative abdominal sepsis. Sci Rep 2015;5. https://doi.org/10.1038/SREP11355.

[164] Guo Q, Qu P, Cui W, Liu M, Zhu H, Chen W, et al. Organism type of infection is associated with prognosis in sepsis: an analysis from the MIMIC-IV database. BMC Infect Dis 2023;23. https://doi.org/10.1186/S12879-023-08387-6.

[165] Fay KT, Ford ML, Coopersmith CM. The intestinal microenvironment in sepsis. Biochim Biophys Acta 2017;1863:2574. https://doi.org/10.1016/J.BBADIS.2017.03.005.

[166] Hotchkiss RS, Swanson PE, Freeman BD, Tinsley KW, Cobb JP, Matuschak GM, et al. Apoptotic cell death in patients with sepsis, shock, and multiple organ dysfunction. Crit Care Med 1999;27:1230–51. https://doi.org/10.1097/00003246-199907000-00002.

[167] McDonald D, Ackermann G, Khailova L, Baird C, Heyland D, Kozar R, et al. Extreme Dysbiosis of the Microbiome in Critical Illness. MSphere 2016;1. https://doi.org/10.1128/MSPHERE.00199-16.

[168] Lankelma JM, van Vught LA, Belzer C, Schultz MJ, van der Poll T, de Vos WM, et al. Critically ill patients demonstrate large interpersonal variation in intestinal microbiota dysregulation: a pilot study. Intensive Care Med 2017;43:59–68. https://doi.org/10.1007/S00134-016-4613-Z/FIGURES/4.

[169] Dofferhoff ASM, de Jong HJ, Bom VJJ, van der Meer J, Limburg PC, de Vries-Hospers HG, et al. Complement activation and the production of inflammatory mediators during the treatment of severe sepsis in humans. Scand J Infect Dis 1992;24:197–204. https://doi.org/10.3109/00365549209052612.

[170] Helling H, Stephan B, Pindur G. Coagulation and complement system in critically ill patients. Clin Hemorheol Microcirc 2015;61:185–93. https://doi.org/10.3233/CH-151993.

[171] Furebring M, Douhan Håkansson L, Nilsson B, Sjölin J. Expression of the C5a receptor (CD88) on granulocytes and monocytes in patients with severe sepsis. 2002.

[172] Arul M, Chinnaiyan MA, Huber-Lang M, Kumar-Sinha C, Barrette TR, Shankar-Sinha S, et al. Molecular signatures of sepsis: Multiorgan gene expression profiles of systemic inflammation. American Journal of Pathology 2001;159:1199–209. https://doi.org/10.1016/s0002-9440(10)62505-9.

[173] Mollnes TE, Brekke O-L, Fung M, Fure H, Christiansen D, Bergseth G, et al. Essential role of the C5a receptor in E coli–induced oxidative burst and phagocytosis revealed by a novel lepirudin-based human whole blood model of inflammation. Blood 2002;100:1869–77. https://doi.org/10.1182/BLOOD.V100.5.1869.H81702001869_1869_1877.

[174] Skjeflo EW, Christiansen D, Espevik T, Nielsen EW, Mollnes TE. Combined Inhibition of Complement and CD14 Efficiently Attenuated the Inflammatory Response Induced by Staphylococcus aureus in a Human Whole Blood Model. The Journal of Immunology 2014;192:2857–64. https://doi.org/10.4049/JIMMUNOL.1300755.

[175] Hellerud BC, Orrem HL, Dybwik K, Pischke SE, Baratt-Due A, Castellheim A, et al. Combined inhibition of C5 and CD14 efficiently attenuated the inflammatory response in a porcine model of meningococcal sepsis. J Intensive Care 2017;5:1–9. https://doi.org/10.1186/S40560-017-0217-0/FIGURES/3.

[176] Skjeflo EW, Sagatun C, Dybwik K, Aam S, Urving SH, Nunn MA, et al. Combined inhibition of complement and CD14 improved outcome in porcine polymicrobial sepsis. Crit Care 2015;19. https://doi.org/10.1186/S13054-015-1129-9.

[177] Sommerfeld O, Medyukhina A, Neugebauer S, Ghait M, Ulferts S, Lupp A, et al. Targeting Complement C5a Receptor 1 for the Treatment of Immunosuppression in Sepsis. Molecular Therapy 2021;29:338. https://doi.org/10.1016/J.YMTHE.2020.09.008.

[178] Flierl MA, Rittirsch D, Nadeau BA, Day DE, Zetoune FS, Sarma JV, et al. Functions of the complement components C3 and C5 during sepsis. The FASEB Journal 2008;22:3483. https://doi.org/10.1096/FJ.08-110595.

[179] Shen XF, Cao K, Jiang JP, Guan WX, Du JF. Neutrophil dysregulation during sepsis: an overview and update. J Cell Mol Med 2017;21:1687–97. https://doi.org/10.1111/jcmm.13112.

[180] Roger T, Calandra T. Interleukin-33 safeguards neutrophils in sepsis. Nat Med 2010;16:638–9. https://doi.org/10.1038/nm0610-638.

[181] Kovach MA, Standiford TJ. The function of neutrophils in sepsis. Curr Opin Infect Dis 2012;25:321–7. https://doi.org/10.1097/QCO.0b013e3283528c9b.

[182] Sônego F, Castanheira FV e. S, Ferreira RG, Kanashiro A, Leite CAVG, Nascimento DC, et al. Paradoxical roles of the neutrophil in sepsis: Protective and deleterious. Front Immunol 2016;7:155. https://doi.org/10.3389/fimmu.2016.00155.

[183] Papayannopoulos V. Neutrophil extracellular traps in immunity and disease. Nature Reviews Immunology 2017 18:2 2017;18:134–47. https://doi.org/10.1038/nri.2017.105.

[184] Margraf S, Lögters T, Reipen J, Altrichter J, Scholz M, Windolf J. Neutrophil-derived circulating free DNA (CF-DNA/NETs): A potential prognostic marker for posttraumatic development of inflammatory second hit and sepsis. Shock 2008;30:352–8. https://doi.org/10.1097/SHK.0B013E31816A6BB1.

[185] Kitchens RL. Role of CD14 in Cellular Recognition of Bacterial Lipopolysaccharides. Chem Immunol 2000;74:61–82. https://doi.org/10.1159/000058750.

[186] Ziegler-Heitbrock L, Ancuta P, Crowe S, Dalod M, Grau V, Hart DN, et al. Nomenclature of monocytes and dendritic cells in blood. Blood 2010;116. https://doi.org/10.1182/BLOOD-2010-02-258558.

[187] Ferreira da Mota NV, Colo Brunialti MK, Santos SS, Machado FR, Assuncao M, Pontes Azevedo LC, et al. Immunophenotyping of monocytes during human sepsis shows impairment in antigen presentation: A shift toward nonclassical differentiation and upregulation of FcGRI-receptor. Shock 2018;50:293–300. https://doi.org/10.1097/SHK.0000000000001078.

[188] Shalova IN, Lim JY, Chittezhath M, Zinkernagel AS, Beasley F, Hernández-Jiménez E, et al. Human monocytes undergo functional re-programming during sepsis mediated by hypoxia-inducible factor-1α. Immunity 2015;42:484–98. https://doi.org/10.1016/J.IMMUNI.2015.02.001/ATTACHMENT/B1453C74-A62C-4FEB-A573-608EE4A2A30B/MMC1.PDF.

[189] Cavaillon JM, Adib-Conquy M. Monocytes/macrophages and sepsis. Crit Care Med 2005;33. https://doi.org/10.1097/01.CCM.0000185502.21012.37.

[190] Davis FM, Schaller MA, Dendekker A, Joshi AD, Kimball AS, Evanoff H, et al. Sepsis Induces Prolonged Epigenetic Modifications in Bone Marrow and Peripheral Macrophages Impairing Inflammation and Wound Healing. Arterioscler Thromb Vasc Biol 2019;39:2353–66. https://doi.org/10.1161/ATVBAHA.119.312754.

[191] van Dissel JT, van Langevelde P, Westendorp RG, Kwappenberg K, Frölich M. Anti-inflammatory cytokine profile and mortality in febrile patients. Lancet 1998;351:950–3. https://doi.org/10.1016/S0140-6736(05)60606-X.

[192] Ertel W, Kremer JP, Kenney J, Steckholzer U, Jarrar D, Trentz O, et al. Downregulation of Proinflammatory Cytokine Release in Whole Blood From Septic Patients. Blood 1995;85:1341–7. https://doi.org/10.1182/BLOOD.V85.5.1341.BLOODJOURNAL8551341.

[193] Munoz C, Carlet J, Fitting C, Misset B, Blériot JP, Cavaillon JM. Dysregulation of in vitro cytokine production by monocytes during sepsis. Journal of Clinical Investigation 1991;88:1747–54. https://doi.org/10.1172/JCI115493.

[194] Steeland S, Libert C, Vandenbroucke RE. A New Venue of TNF Targeting. International Journal of Molecular Sciences 2018, Vol 19, Page 1442 2018;19:1442. https://doi.org/10.3390/IJMS19051442.

[195] Ekaney ML, Otto GP, Sossdorf M, Sponholz C, Boehringer M, Loesche W, et al. Impact of plasma histones in human sepsis and their contribution to cellular injury and inflammation. Crit Care 2014;18:1–9. https://doi.org/10.1186/S13054-014-0543-8/FIGURES/6.

[196] Zhou Y, Dong H, Zhong Y, Huang J, Lv J, Li J. The Cold-Inducible RNA-Binding Protein (CIRP) Level in Peripheral Blood Predicts Sepsis Outcome. PLoS One 2015;10:e0137721. https://doi.org/10.1371/JOURNAL.PONE.0137721.

[197] Sundén-Cullberg J, Norrby-Teglund A, Rouhiainen A, Rauvala H, Herman G, Tracey KJ, et al. Persistent elevation of high mobility group box-1 protein (HMGB1) in patients with severe sepsis and septic shock. Crit Care Med 2005;33:564–73. https://doi.org/10.1097/01.CCM.0000155991.88802.4D.

[198] Sundén-Cullberg J, Norrby-Teglund A, Rouhiainen A, Rauvala H, Herman G, Tracey KJ, et al. Persistent elevation of high mobility group box-1 protein (HMGB1) in patients with severe sepsis and septic shock. Crit Care Med 2005;33:564–73. https://doi.org/10.1097/01.CCM.0000155991.88802.4D.

[199] Lee W, Yuseok O, Yang S, Lee BS, Lee JH, Park EK, et al. JH-4 reduces HMGB1-mediated septic responses and improves survival rate in septic mice. J Cell Biochem 2019;120:6277–89. https://doi.org/10.1002/JCB.27914.

[200] Denning NL, Aziz M, Gurien SD, Wang P. Damps and nets in sepsis. Front Immunol 2019;10:2536. https://doi.org/10.3389/fimmu.2019.02536.

[201] Levi M, van der Poll T. Coagulation and sepsis. Thromb Res 2017;149:38–44. https://doi.org/10.1016/J.THROMRES.2016.11.007.

[202] Greinacher A, Selleng S. How I evaluate and treat thrombocytopenia in the intensive care unit patient. Blood 2016;128:3032–42. https://doi.org/10.1182/BLOOD-2016-09-693655.

[203] Kraemer BF, Campbell RA, Schwertz H, Franks ZG, De Abreu AV, Grundler K, et al. Bacteria differentially induce degradation of Bcl-xL, a survival protein, by human platelets. Blood 2012;120:5014–20. https://doi.org/10.1182/BLOOD-2012-04-420661.

[204] Buchman TG, Hui JJ, Karl IE, Swanson PE, Schmieg RE, Freeman BD, et al. Sepsis-Induced Apoptosis Causes Progressive Profound Depletion of B and CD4+ T Lymphocytes in Humans. The Journal of Immunology 2014;166:6952–63. https://doi.org/10.4049/jimmunol.166.11.6952.

[205] Bosmann M, Ward PA. The inflammatory response in sepsis. Trends Immunol 2013;34:129–36. https://doi.org/10.1016/j.it.2012.09.004.

[206] Boomer JS, To K, Chang KC, Takasu O, Osborne DF, Walton AH, et al. Immunosuppression in patients who die of sepsis and multiple organ failure. JAMA 2011;306:2594–605. https://doi.org/10.1001/jama.2011.1829.

[207] Shankar-Hari M, Rubenfeld GD. Understanding Long-Term Outcomes Following Sepsis: Implications and Challenges. Curr Infect Dis Rep 2016;18:37. https://doi.org/10.1007/s11908-016-0544-7.

[208] Organisation L, Programme NE, Orgnization WH, Organisation IL. Biomarkers and risk assessment : concepts and principles (EHC 155 , 1993) 1993:1–31.

[209] Strimbu K, Tavel JA. What are Biomarkers? Curr Opin HIV AIDS 2010;5:463. https://doi.org/10.1097/COH.0B013E32833ED177.

[210] Choi J-H, Cho S-Y. Biomarkers of Sepsis. Infect Chemother 2014;46:1–12. https://doi.org/10.3947/ic.2014.46.1.1.

[211] Bacconi A, Richmond GS, Baroldi MA, Laffler TG, Blyn LB, Carolan HE, et al. Improved Sensitivity for Molecular Detection of Bacterial and Candida Infections in Blood 2014. https://doi.org/10.1128/JCM.00801-14.

[212] Peri AM, Harris PNA, Paterson DL. Culture-independent detection systems for bloodstream infection. Clin Microbiol Infect 2022;28:195–201. https://doi.org/10.1016/J.CMI.2021.09.039.

[213] Chanput W, Mes JJ, Wichers HJ. THP-1 cell line: An in vitro cell model for immune modulation approach. Int Immunopharmacol 2014;23:37–45. https://doi.org/10.1016/J.INTIMP.2014.08.002.

[214] Chick HM, Rees ME, Lewis ML, Williams LK, Bodger O, Harris LG, et al. Using the Traditional Ex Vivo Whole Blood Model to Discriminate Bacteria by Their Inducible Host Responses. Biomedicines 2024;12:724. https://doi.org/10.3390/BIOMEDICINES12040724/S1.

[215] Bengnér J, Quttineh M, Gäddlin PO, Salomonsson K, Faresjö M. Serum amyloid A – A prime candidate for identification of neonatal sepsis. Clinical Immunology 2021;229:108787. https://doi.org/10.1016/J.CLIM.2021.108787.

[216] Hou T, Huang D, Zeng R, Ye Z, Zhang Y. Accuracy of serum interleukin (IL)-6 in sepsis diagnosis: a systematic review and meta-analysis. Int J Clin Exp Med 2015;8:15238.

[217] Harbarth S, Holeckova K, Froidevaux C, Pittet D, Ricou B, Grau GE, et al. Diagnostic Value of Procalcitonin, Interleukin-6, and Interleukin-8 in Critically Ill Patients Admitted with Suspected Sepsis. Https://DoiOrg/101164/Ajrccm16432009052 2012;164:396–402. https://doi.org/10.1164/AJRCCM.164.3.2009052.

[218] Hoffmann JJML. Neutrophil CD64 as a sepsis biomarker. Biochem Med (Zagreb) 2011;21:282–90. https://doi.org/10.11613/BM.2011.038/FULLARTICLE.

[219] Patnaik R, Azim A, Agarwal V. Neutrophil CD64 a Diagnostic and Prognostic Marker of Sepsis in Adult Critically Ill Patients: A Brief Review. Indian J Crit Care Med 2020;24:1242. https://doi.org/10.5005/JP-JOURNALS-10071-23558.

[220] Moran JL, Santamaria J. Reconsidering lactate as a sepsis risk biomarker. PLoS One 2017;12. https://doi.org/10.1371/JOURNAL.PONE.0185320.

[221] Author C, Won Park D, Eun Kang H. Lactate as a Biomarker for Sepsis Prognosis? Infect Chemother 2016;48:252–3. https://doi.org/10.3947/ic.2016.48.3.252.

[222] Schuetz P, Albrich W, Mueller B. Procalcitonin for diagnosis of infection and guide to antibiotic decisions: Past, present and future. BMC Med 2011;9. https://doi.org/10.1186/1741-7015-9-107.

[223] Choi JJ, McCarthy MW. Novel applications for serum procalcitonin testing in clinical practice. Expert Rev Mol Diagn 2018;18:27–34. https://doi.org/10.1080/14737159.2018.1407244.

[224] Van Der Does Y, Rood PPM, Haagsma JA, Patka P, Van Gorp ECM, Limper M. Procalcitonin-guided therapy for the initiation of antibiotics in the ED: a systematic review. Am J Emerg Med 2016;34:1286–93. https://doi.org/10.1016/J.AJEM.2016.03.065.

[225] Nehring SM, Goyal A, Bansal P, Patel BC. C Reactive Protein 2021.

[226] Sproston NR, Ashworth JJ. Role of C-reactive protein at sites of inflammation and infection. Front Immunol 2018;9:754. https://doi.org/10.3389/fimmu.2018.00754.

[227] Tang Y, Ling N, Li S, Huang J, Zhang W, Zhang A, et al. A panel of urine-derived biomarkers to identify sepsis and distinguish it from systemic inflammatory response syndrome. Scientific Reports 2021 11:1 2021;11:1–11. https://doi.org/10.1038/s41598-021-99595-0.

[228] Wu J, Liang J, An S, Zhang J, Xue Y, Zeng Y, et al. Novel biomarker panel for the diagnosis and prognosis assessment of sepsis based on machine learning. Https://DoiOrg/102217/Bmm-2022-0433 2023;16:1129–38. https://doi.org/10.2217/BMM-2022-0433.

[229] Bradley Z, Bhalla N. Point-of-care diagnostics for sepsis using clinical biomarkers and microfluidic technology. Biosens Bioelectron 2023;227:115181. https://doi.org/10.1016/J.BIOS.2023.115181.

[230] Zelek WM, Harris CL, Morgan BP. Extracting the barbs from complement assays: Identification and optimisation of a safe substitute for traditional buffers. Immunobiology 2018;223:744–9. https://doi.org/10.1016/J.IMBIO.2018.07.016.

[231] GitHub - najoshi/sickle: Windowed Adaptive Trimming for fastq files using quality n.d. https://github.com/najoshi/sickle (accessed September 11, 2023).

[232] GitHub - vsbuffalo/scythe: A 3’-end adapter contaminant trimmer n.d. https://github.com/vsbuffalo/scythe (accessed September 11, 2023).

[233] Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. Curr Protoc Bioinformatics 2020;70:e102. https://doi.org/10.1002/CPBI.102.

[234] GitHub - ablab/quast: Genome assembly evaluation tool n.d. https://github.com/ablab/quast (accessed September 11, 2023).

[235] GitHub - lh3/seqtk: Toolkit for processing sequences in FASTA/Q formats n.d. https://github.com/lh3/seqtk (accessed September 11, 2023).

[236] Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014;30:2068–9. https://doi.org/10.1093/bioinformatics/btu153.

[237] Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 2015;31:3691–3. https://doi.org/10.1093/bioinformatics/btv421.

[238] Ochman H, Selander RK. Standard reference strains of Escherichia coli from natural populations. J Bacteriol 1984;157:690–3. https://doi.org/10.1128/JB.157.2.690-693.1984.

[239] Brynildsrud O, Bohlin J, Scheffer L, Eldholm V. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. Genome Biol 2016;17. https://doi.org/10.1186/S13059-016-1108-8.

[240] GitHub - tseemann/mlst: :id: Scan contig files against PubMLST typing schemes n.d. https://github.com/tseemann/mlst (accessed September 11, 2023).

[241] Jolley KA, Maiden MCJ. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 2010;11:1–11. https://doi.org/10.1186/1471-2105-11-595/FIGURES/4.

[242] GitHub - tseemann/abricate: :mag_right: Mass screening of contigs for antimicrobial and virulence genes n.d. https://github.com/tseemann/abricate (accessed September 11, 2023).

[243] Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res 2017;45:D566–73. https://doi.org/10.1093/NAR/GKW1004.

[244] Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 2012;67:2640–4. https://doi.org/10.1093/JAC/DKS261.

[245] Chen L, Zheng D, Liu B, Yang J, Jin Q. VFDB 2016: hierarchical and refined dataset for big data analysis--10 years on. Nucleic Acids Res 2016;44:D694–7. https://doi.org/10.1093/NAR/GKV1239.

[246] Morpheus n.d. https://software.broadinstitute.org/morpheus/ (accessed September 11, 2023).

[247] Trust S. Representing the UK Sepsis Trust n.d.

[248] Mcpherson D, Griffiths C, Williams M, Baker A, Klodawski E, Jacobson B, et al. Sepsis-associated mortality in England: an analysis of multiple cause of death data from 2001 to 2010. Open 2013;3:2586. https://doi.org/10.1136/bmjopen-2013.

[249] Shahin J, Harrison DA, Rowan KM. Relation between volume and outcome for patients with severe sepsis in United Kingdom: retrospective cohort study. The BMJ 2012;344. https://doi.org/10.1136/BMJ.E3394.

[250] Szakmany T, Lundin RM, Sharif B, Ellis G, Morgan P, Kopczynska M, et al. Sepsis Prevalence and Outcome on the General Wards and Emergency Departments in Wales: Results of a Multi-Centre, Observational, Point Prevalence Study. PLoS One 2016;11. https://doi.org/10.1371/JOURNAL.PONE.0167230.

[251] Szakmany T, Ellis G, Lundin RM, Pignatelli I, Sharif B, Joshi S, et al. Sepsis in Wales on the general wards: Results of a feasibility pilot. Br J Anaesth 2015;114:1000–1. https://doi.org/10.1093/bja/aev133.

[252] Naylor NR, Pouwels KB, Hope R, Green N, Henderson KL, Knight GM, et al. The health and cost burden of antibiotic resistant and susceptible Escherichia coli bacteraemia in the English hospital setting: A national retrospective cohort study. PLoS One 2019;14. https://doi.org/10.1371/JOURNAL.PONE.0221944.

[253] Hywel Dda University Health Board. What will this Annual Report tell you? Our Annual Quality Statement Our Public Health Report. Annual Report 2020.

[254] Simmons MD, Daniel S, Temple M. Sepsis programme successes are responsible for the increased detection of bacteraemia. Journal of Hospital Infection 2019;101:93–9. https://doi.org/10.1016/j.jhin.2018.04.009.

[255] Hywel Dda UHB. Hywel Dda UHB bi-annual infographic n.d.

[256] Estrada-Garcia T, Tarr PI. Escherichia coli Infection. Foodborne Infections and Intoxications 2023:125–63. https://doi.org/10.1016/B978-0-12-819519-2.00018-9.

[257] Bonten M, Johnson JR, Van Den Biggelaar AHJ, Georgalis L, Geurtsen J, De Palacios PI, et al. Epidemiology of Escherichia coli Bacteremia: A Systematic Literature Review. Clinical Infectious Diseases 2021;72:1211–9. https://doi.org/10.1093/CID/CIAA210.

[258] Guliciuc M, Maier AC, Maier IM, Kraft A, Cucuruzac RR, Marinescu M, et al. The Urosepsis—A Literature Review. Medicina (B Aires) 2021;57. https://doi.org/10.3390/MEDICINA57090872.

[259] Spoorenberg V, Prins JM, Opmeer BC, de Reijke TM, Hulscher MEJL, Geerlings SE. The additional value of blood cultures in patients with complicated urinary tract infections. Clin Microbiol Infect 2014;20. https://doi.org/10.1111/1469-0691.12491.

[260] Gomi H, Goto Y, Laopaiboon M, Usui R, Mori R. Routine blood cultures in the management of pyelonephritis in pregnancy for improving outcomes. Cochrane Database Syst Rev 2015;2015. https://doi.org/10.1002/14651858.CD009216.PUB2.

[261] Van Nieuwkoop C, Hoppe BPC, Bonten TN, Van't Wout JW, Aarts NJM, Mertens BJ, et al. Predicting the need for radiologic imaging in adults with febrile urinary tract infection. Clin Infect Dis 2010;51:1266–72. https://doi.org/10.1086/657071.

[262] Echols RM, Tosiello RL, Haverstock DC, Tice AD. Demographic, clinical, and treatment parameters influencing the outcome of acute cystitis. Clin Infect Dis 1999;29:113–9. https://doi.org/10.1086/520138.

[263] Czaja CA, Scholes D, Hooton TM, Stamm WE. Population-based epidemiologic analysis of acute pyelonephritis. Clinical Infectious Diseases 2007;45:273–80. https://doi.org/10.1086/519268/2/45-3-273-TBL004.GIF.

[264] Bai AD, Bonares MJ, Thrall S, Bell CM, Morris AM. Presence of urinary symptoms in bacteremic urinary tract infection: a retrospective cohort study of Escherichia coli bacteremia. BMC Infect Dis 2020;20:1–10. https://doi.org/10.1186/S12879-020-05499-1/TABLES/4.

[265] Woodford HJ, Graham C, Meda M, Miciuleviciene J. Bacteremic urinary tract infection in hospitalized older patients-are any currently available diagnostic criteria sensitive enough? J Am Geriatr Soc 2011;59:567–8. https://doi.org/10.1111/J.1532-5415.2010.03284.X.

[266] Owens CD, Stoessel K. Surgical site infections: epidemiology, microbiology and prevention. Journal of Hospital Infection 2008;70:3–10. https://doi.org/10.1016/S0195-6701(08)60017-1.

[267] Alkaaki A, Al-Radi OO, Khoja A, Alnawawi A, Alnawawi A, Maghrabi A, et al. Surgical site infection following abdominal surgery: a prospective cohort study. Canadian Journal of Surgery 2019;62:111. https://doi.org/10.1503/CJS.004818.

[268] Hecker A, Reichert M, Reuß CJ, Schmoch T, Riedel JG, Schneck E, et al. Intra-abdominal sepsis: new definitions and current clinical standards. Langenbeck's Archives of Surgery 2019 404:3 2019;404:257–71. https://doi.org/10.1007/S00423-019-01752-7.

[269] Daga AP, Koga VL, Soncini JGM, De Matos CM, Perugini MRE, Pelisson M, et al. Escherichia coli Bloodstream Infections in Patients at a University Hospital: Virulence factors and clinical characteristics. Front Cell Infect Microbiol 2019;9:191. https://doi.org/10.3389/FCIMB.2019.00191/BIBTEX.

[270] Melzer M, Toner R, Lacey S, Bettany E, Rait G. Biliary tract infection and bacteraemia: presentation, structural abnormalities, causative organisms and clinical outcomes. Postgrad Med J 2007;83:773–6. https://doi.org/10.1136/PGMJ.2007.064683.

[271] Melzer M, Toner R, Lacey S, Bettany E, Rait G. Biliary tract infection and bacteraemia: presentation, structural abnormalities, causative organisms and clinical outcomes. Postgrad Med J 2007;83:773. https://doi.org/10.1136/PGMJ.2007.064683.

[272] J O. Review on Antimicrobial Resistance Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations. 2014.

[273] Hillock NT, Merlin TL, Turnidge J, Karnon J. Modelling the Future Clinical and Economic Burden of Antimicrobial Resistance: The Feasibility and Value of Models to Inform Policy. Appl Health Econ Health Policy 2022;20:479–86. https://doi.org/10.1007/S40258-022-00728-X/FIGURES/1.

[274] Naylor NR, Pouwels KB, Hope R, Green N, Henderson KL, Knight GM, et al. The health and cost burden of antibiotic resistant and susceptible Escherichia coli bacteraemia in the English hospital setting: A national retrospective cohort study. PLoS One 2019;14. https://doi.org/10.1371/JOURNAL.PONE.0221944.

[275] Worthington RJ, Melander C. Overcoming Resistance to β-Lactam Antibiotics. J Org Chem 2013;78:4207. https://doi.org/10.1021/JO400236F.

[276] Pandey N, Cascella M. Beta Lactam Antibiotics. Antibiotic Discovery and Development 2022;9781461414001:79–117. https://doi.org/10.1007/978-1-4614-1400-1_3.

[277] Sianipar O, Asmara W, Dwiprahasto I, Mulyono B. Mortality risk of bloodstream infection caused by either Escherichia coli or Klebsiella pneumoniae producing extended-spectrum β-lactamase: A prospective cohort study. BMC Res Notes 2019;12:1–7. https://doi.org/10.1186/S13104-019-4751-9/TABLES/2.

[278] Clermont PCR n.d. https://ezclermont.hutton.ac.uk/ (accessed September 6, 2023).

[279] Jolley KA, Maiden MCJ. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 2010;11:1–11. https://doi.org/10.1186/1471-2105-11-595/FIGURES/4.

[280] GitHub - tseemann/mlst: :id: Scan contig files against PubMLST typing schemes n.d. https://github.com/tseemann/mlst (accessed September 6, 2023).

[281] PubMLST - Public databases for molecular typing and microbial genome diversity n.d. https://pubmlst.org/ (accessed September 6, 2023).

[282] Jolley KA, Maiden MCJ. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 2010;11:1–11. https://doi.org/10.1186/1471-2105-11-595/FIGURES/4.

[283] Kahlmeter G, Brown DFJ, Goldstein FW, MacGowan AP, Mouton JW, Odenholt I, et al. European Committee on Antimicrobial Susceptibility Testing (EUCAST) Technical Notes on antimicrobial susceptibility testing. Clinical Microbiology and Infection 2006;12:501–3. https://doi.org/10.1111/j.1469-0691.2006.01454.x.

[284] English surveillance programme for antimicrobial utilisation and resistance (ESPAUR) report 2021 to 2022 2022.

[285] Rodríguez I, Figueiredo AS, Sousa M, Aracil-Gisbert S, Fernández-de-Bobadilla MD, Lanza VF, et al. A 21-Year Survey of Escherichia coli from Bloodstream Infections (BSI) in a Tertiary Hospital Reveals How Community-Hospital Dynamics of B2 Phylogroup Clones Influence Local BSI Rates. MSphere 2021;6. https://doi.org/10.1128/MSPHERE.00868-21/SUPPL_FILE/MSPHERE.00868-21-SF007.PDF.

[286] Daga AP, Koga VL, Soncini JGM, De Matos CM, Perugini MRE, Pelisson M, et al. Escherichia coli Bloodstream Infections in Patients at a University Hospital: Virulence factors and clinical

characteristics. Front Cell Infect Microbiol 2019;9:191. https://doi.org/10.3389/FCIMB.2019.00191/BIBTEX.

[287] Liu Y, Zhu M, Fu X, Cai J, Chen S, Lin Y, et al. Escherichia coli Causing Neonatal Meningitis During 2001–2020: A Study in Eastern China. Int J Gen Med 2021;14:3007. https://doi.org/10.2147/IJGM.S317299.

[288] Ku LC, Boggess KA, Cohen-Wolkowiez M. Bacterial Meningitis in the Infant. Clin Perinatol 2015;42:29. https://doi.org/10.1016/J.CLP.2014.10.004.

[289] Roy S, Datta S, Das P, Gaind R, Pal T, Tapader R, et al. Insight into neonatal septicaemic Escherichia coli from India with respect to phylogroups, serotypes, virulence, extended-spectrum-β-lactamases and association of ST131 clonal group. Epidemiol Infect 2015;143:3266–76. https://doi.org/10.1017/S0950268815000539.

[290] Horner C, Fawley W, Morris K, Parnell P, Denton M, Wilcox M. Escherichia coli bacteraemia: 2 years of prospective regional surveillance (2010–12). Journal of Antimicrobial Chemotherapy 2014;69:91–100. https://doi.org/10.1093/JAC/DKT333.

[291] Doumith M, Day M, Ciesielczuk H, Hope R, Underwood A, Reynolds R, et al. Rapid Identification of Major Escherichia coli Sequence Types Causing Urinary Tract and Bloodstream Infections. J Clin Microbiol 2015;53:160. https://doi.org/10.1128/JCM.02562-14.

[292] Johnson JR, Johnston B, Clabots C, Kuskowski MA, Castanheira M. Escherichia coli sequence type ST131 as the major cause of serious multidrug-resistant E. coli infections in the United States. Clinical Infectious Diseases 2010;51:286–94. https://doi.org/10.1086/653932/2/51-3-286-TBL005.GIF.

[293] Kudinha T, Johnson JR, Andrew SD, Kong F, Anderson P, Gilbert GL. Escherichia coli Sequence Type 131 as a Prominent Cause of Antibiotic Resistance among Urinary Escherichia coli Isolates from Reproductive-Age Women. J Clin Microbiol 2013;51:3270. https://doi.org/10.1128/JCM.01315-13.

[294] Banerjee R, Johnson JR. A new clone sweeps clean: The enigmatic emergence of Escherichia coli sequence type 131. Antimicrob Agents Chemother 2014;58:4997–5004. https://doi.org/10.1128/AAC.02824-14/ASSET/29E1C44F-F77A-4374-87DD-FB85F93DC72A/ASSETS/GRAPHIC/ZAC0081431180001.JPEG.

[295] Whitmer GR, Moorthy G, Arshad M. The pandemic Escherichia coli sequence type 131 strain is acquired even in the absence of antibiotic exposure. PLoS Pathog 2019;15:e1008162. https://doi.org/10.1371/JOURNAL.PPAT.1008162.

[296] Johnson JR, Johnston B, Thuras P, Launer B, Sokurenko E V., Miller LG. Escherichia coli Sequence Type 131 H30 Is the Main Driver of Emerging Extended-Spectrum-β-Lactamase-Producing E. coli at a Tertiary Care Center. MSphere 2016;1. https://doi.org/10.1128/MSPHERE.00314-16.

[297] Abernethy J, Guy R, Sheridan EA, Hopkins S, Kiernan M, Wilcox MH, et al. Epidemiology of Escherichia coli bacteraemia in England: results of an enhanced sentinel surveillance programme. J Hosp Infect 2017;95:365–75. https://doi.org/10.1016/J.JHIN.2016.12.008.

[298] Bou-Antoun S, Davies J, Guy R, Johnson AP, Sheridan EA, Hope RJ. Descriptive epidemiology of Escherichia coli bacteraemia in England, April 2012 to March 2014. Euro Surveill 2016;21. https://doi.org/10.2807/1560-7917.ES.2016.21.35.30329.

[299] Otter JA, Galletly TJ, Davies F, Hitchcock J, Gilchrist MJ, Dyakova E, et al. Planning to halve Gram-negative bloodstream infection: getting to grips with healthcare-associated Escherichia coli bloodstream infection sources. J Hosp Infect 2019;101:129–33. https://doi.org/10.1016/J.JHIN.2018.07.033.

[300] Wang JY, Hsueh PR, Wang JT, Lee LN, Yang PC, Luh KT. Recurrent Infections and Chronic Colonization by an Escherichia coli Clone in the Respiratory Tract of a Patient with Severe Cystic Bronchiectasis. J Clin Microbiol 2000;38:2766. https://doi.org/10.1128/JCM.38.7.2766-2767.2000.

[301] Yahav D, Eliakim-Raz N, Leibovici L, Paul M. Bloodstream infections in older patients. Virulence 2016;7:341. https://doi.org/10.1080/21505594.2015.1132142.

[302] Lark RL, Saint S, Chenoweth C, Zemencuk JK, Lipsky BA, Plorde JJ. Four-year prospective evaluation of community-acquired bacteremia: epidemiology, microbiology, and patient outcome. Diagn Microbiol Infect Dis 2001;41:15–22. https://doi.org/10.1016/S0732-8893(01)00284-X.

[303] Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. Crit Care Med 2001;29:1303–10. https://doi.org/10.1097/00003246-200107000-00002.

[304] Martin GS, Mannino DM, Eaton S, Moss M. The epidemiology of sepsis in the United States from 1979 through 2000. N Engl J Med 2003;348:1546–54. https://doi.org/10.1056/NEJMOA022139.

[305] Umemura Y, Ogura H, Takuma K, Fujishima S, Abe T, Kushimoto S, et al. Current spectrum of causative pathogens in sepsis: A prospective nationwide cohort study in Japan. Int J Infect Dis 2021;103:343–51. https://doi.org/10.1016/J.IJID.2020.11.168.

[306] Bonten M, Johnson JR, Van Den Biggelaar AHJ, Georgalis L, Geurtsen J, De Palacios PI, et al. Epidemiology of Escherichia coli Bacteremia: A Systematic Literature Review. Clin Infect Dis 2021;72:1211–9. https://doi.org/10.1093/CID/CIAA210.

[307] Abernethy JK, Johnson AP, Guy R, Hinton N, Sheridan EA, Hope RJ. Thirty day all-cause mortality in patients with Escherichia coli bacteraemia in England. Clin Microbiol Infect 2015;21:251.e1-251.e8. https://doi.org/10.1016/J.CMI.2015.01.001.

[308] de Lastours V, Laouénan C, Royer G, Carbonnelle E, Lepeule R, Esposito-Farèse M, et al. Mortality in Escherichia coli bloodstream infections: antibiotic resistance still does not make it. Journal of Antimicrobial Chemotherapy 2020;75:2334–43. https://doi.org/10.1093/JAC/DKAA161.

[309] MacKinnon MC, McEwen SA, Pearl DL, Lyytikäinen O, Jacobsson G, Collignon P, et al. Mortality in Escherichia coli bloodstream infections: a multinational population-based cohort study. BMC Infect Dis 2021;21. https://doi.org/10.1186/S12879-021-06326-X.

[310] Pierrakos C, Velissaris D, Bisdorff M, Marshall JC, Vincent JL. Biomarkers of sepsis: time for a reappraisal. Crit Care 2020;24. https://doi.org/10.1186/S13054-020-02993-5.

[311] Cohen J. The immunopathogenesis of sepsis. Nature 2002 420:6917 2002;420:885–91. https://doi.org/10.1038/nature01326.

[312] Cho SY, Choi JH. Biomarkers of Sepsis. Infect Chemother 2014;46:1–12. https://doi.org/10.3947/ic.2014.46.1.1.

[313] Reinhart K, Bauer M, Riedemann NC, Hartog CS. New approaches to sepsis: Molecular diagnostics and biomarkers. Clin Microbiol Rev 2012;25:609–34. https://doi.org/10.1128/CMR.00016-12/ASSET/B378687E-E578-42E4-8FA3-EC26D89A9DBF/ASSETS/GRAPHIC/ZCM9990923950003.JPEG.

[314] Lippi G. Sepsis biomarkers: Past, present and future. Clin Chem Lab Med 2019;57:1281–3. https://doi.org/10.1515/CCLM-2018-1347/MACHINEREADABLECITATION/RIS.

[315] Henriquez-Camacho C, Losa J. Biomarkers for Sepsis. Biomed Res Int 2014;2014. https://doi.org/10.1155/2014/547818.

[316] Mierzchała-Pasierb M, Lipinska-Gediga M. Sepsis diagnosis and monitoring - procalcitonin as standard, but what next? Anaesthesiol Intensive Ther 2019;51:299–305. https://doi.org/10.5114/AIT.2019.88104.

[317] Akova M. Epidemiology of antimicrobial resistance in bloodstream infections. Virulence 2016;7:252–66. https://doi.org/10.1080/21505594.2016.1159366.

[318] Ramsamy Y, Hardcastle TC, Muckart DJJ. Surviving Sepsis in the Intensive Care Unit: The Challenge of Antimicrobial Resistance and the Trauma Patient. World J Surg 2017;41:1165–9. https://doi.org/10.1007/S00268-016-3531-0.

[319] Guillamet MCV, Burnham JP, Kollef MH. Novel Approaches to Hasten Detection of Pathogens and Antimicrobial Resistance in the Intensive Care Unit. Semin Respir Crit Care Med 2019;40:454–64. https://doi.org/10.1055/S-0039-1693160.

[320] Wang J, Zhang H, Yan J, Zhang T. Literature review on the distribution characteristics and antimicrobial resistance of bacterial pathogens in neonatal sepsis. J Matern Fetal Neonatal Med 2022;35:861–70. https://doi.org/10.1080/14767058.2020.1732342.

[321] Dissanayake K, Jayasinghe C, Wanigasekara P, Sominanda A. Potential applicability of cytokines as biomarkers of disease activity in rheumatoid arthritis: Enzyme-linked immunosorbent spot assay-based evaluation of TNF-α, IL-1β, IL-10 and IL-17A. PLoS One 2021;16. https://doi.org/10.1371/JOURNAL.PONE.0246111.

[322] Tanaka T, Narazaki M, Kishimoto T. IL-6 in Inflammation, Immunity, and Disease. Cold Spring Harb Perspect Biol 2014;6:16295–6. https://doi.org/10.1101/CSHPERSPECT.A016295.

[323] Gruys E, Toussaint MJM, Niewold TA, Koopmans SJ. Review: Acute phase reaction and acute phase proteins. J Zhejiang Univ Sci B 2005;6:1045. https://doi.org/10.1631/JZUS.2005.B1045.

[324] Kozak W, Kluger MJ, Soszynski D, Conn CA, Rudolph K, Leon LR, et al. IL-6 and IL-1 beta in fever. Studies using cytokine-deficient (knockout) mice. Ann N Y Acad Sci 1998;856:33–47. https://doi.org/10.1111/J.1749-6632.1998.TB08310.X.

[325] Soares DM, Figueiredo MJ, Martins JM, Machado RR, Sorgi C, Faciolli LH, et al. A crucial role for IL-6 in the CNS of rats during fever induced by the injection of live E. coli. Med Microbiol Immunol 2012;201:47–60. https://doi.org/10.1007/S00430-011-0204-3.

[326] Dalrymple SA, Slattery R, Aud DM, Krishna M, Lucian LA, Murray R. Interleukin-6 is required for a protective immune response to systemic Escherichia coli infection. Infect Immun 1996;64:3231. https://doi.org/10.1128/IAI.64.8.3231-3235.1996.

[327] Jones MR, Quinton LJ, Simms BT, Lupa MM, Kogan MS, Mizgerd JP. Roles of Interleukin-6 in Activation of STAT Proteins and Recruitment of Neutrophils during Escherichia coli Pneumonia. J Infect Dis 2006;193:360–9. https://doi.org/10.1086/499312.

[328] Owusu-Boaitey N, Bauckman KA, Zhang T, Mysorekar IU. Macrophagic control of the response to uropathogenic E. coli infection by regulation of iron retention in an IL-6-dependent manner. Immun Inflamm Dis 2016;4:413–26. https://doi.org/10.1002/IID3.123.

[329] Ching CB, Gupta S, Li B, Cortado H, Mayne N, Jackson AR, et al. Interleukin-6/Stat3 signaling has an essential role in the host antimicrobial response to urinary tract infection. Kidney Int 2018;93:1320–9. https://doi.org/10.1016/J.KINT.2017.12.006.

[330] Guan J, Wang Z, Liu X, Jiang Y, Gao Q, Wu Q, et al. IL-6 and IL-10 Closely Correlate with Bacterial Bloodstream Infection. Iran J Immunol 2020;17:185–203. https://doi.org/10.22034/IJI.2020.87266.1793.

[331] Fu P, Xie S, Zhang X. IL-8 gene locus is associated with risk, severity and 28-day mortality of sepsis in a Chinese population. Clin Exp Med 2019;19:571–6. https://doi.org/10.1007/S10238-019-00584-5/METRICS.

[332] Hack CE, Hart M, Van Schijndel RJMS, Eerenberg AJM, Nuijens JH, Thijs LG, et al. Interleukin-8 in sepsis: Relation to shock and inflammatory mediators. Infect Immun 1992;60:2835–42. https://doi.org/10.1128/IAI.60.7.2835-2842.1992.

[333] de Pablo R, Monserrat J, Prieto A, Álvarez-Mon M. Role of circulating soluble chemokines in septic shock. Medicina Intensiva (English Edition) 2013;37:510–8. https://doi.org/10.1016/J.MEDINE.2012.09.008.

[334] Zhou M, Cheng S, Yu J, Lu Q. Interleukin-8 for Diagnosis of Neonatal Sepsis: A Meta-Analysis 2015. https://doi.org/10.1371/journal.pone.0127170.

[335] Menten P, Wuyts A, Van Damme J. Macrophage inflammatory protein-1. Cytokine Growth Factor Rev 2002;13:455–81. https://doi.org/10.1016/S1359-6101(02)00045-X.

[336] DN C, MA B, TM C, SL K, JF S, IB P, et al. Requirement of MIP-1 alpha for an inflammatory response to viral infection. Science 1995;269:1583–5. https://doi.org/10.1126/SCIENCE.7667639.

[337] A A, SL G, O S, J S, M M, P L, et al. HIV coreceptor downregulation as antiviral principle: SDF-1alpha-dependent internalization of the chemokine receptor CXCR4 contributes to inhibition of HIV replication. J Exp Med 1997;186:139–46. https://doi.org/10.1084/JEM.186.1.139.

[338] Cocchi F, DeVico AL, Yarchoan R, Redfield R, Cleghorn F, Blattner WA, et al. Higher macrophage inflammatory protein (MIP)-1α and MIP-1β levels from CD8+ T cells are associated with asymptomatic HIV-1 infection. Proc Natl Acad Sci U S A 2000;97:13812. https://doi.org/10.1073/PNAS.240469997.

[339] JL L, G L, XZ J, YF L, QY Y, HH J, et al. Assessment of clinical sepsis-associated biomarkers in a septic mouse model. J Int Med Res 2018;46:2410–22. https://doi.org/10.1177/0300060518764717.

[340] S F, J S, Y S, K T, H K, M S, et al. Serum MIP-1 alpha and IL-8 in septic patients. Intensive Care Med 1996;22:1169–75. https://doi.org/10.1007/BF01709331.

[341] Knapp S, Thalhammer F, Locker GJ, Laczika K, Hollenstein U, Frass M, et al. Prognostic Value of MIP-1α, TGF-β2, sELAM-1, and sVCAM-1 in Patients with Gram-Positive Sepsis. Clin Immunol Immunopathol 1998;87:139–44. https://doi.org/10.1006/CLIN.1998.4523.

[342] Ghosh M, Shen Z, Schaefer TM, Fahey J V., Gupta P, Wira CR. CCL20/MIP3a is a Novel Anti-HIV-1 Molecule of the Human Female Reproductive Tract. Am J Reprod Immunol 2009;62:60–71. https://doi.org/10.1111/J.1600-0897.2009.00713.X.

[343] Yang D, Chen Q, Hoover DM, Staley P, Tucker KD, Lubkowski J, et al. Many chemokines including CCL20/MIP-3α display antimicrobial activity. J Leukoc Biol 2003;74:448–55. https://doi.org/10.1189/JLB.0103024.

[344] Chick H. Methods to Identify Biomarkers to Predict Bacterial Sepsis. Swansea University , 2020.

[345] Shi Y, Zhu N, Qiu Y, Tan J, Wang F, Qin L, et al. Resistin-like molecules: a marker, mediator and therapeutic target for multiple diseases. Cell Commun Signal 2023;21. https://doi.org/10.1186/S12964-022-01032-W.

[346] Pang S, Le Y. Role of Resistin in Inflammation and Inflammation-Related Diseases 2006;3.

[347] Kaser S, Kaser A, Sandhofer A, Ebenbichler CF, Tilg H, Patsch JR. Resistin messenger-RNA expression is increased by proinflammatory cytokines in vitro. Biochem Biophys Res Commun 2003;309:286–90. https://doi.org/10.1016/J.BBRC.2003.07.003.

[348] Battu Aruna ‡, Sudip Ghosh ‡, Anil K. Singh §, Shekhar C. Mande ‖, V. Srinivas ⊥, Radha Chauhan ‖ and, et al. Human Recombinant Resistin Protein Displays a Tendency To Aggregate by Forming Intermolecular Disulfide Linkages†. Biochemistry 2003;42:10554–9. https://doi.org/10.1021/BI034782V.

[349] Filková M, Haluzík M, Gay S, Šenolt L. The role of resistin as a regulator of inflammation: Implications for various human pathologies. Clinical Immunology 2009;133:157–70. https://doi.org/10.1016/J.CLIM.2009.07.013.

[350] J S-C, T N, ML L, GE M, L T, J A, et al. Pronounced elevation of resistin correlates with severity of disease in severe sepsis and septic shock. Crit Care Med 2007;35:1536–42. https://doi.org/10.1097/01.CCM.0000266536.14736.03.

[351] Yilmaz TU, Kerem M, Demirtaş CY, Pasaoğlu Ö, Taşcilar Ö, Şakrak Ö, et al. Increased Resistin Levels in Intra-abdominal Sepsis: Correlation with proinflammatory cytokines and Acute Physiology and Chronic Health Evaluation (APACHE) II scores. Sultan Qaboos Univ Med J 2014;14:e506.

[352] Wajant H, Pfizenmaier K, Scheurich P. Tumor necrosis factor signaling. Cell Death & Differentiation 2003 10:1 2003;10:45–65. https://doi.org/10.1038/sj.cdd.4401189.

[353] TNFα and the TNF receptor superfamily: Structure-function relationship(s) - Idriss - 2000 - Microscopy Research and Technique - Wiley Online Library n.d. https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/1097-0029(20000801)50:3%3C184::AID-JEMT2%3E3.0.CO;2-H (accessed September 4, 2022).

[354] Lv B, Huang J, Yuan H, Yan W, Hu G, Wang J. Tumor Necrosis Factor-α as a Diagnostic Marker for Neonatal Sepsis: A Meta-Analysis. The Scientific World Journal 2014;2014. https://doi.org/10.1155/2014/471463.

[355] Burdet C, Clermont O, Bonacorsi S, Laouénan C, Bingen E, Aujard Y, et al. Escherichia coli bacteremia in children age and portal of entry are the main predictors of severity. Pediatric Infectious Disease Journal 2014;33:872–9. https://doi.org/10.1097/INF.0000000000000309.

[356] Lefort A, Panhard X, Clermont O, Woerther PL, Branger C, Mentré F, et al. Host factors and portal of entry outweigh bacterial determinants to predict the severity of Escherichia coli bacteremia. J Clin Microbiol 2011;49:777–83. https://doi.org/10.1128/JCM.01902-10.

[357] Martínez JA, Soto S, Fabrega A, Almela M, Mensa J, Soriano A, et al. Relationship of phylogenetic background, biofilm production, and time to detection of growth in blood culture vials with clinical variables and prognosis associated with Escherichia coli bacteremia. J Clin Microbiol 2006;44:1468–74. https://doi.org/10.1128/JCM.44.4.1468-1474.2006.

[358] Wiles TJ, Norton JP, Russell CW, Dalley BK, Fischer KF, Mulvey MA. Combining Quantitative Genetic Footprinting and Trait Enrichment Analysis to Identify Fitness Determinants of a Bacterial Pathogen. PLoS Genet 2013;9:e1003716. https://doi.org/10.1371/JOURNAL.PGEN.1003716.

[359] Subashchandrabose S, Smith SN, Spurbeck RR, Kole MM, Mobley HLT. Genome-Wide Detection of Fitness Genes in Uropathogenic Escherichia coli during Systemic Infection. PLoS Pathog 2013;9:e1003788. https://doi.org/10.1371/JOURNAL.PPAT.1003788.

[360] Smith SN, Hagan EC, Lane MC, Mobley HLT. Dissemination and systemic colonization of uropathogenic Escherichia coli in a murine model of bacteremia. MBio 2010;1. https://doi.org/10.1128/MBIO.00262-10.

[361] Dale AP, Woodford N. Extra-intestinal pathogenic Escherichia coli (ExPEC): Disease, carriage and clones. Journal of Infection 2015;71:615–26. https://doi.org/10.1016/J.JINF.2015.09.009.

[362] Abe R, Oda S, Sadahiro T, Nakamura M, Hirayama Y, Tateishi Y, et al. Gram-negative bacteremia induces greater magnitude of inflammatory response than Gram-positive bacteremia. Critical Care 2010 14:2 2010;14:1–7. https://doi.org/10.1186/CC8898.

[363] Barber AE, Fleming BA, Mulvey MA. Similarly Lethal Strains of Extraintestinal Pathogenic Escherichia coli Trigger Markedly Diverse Host Responses in a Zebrafish Model of Sepsis . MSphere 2016;1. https://doi.org/10.1128/msphere.00062-16.

[364] M M-R, N F-R, C N-SF, J D-S, MP R-G, FA F, et al. Impact of virulence genes on sepsis severity and survival in Escherichia coli bacteremia. Virulence 2015;6:93–100. https://doi.org/10.4161/21505594.2014.991234.

[365] Kittana H, Gomes-Neto JC, Heck K, Geis AL, Segura Muñoz RR, Cody LA, et al. Commensal Escherichia coliStrains can promote intestinal inflammation via differential interleukin-6 production. Front Immunol 2018;9:2318. https://doi.org/10.3389/FIMMU.2018.02318/BIBTEX.

[366] Wagenlehner FM, Lichtenstern C, Rolfes C, Mayer K, Uhle F, Weidner W, et al. Diagnosis and management for urosepsis. Int J Urol 2013;20:963–70. https://doi.org/10.1111/IJU.12200.

[367] Ramos NL, Lamprokostopoulou A, Chapman TA, Chin JC, Römling U, Brauner A, et al. Virulence characteristics of translocating Escherichia coli and the interleukin-8 response to infection. Microb Pathog 2011;50:81–6. https://doi.org/10.1016/J.MICPATH.2010.11.003.

[368] Chancharoenthana W, Kamolratanakul S, Schultz MJ, Leelahavanichkul A. The leaky gut and the gut microbiome in sepsis - targets in research and treatment. Clin Sci (Lond) 2023;137:645–62. https://doi.org/10.1042/CS20220777.

[369] Chakraborty A, Saralaya V, Adhikari P, Shenoy S, Baliga S, Hegde A. Characterization of Escherichia coli Phylogenetic Groups Associated with Extraintestinal Infections in South Indian Population. Ann Med Health Sci Res 2015;5:241. https://doi.org/10.4103/2141-9248.160192.

[370] Nojoomi F, Ghasemian A. The relation of phylogroups, serogroups, virulence factors and resistance pattern of Escherichia coli isolated from children with septicemia. New Microbes New Infect 2019;29:100517. https://doi.org/10.1016/J.NMNI.2019.100517.

[371] Li F, Zhang Y, Yu B, Zhang Z, Fan Y, Wang L, et al. Evaluation of the diagnostic and prognostic values of serum HSP90α in sepsis patients: a retrospective study. PeerJ 2022;10. https://doi.org/10.7717/PEERJ.12997/SUPP-3.

[372] Chen P, Zhong Z, Jiang H, Chen H, Lyu J, Zhou L. Th17-associated cytokines multiplex testing indicates the potential of macrophage inflammatory protein-3 alpha in the diagnosis of biliary atresia. Cytokine 2019;116:21–6. https://doi.org/10.1016/J.CYTO.2019.01.002.

[373] Fujishima S, Hoffman AR, Vu T, Kim KJ, Zheng H, Daniel D, et al. Regulation of neutrophil interleukin 8 gene expression and protein secretion by LPS, TNF-alpha, and IL-1 beta. J Cell Physiol 1993;154:478–85. https://doi.org/10.1002/JCP.1041540305.

[374] Roelofs JJTH, Rouschop KMA, Teske GJD, Claessen N, Weening JJ, Van Der Poll T, et al. The urokinase plasminogen activator receptor is crucially involved in host defense during acute pyelonephritis. Kidney Int 2006;70:1942–7. https://doi.org/10.1038/SJ.KI.5001947.

[375] Fink MP, Heard SO. Laboratory models of sepsis and septic shock. Journal of Surgical Research 1990;49:186–96. https://doi.org/10.1016/0022-4804(90)90260-9.

[376] Fink MP. Animal models of sepsis and its complications. Kidney Int 2008;74:991–3. https://doi.org/10.1038/KI.2008.442.

[377] Van Der Poll T. Preclinical sepsis models. Surg Infect (Larchmt) 2012;13:287–92. https://doi.org/10.1089/SUR.2012.105.

[378] Glode LM, Mergenhagen SE, Rosenstreich DL. Significant contribution of spleen cells in mediating the lethal effects of endotoxin in vivo. Infect Immun 1976;14:626. https://doi.org/10.1128/IAI.14.3.626-630.1976.

[379] Reynolds K, Novosad B, Hoffhines A, Gipson J, Johnson J, Peters J, et al. Pretreatment with troglitazone decreases lethality during endotoxemia in mice. J Endotoxin Res 2002;8:307–14. https://doi.org/10.1179/096805102125000515.

[380] da Silva AMT, Kaulbach HC, Chuidian FS, Lambert DR, Suffredini AF, Danner RL. Shock and Multiple-Organ Dysfunction after Self-Administration of Salmonella Endotoxin. Https://DoiOrg/101056/NEJM199305203282005 1993;328:1457–60. https://doi.org/10.1056/NEJM199305203282005.

[381] Lowry SF. Human endotoxemia: A model for mechanistic insight and therapeutic targeting. Shock 2005;24:94–100. https://doi.org/10.1097/01.SHK.0000191340.23907.A1.

[382] van Lier D, Geven C, Leijte GP, Pickkers P. Experimental human endotoxemia as a model of systemic inflammation. Biochimie 2019;159:99–106. https://doi.org/10.1016/J.BIOCHI.2018.06.014.

[383] Anel R, Kumar A. Human endotoxemia and human sepsis: Limits to the model. Crit Care 2005;9:151–2. https://doi.org/10.1186/CC3501/METRICS.

[384] Logue GL. Effect of Heparin on Complement Activation and Lysis of Paroxysmal Nocturnal Hemoglobinuria (PNH) Red Cells. Blood 1977;50:239–47. https://doi.org/10.1182/BLOOD.V50.2.239.239.

[385] Markwardt F. Hirudin as alternative anticoagulant--a historical review. Semin Thromb Hemost 2002;28:405–13. https://doi.org/10.1055/S-2002-35292.

[386] Beaudet L, Rodriguez-Suarez R, Venne M-H, Caron M, Bédard J, Brechler V, et al. AlphaLISA immunoassays: the no-wash alternative to ELISAs for research and drug discovery. Nature Methods 2008 5:12 2008;5:an8–9. https://doi.org/10.1038/nmeth.f.230.

[387] Denamur E, Clermont O, Bonacorsi S, Gordon D. The population genetics of pathogenic Escherichia coli. Nature Reviews Microbiology 2020 19:1 2020;19:37–54. https://doi.org/10.1038/s41579-020-0416-x.

[388] Clermont O, Christenson JK, Denamur E, Gordon DM. The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environ Microbiol Rep 2013;5:58–65. https://doi.org/10.1111/1758-2229.12019.

[389] Blount ZD. The unexhausted potential of E. coli. Elife 2015;4. https://doi.org/10.7554/ELIFE.05826.

[390] Rouli L, Merhej V, Fournier PE, Raoult D. The bacterial pangenome as a new tool for analysing pathogenic bacteria. New Microbes New Infect 2015;7:72. https://doi.org/10.1016/J.NMNI.2015.06.005.

[391] Ozer EA, Allen JP, Hauser AR. Characterization of the core and accessory genomes of Pseudomonas aeruginosa using bioinformatic tools Spine and AGEnt. BMC Genomics 2014;15:1–17. https://doi.org/10.1186/1471-2164-15-737/FIGURES/5.

[392] Tantoso E, Eisenhaber B, Kirsch M, Shitov V, Zhao Z, Eisenhaber F. To kill or to be killed: pangenome analysis of Escherichia coli strains reveals a tailocin specific for pandemic ST131. BMC Biology 2022 20:1 2022;20:1–26. https://doi.org/10.1186/S12915-022-01347-7.

[393] Ding W, Baumdicker F, Neher RA. panX: pan-genome analysis and exploration. Nucleic Acids Res 2018;46:E5. https://doi.org/10.1093/NAR/GKX977.

[394] Kaas RS, Friis C, Ussery DW, Aarestrup FM. Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse Escherichia coli genomes. BMC Genomics 2012;13:577. https://doi.org/10.1186/1471-2164-13-577.

[395] Lukjancenko O, Wassenaar TM, Ussery DW. Comparison of 61 Sequenced Escherichia coli Genomes. Microb Ecol 2010;60:708. https://doi.org/10.1007/S00248-010-9717-3.

[396] Rasko DA, Rosovitz MJ, Myers GSA, Mongodin EF, Fricke WF, Gajer P, et al. The Pangenome Structure of Escherichia coli: Comparative Genomic Analysis of E. coli Commensal and Pathogenic Isolates. J Bacteriol 2008;190:6881. https://doi.org/10.1128/JB.00619-08.

[397] Rasko DA, Rosovitz MJ, Myers GSA, Mongodin EF, Fricke WF, Gajer P, et al. The pangenome structure of Escherichia coli: Comparative genomic analysis of E. coli commensal and pathogenic isolates. J Bacteriol 2008;190:6881–93. https://doi.org/10.1128/JB.00619-08/SUPPL_FILE/JB00619_08_V2_SUPPLMENTAL_DATA.ZIP.

[398] Lukjancenko O, Wassenaar TM, Ussery DW. Comparison of 61 sequenced Escherichia coli genomes. Microb Ecol 2010;60:708–20. https://doi.org/10.1007/S00248-010-9717-3.

[399] Schaufler K, Semmler T, Wieler LH, Trott DJ, Pitout J, Peirano G, et al. Genomic and functional analysis of emerging virulent and multidrug-resistant Escherichia coli lineage sequence type 648. Antimicrob Agents Chemother 2019;63. https://doi.org/10.1128/AAC.00243-19/SUPPL_FILE/AAC.00243-19-S0003.XLSX.

[400] Clermont O, Couffignal C, Blanco J, Mentré F, Picard B, Denamur E. Two levels of specialization in bacteraemic Escherichia coli strains revealed by their comparison with commensal strains. Epidemiol Infect 2017;145:872–82. https://doi.org/10.1017/S0950268816003010.

[401] Krawczyk B, Śledzińska A, Szemiako K, Samet A, Nowicki B, Kur J. Characterisation of Escherichia coli isolates from the blood of haematological adult patients with bacteraemia: translocation from gut to blood requires the cooperation of multiple virulence factors. European Journal of Clinical Microbiology and Infectious Diseases 2015;34:1135–43. https://doi.org/10.1007/S10096-015-2331-Z/TABLES/7.

[402] Jasiunaite Z, Hames R, Tang S, Pizza M, Oggioni MR. K1 Escherichia coli form foci in the brain at 24h post infection. Access Microbiol 2022;4:po0262. https://doi.org/10.1099/ACMI.AC2021.PO0200.

[403] Mushtaq N, Redpath MB, Luzio JP, Taylor PW. Prevention and Cure of Systemic Escherichia coli K1 Infection by Modification of the Bacterial Phenotype. Antimicrob Agents Chemother 2004;48:1503. https://doi.org/10.1128/AAC.48.5.1503-1508.2004.

[404] Antoine C, Laforêt F, Goya-Jorge E, Gonza I, Lebrun S, Douny C, et al. Phage Targeting Neonatal Meningitis E. coli K1 In Vitro in the Intestinal Microbiota of Pregnant Donors and Impact on Bacterial Populations. Int J Mol Sci 2023;24:10580. https://doi.org/10.3390/IJMS241310580/S1.

[405] Reygaert WC. An overview of the antimicrobial resistance mechanisms of bacteria. AIMS Microbiol 2018;4:482. https://doi.org/10.3934/MICROBIOL.2018.3.482.

[406] Vestby LK, Grønseth T, Simm R, Nesse LL. Bacterial Biofilm and its Role in the Pathogenesis of Disease. Antibiotics 2020;9. https://doi.org/10.3390/ANTIBIOTICS9020059.

[407] Sharma G, Sharma S, Sharma P, Chandola D, Dang S, Gupta S, et al. Escherichia coli biofilm: development and therapeutic strategies. J Appl Microbiol 2016;121:309–19. https://doi.org/10.1111/JAM.13078.

[408] Jin DJ, Cashel M, Friedman DI, Nakamura Y, Walter WA, Gross CA. Effects of Rifampicin resistant rpoB mutations on antitermination and interaction with nusA in Escherichia coli. J Mol Biol 1988;204:247–61. https://doi.org/10.1016/0022-2836(88)90573-6.

[409] Pandey N, Cascella M. Beta Lactam Antibiotics. Antibiotic Discovery and Development 2022;9781461414001:79–117. https://doi.org/10.1007/978-1-4614-1400-1_3.

[410] Bush K, Bradford PA. β-Lactams and β-Lactamase Inhibitors: An Overview. Cold Spring Harb Perspect Med 2016;6. https://doi.org/10.1101/CSHPERSPECT.A025247.

[411] Day MJ, Hopkins KL, Wareham DW, Toleman MA, Elviss N, Randall L, et al. Extended-spectrum β-lactamase-producing Escherichia coli in human-derived and foodchain-derived samples from England, Wales, and Scotland: an epidemiological surveillance and typing study. Lancet Infect Dis 2019;19:1325–35. https://doi.org/10.1016/S1473-3099(19)30273-7.

[412] Galindo-Méndez M, Galindo-Méndez M. Antimicrobial Resistance in Escherichia coli. E Coli Infections - Importance of Early Diagnosis and Efficient Treatment 2020. https://doi.org/10.5772/INTECHOPEN.93115.

[413] Soto SM. Role of efflux pumps in the antibiotic resistance of bacteria embedded in a biofilm. Virulence 2013;4:223. https://doi.org/10.4161/VIRU.23724.

[414] Yousefian N, Ornik-Cha A, Poussard S, Decossas M, Berbon M, Daury L, et al. Structural characterization of the EmrAB-TolC efflux complex from E. coli. Biochim Biophys Acta Biomembr 2021;1863. https://doi.org/10.1016/J.BBAMEM.2020.183488.

[415] Wen X, Langevin AM, Dunlop MJ. Antibiotic export by efflux pumps affects growth of neighboring bacteria. Scientific Reports 2018 8:1 2018;8:1–9. https://doi.org/10.1038/s41598-018-33275-4.

[416] Dadi BR, Abebe T, Zhang L, Mihret A, Abebe W, Amogne W. Distribution of virulence genes and phylogenetics of uropathogenic Escherichia coli among urinary tract infection patients in Addis Ababa, Ethiopia. BMC Infect Dis 2020;20:1–12. https://doi.org/10.1186/S12879-020-4844-Z/TABLES/7.

[417] Clermont O, Christenson JK, Denamur E, Gordon DM. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environ Microbiol Rep 2013;5:58–65. https://doi.org/10.1111/1758-2229.12019.

[418] Casanova-Hampton K, Carey A, Kassam S, Garner A, Donati GL, Thangamani S, et al. A genome-wide screen reveals the involvement of enterobactin-mediated iron acquisition in Escherichia coli survival during copper stress. Metallomics 2021;13:52. https://doi.org/10.1093/MTOMCS/MFAB052.

[419] Qi B, Han M. Microbial Siderophore Enterobactin Promotes Mitochondrial Iron Uptake and Development of the Host via Interaction with ATP Synthase Article Microbial Siderophore Enterobactin Promotes Mitochondrial Iron Uptake and Development of the Host via Interaction with ATP Synthase. Cell 2018;175:571–82. https://doi.org/10.1016/j.cell.2018.07.032.

[420] Orchard SS, Rostron JE, Segall AM. Escherichia coli enterobactin synthesis and uptake mutants are hypersensitive to an antimicrobial peptide that limits the availability of iron in addition to blocking Holliday junction resolution. Microbiology (N Y) 2012;158:547. https://doi.org/10.1099/MIC.0.054361-0.

[421] Pourahmad Jaktaji R, Jazayeri N. Expression of acrA and acrB Genes in Esherichia coli Mutants with or without marR or acrR Mutations. Iran J Basic Med Sci 2013;16:1254.

[422] Viveiros M, Dupont M, Rodrigues L, Couto I, Davin-Regli A, Martins M, et al. Antibiotic stress, genetic response and altered permeability of E. coli. PLoS One 2007;2. https://doi.org/10.1371/JOURNAL.PONE.0000365.

[423] Piddock LJV. Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria. Clin Microbiol Rev 2006;19:382–402. https://doi.org/10.1128/CMR.19.2.382-402.2006.

[424] Ma D, Cook DN, Alberti M, Pon NG, Nikaido H, Hearst JE. Genes acrA and acrB encode a stress-induced efflux system of Escherichia coli. Mol Microbiol 1995;16:45–55. https://doi.org/10.1111/J.1365-2958.1995.TB02390.X.

[425] Rosenberg EY, Ma D, Nikaido H. AcrD of Escherichia coli is an aminoglycoside efflux pump. J Bacteriol 2000;182:1754–6. https://doi.org/10.1128/JB.182.6.1754-1756.2000.

[426] Spurbeck RR, Dinh PC, Walk ST, Stapleton AE, Hooton TM, Nolan LK, et al. Escherichia coli Isolates That Carry vat, fyuA, chuA, and yfcV Efficiently Colonize the Urinary Tract. Infect Immun 2012;80:4115. https://doi.org/10.1128/IAI.00752-12.

[427] Hancock V, Ferrières L, Klemm P. The ferric yersiniabactin uptake receptor FyuA is required for efficient biofilm formation by urinary tract infectious Escherichia coli in human urine. Microbiology (Reading) 2008;154:167–75. https://doi.org/10.1099/MIC.0.2007/011981-0.

[428] Carniel E, Guiyoule A, Guilvout I, Mercereau-Puijalon O. Molecular cloning, iron-regulation and mutagenesis of the irp2 gene encoding HMWP2, a protein specific for the highly pathogenic Yersinia. Mol Microbiol 1992;6:379–88. https://doi.org/10.1111/J.1365-2958.1992.TB01481.X.

[429] Tu J, Xue T, Qi K, Shao Y, Huang B, Wang X, et al. The irp2 and fyuA genes in High Pathogenicity Islands are involved in the pathogenesis of infections caused by avian pathogenic Escherichia coli (APEC). Pol J Vet Sci 2016;19:21–9. https://doi.org/10.1515/PJVS-2016-0004.

[430] Ling J, Pan H, Gao Q, Xiong L, Zhou Y, Zhang D, et al. Aerobactin Synthesis Genes iucA and iucC Contribute to the Pathogenicity of Avian Pathogenic Escherichia coli O2 Strain E058. PLoS One 2013;8. https://doi.org/10.1371/JOURNAL.PONE.0057794.

[431] Bélanger L, Garenaux A, Harel J, Boulianne M, Nadeau E, Dozois CM. Escherichia coli from animal reservoirs as a potential source of human extraintestinal pathogenic E. coli. FEMS Immunol Med Microbiol 2011;62:1–10. https://doi.org/10.1111/J.1574-695X.2011.00797.X.

[432] Sjöström AE, Balsalobre C, Emödy L, Westerlund-Wikström B, Hacker J, Uhlin BE. The SfaXII protein from newborn meningitis E. coli is involved in regulation of motility and type 1 fimbriae expression. Microb Pathog 2009;46:243–52. https://doi.org/10.1016/J.MICPATH.2009.01.007.

[433] Nsahlai CJ, Silver RP. Purification and characterization of KpsT, the ATP-binding component of the ABC-capsule exporter of Escherichia coli K1. FEMS Microbiol Lett 2003;224:113–8. https://doi.org/10.1016/S0378-1097(03)00428-2.

[434] Zubaidah R. Is there a difference between K1 capsule antigen on E.coli that causes sepsis compared to ESBL- producing E.coli? 2021.

[435] Huang SH, Chen YH, Kong G, Chen SH, Besemer J, Borodovsky M, et al. A novel genetic island of meningitic Escherichia coli K1 containing the ibeA invasion gene (GimA): functional annotation and carbon-source-regulated invasion of human brain microvascular endothelial cells. Funct Integr Genomics 2001;1:312–22. https://doi.org/10.1007/S101420100039.

[436] Bosák J, Kohoutová D, Hrala M, Křenová J, Morávková P, Rejchrt S, et al. Escherichia coli from biopsies differ in virulence genes between patients with colorectal neoplasia and healthy controls. Front Microbiol 2023;14:1141619. https://doi.org/10.3389/FMICB.2023.1141619/BIBTEX.

[437] Liu X, Ma Y, Wang J. Genetic variation and function: revealing potential factors associated with microbial phenotypes. Biophys Rep 2021;7:111–26. https://doi.org/10.52601/BPR.2021.200040.

[438] Chen PE, Shapiro BJ. The advent of genome-wide association studies for bacteria. Curr Opin Microbiol 2015;25:17–24. https://doi.org/10.1016/J.MIB.2015.03.002.

[439] Power RA, Parkhill J, De Oliveira T. Microbial genome-wide association studies: lessons from human GWAS. Nature Reviews Genetics 2016 18:1 2016;18:41–50. https://doi.org/10.1038/nrg.2016.132.

[440] Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, et al. Genome-wide association studies. Nature Reviews Methods Primers 2021 1:1 2021;1:1–21. https://doi.org/10.1038/s43586-021-00056-9.

[441] Wang K, Zhang H, Kugathasan S, Annese V, Bradfield JP, Russell RK, et al. Diverse Genome-wide Association Studies Associate the IL12/IL23 Pathway with Crohn Disease n.d. https://doi.org/10.1016/j.ajhg.2009.01.026.

[442] Siminovitch KA. PTPN22 and autoimmune disease. Nature Genetics 2004 36:12 2004;36:1248–9. https://doi.org/10.1038/ng1204-1248.

[443] Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science (1979) 2007;316:889–94. https://doi.org/10.1126/SCIENCE.1141634/SUPPL_FILE/FRAYLING_SOM.PDF.

[444] Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, et al. Large-Scale Identification, Mapping, and Genotyping of Single-Nucleotide Polymorphisms in the Human Genome. Science (1979) 1998;280:1077–82. https://doi.org/10.1126/SCIENCE.280.5366.1077.

[445] San JE, Baichoo S, Kanzi A, Moosa Y, Lessells R, Fonseca V, et al. Current Affairs of Microbial Genome-Wide Association Studies: Approaches, Bottlenecks and Analytical Pitfalls. Front Microbiol 2020;10:3119. https://doi.org/10.3389/FMICB.2019.03119/BIBTEX.

[446] Méric G, Mageiros L, Pensar J, Laabei M, Yahara K, Pascoe B, et al. Disease-associated genotypes of the commensal skin bacterium Staphylococcus epidermidis. Nat Commun 2018;9. https://doi.org/10.1038/S41467-018-07368-7.

[447] Berthenet E, Yahara K, Thorell K, Pascoe B, Meric G, Mikhail JM, et al. A GWAS on Helicobacter pylori strains points to genetic variants associated with gastric cancer risk. BMC Biol 2018;16. https://doi.org/10.1186/S12915-018-0550-3.

[448] Farhat MR, Freschi L, Calderon R, Ioerger T, Snyder M, Meehan CJ, et al. GWAS for quantitative resistance phenotypes in Mycobacterium tuberculosis reveals resistance genes and regulatory regions. Nat Commun 2019;10. https://doi.org/10.1038/S41467-019-10110-6.

[449] Mageiros L, Méric G, Bayliss SC, Pensar J, Pascoe B, Mourkas E, et al. Genome evolution and the emergence of pathogenicity in avian Escherichia coli. Nat Commun 2021;12. https://doi.org/10.1038/S41467-021-20988-W.

[450] Pleguezuelos-Manzano C, Puschhof J, Rosendahl Huber A, van Hoeck A, Wood HM, Nomburg J, et al. Mutational signature in colorectal cancer caused by genotoxic pks+ E. coli. Nature 2020;580:269. https://doi.org/10.1038/S41586-020-2080-8.

[451] Liu X, Ma Y, Wang J. Genetic variation and function: revealing potential factors associated with microbial phenotypes. Biophys Rep 2021;7:111. https://doi.org/10.52601/BPR.2021.200040.

[452] Lees JA, Bentley SD. Bacterial GWAS: not just gilding the lily. Nat Rev Microbiol 2016;14:406. https://doi.org/10.1038/NRMICRO.2016.82.

[453] Collins C, Didelot X. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. PLoS Comput Biol 2018;14:e1005958. https://doi.org/10.1371/JOURNAL.PCBI.1005958.

[454] Nakashima N, Miyazaki K. Bacterial Cellular Engineering by Genome Editing and Gene Silencing. Int J Mol Sci 2014;15:2773. https://doi.org/10.3390/IJMS15022773.

[455] Liu Y, Feng J, Pan H, Zhang X, Zhang Y. Genetically engineered bacterium: Principles, practices, and prospects. Front Microbiol 2022;13. https://doi.org/10.3389/FMICB.2022.997587.

[456] Sawitzke JA, Thomason LC, Bubunenko M, Li X, Costantino N, Court DL. Recombineering: using drug cassettes to knock out genes in vivo. Methods Enzymol 2013;533:79–102. https://doi.org/10.1016/B978-0-12-420067-8.00007-6.

[457] Buckles EL, Wang X, Lane MC, Lockatell CV, Johnson DE, Rasko DA, et al. Role of the K2 capsule in Escherichia coli urinary tract infection and serum resistance. J Infect Dis 2009;199:1689–97. https://doi.org/10.1086/598524.

[458] Yair Y, Michaux C, Biran D, Bernhard J, Vogel J, Barquist L, et al. Cellular RNA Targets of Cold Shock Proteins CspC and CspE and Their Importance for Serum Resistance in Septicemic Escherichia coli. MSystems 2022;7. https://doi.org/10.1128/MSYSTEMS.00086-22.

[459] Weiser JN, Gotschlich EC. Outer membrane protein A (OmpA) contributes to serum resistance and pathogenicity of Escherichia coli K-1. Infect Immun 1991;59:2252–8. https://doi.org/10.1128/IAI.59.7.2252-2258.1991.

[460] Prasadarao N V., Blom AM, Villoutreix BO, Linsangan LC. A novel interaction of outer membrane protein A with C4b binding protein mediates serum resistance of Escherichia coli K1. J Immunol 2002;169:6352–60. https://doi.org/10.4049/JIMMUNOL.169.11.6352.

[461] Hamilton CM, Aldea M, Washburn BK, Babitzke P, Kushner SR. New method for generating deletions and gene replacements in Escherichia coli. J Bacteriol 1989;171:4617–22. https://doi.org/10.1128/JB.171.9.4617-4622.1989.

[462] Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A 2000;97:6640–5. https://doi.org/10.1073/PNAS.120163297.

[463] Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2006;2:2006.0008. https://doi.org/10.1038/MSB4100050.

[464] Liu A, Tran L, Becket E, Lee K, Chinn L, Park E, et al. Antibiotic sensitivity profiles determined with an Escherichia coli gene knockout collection: Generating an antibiotic bar code. Antimicrob Agents Chemother 2010;54:1393–403. https://doi.org/10.1128/AAC.00906-09/SUPPL_FILE/SUPP_FIGS_REVISED.PDF.

[465] Heo J, Cho K, Kim U, Cho DH, Ko S, Tran QG, et al. Genome-wide high-throughput screening of interactive bacterial metabolite in the algal population using Escherichia coli K-12 Keio collection. Scientific Reports 2020 10:1 2020;10:1–11. https://doi.org/10.1038/s41598-020-67322-w.

[466] Long CP, Antoniewicz MR. Metabolic flux analysis of Escherichia coli knockouts: lessons from the Keio collection and future outlook. Curr Opin Biotechnol 2014;28:127–33. https://doi.org/10.1016/J.COPBIO.2014.02.006.

[467] Sarkar S, Hutton ML, Vagenas D, Ruter R, Schüller S, Lyras D, et al. Intestinal Colonization Traits of Pandemic Multidrug-Resistant Escherichia coli ST131. J Infect Dis 2018;218:979–90. https://doi.org/10.1093/INFDIS/JIY031.

[468] Schwan WR, Ding H. Temporal Regulation of fim Genes in Uropathogenic Escherichia coli during Infection of the Murine Urinary Tract . J Pathog 2017;2017:1–13. https://doi.org/10.1155/2017/8694356.

[469] Gopal N, Chauhan N, Jain U, Dass SK, Kumar S, Chandra R. Designing of a unique bioreceptor and fabrication of an efficient genosensing platform for neonatal sepsis detection. Analytical Methods 2023. https://doi.org/10.1039/D3AY00567D.

[470] Pathania A, Sardesai AA. Distinct Paths for Basic Amino Acid Export in Escherichia coli: YbjE (LysO) Mediates Export of L-Lysine. J Bacteriol 2015;197:2036–47. https://doi.org/10.1128/JB.02505-14.

[471] Dubey S, Majumder P, Penmatsa A, Sardesai AA. Topological analyses of the L-lysine exporter LysO reveal a critical role for a conserved pair of intramembrane solvent-exposed acidic residues. J Biol Chem 2021;297. https://doi.org/10.1016/J.JBC.2021.101168.

[472] Saier MH, Reddy VS, Tamang DG, Västermark Å. The transporter classification database. Nucleic Acids Res 2014;42. https://doi.org/10.1093/NAR/GKT1097.

[473] Salmon KA, Hung SP, Steffen NR, Krupp R, Baldi P, Hatfield GW, et al. Global gene expression profiling in Escherichia coli K12: effects of oxygen availability and ArcA. J Biol Chem 2005;280:15084–96. https://doi.org/10.1074/JBC.M414030200.

[474] Liang H, Mao Y, Sun Y, Gao H. Transcriptional regulator ArcA mediates expression of oligopeptide transport systems both directly and indirectly in Shewanella oneidensis. Scientific Reports 2019 9:1 2019;9:1–12. https://doi.org/10.1038/s41598-019-50201-4.

[475] Eswarappa SM, Panguluri KK, Hensel M, Chakravortty D. The yejABEF operon of Salmonella confers resistance to antimicrobial peptides and contributes to its virulence. Microbiology (Reading) 2008;154:666–78. https://doi.org/10.1099/MIC.0.2007/011114-0.

[476] Wang Z, Bie P, Cheng J, Lu L, Cui B, Wu Q. The ABC transporter YejABEF is required for resistance to antimicrobial peptides and the virulence of Brucella melitensis. Scientific Reports 2016 6:1 2016;6:1–10. https://doi.org/10.1038/srep31876.

[477] Harvey KL, Jarocki VM, Charles IG, Djordjevic SP. The diverse functional roles of elongation factor tu (Ef-tu) in microbial pathogenesis. Front Microbiol 2019;10. https://doi.org/10.3389/FMICB.2019.02351/FULL.

[478] Medicines Agency E. Advice on the designation of antimicrobials or groups of antimicrobials reserved for treatment of certain infections in humans - revision 2022.

[479] Prezioso SM, Brown NE, Goldberg JB. Elfamycins: inhibitors of elongation factor-Tu. Mol Microbiol 2017;106:22–34. https://doi.org/10.1111/MMI.13750.

[480] Nagai K, Domon H, Maekawa T, Hiyoshi T, Tamura H, Yonezawa D, et al. Immunization with pneumococcal elongation factor Tu enhances serotype-independent protection against Streptococcus pneumoniae infection. Vaccine 2019;37:160–8. https://doi.org/10.1016/J.VACCINE.2018.11.015.

[481] Harvey KL, Jarocki VM, Charles IG, Djordjevic SP. The diverse functional roles of elongation factor tu (Ef-tu) in microbial pathogenesis. Front Microbiol 2019;10:485697. https://doi.org/10.3389/FMICB.2019.02351/BIBTEX.

[482] Yu Y, Wang H, Wang J, Feng Z, Wu M, Liu B, et al. Elongation factor thermo unstable (EF-Tu) moonlights as an adhesin on the surface of Mycoplasma hyopneumoniae by binding to fibronectin. Front Microbiol 2018;9:330155. https://doi.org/10.3389/FMICB.2018.00974/BIBTEX.

[483] Widjaja M, Harvey KL, Hagemann L, Berry IJ, Jarocki VM, Raymond BBA, et al. Elongation factor Tu is a multifunctional and processed moonlighting protein. Scientific Reports 2017 7:1 2017;7:1–17. https://doi.org/10.1038/s41598-017-10644-z.

[484] Torres AN, Chamorro-Veloso N, Costa P, Cádiz L, Canto F Del, Venegas SA, et al. Deciphering Additional Roles for the EF-Tu, l-Asparaginase II and OmpT Proteins of Shiga Toxin-Producing Escherichia coli. Microorganisms 2020;8:1–16. https://doi.org/10.3390/MICROORGANISMS8081184.

[485] Sun Y, Wang X, Li J, Xue F, Tang F, Dai J. Extraintestinal pathogenic Escherichia coli utilizes the surface-expressed elongation factor Tu to bind and acquire iron from holo-transferrin. Virulence 2022;13:698–713. https://doi.org/10.1080/21505594.2022.2066274.

[486] Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones JDG, Boller T, et al. Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts Agrobacterium-Mediated Transformation. Cell 2006;125:749–60. https://doi.org/10.1016/j.cell.2006.03.037.

[487] Granato D, Bergonzelli GE, Pridmore RD, Marvin L, Rouvet M, Corthésy-Theulaz IE. Cell surface-associated elongation factor Tu mediates the attachment of Lactobacillus johnsonii NCC533 (La1) to human intestinal cells and mucins. Infect Immun 2004;72:2160–9. https://doi.org/10.1128/IAI.72.4.2160-2169.2004.

[488] Chiu KH, Wang LH, Tsai TT, Lei HY, Liao PC. Secretomic Analysis of Host-Pathogen Interactions Reveals That Elongation Factor-Tu Is a Potential Adherence Factor of Helicobacter pylori during Pathogenesis. J Proteome Res 2017;16:264–73. https://doi.org/10.1021/ACS.JPROTEOME.6B00584.

[489] Strobel L, Johswich KO. Anticoagulants impact on innate immune responses and bacterial survival in whole blood models of Neisseria meningitidis infection. Sci Rep 2018;8:10225. https://doi.org/10.1038/S41598-018-28583-8.

[490] Mollnes TE, Garred P, Bergseth G. Effect of time, temperature and anticoagulants on in vitro complement activation: consequences for collection and preservation of samples to be examined for complement activation. Clin Exp Immunol 1988;73:484.

[491] Huang S, Sandholm K, Jonsson N, Nilsson A, Wieslander A, Grundström G, et al. Low concentrations of citrate reduce complement and granulocyte activation in vitro in human blood. Clin Kidney J 2015;8:31. https://doi.org/10.1093/CKJ/SFU127.

[492] Frawley ER, Fang FC. The Ins and Outs of Bacterial Iron Metabolism. Mol Microbiol 2014;93:609. https://doi.org/10.1111/MMI.12709.

[493] Miethke M, Marahiel MA. Siderophore-Based Iron Acquisition and Pathogen Control. Microbiology and Molecular Biology Reviews 2007;71:413–51. https://doi.org/10.1128/mmbr.00012-07.

[494] Robinson AE, Heffernan JR, Henderson JP. The iron hand of uropathogenic Escherichia coli: the role of transition metal control in virulence. Future Microbiol 2018;13:745. https://doi.org/10.2217/FMB-2017-0295.

[495] Shorr AF, Micek ST, Welch EC, Doherty JA, Reichley RM, Kollef MH. Inappropriate antibiotic therapy in Gram-negative sepsis increases hospital length of stay. Crit Care Med 2011;39:46–51. https://doi.org/10.1097/CCM.0B013E3181FA41A7.

[496] Stupica D, Lusa L, Klevišar MN, Terzić S, Pirš M, Premru MM, et al. Should we consider faecal colonisation with extended-spectrum β-lactamase-producing Enterobacteriaceae in empirical therapy of community-onset sepsis? Int J Antimicrob Agents 2017;50:564–71. https://doi.org/10.1016/J.IJANTIMICAG.2017.06.019.

[497] Smith HL, Elfeki MA, Lowry D, Sabates R, Ropte M, Sittig KR, et al. Comparison of the performance of a clinical classification tree versus clinical gestalt in predicting sepsis with extended-spectrum beta-lactamase-producing gram-negative rods. Antimicrobial Stewardship & Healthcare Epidemiology : ASHE 2022;2. https://doi.org/10.1017/ASH.2021.253.

[498] Fröding I, Valik JK, Bolinder L, Nauclér P, Giske CG. Prediction of bloodstream infection caused by extended-spectrum β-lactamase-producing Enterobacterales in patients with suspected community-onset sepsis. Int J Antimicrob Agents 2019;53:820–9. https://doi.org/10.1016/J.IJANTIMICAG.2019.02.008.

[499] Chen SL, Hung CS, Xu J, Reigstad CS, Magrini V, Sabo A, et al. Identification of genes subject to positive selection in uropathogenic strains of Escherichia coli: A comparative genomics approach. Proc Natl Acad Sci U S A 2006;103:5977. https://doi.org/10.1073/PNAS.0600938103.

[500] Snyder JA, Haugen BJ, Buckles EL, Lockatell CV, Johnson DE, Donnenberg MS, et al. Transcriptome of Uropathogenic Escherichia coli during Urinary Tract Infection. Infect Immun 2004;72:6373. https://doi.org/10.1128/IAI.72.11.6373-6381.2004.

[501] Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. Nat Rev Microbiol 2015;13:269. https://doi.org/10.1038/NRMICRO3432.

[502] Dunne KA, Chaudhuri RR, Rossiter AE, Beriotto I, Browning DF, Squire D, et al. Sequencing a piece of history: Complete genome sequence of the original Escherichia coli strain. Microb Genom 2017;3:mgen000106. https://doi.org/10.1099/MGEN.0.000106/CITE/REFWORKS.

[503] Shea AE, Marzoa J, Himpsl SD, Smith SN, Zhao L, Tran L, et al. Escherichia coli CFT073 Fitness Factors during Urinary Tract Infection: Identification Using an Ordered Transposon Library. Appl Environ Microbiol 2020;86. https://doi.org/10.1128/AEM.00691-20.

[504] Dadi BR, Abebe T, Zhang L, Mihret A, Abebe W, Amogne W. Distribution of virulence genes and phylogenetics of uropathogenic Escherichia coli among urinary tract infection patients in Addis Ababa, Ethiopia. BMC Infect Dis 2020;20:1–12. https://doi.org/10.1186/S12879-020-4844-Z/TABLES/7.

[505] Moez NM, Mashouf RY, Sedighi I, Shokoohizadeh L, Taheri M. Phylogroup classification and investigation the relationships between phylogroups and antibiotic resistance patterns of

uropathogenic E. coli isolated from pediatric urinary tract infection. Gene Rep 2020;20:100758. https://doi.org/10.1016/J.GENREP.2020.100758.

[506] Roubaud Baudron C, Panhard X, Clermont O, Mentré F, Fantin B, Denamur E, et al. Escherichia coli bacteraemia in adults: age-related differences in clinical and bacteriological characteristics, and outcome. Epidemiol Infect 2014;142:2672. https://doi.org/10.1017/S0950268814000211.

[507] Mamani R, Flament-Simon SC, García V, Mora A, Alonso MP, López C, et al. Sequence types, clonotypes, serotypes, and virotypes of extended-spectrum β-lactamase-producing Escherichia coli causing bacteraemia in a Spanish Hospital over a 12-Year Period (2000 to 2011). Front Microbiol 2019;10:1530. https://doi.org/10.3389/FMICB.2019.01530/FULL.

[508] Doumith M, Day M, Ciesielczuk H, Hope R, Underwood A, Reynolds R, et al. Rapid Identification of Major Escherichia coli Sequence Types Causing Urinary Tract and Bloodstream Infections. J Clin Microbiol 2015;53:160. https://doi.org/10.1128/JCM.02562-14.

[509] Tandogdu Z, Wagenlehner FME. Global epidemiology of urinary tract infections. Curr Opin Infect Dis 2016;29:73–9. https://doi.org/10.1097/QCO.0000000000000228.

[510] Daga AP, Koga VL, Soncini JGM, De Matos CM, Perugini MRE, Pelisson M, et al. Escherichia coli Bloodstream Infections in Patients at a University Hospital: Virulence Factors and Clinical Characteristics. Front Cell Infect Microbiol 2019;9. https://doi.org/10.3389/FCIMB.2019.00191.

[511] Achmad Ali Fikri, Syamsul Arifin MFF. Global Antimicrobial Resistance and Use Surveillance System (GLASS) Report 2022. World Health Organization 2022;2:2003–5.

[512] Kraft R, Herndon DN, Finnerty CC, Cox RA, Song J, Jeschke MG. Predictive Value of IL-8 for Sepsis and Severe Infections After Burn Injury: A Clinical Study. Shock 2015;43:222–7. https://doi.org/10.1097/SHK.0000000000000294.

[513] Zeng G, Chen D, Zhou R, Zhao X, Ye C, Tao H, et al. Combination of C-reactive protein, procalcitonin, IL-6, IL-8, and IL-10 for early diagnosis of hyperinflammatory state and organ dysfunction in pediatric sepsis. J Clin Lab Anal 2022;36. https://doi.org/10.1002/JCLA.24505.

[514] Ranjan A, Shaik S, Nandanwar N, Hussain A, Tiwari SK, Semmler T, et al. Comparative Genomics of <i>Escherichia coli</i> Isolated from Skin and Soft Tissue and Other Extraintestinal Infections. MBio 2017;8. https://doi.org/10.1128/mBio.01070-17.

[515] Micenková L, Beňová A, Frankovičová L, Bosák J, Vrba M, Ševčíková A, et al. Human Escherichia coli isolates from hemocultures: Septicemia linked to urogenital tract infections is caused by isolates harboring more virulence genes than bacteraemia linked to other conditions. Int J Med Microbiol 2017;307:182–9. https://doi.org/10.1016/J.IJMM.2017.02.003.

[516] Micenková L, Bosák J, Vrba M, Ševčíková A, Šmajs D. Human extraintestinal pathogenic Escherichia coli strains differ in prevalence of virulence factors, phylogroups, and bacteriocin determinants 2016. https://doi.org/10.1186/s12866-016-0835-z.

[517] Miajlovic H, Smith SG. Bacterial self-defence: how Escherichia coli evades serum killing. FEMS Microbiol Lett 2014;354:1–9. https://doi.org/10.1111/1574-6968.12419.

[518] Ma J, An C, Jiang F, Yao H, Logue C, Nolan LK, et al. Extraintestinal pathogenic Escherichia coli increase extracytoplasmic polysaccharide biosynthesis for serum resistance in response to bloodstream signals. Mol Microbiol 2018;110:689–706. https://doi.org/10.1111/MMI.13987.

[519] Freire CA, Silva RM, Ruiz RC, Pimenta DC, Bryant JA, Henderson IR, et al. Secreted Autotransporter Toxin (Sat) Mediates Innate Immune System Evasion. Front Immunol 2022;13:1. https://doi.org/10.3389/FIMMU.2022.844878/FULL.

[520] Wang CY, Wang SW, Huang WC, Kim KS, Chang NS, Wang YH, et al. Prc Contributes to Escherichia coli Evasion of Classical Complement-Mediated Serum Killing. Infect Immun 2012;80:3399. https://doi.org/10.1128/IAI.00321-12.

[521] Pierrakos C, Vincent JL. Sepsis biomarkers: A review. Crit Care 2010;14:1–18. https://doi.org/10.1186/CC8872/TABLES/10.

[522] Zhou Y, Zhang Y, Johnson A, Venable A, Griswold J, Pappas D. Combined CD25, CD64, and CD69 biomarker panel for flow cytometry diagnosis of sepsis. Talanta 2019;191:216–21. https://doi.org/10.1016/J.TALANTA.2018.08.058.

[523] Shapiro NI, Trzeciak S, Hollander JE, Birkhahn R, Otero R, Osborn TM, et al. A prospective, multicenter derivation of a biomarker panel to assess risk of organ dysfunction, shock, and death in emergency department patients with suspected sepsis. Crit Care Med 2009;37:96–104. https://doi.org/10.1097/CCM.0B013E318192FD9D.

[524] Garénaux A, Caza M, Dozois CM. The Ins and Outs of siderophore mediated iron uptake by extra-intestinal pathogenic Escherichia coli. Vet Microbiol 2011;153:89–98. https://doi.org/10.1016/J.VETMIC.2011.05.023.

[525] Gao Q, Wang X, Xu H, Xu Y, Ling J, Zhang D, et al. Roles of iron acquisition systems in virulence of extraintestinal pathogenic Escherichia coli: salmochelin and aerobactin contribute more to virulence than heme in a chicken infection model. BMC Microbiol 2012;12:143. https://doi.org/10.1186/1471-2180-12-143.

[526] Weiser JN, Gotschlich EC. Outer membrane protein A (OmpA) contributes to serum resistance and pathogenicity of Escherichia coli K-1. Infect Immun 1991;59:2252–8. https://doi.org/10.1128/IAI.59.7.2252-2258.1991.

[527] Scheller D, Twittenhoff C, Becker F, Holler M, Narberhaus F. OmpA, a Common Virulence Factor, Is Under RNA Thermometer Control in Yersinia pseudotuberculosis. Front Microbiol 2021;12:687260. https://doi.org/10.3389/FMICB.2021.687260/BIBTEX.

[528] Massella E, Giacometti F, Bonilauri P, Reid CJ, Djordjevic SP, Merialdi G, et al. Antimicrobial resistance profile and expec virulence potential in commensal escherichia coli of multiple sources. Antibiotics 2021;10. https://doi.org/10.3390/ANTIBIOTICS10040351/S1.

[529] Hallstrom KN, McCormick BA. Pathogenicity Islands: Origins, Structure, and Roles in Bacterial Pathogenesis. Molecular Medical Microbiology 2014:303–14. https://doi.org/10.1016/B978-0-12-397169-2.00016-0.

[530] Novick RP. Pathogenicity and Other Genomic Islands. Brenner's Encyclopedia of Genetics: Second Edition 2013:240–2. https://doi.org/10.1016/B978-0-12-374984-0.01126-8.

[531] Gal-Mor O, Finlay BB. Pathogenicity islands: a molecular toolbox for bacterial virulence. Cell Microbiol 2006;8:1707–19. https://doi.org/10.1111/J.1462-5822.2006.00794.X.

[532] Wang B, Lin AE, Yuan J, Novak KE, Koch MD, Wingreen NS, et al. Single-cell massively-parallel multiplexed microbial sequencing (M3-seq) identifies rare bacterial populations and profiles phage infection. Nature Microbiology 2023 2023:1–17. https://doi.org/10.1038/s41564-023-01462-3.

[533] AL-Ishaq R, Armstrong J, Gregory M, O'Hara M, Phiri K, Harris LG, et al. Effects of polysaccharide intercellular adhesin (PIA) in an ex vivo model of whole blood killing and in prosthetic joint infection (PJI): A role for C5a. Int J Med Microbiol 2015;305:948–56. https://doi.org/10.1016/J.IJMM.2015.08.005.

[534] Matsumoto H, Ogura H, Shimizu K, Ikeda M, Hirose T, Matsuura H, et al. The clinical importance of a cytokine network in the acute phase of sepsis. Scientific Reports 2018 8:1 2018;8:1–11. https://doi.org/10.1038/s41598-018-32275-8.

[535] Mera S, Tatulescu D, Cismaru C, Bondor C, Slavcovici A, Zanc V, et al. Multiplex cytokine profiling in patients with sepsis. APMIS 2011;119:155–63. https://doi.org/10.1111/J.1600-0463.2010.02705.X.

[536] Vital JS, Tanoeiro L, Lopes-Oliveira R, Vale FF. Biomarker Characterization and Prediction of Virulence and Antibiotic Resistance from Helicobacter pylori Next Generation Sequencing Data. Biomolecules 2022;12. https://doi.org/10.3390/BIOM12050691.

[537] Sabat AJ, Pantano D, Akkerboom V, Bathoorn E, Friedrich AW. Pseudomonas aeruginosa and Staphylococcus aureus virulence factors as biomarkers of infection. Biol Chem 2021;402:1565–73. https://doi.org/10.1515/HSZ-2021-0243.

[538] Zhou M, Wu Y, Kudinha T, Jia P, Wang L, Xu Y, et al. Comprehensive Pathogen Identification, Antibiotic Resistance, and Virulence Genes Prediction Directly From Simulated Blood Samples and Positive Blood Cultures by Nanopore Metagenomic Sequencing. Front Genet 2021;12:620009. https://doi.org/10.3389/FGENE.2021.620009/BIBTEX.

[539] Avershina E, Frye SA, Ali J, Taxt AM, Ahmad R. Ultrafast and Cost-Effective Pathogen Identification and Resistance Gene Detection in a Clinical Setting Using Nanopore Flongle Sequencing. Front Microbiol 2022;13. https://doi.org/10.3389/FMICB.2022.822402/FULL.

[540] Wolff N, Hendling M, Schroeder F, Schönthaler S, Geiss AF, Bedenic B, et al. Full pathogen characterisation: species identification including the detection of virulence factors and antibiotic resistance genes via multiplex DNA-assays. Scientific Reports 2021 11:1 2021;11:1–14. https://doi.org/10.1038/s41598-021-85438-5.

[541] Gautreau G, Bazin A, Gachet M, Planel R, Burlot L, Dubois M, et al. PPanGGOLiN: Depicting microbial diversity via a partitioned pangenome graph. PLoS Comput Biol 2020;16. https://doi.org/10.1371/JOURNAL.PCBI.1007732.

[542] GitHub - SionBayliss/PIRATE: A toolbox for pangenome analysis and threshold evaluation. n.d. https://github.com/SionBayliss/PIRATE (accessed September 15, 2023).

[543] Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G, Lees JA, et al. Producing polished prokaryotic pangenomes with the Panaroo pipeline. Genome Biol 2020;21. https://doi.org/10.1186/S13059-020-02090-4.

[544] Urhan A, Abeel T. A comparative study of pan-genome methods for microbial organisms: Acinetobacter baumannii pan-genome reveals structural variation in antimicrobial resistance-carrying plasmids. Microb Genom 2021;7:000690. https://doi.org/10.1099/MGEN.0.000690/CITE/REFWORKS.

[545] Roder T, Pimentel G, Fuchsmann P, Stern MT, Ah U von, Vergères G, et al. Scoary2: Rapid association of phenotypic multi-omics data with microbial pan-genomes. BioRxiv 2023:2023.04.19.537353. https://doi.org/10.1101/2023.04.19.537353.

[546] Drouin A, Giguère S, Déraspe M, Marchand M, Tyers M, Loo VG, et al. Predictive computational phenotyping and biomarker discovery using reference-free genome comparisons. BMC Genomics 2016;17. https://doi.org/10.1186/S12864-016-2889-6.

[547] Lees JA, Galardini M, Bentley SD, Weiser JN, Corander J. pyseer: a comprehensive tool for microbial pangenome-wide association studies. Bioinformatics 2018;34:4310–2. https://doi.org/10.1093/BIOINFORMATICS/BTY539.

[548] GitHub - bioinfo-ut/PhenotypeSeeker: Identify phenotype-specific k-mers and predict phenotype using sequenced bacterial strains n.d. https://github.com/bioinfo-ut/PhenotypeSeeker (accessed September 15, 2023).

[549] Strobel L, Johswich KO. Anticoagulants impact on innate immune responses and bacterial survival in whole blood models of Neisseria meningitidis infection. Sci Rep 2018;8:10225. https://doi.org/10.1038/S41598-018-28583-8.

[550] Massella E, Giacometti F, Bonilauri P, Reid CJ, Djordjevic SP, Merialdi G, et al. Antimicrobial resistance profile and expec virulence potential in commensal escherichia coli of multiple sources. Antibiotics 2021;10:351. https://doi.org/10.3390/ANTIBIOTICS10040351/S1.

[551] Bunduki GK, Heinz E, Phiri VS, Noah P, Feasey N, Musaya J. Virulence factors and antimicrobial resistance of uropathogenic Escherichia coli (UPEC) isolated from urinary tract infections: a systematic review and meta-analysis. BMC Infect Dis 2021;21. https://doi.org/10.1186/S12879-021-06435-7.

[552] Lloyd AL, Smith SN, Eaton KA, Mobley HLT. Uropathogenic Escherichia coli Suppresses the Host Inflammatory Response via Pathogenicity Island Genes sisA and sisB. Infect Immun 2009;77:5322. https://doi.org/10.1128/IAI.00779-09.

[553] Bunduki GK, Heinz E, Phiri VS, Noah P, Feasey N, Musaya J. Virulence factors and antimicrobial resistance of uropathogenic Escherichia coli (UPEC) isolated from urinary tract infections: a systematic review and meta-analysis. BMC Infect Dis 2021;21. https://doi.org/10.1186/S12879-021-06435-7.

[554] Dale AP, Pandey AK, Hesp RJ, Belogiannis K, Laver JR, Shone CC, et al. Genomes of Escherichia coli bacteraemia isolates originating from urinary tract foci contain more virulence-associated genes than those from non-urinary foci and neutropaenic hosts. J Infect 2018;77:534. https://doi.org/10.1016/J.JINF.2018.10.011.

[555] Sabri M, Léveillé S, Dozois CM. A SitABCD homologue from an avian pathogenic Escherichia coli strain mediates transport of iron and manganese and resistance to hydrogen peroxide. Microbiology (N Y) 2006;152:745–58. https://doi.org/10.1099/MIC.0.28682-0/CITE/REFWORKS.

[556] Johnson TJ, Wannemuehler YM, Nolan LK. Evolution of the iss gene in Escherichia coli. Appl Environ Microbiol 2008;74:2360–9. https://doi.org/10.1128/AEM.02634-07/SUPPL_FILE/SUPPLEMENTARY_TABLE_1.XLS.

[557] Koga VL, Tomazetto G, Cyoia PS, Neves MS, Vidotto MC, Nakazato G, et al. Molecular screening of virulence genes in extraintestinal pathogenic Escherichia coli isolated from human blood culture in Brazil. Biomed Res Int 2014;2014. https://doi.org/10.1155/2014/465054.

[558] Meena PR, Yadav P, Hemlata H, Tejavath KK, Singh AP. Poultry-origin extraintestinal Escherichia coli strains carrying the traits associated with urinary tract infection, sepsis, meningitis and avian colibacillosis in India. J Appl Microbiol 2021;130:2087–101. https://doi.org/10.1111/JAM.14905.

[559] Breland EJ, Eberly AR, Hadjifrangiskou M. An Overview of Two-Component Signal Transduction Systems Implicated in Extra-Intestinal Pathogenic E. coli Infections. Front Cell Infect Microbiol 2017;7:162. https://doi.org/10.3389/FCIMB.2017.00162.

[560] Bunduki GK, Heinz E, Phiri VS, Noah P, Feasey N, Musaya J. Virulence factors and antimicrobial resistance of uropathogenic Escherichia coli (UPEC) isolated from urinary tract infections: a systematic review and meta-analysis. BMC Infect Dis 2021;21. https://doi.org/10.1186/S12879-021-06435-7.

[561] Schierack P, Walk N, Ewers C, Wilking H, Steinrück H, Filter M, et al. ExPEC-typical virulence-associated genes correlate with successful colonization by intestinal E. coli in a small piglet group. Environ Microbiol 2008;10:1742–51. https://doi.org/10.1111/J.1462-2920.2008.01595.X.

[562] Meena PR, Yadav P, Hemlata H, Tejavath KK, Singh AP. Poultry-origin extraintestinal Escherichia coli strains carrying the traits associated with urinary tract infection, sepsis, meningitis and avian colibacillosis in India. J Appl Microbiol 2021;130:2087–101. https://doi.org/10.1111/JAM.14905.

[563] Lüthje P, Brauner A. Virulence Factors of Uropathogenic E. coli and Their Interaction with the Host. Adv Microb Physiol 2014;65:337–72. https://doi.org/10.1016/BS.AMPBS.2014.08.006.

[564] Singer RS. Urinary tract infections attributed to diverse ExPEC strains in food animals: evidence and data gaps. Front Microbiol 2015;6. https://doi.org/10.3389/FMICB.2015.00028.

Appendices

Appendix 2

Appendix 2.1 List of *E. coli* isolates obtain from the Hywel Dda University Health Board including lab ID, Hospital ID and source of infection.

Lab ID	Hospital ID	Source
1	B5934	Urinary
2	B5946	Intra-abdominal
4	B4019	Intra-abdominal
5	B4060	Urinary
6	B4053	Intra-abdominal
7	B17EOO4082	Unknown
8	B17E5004141	Intra-abdominal
9	BHEE00 4183	Biliary
10	B17E009094	Intra-abdominal
11	B4297	Urinary
12	B4298	Intra-abdominal
13	B174376	Biliary
14	B17E00 4376	Urinary
15	B17E005868	Biliary
16	B17E00 4216	Unknown
17	B17E00 4224	Unknown
18	B17E00 4409	Urinary
19	17/119	Biliary
20	17/123	Other
21	17/122	Urinary
22	17/124	Urinary
23	17/127	Urinary
24	17/128	Unknown
25	17/129	Urinary
26	17/131	Urinary
27	3729	Intra-abdominal
28	3736	Urinary
29	3737	Urinary
30	3743	Intra-abdominal
31	3747	Intra-abdominal
32	3757	Intra-abdominal
33	17/104	Urinary
34	17/106	Intra-abdominal
35	17/107	Urinary
36	17/108	Biliary
37	3710	Urinary
38	3697	Intra-abdominal
39	3701	Biliary
40	3693	Urinary
41	3714	Urinary
42	3718	Urinary

43	3717	Intra-abdominal
44	3708	Urinary
45	3690	Urinary
46	3730	Urinary
47	3731	Urinary
48	3744	Unknown
49	3745	Intra-abdominal
50	3756	Unknown
51	3752	Urinary
52	B17E008763	Intra-abdominal
53	B17E00 8876	Intra-abdominal
54	B17E00 8958	Urinary
55	B17E00 9042	Urinary
56	B17E00 9451	Abdominal
57	B17E00 8980	Other
58	B17E00 9003	Urinary
59	B17E00 9046	Urinary
60	1124046916	Abdominal
61	1124046948	Abdominal
62	1124046825	Urinary
63	1124046973	Urinary
64	11240555539	Abdominal
65	112405550	Urinary
66	1124055651	Abdominal
67	1124055426	Urinary
68	1124055402	Other
69	B17A002373	Urinary
70	1123209048	Abdominal
71	B17A002400	Urinary
72	B17A002504	Urinary
73	B17A002552	Urinary
74	B17A002624	Urinary
75	B18A000069	Urinary
76	B18A0000136	Abdominal
77	B18A000099	Urinary
78	B18A000001	Urinary
81	B18A000291	Urinary
82	B17E009423	Urinary
84	B18E000850	Urinary
85	B18E000828	Urinary
86	B18E000796	Urinary
87	B18E000711	Urinary
88	B18E000666	Urinary
89	B18E000643	Respiratory
90	E17E505558	Water
91	E17E505761	Water

93	E17E505979	Water
94	E18E500166	Water
96	E18E500522	Water
97	E17E506268	Water
98	E17E505024	Water
102	B19E006314	Urinary
103	B19E006384	Urinary
104	B19E006350	Unknown
105	B19E006332	Urinary
106	B19E006071	Biliary
107	B19E006491	Unknown
108	B19E006385	Abdominal
109	B19E006092	Abdominal
110	B19E006458	?Abdominal
111	B19E006433	Unknown
112	B19E006844	Unknown
113	B19E006453	?Abdominal
114	B19E006780	Urinary
115	B19E006598	Urinary
116	B19E006597	Urinary
117	B19E006406	Urinary
118	B19E006615	Abdominal
119	B20E008505	Biliary
120	B19E008446	Urinary
121	B19E008491(2)	?Abdominal
122	B19E008826	Unknown
123	B19E008803	Urinary
124	B19E008822	Urinary
125	B19E008368	Unknown
126	B19E008779	Urinary
127	B19E008782	Abdominal
128	B19E008766	Urinary
129	B20E000718	Other
130	B20E000313	Urinary/Abdominal
131	B20E000781	?Urinary
132	B20E000572	Unknown
133	B20E000211	Urinary
135	B20E000673	Urinary
137	B20E000048	Unknown
138	B20E000204	Urinary/Abdominal
139	B20E000923	Urinary
140	B20E000890	Abdominal
141	B20E000871	Urinary
142	B20E001499	Unknown
143	B20E001467	Urinary
144	B20E002542	Unclear - asymptomatic

145	B20E004109	Unknown
146	B20E002899	Urinary
147	B20E004076	Urinary
148	B20E002584	Unknown
149	B20E001641	Urinary
150	B20E001561	?Urinary
151	B20E001607	Urinary
152	B20E001648	Intra-abdominal
153	B20E008505	Intra-abdominal
154	B20E001587	Abdominal
155	B20E001624	Respiratory
156	B20E008550	Biliary
157	B20E008557	Unknown
158	B19E008017	?Respiratory
159	B19E007568	Urinary
160	B19E008156	Urinary
161	B19E008200	Urinary
162	B19E007970	Urinary/Biliary
163	B19E007324	Urinary
164	B19E008159	Abdominal
165	B19E008193	Unknown
166	B19E008123	Urinary
167	B19E008195	?Urinary
168	B19E008055	Biliary
169	B19E008227	Unknown
170	B20E000880	Urinary
U1	U20E258097	ANC Urine
U2	U20E240114	ANC Urine
U3	U20E240117	ANC Urine
U4	U20E282052	ANC Urine
U5	U20E291091	ANC Urine
U6	M21T003379	ANC Urine

Appendix table 2.1 Additional appendix files (attached separately end of document)

Appendix ID	Attached file
Appendix 2.2	Favourable opinion letter
Appendix 2.3	Letter of HRA approval
Appendix 2.4	Biomarkers in sepsis protocol
Appendix 2.5	Biomarkers in sepsis appendix

Appendix 3

Chapter 3 appendix

Appendix 3.1 Sequenced *E. coli* isolates source of infection, Abdominal includes both abdominal and intra-abdominal isolates.

Source of infection	Number of sequenced isolates
Urinary	43
Abdominal	20
Biliary	8
Unknown/other	21

Appendix 3.2 Antibiotic tests sets A1, A2a and H7: all antibiotic testing was performed at the hospital

A1 & A2a - Blood cultures	H7 - abx test set: URINE
Ampicillin	Ampicillin
Ceftazidime	Cefixime
Ciprofloxacin	Cephalexin
Cefotaxime	
Ertapenem	Ceftriaxone
Gentamicin	Amicacin
Imipenem	Amoxacillin
Meropenem	Augmentin
Piperacillin/Tazobactam	Ciprofloxacin
Amicacin	
Amoxiciccin	
Augmentin:	
Amoxicillin/Clavulanate	
Co-trimoxazole	

Appendix 3.3: Antibiotic abbreviations *note, there were some antibiotics listed that were not tested on all of the isolates or which were used for treatment but not tested.

Abx abbreviations	Antibiotic
ACYC	Acyclovir
AMI	Amicacin
AMI	Amicacin
AMP	Ampicillin
AMX	Amoxicillin
AUG	Co-amoxiclav/Augmentin
CAZ	Ceftazidine
CEP	Cefalexin
CET	Ceftriaxone
CEFP	Cefpodoxime
CIP	Ciprofloxacin
CLA	Clarithromycin
CTX	Cefotaxime
CXM	Cefuroxime
ERT	Ertapenem
FOS	Fosfomycin
GENT	Gentamicin
IMP	Imipenem
MEC	Mecillinam
MEM	Meropenem
MTZ	Metronidazole
PIPT	Piperacillin/Tazobactam
PIV	Pivmecillinam
SXT	Cotrimoxazole
TAZ	Tazocin
TEMO	Temocillin
VANC	Vancomycin

Appendix 3.4 Patient data as separate excel file

ID	Hospital ID	Source	Phylogroup	Sequence type	Survival in serum	Sepsis	Mortality	Nosocomial
1	B5934	Urinary			1	1	0	0
2	B5946	Intra-abdominal			1	1	1	1
4	B4019	Intra-abdominal	B2	73	0	1	0	0
5	B4060	Urinary	E	118	0	1	N/A	1
6	B4053	Intra-abdominal	B2	95	0	1	N/A	0
7	B17EOO4082	Unknown			1	1	0	0
8	B17E5004141	Intra-abdominal			1	0	0	0
9	BHEE00 4183	Biliary			0	0	0	0
10	B17E009094	Intra-abdominal			0	0	0	0
11	B4297	Urinary			1	0	0	0
12	B4298	Intra-abdominal			1	0	0	1
13	B174376	Biliary			1	1	0	0
14	B17E00 4376	Urinary	B2	73	1	0	0	1
15	B17E005868	Biliary	B2	73	1	0	0	0
16	B17E00 4216	Unknown	B2	131	1	0	1	1
17	B17E00 4224	Unknown	B2	12	1	0	0	0
18	B17E00 4409	Urinary	B2	131	1	0	0	0
19	17/119	Biliary	B2	95	1	0	1	1
20	17/123	Other	B1	2628	1	0	0	0
21	17/122	Urinary			1	0	0	0
22	17/124	Urinary	D		1	0	0	0
23	17/127	Urinary	B1		1	0	0	1
24	17/128	Unknown			1	0	0	0
25	17/129	Urinary			1	0	0	1
26	17/131	Urinary	A		1	1	0	0
27	3729	Intra-abdominal	B2		0	0	0	0
28	3736	Urinary			1	0	0	0
29	3737	Urinary			0	0	0	0
30	3743	Intra-abdominal			1	1	0	0
31	3747	Intra-abdominal	B2		1	0	0	0
32	3757	Intra-abdominal			1	1	0	1
33	17/104	Urinary	B2	73	1	0	0	0
34	17/106	Intra-abdominal	B2	95	1	0	0	0
35	17/107	Urinary	F	62	1	0	0	0
36	17/108	Biliary	B2	131	1	0	1	0
37	3710	Urinary	B2	978	1	0	0	0
38	3697	Intra-abdominal	B2	73	0	0	0	0
39	3701	Biliary	B2	-	1	0	0	0
40	3693	Urinary	D	62	1	1	0	0
41	3714	Urinary	B2	88	0	1	0	0
42	3718	Urinary	C	73	1	1	0	0
43	3717	Intra-abdominal	B2	144	0	1	0	0

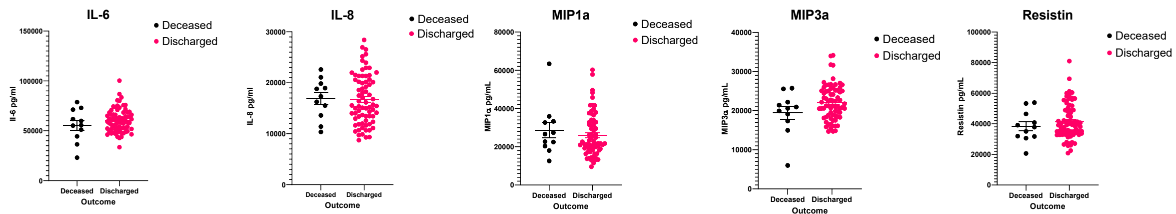
44	3708	Urinary	-		0	1	0	0
45	3690	Urinary	B2	131	1	1	0	0
46	3730	Urinary	B2	131	1	1	0	1
47	3731	Urinary	B2	64	0	1	0	1
48	3744	Unknown	B2	350	0	0	1	N/A
49	3745	Intra-abdominal			1	1	0	0
50	3756	Unknown			1	1	0	1
51	3752	Urinary	-		1	0	0	0
52	B17E008763	Intra-abdominal	B2	75	0	1	1	1
53	B17E00 8876	Intra-abdominal	B2	75	0	0	1	1
54	B17E00 8958	Urinary	U/Cryptic	12	1	0	0	0
55	B17E00 9042	Urinary	B2	79	0	0	0	0
56	B17E00 9451	Abdominal	B2	79	1	0	1	1
57	B17E00 8980	Other	B2	80	1	0	0	0
58	B17E00 9003	Urinary	B2	14	1	0	0	0
59	B17E00 9046	Urinary	B2	10	1	1	0	1
60	1124046916	Abdominal	No match		0	0	1	0
61	1124046948	Abdominal	B2	429	0	1	0	1
62	1124046825	Urinary	B2	73	1	1	0	0
63	1124046973	Urinary	B2	131	1	0	0	0
64	11240555539	Abdominal	-		0	0	0	1
65	112405550	Urinary	B2	12	1	1	1	0
66	1124055651	Abdominal	B2	127	1	1	0	0
67	1124055426	Urinary	B2	429	1	1	0	1
68	1124055402	Other	A	6318	0	0	0	0
69	B17A002373	Urinary	B2	12	1	0	0	0
70	1123209048	Abdominal	B2	73	1	0	0	0
71	B17A002400	Urinary	B2	12	1	0	0	0
72	B17A002504	Urinary	B2	1838	1	0	1	1
73	B17A002552	Urinary	A	69	1	0	1	0
74	B17A002624	Urinary	D	127	1	1	0	0
75	B18A000069	Urinary	B2	38	1	1	0	0
76	B18A0000136	Abdominal	-		1	0	0	0
77	B18A000099	Urinary	B2	-	1	0	0	0
78	B18A000001	Urinary	B2	127	1	1	0	1
81	B18A000291	Urinary	B2	88	1	0	0	1
82	B17E009423	Urinary	B2	-	1	0	0	0
84	B18E000850	Urinary	B2	1064	1	1	0	0
85	B18E000828	Urinary	B2	12	0	1	0	0
86	B18E000796	Urinary	B2	404	0	1	0	1
87	B18E000711	Urinary	B2	404	1	1	0	0
88	B18E000666	Urinary	B2	10	1	0	0	0
89	B18E000643	Respiratory			1	0	0	1
90	E17E505558	Water	B1	4118	1	N/A	N/A	N/A
91	E17E505761	Water	D	-	1	N/A	N/A	N/A
93	E17E505979	Water	A	1205	1	N/A	N/A	N/A

94	E18E500166	Water	B1	-	1	N/A	N/A	N/A
96	E18E500522	Water	Cryptic		1	N/A	N/A	N/A
97	E17E506268	Water	U/Cryptic	-	1	N/A	N/A	N/A
98	E17E505024	Water	U/Cryptic	43	0	N/A	N/A	N/A
C1 (REF) A1	ATCC 35218	Reference			1	N/A	N/A	N/A
C2 (REF) A2	ATCC 25922	Reference			1	N/A	N/A	N/A
C3 (C1)	NCTC 1093	Intestinal Isolate			1	N/A	N/A	N/A
C4 (C2)	NCTC 9001	Urinary Isolate			1	N/A	N/A	N/A
K12 (lab)		Lab strain			0	N/A	N/A	N/A
K12 (parent)	BW25113	Keoi knockout parent			0	N/A	N/A	N/A
102	B19E006314	Urinary	B2	131	0	N/A	N/A	N/A
103	B19E006384	Urinary	B2	131	1	1	0	0
104	B19E006350	Unknown			1	0	0	1
105	B19E006332	Urinary			1	1	1	0
106	B19E006071	Biliary	A	10	0	1	0	N/A
107	B19E006491	Unknown			1	0	0	1
108	B19E006385	Abdominal			N/A	0	0	1
109	B19E006092	Abdominal	B2	95	0	1	0	0
110	B19E006458	Abdominal	B2	131	0	1	1	1
111	B19E006433	Unknown			1	1	0	N/A
112	B19E006844	Unknown	B2	131	1	0	1	0
113	B19E006453	Abdominal	B2	131	0	1	1	0
114	B19E006780	Urinary			0	1	0	0
115	B19E006598	Urinary			0	1	0	0
116	B19E006597	Urinary			0	0	0	0
117	B19E006406	Urinary			1	0	0	1
118	B19E006615	Abdominal	B2	73	1	1	0	0
119	B20E008505	Biliary	B2	550	1	1	1	1
120	B19E008446	Urinary			0	1	0	1
121	B19E008491(2)	Abdominal			1	1	0	0
122	B19E008826	Unknown			1	0	0	1
123	B19E008803	Urinary			1	1	0	0
124	B19E008822	Urinary			1	1	0	0
125	B19E008368	Unknown			1	1	0	0
126	B19E008779	Urinary			1	1	0	1
127	B19E008782	Abdominal			0	1	0	1
128	B19E008766	Urinary			1	1	0	0
129	B20E000718	Other	B2	131	1	0	0	0
130	B20E000313	Urinary/Abdominal	B2	7514	1	1	1	0
131	B20E000781	Urinary	B2	131	0	1	0	0
132	B20E000572	Unknown	B2	131	0	1	0	0
133	B20E000211	Urinary			0	0	0	0
135	B20E000673	Urinary			0	0	N/A	0
137	B20E000048	Unknown			1	1	0	0

138	B20E000204	Urinary/Abdominal	B2	131	0	0	0	0
139	B20E000923	Urinary			1	1	0	0
140	B20E000890	Abdominal			0	0	0	1
141	B20E000871	Urinary			1	1	0	0
142	B20E001499	Unknown			0	N/A	N/A	N/A
143	B20E001467	Urinary			0	1	0	0
144	B20E002542	Unclear - asymptomatic	B2	131	1	0	0	0
145	B20E004109	Unknown			0	0	1	0
146	B20E002899	Urinary			1	0	0	0
147	B20E004076	Urinary			0	0	0	1
148	B20E002584	Unknown	B2	131	0	1	0	0
149	B20E001641	Urinary			1	1	0	N/A
150	B20E001561	?Urinary			1	0	0	N/A
151	B20E001607	Urinary			1	1	0	0
152	B20E001648	Intra-abdominal	D	69	1	1	0	0
153	B20E008505	Intra-abdominal			1	1	1	N/A
154	B20E001587	Abdominal			1	1	N/A	N/A
155	B20E001624	Respiratory	B2	73	0	1	0	1
156	B20E008550	Biliary	B2	131	0	1	0	0
157	B20E008557	Unknown	B2	12	0	0	1	0
158	B19E008017	?Respiratory			0	1	0	0
159	B19E007568	Urinary			0	0	0	0
160	B19E008156	Urinary			1	1	0	0
161	B19E008200	Urinary			1	1	0	0
162	B19E007970	Urinary/Biliary	A	744	1	0	0	0
163	B19E007324	Urinary			1	1	0	0
164	B19E008159	Abdominal	B2	80	0	0	0	0
165	B19E008193	Unknown			0	0	0	0
166	B19E008123	Urinary			1	1	0	0
167	B19E008195	Urinary	B2	420	1	1	1	0
168	B19E008055	Biliary	B1	58	0	1	0	1
169	B19E008227	Unknown	B2	131	0	0	0	0
170	B20E000880	Urinary	B2	73	1	0	0	1
U1		ANC Urine	B2		1	N/A	N/A	N/A
U2		ANC Urine	B2		1	N/A	N/A	N/A
U3		ANC Urine	B2		0	N/A	N/A	N/A
U4		ANC Urine	B2			N/A	N/A	N/A
U5		ANC Urine	D			N/A	N/A	N/A
U6		ANC Urine	B2			N/A	N/A	N/A

Appendix 4

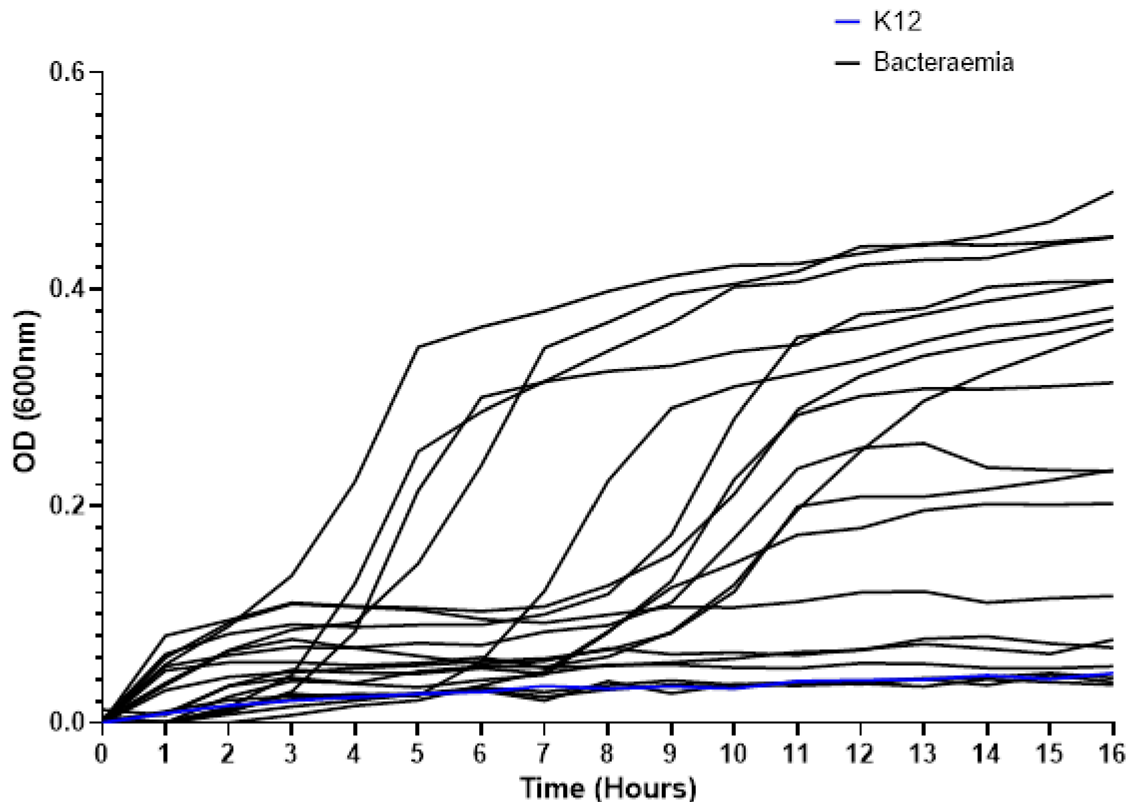
Chapter 4 appendix



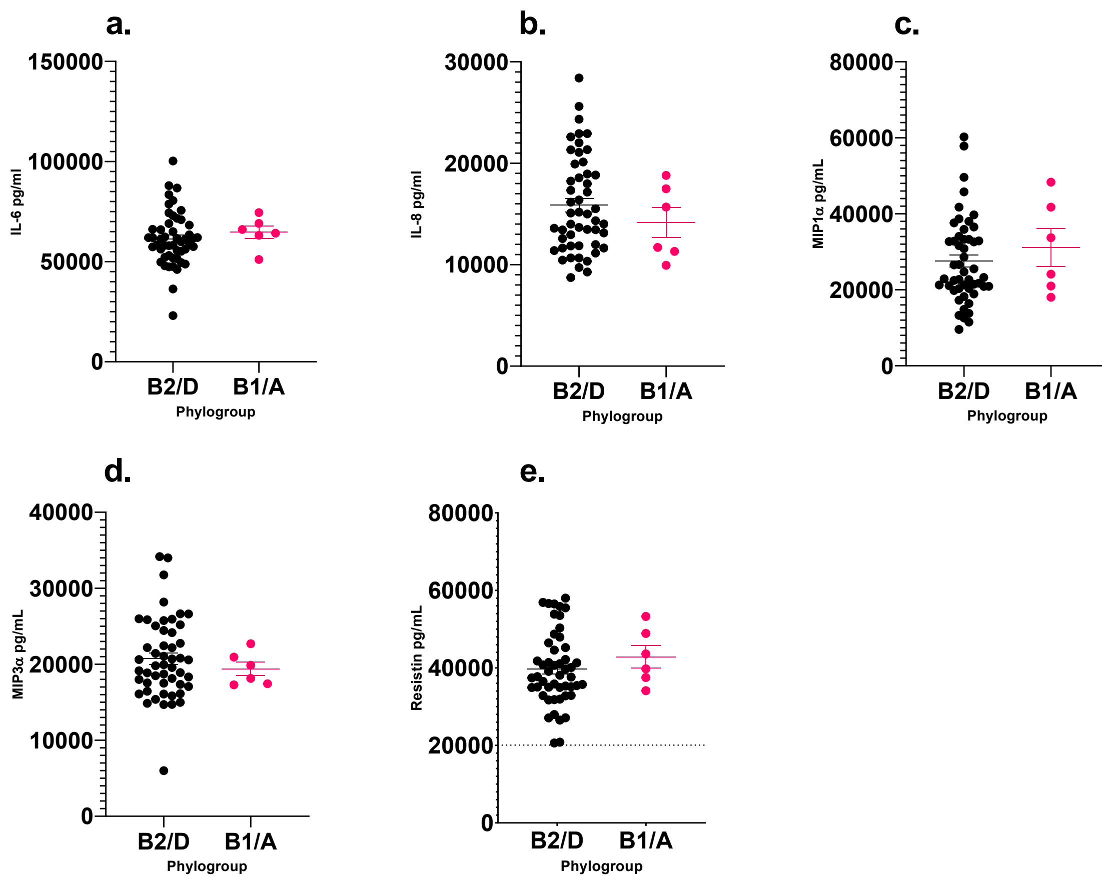
Appendix 4.1 Cytokine response grouped by disease outcome

Whole blood from healthy volunteers was infected with *E. coli* isolates from either bacteraemia or non-pathogenic sources for 4 hours and the platelet poor plasma isolated prior to host cytokine determination by ELISA. Isolates were then grouped according to disease outcome as determined in Chapter 3. Data presented as mean ± SEM of each cytokine. Each dot represents the average of three different blood donors per bacterial strain. No significant difference was observed between groups (p>0.05, Mann Whitney U test).

Appe



Appendix 4.2 Growth of *E. coli* bacteraemia isolates in serum.
Example growth curves of *E. coli* bacteraemia isolates. *E. coli* bacteraemia isolates (x19 strains, black) and K12 lab strain (blue), were grown in 90% human serum (gold top blood collection tubes) over 16 hours . Each line is the average of 2 independent experiments.



Appendix 4.3 *Ex-vivo* whole blood cytokine response to *E. coli* isolates grouped by phylogroup.
Whole blood from healthy volunteers was infected with *E. coli* isolates from either bacteraemia or non-pathogenic sources for 4 hours and the platelet poor plasma isolated prior to host cytokine determination by ELISA. *E. coli* isolates were grouped by phylogroups either B2/D or B1/A as determined in chapter 3. Each point on the graph represents the average of three independent measurements. Data shows the mean ± SEM. No significant difference between groups (Mann-Whitney test).

Appendix 5

Appendix 5.1 Common *E. coli* virulence factors and associated pathotype[550–564] Common virulence factors (Table 1.4) were subjected to literature search and associations with pathotypes noted. VFs

Virulence factor name	Group	Description	Function	ExPEC pathotype
bmaE	Adhesin	M-agglutinin subunit		
papA	Adhesin	P fimbriae	Stimulate t cell cytokine production, colonisation factor	upec, sepec, apec
papC	Adhesin	P fimbriae	Stimulate t cell cytokine production, colonisation factor	upec, nmec, apec
papG	Adhesin	P fimbriae		Upec
sfa	Adhesin	S fimbrial subunits	adhesion to host cells, facilitate entry into tissues	
foc	Adhesin	F1c fimbrial subunit	adhesion to renal epithelial cells, and bladder + kidney epithelial cells	upec
afa	Adhesin	Adhesins		upec
iha	Adhesin		iron-regulated adhesion	upec
mat	Adhesin		meningitis associated and temperature regulated fimbriae	nmec
dra	Adhesin	binding antigen Dr		upec
crl, csg	Adhesin	Curli fibre gene	Biofilm formation, pathogenicity promoter, induces strong immune response	upec, sepec, apec
agn43(flu)	Adhesin	Antigen 43	Autotransporter family, adhesion and biofilm formation	upec
hra	Adhesin	Heat resistant Haemagglutinin		
tsh	Adhesin	Temperature sensitive hemagglutinin		
fimA/H	Adhesin	Type 1 fimbriae	colonisation in extraintestinal infections, biofilm formation	upec, nmec, sepec, apec
papG	Adhesin	P fimbriae		upec, apec
papC	Adhesin	P fimbriae		upec, sepec, apec
fimP				upec
tia	Adhesin			apec
yqi	Adhesin	Yqi pili	Key adhesion factor during infection of lung	apec
kpsMT II	Protectins/serum resistance	Group 2 capsular polysaccharides		
traT	Protectins/serum resistance	Transfer protein	Inhibition of classical complement activation	nmec,sepec, apec
neuC	Protectins/serum resistance	K1 capsular polysaccharide		nmec
cvi/cva	Protectins/serum resistance	Structural genes of colicin V operon		
iss			Increased serum survival	nmec, sepec, apec
chuA	Iron acquisition	Heme receptor gene		
iutA	Iron acquisition	aerobactin receptor		
irp	Iron acquisition	Iron repressible protein	Yersiniabactin synthesis	nmec
iuc ABCD	Iron acquisition	aerobactin		nmec,upec,apec

sit ABCD	Iron acqusition		Iron transport	nmec, upec,apec
iro BCDEN	Iron acqusition	salmochelin	Siderophore receptor	nmec, upec, apec, sepec
ireA	Iron acquisition	Iron responsive element	Putative catecholate siderophore receptor	
EitA-D	Iron acqusition	Iron transport system		apec
iss	Outer membrane protein	serum survival, outer membrane endoprotease		nmec, apec, sepec
omp A and T	Outer membrane protein	outer membrane proteins	intracellular survival, evasion from host immune response	apec, nmec, upec
colV, cvaC	Outer membrane protein		Colonisation facilitation	nmec, sepec, apec
AatA	Outer membrane protein			apec
hly	Toxin	heamolysin	pore formation in host cells	upec
hlyF				apec
irp		Yersiniabactin synthesis		nmec
ibe ABCD			Cell invasion into host tissues	nmec, sepec, apec
pic	Toxin	Serine protease autotransporter	Degrades mucins, epithelium colonisation, cell membrane damage	upec
sat	Toxin	Secreted autotransporter toxin	influences cell vacuolization	upec, sepec
vat	Toxin	Vacuolating auto transported toxin	induces cell vacuolization	upec
cnf	Toxin	Cytotoxic necrotizing factor	creating of pores in host cell membranes	upec
cdt	Toxin	Cytolethal distending toxin	Cytolethal distending factor	sepec
usp	Toxin	uropathogenic specific protein		
astA	Toxin	EAST1	Heat stable cytotoxin associated with enteroaggregative *E. coli*	
ibeA	Toxin		Invasion of brain endothelium	
malX	Misc	Pathogenicity associated island marker CFT073		
YjaA	Unknown			
fyuA	Iron acquisition	Yersiniabactin receptor		upec
shiA	Immune suppression			
sisA	Immune suppression			
sisB	Immune suppression			
sivH	Immune suppression			
Eco274	Immune suppression			
gimB	Invasin	Genetic island associated with newborn meningitis		nmec

Appendix 5.2: Top *E. coli* virulence factors identified using abricate with the vfdb

Virulence factor	Product	Number of isolates
csgF	Curli assembly component	179
entD	Enterobactin synthesis	179
fepD	Ferric enterobactin transport system	179
fepG	Ferric enterobactin transport system	179
ompA	Outer membrane protein	179
csgB	Minor curli subunit	178
csgG	Curlin production/assembly	178
entA	Enterobactin synthesis	178
entE	Enterobactin synthesis	178
entS	Enterobactin exporter	178
fepB	Enterobactin binding protein	178
fepC	Ferric enterobactin transport system	178
entB	Enterobactin synthesis	177
fes	Enterochelin enzyme (enterobactin synthesis)	177
entC	Enterobactin gene cluster	176
fepA	Ferric enterobactin binding protein	176
csgD	Curli master regulator (biofilm)	175
fimH	Fimbrial adhesin	173
entF	Enterobactin synthase component	172
yagV/ecpE	*E. coli* common pilus	172
yagW/ecpD	*E. coli* common pilus	171
yagY/ecpB	*E. coli* common pilus	171
fimF	Fimbrial adhesin adapter	170
fimG	Fimbrial adhesin regulator	170
gspM	Type II secretion system protein	170
yagZ/ecpA	*E. coli* common pilus	170
fimC	Fimbrial adhesin chaperone	168
fimD	Fimbrial adhesin usher protein	168
fimE	Fimbrial adhesin recombinase	168
ykgK/ecpR	*E. coli* common pilus	168
fimA	Fimbrial adhesin subunit	167
fimI	Fimbrial adhesin cluster	167
yagX/ecpC	*E. coli* common pilus	167
fdeC	Intimin like adhesin	166
gspL	Inner membrane protein (Type II secretion system)	165
fimB	Fimbrial adhesin regulator	152
aslA	Ser type sulfatase	146
gspG	Type II secretion system core protein	145
gspK	Type II secretion system protein	145
gspC	Type II secretion system protein	144
gspH	Type II secretion system protein	144

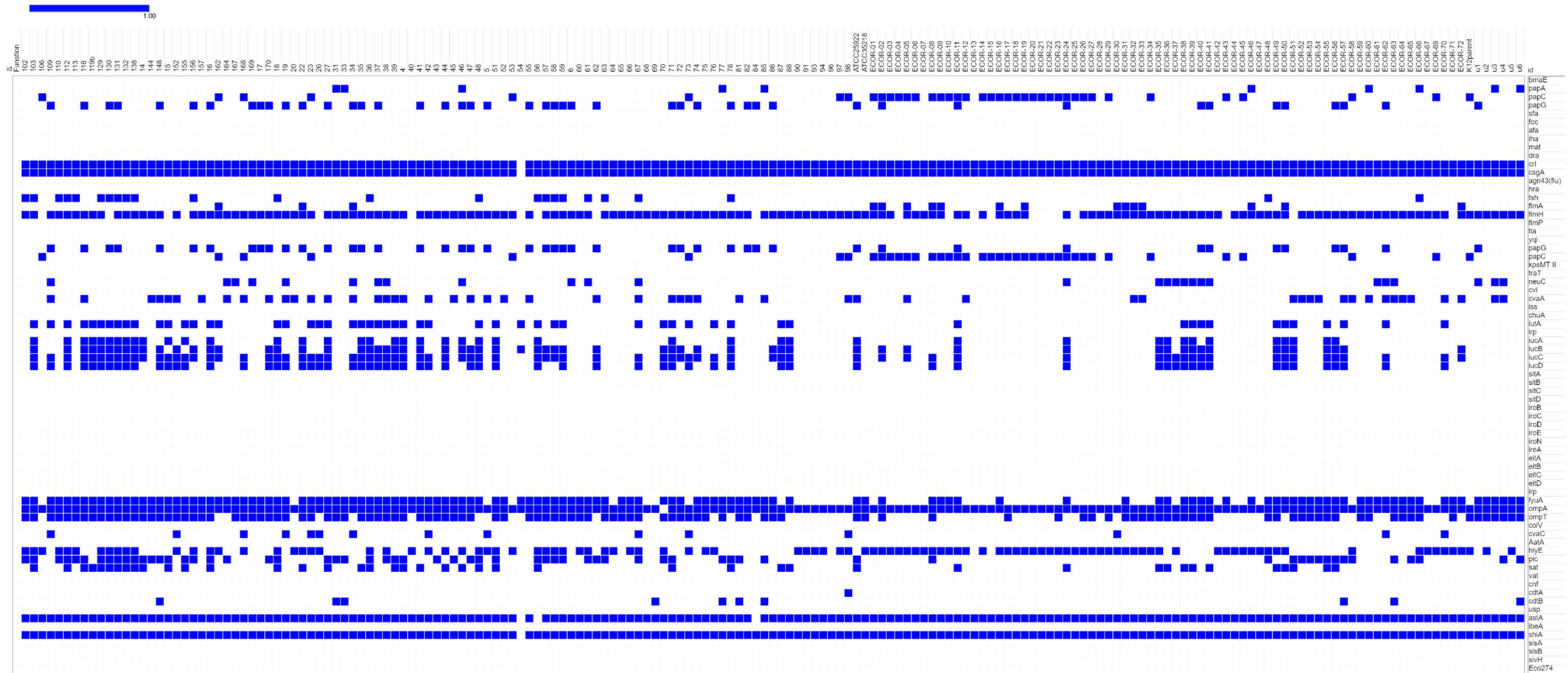
gspI	Type II secretion system psuodopilin	144
gspJ	Type II secretion system protein	144
gspF	General secretion pathway protein	142
gspD	Type II secretion system protein	141
gspE	Type II secretion system protein	141
ybtU	Siderophore synthesis	129
ybtX	Siderophore synthesis	129
ybtE	Siderophore synthesis	128
fyuA	Yersiniabactin receptor	127

Appendix 5.3AMR resistance genes in the *E. coli* collection including ECOR and reference isolates. Identified using abricate with the database card.

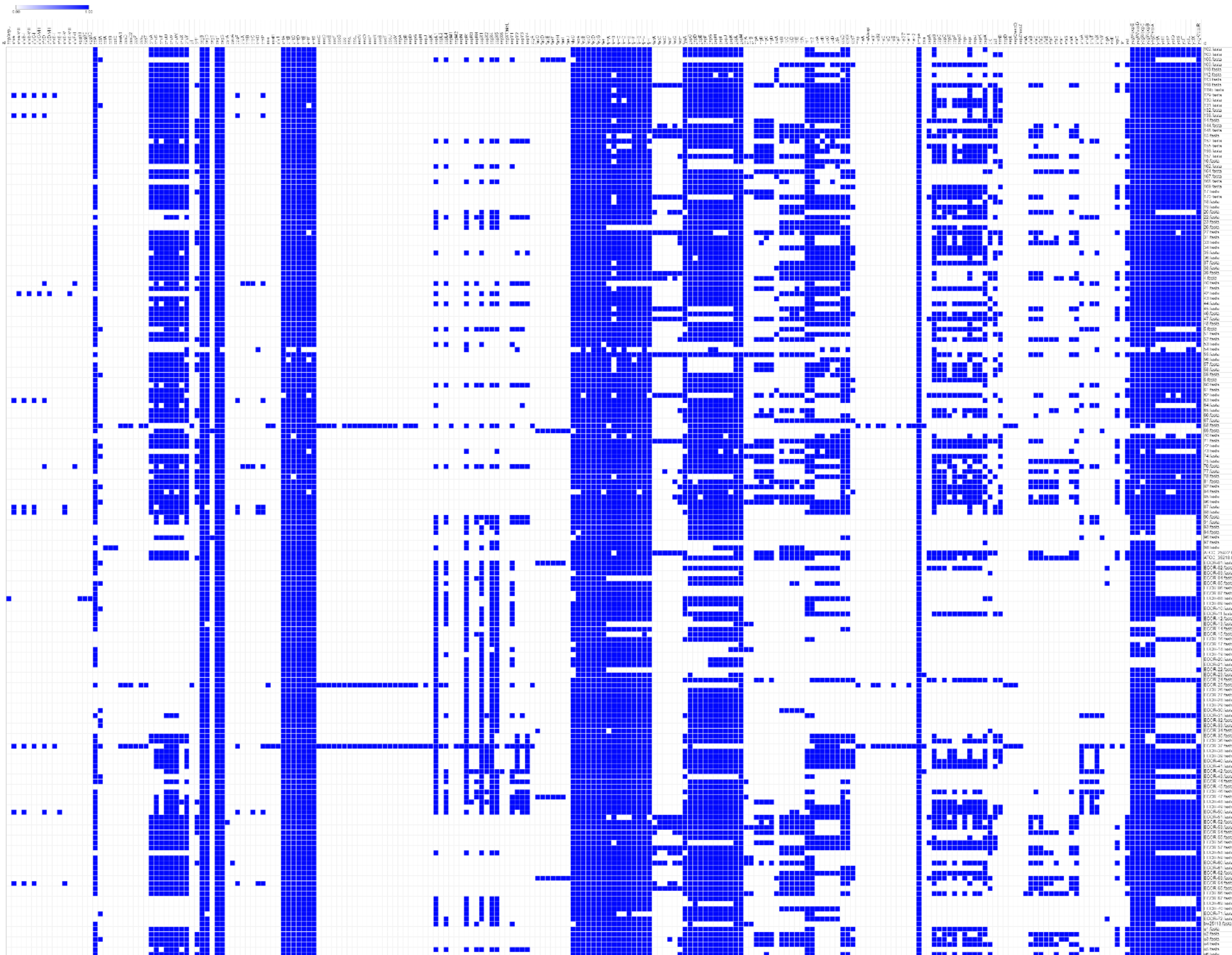
Gene name	Product	Number of isolates
acrB	Inner membrane transporter	178
acrD	Inner membrane transporter	178
bacA	Resistance to bacitracin	178
baeS	Sensor kinase	178
CRP	Global efflux pump regulator	178
emrA	Membrane fusion protein	178
eptA	Phosphoethanolamine transferase	178
Escherichia_coli_ampH	Penicillin binding protein	178
H-NS	DNA binding protein	178
mdtA	Multidrug resistance protein	178
tolC	Outer membrane efflux protein	178
baeR	Transcriptional regulatory protein	177
emrB	Multidrug export protein	177
Escherichia_coli_acrA	Aerobic respiration control protein	177
gadX	Transcriptional regulator	177
marA	Transcriptional activator	177
mdtG	Multidrug resistance protein	177
mdtH	Multidrug resistance protein	177
msbA	ABC transporter	177
pmrF	Transporter	177
acrF	Multidrug export protein	176
acrS	Multidrug efflux transport repressor	176
cpxA	Histidine kinase	176
emrR	Multidrug resistance pump regulator	176
Escherichia_coli_ampC	AmpC β-lactamase	176
gadW	Glutamic acid decarboxylase	176
mdtE	Multidrug efflux pump	176
mdtF	Multidrug efflux pump	176
mdtN	Multidrug resistance protein	176
mdtP	Multidrug resistance protein	176

yojI	ABC transporter	176
acrE	Multidrug export protein	175
emrK	Multidrug resistance protein	175
Escherichia_coli_mdfA	Multidrug resistance protein	175
evgS	Sensor protein	175
mdtC	Multidrug efflux pump	175
emrY	Multidrug efflux protein	174
kdpE	Transcriptional regulatory protein	174
mdtB	Multidrug resistance protein	174
mdtO	Multidrug resistance protein	174
evgA	DNA binding transcriptional activator	173
ugd	Glucose dehydrogenase	160
Escherichia_coli_emrE	Methyltransferase	155
mdtM	Multidrug resistance protein	116
mphB	Macrolide resistance gene	82
Escherichia_coli_ampC1_beta-lactamase	β-lactamase	75
TEM-1	TEM-1 β-lactamase	36
APH(6)-Id	Streptomycin phosphotransferase	28
sul1	Sulphonamide resistance synthase	28
APH(3'')-Ib	Aminoglycoside phosphotransferase	27
sul2	Sulphonamide resistance synthase	27
dfrA17	Dihydrofolate reductase	16
tet(A)	Tetracycline resistance protein	16
mphA	Azithromycin resistance enzyme	15
tet(B)	Tetracycline resistance protein	15
ANT(3'')-IIa	Streptomycin nucleotidyltransferase	14
AAC(6')-Ib-cr	Fluoroquinolone acetylating aminoglycoside-(6)-N-acetyltransferase	10
aadA5	Aminoglycoside nucleotidyltransferase	10
OXA-1	Penicillinase	10

Appendix 5.4 Common virulence factors gene presence/absence Abricate output was uploaded onto Morpheus to generate heatmap



Appendix 5.5 Gene presence/absence vfdb abricate output was uploaded onto Morpheus to generate heatmap



Appendix 5.6 CARD gene presence/absence heatmap abricate output was uploaded onto Morpheus to generate heatmap

Appendix 5.7 Virulence factors identified in bacteraemia isolates output from abricate was used to generate a list of most common virulence factors in *E. coli* bacteraemia isolates. ECOR, reference and water isolates excluded from analysis.

Gene	**Number of isolates**
csgF	92
entD	92
fepD	92
fepG	92
ompA	92
csgB	91
csgG	91
entA	91
entE	91
entS	91
fepB	91
fepC	91
yagY/ecpB	91
csgD	90
entB	90
fepA	90
fes	90
fimH	90
yagV/ecpE	90
yagZ/ecpA	90
ykgK/ecpR	90
entC	89
gspL	89
gspM	89
yagW/ecpD	89
fdeC	88
fimF	88
fimG	88
fimC	87
fimD	87
yagX/ecpC	87
fimA	86
fimE	86
fimI	86
entF	85
aslA	84
ybtU	84
ybtX	84
ybtE	83
fyuA	82
ybtA	82
ybtQ	82

ybtT	82
chuV	81
ybtP	81
chuU	80
chuW	80
chuY	79
ybtS	79
chuS	78
gspG	78
irp2	78
gspC	77
gspF	77
gspE	76
gspH	76
gspI	76
gspJ	76
irp1	76
gspD	75
gspK	75
kpsM	75
kpsD	74
papI	74
chuT	73
chuX	72
fimB	71
chuA	70
papB	67
papX	62
iucA	55
iucB	54
iucD	53
iutA	53
papF	52
iucC	50
papJ	50
papK	50
papD	49
papC	48
papH	47
vat	47
sat	37
iroB	36
iroD	36
hlyB	34
iroC	34
iroE	34

papG	34
senB	34
hlyD	33
iroN	33
cnf1	32
hlyA	31
hlyC	31
papE	31
sfaC	30
tcpC	29
sfaD	25
sfaY	25
focH	21
sfaB	21
pic	20
sfaX	20
espL1	18
espR1	18
espX1	16
espX5	16
focG	16
espX4	14
focD	14
kpsT	14
sfaG	14
focA	13
focC	13
focF	12
ibeA	12
espY1	11
sfaE	11
sfaF	11
espL4	10
shuA	10
shuX	10
astA	9
espY2	9
espY3	8
gtrA	8
gtrB	8
shuT	8
draP	7
espY4	7
papA	7
espR4	6
afaA	5

afaB-I	5
afaC-I	5
afaD	5
daaF	5
draD	3
sfaH	3
afaE-V	2
afaF-III	2
draA	2
draB	2
draC	2
faeD	2
faeE	2
faeF	2
faeH	2
faeI	2
nleF	2
shuS	2
afaA-VIII	1
afaB-VIII	1
afaC-VIII	1
afaD-VIII	1
afaE-I	1
afaE-VIII	1
cesAB	1
cesD	1
cesD2	1
cesL	1
cesT	1
cif	1
csgE	1
eae	1
east1	1
escC	1
escD	1
escE	1
escF	1
escI	1
escJ	1
escL	1
escN	1
escO	1
escP	1
escR	1
escS	1
escT	1

escU	1
escV	1
espB	1
espD	1
espG	1
espM1	1
espX2	1
etgA	1
faeC	1
faeJ	1
map	1
nleA/espI	1
nleB2	1
nleC	1
nleH1	1
nleH2	1
paa	1
sepD	1
sepL	1
sepQ/escQ	1
sfaS	1
shuY	1

Appendix 6

Appendix 6.1 Full list of genes associated with bacteraemia

Full list of genes identified by scoary when isolates from a bacteraemia source were compared to the ECOR collection and non-pathogenic *E. coli* strains (Section 6.4.1).

Gene	Non-unique Gene name	Annotation	Sensitivity	Specificity	Naive_p	Bonferroni_p
srpC		putative chromate transport protein	19.56522	100	3.40E-06	0.101327076
group_14844		Transcriptional regulator	19.56522	100	3.40E-06	0.101327076
group_506		hypothetical protein	19.56522	100	3.40E-06	0.101327076
group_14853		hypothetical protein	18.47826	100	7.46E-06	0.222723879
pinE_1		e14 prophage; site-specific DNA recombinase	18.47826	100	7.46E-06	0.222723879
tap_2		Multidrug efflux pump Tap	17.3913	100	1.63E-05	0.486423295
group_12279		hypothetical protein	16.30435	100	3.54E-05	1
group_14839		hypothetical protein	16.30435	100	3.54E-05	1
group_1176		hypothetical protein	14.13043	100	0.000164	1
group_8825	ycaO	protein involved in beta-methyl thiolation of ribosomal protein S12	14.13043	100	0.000164	1
group_14940		hypothetical protein	11.95652	100	0.000739	1
group_2979		hypothetical protein	11.95652	100	0.000739	1
group_12267		hypothetical protein	10.86957	100	0.001556	1
group_15734	traV	Protein TraV	10.86957	100	0.001556	1
group_1869	mprA_3	MprA-CCCP	10.86957	100	0.001556	1
group_329		hypothetical protein	10.86957	100	0.001556	1
group_8829		hypothetical protein	10.86957	100	0.001556	1
group_12428		hypothetical protein	9.782609	100	0.00326	1

group_16145	repA	Regulatory protein RepA	9.782609	100	0.00326	1
group_3203	papC	Outer membrane usher protein PapC	9.782609	100	0.00326	1
group_6320	elfC	putative outer membrane usher protein	9.782609	100	0.00326	1
group_8873		hypothetical protein	9.782609	100	0.00326	1
group_10291		hypothetical protein	8.695652	100	0.006793	1
group_6017	speC	ornithine decarboxylase, biosynthetic	8.695652	100	0.006793	1
group_13467		hypothetical protein	7.608696	100	0.014077	1
group_3608	ccdB	hypothetical protein	7.608696	100	0.014077	1
group_5651	ydjH_2	putative kinase	7.608696	100	0.014077	1
group_7619		hypothetical protein	7.608696	100	0.014077	1
group_12478		hypothetical protein	6.521739	100	0.029017	1
group_16277		hypothetical protein	6.521739	100	0.029017	1
group_16278		hypothetical protein	6.521739	100	0.029017	1
group_3124		Tn3 family transposase Tn3	6.521739	100	0.029017	1
group_3617	yiaK	2,3-diketo-L-gulonate reductase monomer	6.521739	100	0.029017	1
group_4286		hypothetical protein	6.521739	100	0.029017	1
group_7794	pgi	phosphoglucose isomerase	6.521739	100	0.029017	1
group_8826		hypothetical protein	6.521739	100	0.029017	1
group_7124		hypothetical protein	23.91304	98.85057	1.81E-06	0.054098331
group_857	ydfB	Qin prophage; small protein	23.91304	98.85057	1.81E-06	0.054098331
zntR_2		ZntR transcriptional activator	20.65217	98.85057	1.76E-05	0.525290724
group_2523		hypothetical protein	20.65217	98.85057	1.76E-05	0.525290724
merA		Mercuric reductase	16.30435	98.85057	0.000321	1

group_15932		hypothetical protein		16.30435	98.85057	0.000321	1
merC		Mercuric transport protein MerC		16.30435	98.85057	0.000321	1
merT		Mercuric transport protein MerT		16.30435	98.85057	0.000321	1
merP		Mercuric transport protein periplasmic component		16.30435	98.85057	0.000321	1
group_4018	csgD	CsgD DNA-binding transcriptional dual regulator		16.30435	98.85057	0.000321	1
group_12281		hypothetical protein		14.13043	98.85057	0.001298	1
dosP_2		c-di-GMP phosphodiesterase, heme-regulated		13.04348	98.85057	0.002569	1
group_15931		hypothetical protein		13.04348	98.85057	0.002569	1
group_5829	poxB	pyruvate oxidase monomer		11.95652	98.85057	0.005029	1
group_8776		hypothetical protein		11.95652	98.85057	0.005029	1
group_4829		hypothetical protein		9.782609	98.85057	0.018597	1
group_11724	corA	Ni2+ / Co2+ / Mg2+ transporter		8.695652	98.85057	0.035054	1
group_5030	epmC	EF-P-Lys34 hydroxylase		8.695652	98.85057	0.035054	1
group_7374		hypothetical protein		8.695652	98.85057	0.035054	1
group_9159		IS21 family transposase ISEc10		26.08696	97.70115	2.88E-06	0.086094101
group_2611	insC-1_2	IS2 element protein InsA		25	97.70115	6.05E-06	0.180643045
group_2976		hypothetical protein		25	97.70115	6.05E-06	0.180643045
group_12115		hypothetical protein		16.30435	97.70115	0.00155	1
group_9245		hypothetical protein		16.30435	97.70115	0.00155	1
eno_2		degradosome		15.21739	97.70115	0.002943	1
group_15377		hypothetical protein		15.21739	97.70115	0.002943	1
crcB_2		F- efflux transporter / involved in resistance to camphor-induced chromosome decondensation		14.13043	97.70115	0.005516	1
group_8124	yfdV	YfdV AEC Transporter		14.13043	97.70115	0.005516	1

group_87 39		hypothetical protein	14.1304 3	97.7011 5	0.0055 16	1
group_15 740		hypothetical protein	13.0434 8	97.7011 5	0.0101 94	1
emrK_2		EmrKY-TolC multidrug efflux transport system - membrane fusion protein	13.0434 8	97.7011 5	0.0101 94	1
group_56 76	yqiH	putative membrane protein	11.9565 2	97.7011 5	0.0185 59	1
pemI		Antitoxin PemI	38.0434 8	96.5517 2	3.26E-09	9.73E-05
group_22 32		hypothetical protein	36.9565 2	96.5517 2	7.43E-09	0.0002217 95
group_98 47	chpB	ChpB toxin of the ChpB-ChpS toxin-antitoxin system	35.8695 7	96.5517 2	1.67E-08	0.0004998 79
insN-1_1		CP4-6 prophage; partial regulator of insertion element IS911A	32.6087	96.5517 2	1.79E-07	0.0053416 4
folP_2		dihydropteroate synthase	29.3478 3	96.5517 2	1.73E-06	0.0515642 49
xerC_4		Tyrosine recombinase XerC	29.3478 3	96.5517 2	1.73E-06	0.0515642 49
group_28 73	umuC_2	SOS mutagenesis and repair	19.5652 2	96.5517 2	0.0008 39	1
znuB_3		Zn2+ ABC transporter - membrane subunit	18.4782 6	96.5517 2	0.0015 69	1
group_14 49	ydfU_1	Qin prophage; predicted protein	16.3043 5	96.5517 2	0.0052 69	1
group_15 357		Sugar transporter SemiSWEET	14.1304 3	96.5517 2	0.0166 65	1
ntdC		hypothetical protein	14.1304 3	96.5517 2	0.0166 65	1
traY		Relaxosome protein TraY	29.3478 3	95.4023	7.85E-06	0.2343138 54
group_99 87	traJ	Protein TraJ	27.1739 1	95.4023	3.12E-05	0.9314581 34
group_12 109	yciC_2	Putative metal chaperone YciC	15.2173 9	95.4023	0.0240 62	1
bla		Beta-lactamase TEM	42.3913	94.2528 7	6.97E-09	0.0002080 96
group_32 27		hypothetical protein	33.6956 5	94.2528 7	1.84E-06	0.0550410 55
group_98 28		hypothetical protein	33.6956 5	94.2528 7	1.84E-06	0.0550410 55
group_97 99		hypothetical protein	31.5217 4	94.2528 7	7.55E-06	0.2252753 16
pinE_2		e14 prophage; site-specific DNA recombinase	29.3478 3	94.2528 7	2.94E-05	0.8764045 64

group_9839		IS6 family transposase IS26	22.82609	94.25287	0.001263	1
group_14884		hypothetical protein	43.47826	93.10345	1.36E-08	0.000406239
agp_2		3-phytase / glucose-1-phosphatase	35.86957	93.10345	1.67E-06	0.049851372
group_7134	traA	Pilin	32.6087	93.10345	1.33E-05	0.397146164
traM		Relaxosome protein TraM	31.52174	93.10345	2.59E-05	0.77248957
group_2877		hypothetical protein	22.82609	93.10345	0.003239	1
group_5906		hypothetical protein	19.56522	93.10345	0.015504	1
group_889		hypothetical protein	18.47826	93.10345	0.025265	1
ykgG_2		putative transporter	0	93.10345	0.01203	1
group_5598		hypothetical protein	0	93.10345	0.01203	1
nqrC		Na(+)-translocating NADH-quinone reductase subunit C	38.04348	91.95402	2.40E-06	0.071693246
group_1890	yjgL	putative protein	0	91.95402	0.005632	1
insH-6		CP4-44 prophage; IS5 transposase and trans-activator	0	91.95402	0.005632	1
group_4768		hypothetical protein	0	91.95402	0.005632	1
group_7558	yraI	putative pilin chaperone	0	91.95402	0.005632	1
group_994		hypothetical protein	0	91.95402	0.005632	1
tpd		34 kDa membrane antigen	42.3913	90.8046	4.20E-07	0.012527574
efeU_2		hypothetical protein	42.3913	90.8046	4.20E-07	0.012527574
group_8389		hypothetical protein	42.3913	90.8046	4.20E-07	0.012527574
group_2393	ydbC	putative oxidoreductase, NAD(P)-binding	39.13043	90.8046	3.72E-06	0.111078343
group_11833	folP_2	dihydropteroate synthase	21.73913	90.8046	0.024078	1
ydeO_1		YdeO DNA-binding transcriptional dual regulator	0	90.8046	0.00262	1
ykgE_2		putative oxidoreductase	0	90.8046	0.00262	1

ykgF_2		putative amino acid dehydrogenase with NAD(P)-binding domain and ferridoxin-like domain protein	0	90.8046	0.00262	1
group_1416		hypothetical protein	0	90.8046	0.00262	1
group_4210	insD	Qin prophage; predicted transposase	0	90.8046	0.00262	1
lldP_2		(R)-lactate / (S)-lactate / glycolate:H+ symporter LldP	0	90.8046	0.00262	1
group_3099		hypothetical protein	39.13043	89.65517	1.09E-05	0.324515756
group_9884		hypothetical protein	36.95652	89.65517	4.07E-05	1
group_21459	traM	Relaxosome protein TraM	1.086957	89.65517	0.008106	1
group_675	tfaQ_2	Qin prophage; predicted tail fibre assembly protein	0	89.65517	0.00121	1
group_13358		hypothetical protein	1.086957	88.50575	0.004064	1
group_24229	yfcO	putative protein	0	88.50575	0.000555	1
group_7681		hypothetical protein	0	88.50575	0.000555	1
sopB		Protein SopB	34.78261	87.35632	0.000764	1
group_9993		hypothetical protein	34.78261	87.35632	0.000764	1
group_4413	ygiS	putative transporter subunit	2.173913	87.35632	0.008458	1
group_10903		hypothetical protein	0	87.35632	0.000253	1
iutA		Ferric aerobactin receptor	47.82609	86.2069	8.88E-07	0.026515563
group_7528		hypothetical protein	2.173913	85.05747	0.002299	1
elfC_2		putative outer membrane usher protein	1.086957	85.05747	0.000475	1
group_469	yeeJ_2	adhesin	1.086957	85.05747	0.000475	1
group_6302	nmpC	outer membrane porin protein; locus of qsr prophage	1.086957	85.05747	0.000475	1
ynjI_2		putative inner membrane protein	0	83.90805	2.29E-05	0.683891258
group_7652	rzpD	DLP12 prophage; predicted murein endopeptidase	2.173913	82.75862	0.000589	1
group_475	yeeJ_1	adhesin	0	82.75862	1.01E-05	0.302570072

noc		Nucleoid occlusion protein	35.86957	81.6092	0.011605	1
ydfN		hypothetical protein	2.173913	80.45977	0.000143	1
group_157	vgrG1_2	Actin cross-linking toxin VgrG1	2.173913	78.16092	3.31E-05	0.988913131
group_14975		hypothetical protein	1.086957	77.01149	2.16E-06	0.064500651
group_21728	dicA	DicA DNA-binding transcriptional repressor	1.086957	73.56322	1.83E-07	0.005475585
group_21729	dicC	Qin prophage; DNA-binding transcriptional regulator for DicB	1.086957	73.56322	1.83E-07	0.005475585
tufB		elongation factor Tu	91.30435	65.51724	4.63E-16	1.38E-11
php		putative hydrolase	17.3913	52.87356	2.49E-05	0.741934535
php_1		putative hydrolase	82.6087	50.57471	3.26E-06	0.097254163
ybjE		putative transporter	95.65217	49.42529	1.03E-12	3.09E-08
group_4156	ybjE	putative transporter	3.26087	49.42529	8.40E-14	2.51E-09
group_5905		hypothetical protein	79.34783	5.747126	0.003955	1
yeeO		YeeO MATE transporter	83.69565	4.597701	0.014169	1
yicO		putative membrane protein with possible relationship to novobiocin and deoxycholate resistance	91.30435	1.149425	0.035054	1
yfcJ		putative transport protein YfcJ	89.13043	1.149425	0.009732	1
poxB		pyruvate oxidase monomer	86.95652	1.149425	0.002569	1
group_11265	yedW_3	putative DNA-binding response regulator in two-component system with YedV	81.52174	1.149425	7.66E-05	1
ybdK		carboxylate-amine ligase	93.47826	0	0.029017	1
xylG		xylose ABC transporter - ATP binding subunit	93.47826	0	0.029017	1
group_5823		hypothetical protein	93.47826	0	0.029017	1
potI		putrescine ABC transporter - membrane subunit	93.47826	0	0.029017	1
glpQ		glycerophosphoryl diester phosphodiesterase, periplasmic	93.47826	0	0.029017	1
flhD		FlhDC DNA-binding transcriptional dual regulator	93.47826	0	0.029017	1

ptsI		PTS enzyme I	93.47826	0	0.029017	1
ccmF		cytochrome c-type biogenesis protein	92.3913	0	0.014077	1
mdtB		MdtABC-TolC multidrug efflux transport system - membrane subunit	92.3913	0	0.014077	1
tqsA		quorum signal AI-2 exporter	92.3913	0	0.014077	1
pspF		PspF transcriptional dual regulator	92.3913	0	0.014077	1
pgi		phosphoglucose isomerase	92.3913	0	0.014077	1
nuoF		NADH:ubiquinone oxidoreductase, chain F	92.3913	0	0.014077	1
clpA		ClpAXP	91.30435	0	0.006793	1
ycbB		L,D-transpeptidase YcbB	91.30435	0	0.006793	1
speC		ornithine decarboxylase, biosynthetic	91.30435	0	0.006793	1
group_4480		hypothetical protein	90.21739	0	0.00326	1
eutE		putative aldehyde dehydrogenase, ethanolamine utilization protein	89.13043	0	0.001556	1
adrA		putative diguanylate cyclase	86.95652	0	0.000349	1
csgD		CsgD DNA-binding transcriptional dual regulator	83.69565	0	3.54E-05	1
ycaO		protein involved in beta-methylthiolation of ribosomal protein S12	83.69565	0	3.54E-05	1
group_2522		hypothetical protein	78.26087	0	6.89E-07	0.020559693

Appendix 6.2 Full list of genes associated with urinary source of infection

Full list of genes identified by scoary when urinary bacteraemia isolate genomes were compared to all other *E. coli* used in the genetic analysis (Section 6.4.2).

Gene	Non-unique Gene name	Annotation	Sensitivity	Specificity	Naive_p	Bonferroni_p
group_4156	ybjE	putative transporter	2.222222	64.34109	1.55E-06	0.046089
tufB		elongation factor Tu	91.11111	47.28682	1.81E-06	0.053777
flxA		Qin prophage; predicted protein	2.222222	67.44186	1.18E-05	0.351956
glvG		putative phospho-glucosidase, truncated	0	73.64341	1.45E-05	0.432292
ybjE		putative transporter	95.55556	34.88372	2.57E-05	0.763654
znuB_2		Zn2+ ABC transporter - membrane subunit	93.33333	37.9845	4.57E-05	1
group_7201	insG	IS4 family transposase ISCro3	0	75.96899	5.52E-05	1
group_11992		hypothetical protein	2.222222	72.09302	8.82E-05	1
group_9903		hypothetical protein	0	77.51938	0.000105	1
group_11459		hypothetical protein	2.222222	72.86822	0.00017	1
group_11899		hypothetical protein	2.222222	72.86822	0.00017	1
yzgL		putative protein	0	78.29457	0.00021	1
group_5823		hypothetical protein	86.66667	0	0.000231	1
fhuC_2		iron (III) hydroxamate ABC transporter - ATP binding subunit	88.88889	39.53488	0.000355	1
group_21728	dicA	DicA DNA-binding transcriptional repressor	0	81.39535	0.000681	1
group_21729	dicC	Qin prophage; DNA-binding transcriptional regulator for DicB	0	81.39535	0.000681	1
yabP		hypothetical protein	0	81.39535	0.000681	1
group_3588	ydeP	acid resistance protein	6.666667	69.76744	0.001033	1
group_17806	fliD	flagellar cap protein FliD; filament capping protein; enables filament assembly	0	82.94574	0.001238	1
rem		Qin prophage; predicted protein	2.222222	78.29457	0.001881	1
gadA		glutamate decarboxylase A subunit	17.77778	55.81395	0.002107	1
group_5440	ybeF	putative DNA-binding transcriptional regulator, LYSR-type	0	84.49612	0.002267	1
group_14975		hypothetical protein	0	83.72093	0.00232	1
group_834	yghJ	putative lipoprotein	0	86.82171	0.007101	1
yohF		putative oxidoreductase with NAD(P)-binding Rossmann-fold domain protein	86.66667	34.10853	0.007551	1
group_12949		hypothetical protein	2.222222	82.94574	0.009637	1
group_6886	xapR	XapR transcriptional activator	2.222222	82.94574	0.009637	1
group_8479		hypothetical protein	2.222222	82.94574	0.009637	1
group_5603		hypothetical protein	4.444444	79.06977	0.009854	1
ydeP		acid resistance protein	88.88889	30.23256	0.01007	1
group_10441	insE-1	IS3 element protein InsE	2.222222	82.17054	0.01007	1

group_1933	insEF-1	IS3 element transposase	2.222222	82.17054	0.01007	1
group_475	yeeJ_1	adhesin	0	88.37209	0.012592	1
rusA_1		endodeoxyribonuclease RUS (Holliday junction resolvase)	75.55556	46.51163	0.013153	1
group_8302	essD_1	DLP12 prophage; predicted phage lysis protein	2.222222	84.49612	0.016469	1
group_3075	insI-2	transposase of IS30	2.222222	83.72093	0.016642	1
group_20573	ybcO	DLP12 prophage; predicted protein	0	89.14729	0.022131	1
elfC_2		putative outer membrane usher protein	0	89.14729	0.022131	1
group_3289	yggP_1	putative dehydrogenase	0	89.14729	0.022131	1
group_6302	nmpC	outer membrane porin protein; locus of qsr prophage	0	89.14729	0.022131	1
group_8880		hypothetical protein	0	89.92248	0.02249	1
group_999		hypothetical protein	2.222222	85.27132	0.027434	1
group_17864		hypothetical protein	2.222222	86.04651	0.027949	1
group_17865		hypothetical protein	2.222222	86.04651	0.027949	1
ydfN		hypothetical protein	2.222222	86.04651	0.027949	1
group_5095	yjjQ	putative DNA-binding transcriptional regulator	15.55556	67.44186	0.034134	1
yfdT		CPS-53 (KpLE1) prophage; predicted protein	6.666667	79.06977	0.037508	1
yfdS		CPS-53 (KpLE1) prophage; predicted protein	6.666667	79.84496	0.038043	1
ydeQ		putative fimbrial-like adhesin protein	88.88889	26.35659	0.03846	1
group_5114	ecpD	hypothetical protein	2.222222	86.82171	0.04504	1
group_6		hypothetical protein	2.222222	86.82171	0.04504	1
group_6544	torA	trimethylamine N-oxide reductase, catalytic subunit	2.222222	86.82171	0.04504	1

Appendix 6.3 Full list of genes associated with an abdominal source of infection

Full list of genes identified by scoary when genomes from abdominal bacteraemia isolates were compared against all other isolates (Section 6.4.2).

Gene	Non-unique Gene name	Annotation	Sensitivity	Specificity	Odds_ratio	Naive_p	Bonferroni_p
group_11920		hypothetical protein	0	66.66667	0	0.000335	1
group_11805		hypothetical protein	90.90909	43.13725	7.586207	0.001879	1
hpcE		Homoprotocatechuate catabolism bifunctional isomerase/decarboxylase	0	72.54902	0	0.002416	1
puuC_1		gamma-glutamyl-gamma-aminobutyraldehyde dehydrogenase	0	72.54902	0	0.002416	1
hpcB		3,4-dihydroxyphenylacetate 2,3-dioxygenase	0	72.54902	0	0.002416	1
farR		HTH-type transcriptional regulator FarR	0	73.20261	0	0.002498	1
mhpD_2		2-hydroxypentadienoate hydratase	0	73.20261	0	0.002498	1
yfaU_2		2-keto-3-deoxy-L-rhamnonate aldolase	0	73.20261	0	0.002498	1
hpaB		4-hydroxyphenylacetate 3-monooxygenase oxygenase component	0	73.20261	0	0.002498	1
rutF_1		flavin reductase	0	73.20261	0	0.002498	1
group_1772		hypothetical protein	0	75.81699	0	0.004839	1
group_1224		hypothetical protein	9.090909	62.7451	0.168421	0.007785	1
group_6219	envY	EnvY DNA-binding transcriptional activator	0	79.08497	0	0.01537	1
ydfA		Qin prophage; predicted protein	0	79.73856	0	0.015387	1
hpcD		5-carboxymethyl-2-hydroxymuconate Delta-isomerase	0	78.43137	0	0.015898	1
ybcO		DLP12 prophage; predicted protein	18.18182	55.55556	0.277778	0.020909	1
insC-1		IS2 element protein InsA	9.090909	67.97386	0.212245	0.025586	1
group_3082	fimA_2	major type 1 subunit fimbrin (pilin)	9.090909	67.97386	0.212245	0.025586	1

insCD-1		IS2 element transposase InsAB'	9.090909	67.97386	0.212245	0.025586	1
yfaA		putative protein	100	18.30065	inf	0.027012	1
insO-1_3		hypothetical protein	0	81.69935	0	0.027012	1
group_9483		hypothetical protein	0	82.35294	0	0.027497	1
group_5365		hypothetical protein	0	82.35294	0	0.027497	1
group_3745		IS3 family transposase IS3411	4.545455	73.85621	0.134524	0.028952	1
ydfT		Qin prophage; predicted antitermination protein Q	18.18182	58.16993	0.309028	0.036672	1
glcA		glycolate / lactate:H+ symporter	90.90909	30.06536	4.299065	0.042128	1
group_37		hypothetical protein	0	83.66013	0	0.046855	1
group_4977	yfjI_1	CP4-57 prophage; predicted protein	0	83.66013	0	0.046855	1
group_1360		hypothetical protein	100	15.68627	inf	0.046979	1
group_578		IS66 family transposase ISEc23	0	84.31373	0	0.046979	1
group_12516		hypothetical protein	0	84.31373	0	0.046979	1
group_1346		hypothetical protein	0	84.31373	0	0.046979	1
group_4415	yghT	putative protein with nucleoside triphosphate hydrolase domain protein	0	84.31373	0	0.046979	1
group_6131		hypothetical protein	0	84.31373	0	0.046979	1
ydfB		Qin prophage; small protein	0	83.00654	0	0.048168	1
group_1618	rzpQ	Qin prophage; predicted protein	0	83.00654	0	0.048168	1
group_3080	insF-1_2	IS3 element protein InsF	0	83.00654	0	0.048168	1
group_3591	yfaA	putative protein	0	83.00654	0	0.048168	1
group_5406	xerC_1	Tyrosine recombinase XerC	4.545455	75.81699	0.149292	0.04941	1
group_10367		hypothetical protein	4.545455	77.12418	0.160544	0.049924	1
group_2250		hypothetical protein	4.545455	77.12418	0.160544	0.049924	1

Appendix 6.4 Full list of genes associated with urinary source of bacteraemia

Full list of genes identified by scoary analysis when genomes from a urinary source of bacteraemia were compared to genomes from all other bacteraemia sources (Section 6.4.3).

Gene	Annotation	Sensitivity	Specificity	Naive_p	Bonferroni_p
group_16170 (iutA)	Ferric aerobactin receptor	22.22222	100	0.001165	1
group_16178	hypothetical protein	22.22222	100	0.001165	1
group_6577 (agaC_1)	galactosamine PTS permease - cryptic	15.55556	100	0.012383	1
yqiG_2	putative membrane protein	55.55556	71.42857	0.016643	1
group_1831	hypothetical protein	13.33333	100	0.026518	1
group_3095	hypothetical protein	13.33333	100	0.026518	1
xapB	xanthosine:H+ symporter XapB	75.55556	7.142857	0.040197	1

Appendix 6.5 Full list of genes associated with Mortality

Full list of genes identified by scoary as being associated with mortality as identified in patient data (Chapter 3).

Gene	Non-unique Gene name	Annotation	Sensitivity	Specificity	Odds_ratio	Naive_p	Bonferroni_p
group_4139		hypothetical protein	17.64706	34.72222	0.113982	0.0007	1
group_3099		hypothetical protein	5.882353	51.38889	0.066071	0.000876	1
yhgE		putative transport protein	100	38.88889	inf	0.000975	1
traQ		Protein TraQ	11.76471	44.44444	0.106667	0.001092	1
group_1386	yhgE	putative transport protein	0	62.5	0	0.00114	1
arfB		peptidyl-tRNA hydrolase, ribosome rescue factor	100	34.72222	inf	0.002318	1
traD		Coupling protein TraD	5.882353	55.55556	0.078125	0.004041	1
yhdJ_2		DNA adenine methyltransferase	5.882353	55.55556	0.078125	0.004041	1
group_5082	arfB	peptidyl-tRNA hydrolase, ribosome rescue factor	0	66.66667	0	0.00457	1
higB-2_1		Toxin HigB-2	0	68.05556	0	0.004674	1
group_6738		hypothetical protein	0	68.05556	0	0.004674	1
yqiG_2		putative membrane protein	11.76471	50	0.133333	0.005415	1
group_4140		hypothetical protein	17.64706	45.83333	0.181319	0.007398	1
group_7134	traA	Pilin	5.882353	59.72222	0.092672	0.008606	1
group_11824		hypothetical protein	5.882353	61.11111	0.098214	0.008942	1
cpsB		mannose-1-phosphate guanylyltransferase	100	29.16667	inf	0.009304	1
group_12229		hypothetical protein	0	70.83333	0	0.009304	1
group_2477		hypothetical protein	0	70.83333	0	0.009304	1

idnO_2		5-keto-D-gluconate 5-reductase	0	70.83333	0	0.009304	1
group_15693		hypothetical protein	0	70.83333	0	0.009304	1
yagG		YagG GPH Transporter	0	70.83333	0	0.009304	1
group_2123		hypothetical protein	0	70.83333	0	0.009304	1
group_15696		hypothetical protein	0	70.83333	0	0.009304	1
group_15743		hypothetical protein	0	70.83333	0	0.009304	1
group_12251		hypothetical protein	0	69.44444	0	0.009572	1
ogl		Oligogalacturonate lyase	0	69.44444	0	0.009572	1
group_3718		hypothetical protein	0	69.44444	0	0.009572	1
group_7128		hypothetical protein	0	72.22222	0	0.010014	1
group_6396		hypothetical protein	0	72.22222	0	0.010014	1
umuD_2		SOS mutagenesis; error-prone repair; processed to UmuD'; forms complex with UmuC	5.882353	62.5	0.104167	0.01004	1
group_4738		hypothetical protein	11.76471	52.77778	0.14902	0.011703	1
yedZ1		Putative protein-methionine-sulfoxide reductase subunit YedZ1	11.76471	55.55556	0.166667	0.013219	1
yedW_11		putative DNA-binding response regulator in two-component system with YedV	11.76471	55.55556	0.166667	0.013219	1
group_15641		hypothetical protein	11.76471	55.55556	0.166667	0.013219	1
yedY_2		reductase	11.76471	55.55556	0.166667	0.013219	1
group_11966		hypothetical protein	11.76471	55.55556	0.166667	0.013219	1
group_1175		hypothetical protein	5.882353	63.88889	0.110577	0.017409	1
finO		Fertility inhibition protein	5.882353	63.88889	0.110577	0.017409	1
group_509		hypothetical protein	5.882353	65.27778	0.1175	0.018574	1
traI		Multifunctional conjugation protein TraI	5.882353	65.27778	0.1175	0.018574	1
intA		CP4-57 prophage; integrase	0	75	0	0.018731	1
group_5387	ssb_2	ssDNA-binding protein	0	75	0	0.018731	1
group_3381		hypothetical protein	11.76471	58.33333	0.186667	0.024689	1
group_2635		hypothetical protein	11.76471	58.33333	0.186667	0.024689	1
hha_2		haemolysin expression modulating protein	5.882353	66.66667	0.125	0.033041	1
gmm		GDP-mannose mannosyl hydrolase	94.11765	31.94444	7.510204	0.033798	1
group_5075		hypothetical protein	5.882353	68.05556	0.133152	0.033798	1
yfjI_1		CP4-57 prophage; predicted protein	5.882353	68.05556	0.133152	0.033798	1
fimC_4		periplasmic chaperone, required for type 1 fimbriae	5.882353	68.05556	0.133152	0.033798	1
hlyA		Hemolysin, chromosomal	5.882353	68.05556	0.133152	0.033798	1
yhcR		Endonuclease YhcR	5.882353	68.05556	0.133152	0.033798	1
group_11821		hypothetical protein	5.882353	68.05556	0.133152	0.033798	1
group_8823		IS110 family transposase ISEc21	0	77.77778	0	0.034569	1
insO-2		hypothetical protein	11.76471	59.72222	0.197701	0.044881	1
group_9828		hypothetical protein	11.76471	59.72222	0.197701	0.044881	1
group_9840		hypothetical protein	11.76471	61.11111	0.209524	0.045027	1
group_8380		hypothetical protein	11.76471	62.5	0.222222	0.047643	1

Appendix 6.6 Full list of genes associated with Nosocomial

Full list of genes identified by scoary analysis as being associated with nosocomial infections as identified by patient data (Chapter 3).

Gene	Non-unique Gene name	Annotation	Sensitivity	Specificity	Odds_ratio	Naive_p	Bonferroni_p
yhdJ_2		DNA adenine methyltransferase	13.04348	59.64912	0.221739	0.019549	1
allD		ureidoglycolate dehydrogenase	100	19.29825	inf	0.028352	1
aat		leucyl, phenylalanyl-tRNA-protein transferase	91.30435	31.57895	4.846154	0.044836	1
group_5933	aat	leucyl, phenylalanyl-tRNA-protein transferase	8.695652	68.42105	0.206349	0.044836	1
group_2526	yicI	alpha-xylosidase	8.695652	68.42105	0.206349	0.044836	1

Appendix 6.7 Full list of genes associated with plasma resistance

Full list of genes identified by scoary as being associated with plasma resistance in *E. coli* isolate collection as identified in chapter 4.

Gene	Non-unique Gene name	Annotation	Sensitivity	Specificity	Odds_ratio	Naive_p	Bonferroni_p
group_3464		hypothetical protein	22.91667	100	inf	0.000705	1
group_8772		hypothetical protein	20.83333	100	inf	0.001507	1
group_12789		hypothetical protein	0	85.36585	0	0.007738	1
group_7466		IS1380 family transposase ISEcp1	0	85.36585	0	0.007738	1
yejF		peptide ABC transporter - ATP binding subunit	97.91667	19.5122	11.39394	0.010405	1
group_7879	yejF	peptide ABC transporter - ATP binding subunit	2.083333	80.4878	0.087766	0.010405	1
agaC_1		galactosamine PTS permease - cryptic	85.41667	0	0	0.013949	1
group_1050		hypothetical protein	14.58333	100	inf	0.013949	1
group_3530	agaA	N-acetylgalactosamine-6-phosphate deacetylase	14.58333	100	inf	0.013949	1
group_6273	yfjR	CP4-57 prophage; predicted DNA-binding transcriptional regulator	14.58333	100	inf	0.013949	1
group_6577	agaC_1	galactosamine PTS permease - cryptic	14.58333	100	inf	0.013949	1
group_10464		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_12558		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_12562		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_16543		hypothetical protein	18.75	97.56098	9.230769	0.018336	1

group_16544		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_16545		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_16547		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_16548		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_16549		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_16550		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_16554		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_16557		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_5584		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
group_6472		hypothetical protein	18.75	97.56098	9.230769	0.018336	1
ampG_2		muropeptide:H+ symporter	97.91667	14.63415	8.057143	0.045272	1
group_9771	msbA	Putative multidrug export ATP-binding/permease protein	97.91667	14.63415	8.057143	0.045272	1

Appendix 6.8 Full list of genes associated with low cytokine <1000pg/mL in THP-1 cells

Full list of genes identified by scoary analysis as being associated with IL-8 levels <1000pg/mL. Note, although IL-8 was used for comparison isolates with low IL-8 also induced low TNFα. As identified in chapter 4.

Gene	Non-unique Gene name	Annotation	Sensitivity	Specificity	Odds_ratio	Naive_p	Bonferroni_p
flgI		flagellar P-ring protein FlgI	100	26.82927	inf	0.001643	1
group_8081	flgI	flagellar P-ring protein FlgI	0	73.17073	0	0.001643	1
group_4006	prpE	propionyl-CoA synthetase	0	75.60976	0	0.003228	1
group_5152	flgK	flagellar biosynthesis, hook-filament junction protein 1	0	76.82927	0	0.00572	1
insAB-1		InsAB' transposase	4	71.95122	0.106884	0.012261	1
group_11877		hypothetical protein	0	80.4878	0	0.020602	1
yjhB		YjhB MFS transporter	4	74.39024	0.121032	0.022077	1
yjhC		KpLE2 phage-like element; predicted oxidoreductase	4	74.39024	0.121032	0.022077	1
group_447		hypothetical protein	4	75.60976	0.129167	0.02327	1
group_4697	yjhR	KpLE2 phage-like element; predicted frameshift suppressor	8	70.73171	0.210145	0.033729	1
group_5406	xerC_1	Tyrosine recombinase XerC	0	84.14634	0	0.035763	1
group_10367		hypothetical protein	0	84.14634	0	0.035763	1
ogrK		DNA-binding transcriptional regulator, prophage P2 remnant	0	84.14634	0	0.035763	1
group_7462		hypothetical protein	0	84.14634	0	0.035763	1
dam_2		DNA adenine methyltransferase	0	82.92683	0	0.036726	1
eutJ		putative chaperonin, ethanolamine utilization protein	96	21.95122	6.75	0.040671	1
group_305		hypothetical protein	4	78.04878	0.148148	0.040671	1
umuC_2		SOS mutagenesis and repair	4	78.04878	0.148148	0.040671	1

Appendix 6.9 Full list of genes associated with low cytokine <500pg/mL in THP-1 cells

Full list of genes identified by scoary analysis as being associated with IL-8 levels <500pg/mL. Note, although IL-8 was used for comparison isolates with low IL-8 also induced low TNFα. As identified in chapter 4.

Gene	Non-unique Gene name	Annotation	Sensitivity	Specificity	Odds_ratio	Naive_p	Bonferroni_p	Benjamini_H_p
ynbC		putative hydrolase	100	55.95238	inf	6.24E-06	0.139319	0.001443
group_10095	ydfD	Qin prophage; predicted protein	0	61.90476	0	0.001104	1	0.056802
umuC_2		SOS mutagenesis and repair	0	78.57143	0	0.037229	1	0.705011

Appendix 6.10 Full list of genes associated with sepsis/non-sepsis

Full list of genes identified by scaory as being associated with sepsis causing isolates as identified through patient data (Chapter 3).

Gene	Non-unique Gene name	Annotation	Sensitivity	Specificity	Odds_ratio	Naive_p	Bonferroni_p
group_2526	yicI	alpha-xylosidase	9.302326	59.18367	0.148718	0.000714	1
group_4562	yicJ	YicJ GPH transporter	53.48837	75.5102	3.545833	0.00537	1
yjjQ		putative DNA-binding transcriptional regulator	90.69767	32.65306	4.727273	0.01028	1
group_5095	yjjQ	putative DNA-binding transcriptional regulator	9.302326	67.34694	0.211538	0.01028	1
insD		Qin prophage; predicted transposase	0	85.71429	0	0.013484	1
group_6273	yfjR	CP4-57 prophage; predicted DNA-binding transcriptional regulator	0	85.71429	0	0.013484	1
group_7475		hypothetical protein	0	85.71429	0	0.013484	1
group_6448		hypothetical protein	2.325581	81.63265	0.10582	0.017489	1
group_10025		hypothetical protein	51.16279	73.46939	2.901099	0.018751	1
yehY		YehW/YehX/YehY/YehZ ABC transporter	100	12.2449	inf	0.02816	1
group_1864	mprA_3	MprA-CCCP	0	87.7551	0	0.02816	1
group_5711	sdsR	SdsRQP multidrug efflux transport system - predicted membrane fusion protein	0	87.7551	0	0.02816	1
group_12032		hypothetical protein	46.51163	75.5102	2.681159	0.030775	1
group_4965	gspD_1	putative protein secretion protein for export	2.325581	83.67347	0.122024	0.033514	1
yicI		alpha-xylosidase	44.18605	77.55102	2.734848	0.043964	1

Appendix 6.11 Final gene target list.

Isolates with the best sensitivity/specificity and lowest p value were selected for the final candidate list to be screened against available K12 mutants.

Gene Name	Annotation	Association	Naïve p value
eutJ	putative chaperonin, ethanolamine utilization protein	Less than 1000pg IL-8	0.040671398
ynbC	putative hydrolase	Low IL-8 < 500	6.24E-06
gmm	GDP-mannose mannosyl hydrolase	Mortality	0.03379756
yhgE	putative transport protein	Mortality	0.000974849
cpsB	mannose-1-phosphate guanylyltransferase	Mortality	0.009303852
yhdJ_2	DNA adenine methyltransferase	Nosocomial	0.019549098
allD	ureidoglycolate dehydrogenase	Nosocomial	0.028352288
aat	leucyl, phenylalanyl-tRNA-protein transferase	Nosocomial	0.044836124
group_11805	hypothetical protein	Abdominal	0.00187874
glcA	glycolate / lactate:H+ symporter	Abdominal	0.042128077
yfaA	putative protein	Abdominal	0.027011704
group_1360	hypothetical protein	Abdominal vs other bacteraemia's	0.032642468
tufB	elongation factor Tu	Bacteraemia	4.63E-16
group_4156 (non-unique ybjE	putative transporter	Bacteraemia	8.40E-14
ybjE	putative transporter	Bacteraemia	1.03E-12
pemI	Antitoxin PemI	Bacteraemia	3.26E-09
bla	Beta-lactamase TEM	Bacteraemia	6.97E-09
group_9847(non-unique chpB)	ChpB toxin of the ChpB-ChpS toxin-antitoxin system	Bacteraemia	1.67E-08
tpd	34 kDa membrane antigen	Bacteraemia	4.20E-07
iutA	Ferric aerobactin receptor	Bacteraemia	8.88E-07

agp_2	3-phytase / glucose-1-phosphatase	Bacteraemia	1.67E-06
nqrC	Na(+)-translocating NADH-quinone reductase subunit C	Bacteraemia	2.40E-06
group_7134 (non-unique traA	Pilin	Bacteraemia	1.09E-05
yejF	peptide ABC transporter - ATP binding subunit	Serum resistance	0.010405101
group_7879 non unique (yejF)	peptide ABC transporter - ATP binding subunit	Serum resistance	0.010405101
allS	AllS transcriptional activator	Serum resistance	0.02117509
group16170 (iutA	Ferric aerobactin receptor	Urinary vs other bacteraemias	0.001165189
yqiG_2	putative membrane protein	Urinary vs other bacteraemias	0.016642661
tufB	elongation factor Tu	Urinary	1.81E-06
znuB_2	Zn2+ ABC transporter - membrane subunit	Urinary	4.57E-05
fhuC_2	iron (III) hydroxamate ABC transporter - ATP binding subunit	Urinary	0.000354559
yohF	putative oxidoreductase with NAD(P)-binding Rossmann-fold domain protein	Urinary	0.007551123
ydeP	acid resistance protein	Urinary	0.010069958
rusA_1	endodeoxyribonuclease RUS (Holliday junction resolvase)	Urinary	0.013152763
ydeQ	putative fimbrial-like adhesin protein	Urinary	0.038459844