

Manuscript Number: THELANCETRM-D-16-00682R1

Title: Enhanced resolution of the host-pathogen interaction in sepsis: new opportunities for -omic approaches.

Article Type: Review (Post author-enquiry)

Keywords: sepsis; infection; host-pathogen interactions; genomics; gene expression; gene expression profiling; gene expression regulation; epigenomics; proteomics; metabolomics; metagenomics; microbiota; high-throughput nucleotide sequencing

Corresponding Author: Professor Julian Charles Knight, MA MBChB DPhil

Corresponding Author's Institution: Oxford University

First Author: Cyndi Goh, MA MBBChir FFICM FRCA MRCP

Order of Authors: Cyndi Goh, MA MBBChir FFICM FRCA MRCP; Julian Charles Knight, MA MBChB DPhil

Manuscript Region of Origin: UNITED KINGDOM

Abstract: Progress in sepsis research has been severely hampered by a heterogeneous disease phenotype, limiting interpretation of clinical trials and development of effective therapeutic interventions. Application of omics-based methodologies is advancing understanding of the dysregulated host immune response to infection in sepsis. However the frequently elusive nature of the infecting organism in sepsis has limited efforts to understand the impact of heterogeneity involving the pathogen. Recent advances in nucleic acid sequencing based pathogen analysis provide the opportunity for more accurate and comprehensive microbiological diagnosis. In this Review, we explore how better resolution of the host-pathogen interaction can significantly enhance, and in turn benefit from, current and future application of omics-based approaches to understanding the host response in sepsis. We illustrate this using recent work accounting for variation involving the pathogen. We propose there is a timely opportunity to further resolve sepsis heterogeneity by considering host-pathogen interactions, enabling progress towards a precision medicine approach.

## **Title**

Enhanced resolution of the host-pathogen interaction in sepsis: new opportunities for –omic approaches

## **Authors**

Cyndi Goh FFICM<sup>1</sup> and Julian C Knight DPhil<sup>1\*</sup>

## **Affiliations**

<sup>1</sup>Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford UK

\*Correspondence to: Julian Knight, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK; email [julian@well.ox.ac.uk](mailto:julian@well.ox.ac.uk); tel 01865 287651

## **Summary**

Progress in sepsis research has been severely hampered by a heterogeneous disease phenotype, limiting interpretation of clinical trials and development of effective therapeutic interventions. Application of omics-based methodologies is advancing understanding of the dysregulated host immune response to infection in sepsis. However the frequently elusive nature of the infecting organism in sepsis has limited efforts to understand the impact of heterogeneity involving the pathogen. Recent advances in nucleic acid sequencing based pathogen analysis provide the opportunity for more accurate and comprehensive microbiological diagnosis. In this Review, we explore how better resolution of the host-pathogen interaction can significantly enhance, and in turn benefit from, current and future application of omics-based approaches to understanding the host response in sepsis. We illustrate this using recent work accounting for variation involving the pathogen. We propose there is a timely opportunity to further resolve sepsis heterogeneity by considering host-pathogen interactions, enabling progress towards a precision medicine approach.

## **Key messages**

- The complexity of the sepsis phenotype can be observed at both a whole-organism level (e.g. disease course) and at a molecular level (e.g. gene expression patterns)
- Omics-based approaches have facilitated our understanding of this heterogeneous disease phenotype from a variety of complementary perspectives but are themselves currently limited by this heterogeneity
- Evidence suggests that unique characteristics attributable to individual pathogens impact on the nature and consequences of interaction with the host in sepsis but this is often not accounted for in omics-based studies of sepsis
- New developments in clinical microbiology provide unique opportunities beyond pathogen identification, enabling us to more accurately characterise the host-pathogen interaction

- Achieving our goal of personalised therapy for individuals with sepsis will require researchers and clinicians to integrate studies of both pathogen and host

## **Introduction**

Clinicians and researchers are confronted with an intriguing issue when considering the patient with sepsis, namely how much importance to attribute to the pathogen in the context of the dramatic dysregulated host immune response to infection. Given the current limitations of clinical microbiology and lack of specific treatment options for sepsis, it is unsurprising that haemodynamic resuscitation and supportive therapy for host organ systems are generally the focus after antimicrobial therapy has been instituted, with the pathogen having minimal impact on clinical decision-making beyond antimicrobial choice and source control. The new consensus definitions for sepsis published earlier this year rightly focus our attention on early recognition of organ dysfunction.<sup>1</sup> This is much needed, not just because early treatment saves lives but also since it is often this onset of organ dysfunction in the clinical setting that triggers a search for the underlying infection.<sup>2</sup>

Previous emphasis on the systemic inflammatory response syndrome (SIRS) may be partly responsible for scepticism regarding the unique role or relative importance of the pathogen in sepsis. There is a widely held view that the inflammatory response and corresponding “genomic storm” initiated by the pathogen is more similar than dissimilar to that triggered by severe trauma<sup>3</sup> or major surgery. Certainly, endogenous alarmins (released in response to trauma or tissue damage) can initiate innate and adaptive immune responses in the same way as exogenous pathogen-associated molecular patterns. However, we would argue that to simplify the host-pathogen interaction to a mere switching on of a systemic inflammatory response would be to ignore a richly complex and informative biological process. The coexistence of host and pathogen over millennia reflects an evolutionary phenomenon where each moulds the genetic diversity of the other. There are numerous examples of host

genomics influencing infection susceptibility,<sup>4</sup> the shaping of host genetic evolution by pathogens,<sup>5</sup> different patterns of host gene expression within the same disease phenotype,<sup>6</sup> pathogens evolving to escape host immune responses,<sup>7</sup> and infection with one pathogen leading to increased susceptibility to another pathogen.<sup>8</sup>

Observational studies describe interesting associations between basic host demographics and susceptibility to specific pathogen types; male sex and black ethnicity for example is associated with Gram-positive sepsis,<sup>9</sup> whilst the elderly demonstrate increased risk for Gram-negative sepsis.<sup>10</sup> Even the association between low socio-economic status and poor sepsis outcome may have deep biological roots; rhesus macaque monkeys of lower social ranking are observed to display chronic, pro-inflammatory patterns of gene expression which are heightened by infective stimuli.<sup>11</sup> These observations only scratch the surface of the way in which host, pathogen, environment and microbiome interact to contribute to the complex heterogeneity of the sepsis syndrome.<sup>12</sup>

In an effort to dissect sepsis heterogeneity, omics-based methods have applied high-throughput techniques to understanding the molecular mechanisms of dysregulated host immunity from the level of gene to biological phenotype (figure 1). Genomics,<sup>13,14</sup> transcriptomics,<sup>15,16</sup> epigenomics,<sup>17</sup> proteomics,<sup>18</sup> metabolomics,<sup>18</sup> and metagenomics<sup>19</sup> have contributed to our understanding of sepsis pathophysiology with evidence that these approaches enable increased resolution of the sepsis phenotype including, for example, definition of specific subgroups dependent on response state.<sup>15</sup> However, translation to the clinical setting remains elusive.

In this Review, we will argue that omics-based approaches as applied to sepsis will remain limited whilst such studies focus on the host to the exclusion of the pathogen. We will use examples from sepsis and other infectious diseases to illustrate where efforts to account for

the degree of heterogeneity involving the pathogen have revealed valuable insights into host biology. We will also emphasise the necessity of an improved understanding of the host-pathogen interaction to realising the aim of precision medicine in sepsis, identified by a recent Commission as a key priority for future sepsis care and research.<sup>12</sup>

### **New technologies in clinical microbiology provide opportunities for enhanced resolution**

A recent epidemiological study of nearly seven million patients admitted with sepsis in the USA estimated the incidence of culture negative sepsis to be 47%, with culture negativity identified as an independent predictor of mortality.<sup>20</sup> Obtaining a microbiological diagnosis is a key clinical priority, enabling effective antimicrobial therapy within the framework of responsible stewardship. Emerging new technologies bring the potential of more rapid and detailed resolution of microbiology in sepsis,<sup>21</sup> and can be classified into one of three main approaches (table 1): (a) PCR-based techniques, (b) mass spectrometry (MS) based methods, and (c) nucleic acid sequencing (next-generation sequencing, NGS) methods. There has been a focus on PCR and MS based techniques as applied to sepsis in the current literature,<sup>22,23</sup> these techniques have the potential to deliver clinical benefits in the short term. However, they are being increasingly overshadowed by NGS and the wealth of information it can deliver to the clinical setting.

Machines such as the bench-top Illumina MiSeq and higher-capacity HiSeq are now established platforms with the potential to provide high-resolution sequence information on the community of organisms present in a clinical specimen with or without preceding culture.<sup>24</sup> This metagenomic approach is particularly attractive in sepsis since NGS does not require prior expectation of a particular organism. The simultaneous detection of multiple pathogens in a sample is also advantageous; coinfection may be more relevant to sepsis than we currently appreciate and be significantly under diagnosed. In addition, the ability to

obtain whole-genome sequence data, enables detailed profiling of antimicrobial resistance and virulence factors, e.g. in *Staphylococcus aureus* infection,<sup>25</sup> where targeted approaches are limited. In the emerging threat of carbapenem-resistant Enterobacteriaceae, NGS enables identification of specific extended spectrum beta-lactamase producing isolates where phenotypic testing and targeted nucleic acid methods are limited.<sup>26</sup>

Newly developed platforms such as the Oxford Nanopore Technologies MinION, which weighs just 103 grams, are breaking the bounds of what was previously thought possible. The MinION has facilitated epidemiological outbreak surveillance of Ebola virus in West Africa<sup>27</sup> and *Salmonella* in a UK-based inpatient setting.<sup>28</sup> However, the high sequencing error rate (5-30%), particularly in the repetitive regions, remains a current limitation.

Although the challenges are not to be underestimated, we are beginning to see progress in translating the benefits of NGS to the clinical setting. Previously, laborious and expensive library preparation stages have been vital to making a clinical sample suitable for sequencing. However, automated processes are becoming increasingly commonplace and new technologies such as the MinION require less pre-processing. Computational tools developed specifically for clinicians also enable genotype-based species identification and determination of antimicrobial resistance profiles for pathogens such as *S. aureus*.<sup>25</sup> In our opinion, the interpretation of “positive” results remains one of the most significant challenges associated with NGS technologies that are highly sensitive to low pathogen burdens, especially in the context of sample contamination.<sup>29</sup> One strategy, adopted in the analysis of sequencing data from the plasma of ICU sepsis patients, involves assigning a sepsis indicating quantifier (SIQ) score to each identified microbe in a sample, to indicate the probability of it being a true positive finding (based on the number of reads mapping to the microbe’s reference genome in the patient sample relative to healthy individuals).<sup>30</sup> Another challenge is inferring causality, since the presence of nucleic acid corresponding to a particular species may be due to a non-viable organism or reflect the normal microbiome of

a particular environment. For example, systematic PCR testing of nasopharyngeal and/or bronchoalveolar lavage specimens from ICU patients admitted with pneumonia identified respiratory viruses in over a third of patients, the relevance of which remains uncertain.<sup>31</sup> Strategies to address this issue, which is not exclusive to NGS, are helpfully covered in another Review.<sup>32</sup>

## **Applying different omics-based approaches**

### Host genomics: associations dependent on pathogen

The surprising observation that premature mortality from severe infection is strongly heritable was documented in a noteworthy epidemiological study nearly three decades ago.<sup>33</sup> Since then, a number of genomic approaches have been applied to studying the association between host genetics and sepsis susceptibility or disease outcome, with often conflicting or underwhelming results. A systematic review<sup>34</sup> of 76 candidate gene studies in sepsis assessed the majority of studies to be of low to moderate quality. Using a Bayesian method, the authors calculated that 30 of the 32 single-nucleotide polymorphisms (SNPs) identified to be associated with sepsis had at least a 50% probability of being a false positive finding if the prior probability of a true association was set at a generous cut-off of 0.01.

We are aware of only two reported genome-wide association studies (GWAS) of sepsis outcome in the current literature,<sup>13,14</sup> each of which identified different, non-overlapping associations with 28-day survival. In the first GWAS,<sup>13</sup> the reported genome-wide significant association involved a SNP (rs4957796) located in an intron of the *Fps/Fes* related tyrosine kinase (*FER*) gene of chromosome 5. Notably, this association reached significance only when patients with sepsis from pneumonia were analysed independently from those with intra-abdominal infection ( $p$  combined =  $5.6 \times 10^{-8}$ ), with effect size increasing when bacterial cases of pneumonia were analysed separately (odds ratio reduced from 0.56 to 0.4, indicating increased effect size). The *FER* gene is implicated in sepsis-relevant pathways involving neutrophil chemotaxis and endothelial permeability; electroporation-mediated gene



therapy of a plasmid containing the protective SNP has been shown to improve survival in murine models of traumatic lung contusion and pneumonia.<sup>35</sup> However, two independent studies (a second GWAS<sup>14</sup> and a targeted validation study<sup>36</sup>) have failed to replicate this association involving the *FER* gene. The second GWAS<sup>14</sup> analysed a cohort of patients with a diverse range of aetiologies (lung, abdominal, urogenital, bone, soft tissue and wound infections) and identified a different association (rs117983287) localising to the vacuolar protein sorting 13 homolog A (*VPS13A*) gene on chromosome 9 ( $p=8.16 \times 10^{-8}$ ). The failure of replication between these GWAS, despite similar methods and outcome of interest, is most likely attributable to the different sepsis aetiologies studied in each. When the validation study<sup>36</sup> limited their analysis to only patients with pneumonia, their small sample size of 298 individuals resulted in low statistical power (49%) to detect an association at a significance level of 0.05. To our knowledge, no independent validations of the association between the *VPS13A* gene and sepsis outcome have yet been published.

Considerably better progress has been made in other infectious diseases, such as Hepatitis C virus infection where a GWAS-identified SNP (rs12879860) in the promoter region of the *IL28B* gene is associated with spontaneous viral clearance,<sup>37</sup> response to pegylated interferon and ribavirin,<sup>38</sup> and interferon-free regimes (e.g. faldaprevir, deleobuvir and ribavirin).<sup>39</sup> A similar application of pharmacogenomics to sepsis therapeutics is particularly appealing as heterogeneity in underlying host genetics may explain some of the overwhelming failure of over 100 phase two and three clinical trials of drugs for sepsis.<sup>40</sup> In a GWAS of treatment response in sepsis patients receiving drotrecogin alfa, a number of SNP combinations in biologically relevant genes were identified to be associated.<sup>41</sup> In the top “combination marker” of three SNPs, observed in 26% of the cohort, a dramatic absolute risk reduction in 28-day mortality of 41.7% was observed (compared with 6.1% in the whole cohort). This result needs to be treated with caution, given the absence of a replication cohort. Nevertheless, the principle underlying the study is an important one; subgroups of individuals with certain genotypes may benefit from therapies we have evaluated as non-

efficacious in a heterogeneous cohort. An improved understanding of the role of host genomics in sepsis has the potential to enable more intelligent clinical trial design and repurpose existing therapy.

#### Host transcriptomic signatures of different classes of pathogen

The aim of improving diagnosis and disease stratification through the definition of sub-phenotypes with distinct host response features has been limited by challenges in dissecting the various sources of disease heterogeneity. As such, recent efforts to apply studies of host transcriptomics to sepsis biomarker development have not yet translated to validation in the clinical setting.

Studies comparing host transcriptomic signatures between different microbiological aetiologies have yielded varying results (table 2). One early study comparing gene expression in ICU patients with Gram-positive and Gram-negative sepsis did not identify significant differences.<sup>42</sup> This is perhaps unsurprising given the small patient numbers (18 Gram positive patients, 25 Gram negative patients) and diverse range of sepsis aetiologies studied. For example, the Gram-negative group included many infection types (respiratory, intra-abdominal, urinary, CNS, other) from ten different bacterial species. It is possible the authors of this study may have observed a unique transcriptome profile if they had refined their analysis to a particular infection (e.g. respiratory) or pathogen type.

A similar approach applied to genome-wide gene expression profiling of peripheral whole blood in sepsis from acute respiratory illness enabled definition of three gene sets which classified patients into bacterial, viral, co-infected and non-infected aetiologies with an accuracy superior to procalcitonin (87% vs. 78%;  $p < 0.03$ ).<sup>6</sup> In critically ill sepsis patients with proven *E. Coli* bacteraemia and *Candida* fungaemia, analysis of peripheral blood gene expression microarray data demonstrates that the large majority of genes (5,977) with significant differential expression from healthy volunteers (adjusted p value,  $< 0.05$ ) are

common to both pathologies<sup>16</sup> with genes encoding products within pathways relevant to glycolysis and oxidative metabolism upregulated in both groups. However, a unique transcriptional response specific to each pathogen was also seen (*E. coli* bacteraemia, 1718 genes; *Candida* fungaemia, 830 genes) (figure 2). Multicohort analysis of publicly available transcriptomic data from patients with acute infection (including, but not exclusively sepsis), provide further evidence for distinct gene expression patterns depending on pathogen.<sup>43</sup> Analysis of datasets from eight cohorts of patients enabled derivation of a seven-gene set, which discriminated bacterial from viral infection when applied to 30 independent cohorts. In combination with a separate 11-gene set that discriminates infection from no infection,<sup>44</sup> the authors derived an antibiotics decision model which had 94% sensitivity and 59.8% specificity for bacterial infection. These observations highlight the importance of accounting for underlying microbiological heterogeneity in studies of host gene expression in sepsis, such that any differences between study groups can be clearly distinguished from the signature attributable to the pathogen.

The bioinformatic challenges of comparing results between studies should not be underestimated and may also contribute to the inconsistent results observed in table 2. Even analyses performed using the same platform require careful normalisation to account for batch-to-batch variation. With RNASeq now superseding microarray technologies, comparing results obtained from different technologies is not straightforward. Guidelines provided by collaborative groups such as ENCODE (Encyclopedia of DNA Elements) provide valuable recommendations for standardisation of experimental and analytic methods.<sup>45</sup> Whilst these new technologies provide unique opportunities, such as the ability to study both human and pathogen gene expression simultaneously,<sup>46</sup> the academic and clinical community will need to evaluate results arising from these studies with a discerning eye.

### Regulation of gene expression: the impact of the pathogen

The regulation of gene expression is a dynamic process, involving a complex interplay of highly coordinated mechanisms from transcription to post-translational stages. A comprehensive review of published GWAS highlighted that the majority of trait/disease-associated SNPs involved non-protein coding regions of the genome, with 88% located in intergenic or intronic regions.<sup>47</sup> These genetic variants are believed to contribute to the heritability of inter-individual variation in gene expression. Indeed, SNPs associated with complex traits are significantly more likely to be expression quantitative trait loci (eQTLs) than frequency-matched SNPs from GWAS platforms.<sup>48</sup> Given that eQTLs demonstrate strong heritability,<sup>49</sup> these regulatory genetic variants may reflect the major selective pressure that pathogens have exerted on human genetics.

The application of a quantitative trait loci mapping approach to define the association of a genetic variant with an intermediate phenotype (e.g. levels of a transcript (eQTL), protein or metabolite) has convincingly demonstrated the importance of cellular context and environment on the activity of a regulatory variant. In vitro studies of human monocytes exposed to biological stimuli relevant to bacterial (lipopolysaccharide) and mycobacterial/viral (interferon-gamma) infections show that not only is exposure to these stimuli necessary to demonstrate the association of SNPs on the expression of innate immune system related genes but also that the direction of effect on gene expression may differ between different stimuli.<sup>50</sup> A similar approach applied to human primary dendritic cells exposed to mycobacterium tuberculosis identified 198 loci associated with gene expression that were not observed in unstimulated cells, reflecting the effect of host-pathogen interaction.<sup>51</sup> These observations are highly relevant to studies of patients with sepsis where we must consider the biological context set by a particular infecting pathogen and its influence on underlying host genomic modulation of gene expression. In a study of ICU patients with sepsis from community acquired pneumonia (CAP), there was evidence of eQTL specific to sepsis response state of patients and of significant enrichment for eQTL in

gene sets associated with viral respiratory infection, even though only 9% of patients within this cohort had CAP of confirmed viral aetiology.<sup>15</sup> This analysis would have been enhanced by higher resolution microbiological phenotyping of the cohort.

The relevance of non-coding RNAs in the regulation of gene expression has become increasingly apparent, with microRNAs (miRNAs) representing a focus of attention for the identification of potential biomarkers in sepsis. However, conflicting results represent a significant challenge for the interpretation and application of these studies. For example, significantly reduced miR-223 levels were observed in one cohort of patients with sepsis compared with patients with SIRS and healthy individuals<sup>52</sup> whilst elevated miR-223 levels were observed in another cohort of sepsis patients compared with healthy individuals.<sup>53</sup> In a third study, no differences in miR-223 levels were found between individuals with and without sepsis.<sup>54</sup> The authors of the latter study attribute these conflicting results to important differences in methods of normalisation. However, we would argue that differences in infection aetiology deserve consideration. In these three studies, the first included only surgical ICU patients with intra-abdominal and trauma-related infection whilst the other two recruited patients with respiratory infection and a broader range of infective pathologies. Even within the context of different strains of the same virus, significant differences in miR-223 expression can be observed. In mice infected with a reconstructed pandemic 1918 H1N1 influenza A virus, distinct patterns of lung tissue microRNA expression were observed compared to those infected with a seasonal H1N1 influenza A virus (A/Texas/36/91). Specifically, levels of miR-223 were nearly 3 times higher with the pandemic strain compared to the seasonal strain.<sup>55</sup>

#### Insights from proteomics and metabolomics

Large-scale analysis of the protein and metabolite composition of biological samples afforded by advances in MS based methods presents an attractive and versatile opportunity for sepsis research. Since there is no direct relationship between mRNA expression and

protein/metabolite levels, these approaches have the potential to provide unique functional insights beyond that afforded by genomics and transcriptomics.

In a comprehensive study of patients presenting to the emergency department with suspected community-acquired sepsis, unique plasma proteomic and metabolomic profiles were seen in sepsis survivors, sepsis nonsurvivors and individuals with a non-infective SIRS.<sup>18</sup> Nonsurvivors demonstrated profiles consistent with impaired mitochondrial fatty acid  $\beta$ -oxidation and the authors of this study developed a prognostic logistic regression model based on seven parameters (four carnitine esters, lactate, age, and haematocrit). This model was validated in two separate cohorts and successfully classified sepsis survivors and nonsurvivors with an accuracy of 85%, superior to that of lactate, Sequential Organ Failure Assessment (SOFA), or Acute Physiology and Chronic Health Evaluation II (APACHE II) scores. Interestingly, no major differences were seen between the plasma metabolome and proteome of patients with sepsis due to *Streptococcus pneumoniae*, *S. aureus* and *E. coli*. The authors postulated that this might have been due to heterogeneity with respect to infection site and the possibility that subtle differences were overwhelmed by a generalised septic response. Indeed, a study performing ELISA-based quantification of a more limited panel of eleven plasma cytokines observed higher IL-1 $\beta$ , IL-6, and IL-18 concentrations in patients with Gram-positive sepsis when compared to Gram-negative sepsis.<sup>56</sup> In another study recruiting a more homogeneous cohort of Gambian children with pneumonia, an MS based proteomic analysis identified 42 proteins which differentiated severe from non-severe pneumonia and non-severe pneumonia from controls.<sup>57</sup> One of these proteins was lipocalin-2, which discriminates pneumonia of probable bacterial from viral aetiology. Children with plasma concentrations of lipocalin-2 over 163 ng/ml were nine times more likely to have a positive blood culture with a clinically significant isolate.

Another important consideration in evaluating proteomic/metabolomic studies is choice of technology. MS, which is less sensitive for the measurement of smaller, lower abundance

proteins was used in the study of emergency department patients.<sup>18</sup> When the proteome of premature neonates with infection was analysed utilising antibody microarrays as an alternative technique, differential expression of eight serum proteins was observed between patients with cultures positive for coagulase-negative Staphylococci, *E. coli*, *Candida* and group B Streptococci.<sup>58</sup> Proteomic and metabolomic studies of sepsis represent a rich, untapped source of potential disease biomarkers for the clinical setting.

#### Pathogen-dependent modulation of host epigenetic regulation

From an evolutionary biology perspective, the co-existence of pathogen and host is a major selective pressure on the genetic diversity of both organisms.<sup>5</sup> Whilst random point mutations in DNA sequence are a key source of diversity as substrate for this, the role of epigenetic changes in contributing to more rapid changes to organism phenotype is being increasingly recognised.<sup>59</sup> Both viruses<sup>60</sup> and bacteria<sup>61</sup> are able to effectively induce epigenetic changes in humans, modulating the biological interaction between pathogen and host with significant consequences to disease course.

For example, the H3N2 influenza A virus possesses a histone-like sequence in its non-structural protein 1 (NS1) tail, allowing it to interact with the host epigenome.<sup>62</sup> The binding of the NS1 protein to the human polymerase associated factor 1 (PAF1) transcription elongation complex allows the virus to target sites of actively transcribed antiviral genes, suppressing the antiviral response. Interestingly, the H1N1 influenza A virus does not possess this histone-like NS1 tail, which may explain the varying disease phenotypes seen between different strains.

There are few studies of epigenetics in sepsis but significant corroborative evidence to support this as a promising approach by which to advance our understanding of the disease. In an animal model of concomitant acute lung injury and sepsis, anaesthetised mice were subjected to pulmonary aspiration of a *S. aureus* culture and mechanical ventilation.<sup>17</sup> After

only six hours, there was decreased expression of the angiogenic genes *ANGPT1*, *TEK*, and *KDR* genes in the lung, kidney, and liver. Chromatin immunoprecipitation (ChIP) assays demonstrating a decrease in RNA Polymerase II levels and histone deacetylation at these genes provide evidence for the role of epigenetic changes in contributing to sepsis-induced endothelial dysfunction.

In the previously mentioned study of gene expression in adult ICU patients with sepsis from CAP,<sup>15</sup> the locations of the observed sepsis eQTL showed significant overlap with epigenetic marks observed in monocytes stimulated with bacterial lipopolysaccharide. This included DNase I hypersensitive sites (regions of chromatin where DNase I activity results in the DNA being accessible to transcription factor binding) and histone marks associated with enhancer and promoter regions (H3K27ac, H3K4me1, H3K4me3).

In summary, studies in animal models and humans with sepsis suggest a strong role for epigenetic changes in contributing to sepsis pathophysiology. However, for these to translate to patient benefit (e.g. identification of novel therapeutic targets), these epigenetic mechanisms will need to be dissected further. We propose that studies that take account of specific infection types will be a powerful way by which researchers can achieve this aim.

#### Metagenomics reveals the impact of the microbiome

Thus far, we have argued that features unique to individual pathogens and their host interact to result in the heterogeneous sepsis immune response. Yet, to consider this host-pathogen interaction in isolation from the microbiome would be to disregard the rich microbial context in which this interaction occurs. The term “pathobiome” helpfully emphasises this concept - that a microorganism’s pathogenicity depends not just on its specific virulence factors, but also on host factors, environmental factors and the microbial community of which it is a part (figure 1).<sup>63</sup>



The increasing affordability of high-throughput sequencing has enabled advances in metagenomics, the culture-independent, sequence-based analysis of genetic material recovered from the microbial community of a sample. Of relevance to sepsis is the way in which one microorganism (or a community of microorganisms) may influence the interaction of another microorganism with the host. This synergism between multiple pathogens is well-recognised in influenza virus infection, where adherence of *S. pneumoniae* to the respiratory epithelium is facilitated by viral neuraminidase.<sup>64</sup> However, the increased sensitivity and unbiased nature of metagenomics now enables us to consider synergism beyond a handful of organisms, by revealing new mechanisms through which the whole microbiome influences host susceptibility to infection.

Mucosal surfaces represent a critical interface for the microbiome-host interaction; microorganisms previously considered to be mere commensals act both directly and indirectly at sites such as the intestinal and respiratory epithelium to protect the host from potentially pathogenic organisms. Directly, the intact microbiome provides a “barrier effect” by limiting the supply of essential resources at these sites. Indirectly, the microbiome interacts with immune cells at mucosal surfaces leading to modulation of key innate and adaptive immune pathways.<sup>65</sup> In murine models, the bacterial component of the gut microbiome has been shown to play a critical role in enabling persistent norovirus infection by limiting the efficacy of interferon-gamma mediated innate immunity.<sup>66</sup> Adaptive immunity is also mediated by the microbiome, with the response to influenza virus shown to depend critically on a healthy lung bacterial microbiome. Mice receiving a four-week course of oral antibiotics developed marked dysbiosis of the lung microbiome, leading to defective CD4 T-cell, CD8 T-cell and B-cell mediated immunity to subsequent intranasal influenza virus challenge.<sup>67</sup> This suggests that antibiotics may be not just unnecessary, but actually harmful in cases of viral respiratory infection with potentially critical implications in patients with chronic lung conditions, who frequently receive antibiotics as prophylaxis or treatment of exacerbations. For example, quorum sensing enables bacterial populations to both modulate

expression of virulence genes in response to their environment and thus allow isolates of the same organism to vary in their interaction with the host; and influence microbe-microbe cross-talk and thus host interactions as illustrated by introduction of a mutant *E. coli* strain able to overproduce the quorum sensing signalling molecule autoinducer 2 (AI-2) to the gut of antibiotic-treated mice, restoring the depleted Firmicutes phyla in the gut microbiota.<sup>68</sup>

Large-scale endeavours such as the NIH Human Microbiome Project have revealed the extent to which the microbiome differs between healthy individuals.<sup>69</sup> This diversity is often lost in critical illnesses such as sepsis where both disrupted host pathophysiology and clinical interventions to manage the underlying condition result in striking changes to the microbiome.<sup>70</sup> There is convincing evidence to suggest that host genetic variants critically interact with the microbiome to result in the development of immune-related disease; K/BxN transgenic mice do not develop inflammatory arthritis in a germ-free environment since segmented filamentous bacteria in the gut are essential to Th17 cell activation and subsequent joint inflammation.<sup>71</sup> This may have implications in sepsis where the relevance of host genetics to disease may be better understood in the context of interindividual differences in the microbiome.

There is the potential for metagenomic analysis of the microbiome to yield prognostic information in sepsis: in conditions as different as bronchiectasis<sup>72</sup> and metabolic syndrome,<sup>73</sup> the microbiome has been shown to correlate with disease phenotype and outcome. Metagenome-wide association studies (MWAS) which identify associations between microbial genes and disease traits have found that restoration of dysbiosis in the dental microbiome is associated with good response to disease-modifying antirheumatic drugs in rheumatoid arthritis.<sup>74</sup> It is not inconceivable that microbiome-based tests could similarly be identified as biomarkers for early diagnosis, patient stratification and evaluation of treatment response in sepsis. In addition, the microbiome represents an attractive therapeutic target in sepsis with murine models showing a reduction in circulating aged

neutrophils following antibiotic-depletion of the gut microbiota, with corresponding improvements in survival from endotoxin-induced septic shock.<sup>19</sup> As an important source of heterogeneity in sepsis, characterisation of the microbiome and integration of this “second genome”<sup>75</sup> into other omics-based approaches should be a key priority in further sepsis research.

## **Conclusion and future opportunities**

Having been confronted by a relentless succession of negative trials, the critical care community is all too aware of the need to shift from syndromes, pure organ support and all-inclusive trials to specific disease phenotypes, targeted treatment and well-defined study populations. New definitions for the archetypal ICU syndromes of acute respiratory distress syndrome and sepsis represent an excellent starting point but disease heterogeneity needs to be considered at all stages from bench to bedside to enable omics-based studies to translate to meaningful patient benefit (panel 1).

In this Review we have stressed the importance of careful consideration of the host-pathogen interaction in applying omics-based approaches to sepsis. The number of factors influencing the host sepsis response is considerable (figure 1), accounting for these extensive and dynamically changing variables will require collaborative efforts involving biorepositories with both detailed clinical and biological phenotyping, such as that established by the Molecular Diagnosis and Risk Stratification of Sepsis (MARS) consortium, which has recruited approximately 7,000 ICU patients with sepsis to date.<sup>76</sup>

Although consideration of the host-pathogen interaction involves far more than pathogen species identification, incorporating more accurate microbiological phenotyping into sepsis research is a good place from which to start. When defining study inclusion criteria, there is a need to move away from all-cause sepsis to better phenotyped groups. As a crude starting point, study cohorts defined on the basis of anatomical location of infection would helpfully

restrict the number of potentially causative microorganisms. However, more effective use of clinical microbiology, new microbiological diagnostic tools,<sup>21</sup> and disease biomarkers<sup>77</sup> would further enhance definition of sepsis study cohorts. Although this may limit the number of individuals it is possible to recruit to a study, stricter inclusion criteria may actually enable improved sensitivity of detection for previously diluted biological signals.

There are also important ramifications for the timing of patient recruitment and biological sample collection. The initial stages of sepsis are a key disease time-point when pathogen burden is at its highest, prior to the successful activation of host defence mechanisms and the institution of antimicrobial therapy. Thus, any delays in recruitment may limit our ability to identify the impact of the pathogen on -omics profiles. Later recruitment may also suffer from the confounders of secondary and nosocomial infections, which are commonly seen in the immunosuppressed stages of sepsis. Evaluation of -omic datasets between patients, who have varying disease courses and present at different stages remains a significant challenge (figure 1); endeavours to establish serial sampling and biorepositories which initiate sample collection in at-risk patients prior to clinical presentation of sepsis are to be encouraged.

The importance of well-curated clinical microbiology data is also paramount. From our experience with the UK Genomic Advances in Sepsis (GAinS) study<sup>13,15</sup> (which has recruited over two thousand patients with severe sepsis from CAP or faecal peritonitis), interpreting the clinical relevance of a documented positive culture result is difficult without details of timing, type of culture and any negative cultures obtained. However, our ability to really propel our understanding of the host-pathogen interaction in sepsis will be necessarily limited if sepsis research does not incorporate new developments to microbiological diagnostics. In the immediate future, the use of techniques such as electrospray ionisation-mass spectrometry<sup>22</sup> and multiplex PCR should be more routinely incorporated into both clinical and research settings. However, we would strongly argue that high-throughput

sequencing techniques<sup>24</sup> remain an enormous potential in the medium to long-term. Comparative studies of the various techniques in sepsis are urgently needed.

Finally, we would advocate the prioritisation of studies integrating metagenomic profiling of the microbiome in sepsis patients with other omics-based approaches, since it is within this important context that the host-pathogen interaction in sepsis occurs. The microbiome has the potential to emerge not just as a biomarker of disease phenotype in sepsis, but also as a therapeutic target.

There are many sources and manifestations of disease heterogeneity in sepsis, of which the pathogen plays a previously underappreciated part. We now have the tools to more successfully understand the relevance of individual pathogens for the host immune response in sepsis. New efforts to apply omics-based approaches to sepsis would be greatly enhanced by more careful consideration of the role of the pathogen in disease phenotype.

### **Search strategy and selection criteria**

References for this review were identified through searches of PubMed for articles published from January, 1971, to September, 2016, by use of the MeSH terms “Sepsis”, “Infection”, “Host-Pathogen Interactions”, “Genomics”, “Gene Expression”, “Gene Expression Profiling”, “Gene Expression Regulation”, “Epigenomics”, “Proteomics”, “Metabolomics”, “Metagenomics”, “Microbiota” and “High-Throughput Nucleotide Sequencing”. Articles resulting from these searches and relevant references cited in those articles were reviewed. Only English-language articles were included.

## **Contributors**

CG and JK contributed equally to the conception and planning of this Review, literature search, and writing of the manuscript.

## **Declaration of interests**

We declare no competing interests.

## **Acknowledgements**

CG was supported by a Medical Research Council/British Journal of Anaesthesia/Royal College of Anaesthetists Fellowship (MR/M018725/1) during the conduct of this Review. JCK is supported by grants from the Wellcome Trust; European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) (ERC Grant agreement no. 281824); and Arthritis Research UK (20773). Both authors are supported by the Wellcome Trust (090532/Z/09/Z [core facilities Wellcome Trust Centre for Human Genetics including High-Throughput Genomics Group]). The organizations funding this study had no role in the design or conduct, preparation, review, or approval of the manuscript.

## References

- 1 Singer M, Deutschman CS, Seymour CW, *et al.* The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016; **315**: 801–10.
- 2 Vincent JL, Mira JP, Antonelli M. Sepsis: Older and newer concepts. *Lancet Respir Med* 2016; **4**: 237–40.
- 3 Xiao W, Mindrinos MN, Seok J, *et al.* A genomic storm in critically injured humans. *J Exp Med* 2011; **208**: 2581–90.
- 4 Allison AC. Protection afforded by sickle-cell trait against subtertian malarial infection. *Br Med J* 1954; **1**: 290–4.
- 5 Sironi M, Cagliani R, Forni D, Clerici M. Evolutionary insights into host–pathogen interactions from mammalian sequence data. *Nat Rev Genet* 2015; **16**: 224–36.
- 6 Tsalik EL, Henao R, Nichols M, *et al.* Host gene expression classifiers diagnose acute respiratory illness etiology. *Sci Transl Med* 2016; **8**: 322ra11–322ra11.
- 7 Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. Evolution and ecology of influenza A viruses. *Microbiol Rev* 1992; **56**: 152–79.
- 8 Peltola VT, Murti KG, McCullers JA. Influenza virus neuraminidase contributes to secondary bacterial pneumonia. *J Infect Dis* 2005; **192**: 249–57.
- 9 Esper AM, Moss M, Lewis CA, Nisbet R, Mannino DM, Martin GS. The role of infection and comorbidity: Factors that influence disparities in sepsis. *Crit Care Med* 2006; **34**: 2576–82.
- 10 Martin GS, Mannino DM, Moss M. The effect of age on the development and outcome of adult sepsis. *Crit Care Med* 2006; **34**: 15–21.
- 11 Snyder-Mackler N, Sanz J, Kohn JN, *et al.* Social status alters immune regulation and response to infection in macaques. *Science* 2016; **354**: 1041–5.
- 12 Cohen J, Vincent JL, Adhikari NKJ, *et al.* Sepsis: A roadmap for future research. *Lancet Infect Dis* 2015; **15**: 581–614.
- 13 Rautanen A, Mills TC, Gordon AC, *et al.* Genome-wide association study of survival from sepsis due to pneumonia: An observational cohort study. *Lancet Respir Med*

- 2015; **3**: 53–60.
- 14 Scherag A, Schoneweck F, Kesselmeier M, *et al.* Genetic Factors of the Disease Course after Sepsis: A Genome-Wide Study for 28Day Mortality. *EBioMedicine* 2016; published online Sept. DOI:10.1016/j.ebiom.2016.08.043.
  - 15 Davenport EE, Burnham KL, Radhakrishnan J, *et al.* Genomic landscape of the individual host response and outcomes in sepsis: A prospective cohort study. *Lancet Respir Med* 2016; **4**: 259–71.
  - 16 Cheng S-C, Scicluna BP, Arts RJW, *et al.* Broad defects in the energy metabolism of leukocytes underlie immunoparalysis in sepsis. *Nat Immunol* 2016; **17**: 406–13.
  - 17 Bomsztyk K, Mar D, An D, *et al.* Experimental acute lung injury induces multi-organ epigenetic modifications in key angiogenic genes implicated in sepsis-associated endothelial dysfunction. *Crit Care* 2015; **19**: 225.
  - 18 Langley RJ, Tsalik EL, van Velkinburgh JC, *et al.* An integrated clinico-metabolomic model improves prediction of death in sepsis. *Sci Transl Med* 2013; **5**: 195ra95.
  - 19 Zhang D, Chen G, Manwani D, *et al.* Neutrophil ageing is regulated by the microbiome. *Nature* 2015; **525**: 528–32.
  - 20 Gupta S, Sakhuja A, Kumar G, McGrath E, Nanchal RS, Kashani KB. Culture Negative Severe Sepsis – Nationwide Trends and Outcomes. *Chest* DOI:<http://dx.doi.org/10.1016/j.chest.2016.08.1460>.
  - 21 Buchan BW, Ledeboer NA. Emerging technologies for the clinical microbiology laboratory. *Clin Microbiol Rev* 2014; **27**: 783–822.
  - 22 Vincent J-L, Brealey D, Libert N, *et al.* Rapid Diagnosis of Infection in the Critically Ill, a Multicenter Study of Molecular Detection in Bloodstream Infections, Pneumonia, and Sterile Site Infections. *Crit Care Med* 2015; **43**: 2283–91.
  - 23 Personne Y, Ozongwu C, Platt G, *et al.* 'Sample-in, answer-out'? Evaluation and comprehensive analysis of the Unyvero P50 pneumonia assay. *Diagn Microbiol Infect Dis* 2016; **86**: 5–10.
  - 24 Frey KG, Herrera-Galeano J, Redden CL, *et al.* Comparison of three next-generation



- sequencing platforms for metagenomic sequencing and identification of pathogens in blood. *BMC Genomics* 2014; **15**: 96.
- 25 Bradley P, Gordon NC, Walker TM, *et al.* Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *Nat Commun* 2015; **6**: 10063.
  - 26 Pitout JD, Laupland KB. Extended-spectrum  $\beta$ -lactamase-producing Enterobacteriaceae: an emerging public-health concern. *Lancet Infect Dis* 2008; **8**: 159–66.
  - 27 Quick J, Loman NJ, Duraffour S, *et al.* Real-time, portable genome sequencing for Ebola surveillance. *Nature* 2016; **530**: 228–32.
  - 28 Quick J, Ashton P, Calus S, *et al.* Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of *Salmonella*. *Genome Biol* 2015; **16**: 114.
  - 29 Salter SJ, Cox MJ, Turek EM, *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 2014; **12**: 87.
  - 30 Grumaz S, Stevens P, Grumaz C, *et al.* Next-generation sequencing diagnostics of bacteremia in septic patients. *Genome Med* 2016; **8**: 73.
  - 31 Choi S-H, Hong S-B, Ko G-B, *et al.* Viral infection in patients with severe pneumonia requiring intensive care unit admission. *Am J Respir Crit Care Med* 2012; **186**: 325–32.
  - 32 Lipkin WI. Microbe hunting. *Microbiol Mol Biol Rev* 2010; **74**: 363–77.
  - 33 Sørensen TI, Nielsen GG, Andersen PK, Teasdale TW. Genetic and environmental influences on premature death in adult adoptees. *N Engl J Med* 1988; **318**: 727–32.
  - 34 Clark MF, Baudouin SV. A systematic review of the quality of genetic association studies in human sepsis. *Intensive Care Med* 2006; **32**: 1706–12.
  - 35 Dolgachev VA, Goldberg R, Suresh M V, *et al.* Electroporation-mediated delivery of the FER gene in the resolution of trauma-related fatal pneumonia. *Gene Ther* 2016; published online Aug. DOI:10.1038/gt.2016.58.
  - 36 Schönebeck F, Kuhnt E, Scholz M, Brunkhorst FM, Scherag A. Common genomic

- variation in the FER gene: useful to stratify patients with sepsis due to pneumonia? *Intensive Care Med* 2015; **41**: 1379–81.
- 37 Thomas DL, Thio CL, Martin MP, *et al.* Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature* 2009; **461**: 798–801.
  - 38 Ge D, Fellay J, Thompson AJ, Simon JS, Shianna K V. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 2009.  
DOI:10.1038/nature08309.
  - 39 Zeuzem S, Soriano V, Asselah T, *et al.* Faldaprevir and Deleobuvir for HCV Genotype 1 Infection. *N Engl J Med* 2013; **369**: 630–9.
  - 40 Marshall JC. Why have clinical trials in sepsis failed? *Trends Mol Med* 2014; **20**: 195–203.
  - 41 Man M, Close SL, Shaw AD, *et al.* Beyond single-marker analyses: mining whole genome scans for insights into treatment responses in severe sepsis. *Pharmacogenomics J* 2013; **13**: 218–26.
  - 42 Tang BMP, McLean AS, Dawes IW, Huang SJ, Cowley MJ, Lin RCY. Gene-expression profiling of gram-positive and gram-negative sepsis in critically ill patients. *Crit Care Med* 2008; **36**: 1125–8.
  - 43 Sweeney TE, Wong HR, Khatri P. Robust classification of bacterial and viral infections via integrated host gene expression diagnostics. *Sci Transl Med* 2016; **8**: 346ra91.
  - 44 Sweeney TE, Shidham A, Wong HR, Khatri P. A comprehensive time-course-based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. *Sci Transl Med* 2015; **7**: 287ra71.
  - 45 ENCODE Project. Standards, Guidelines and Best Practices for RNA-Seq.  
[https://www.encodeproject.org/documents/91494746-0ffe-4931-b219-a09802ce1cfa/@@download/attachment/RNA\\_standards\\_v1\\_2011\\_May.pdf](https://www.encodeproject.org/documents/91494746-0ffe-4931-b219-a09802ce1cfa/@@download/attachment/RNA_standards_v1_2011_May.pdf)  
(accessed Dec 12, 2016).
  - 46 Westermann AJ, Gorski S a., Vogel J. Dual RNA-seq of pathogen and host. *Nat Rev Microbiol* 2012; **10**: 618–30.

- 47 Hindorff LA, Sethupathy P, Junkins HA, *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci* 2009; **106**: 9362–7.
- 48 Nicolae DL, Gamazon E, Zhang W, Duan S, Eileen Dolan M, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: Annotation to enhance discovery from GWAS. *PLoS Genet* 2010; **6**. DOI:10.1371/journal.pgen.1000888.
- 49 Wright FA, Sullivan PF, Brooks AI, *et al.* Heritability and genomics of gene expression in peripheral blood. *Nat Genet* 2014; **46**: 430–7.
- 50 Fairfax BP, Humburg P, Makino S, *et al.* Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* 2014; **343**: 1246949.
- 51 Barreiro LB, Tailleux L, Pai AA, Gicquel B, Marioni JC, Gilad Y. Deciphering the genetic architecture of variation in the immune response to Mycobacterium tuberculosis infection. *Proc Natl Acad Sci U S A* 2011; **109**: 1204–9.
- 52 Wang J feng, Yu M li, Yu G, *et al.* Serum miR-146a and miR-223 as potential new biomarkers for sepsis. *Biochem Biophys Res Commun* 2010; **394**: 184–8.
- 53 Wang H, Zhang P, Chen W, Feng D, Jia Y, Xie L. Four serum microRNAs identified as diagnostic biomarkers of sepsis. *J Trauma Acute Care Surg* 2012; **73**: 850–4.
- 54 Benz F, Tacke F, Luedde M, *et al.* Circulating MicroRNA-223 serum levels do not predict sepsis or survival in patients with critical illness. *Dis Markers* 2015. DOI:10.1155/2015/384208.
- 55 Li Y, Chan EY, Li J, *et al.* MicroRNA expression and virulence in pandemic influenza virus-infected mice. *J Virol* 2010; **84**: 3023–32.
- 56 Feezor RJ, Oberholzer C, Baker H V, *et al.* Molecular characterization of the acute inflammatory response to infections with gram-negative versus gram-positive bacteria. *Infect Immun* 2003; **71**: 5803–13.
- 57 Huang H, Ideh RC, Gitau E, *et al.* Discovery and validation of biomarkers to guide clinical management of pneumonia in african children. *Clin Infect Dis* 2014; **58**: 1707–15.

- 58 Kingsmore SF, Kennedy N, Halliday HL, *et al.* Identification of diagnostic biomarkers for infection in premature neonates. *Mol Cell Proteomics* 2008; **7**: 1863–75.
- 59 Rando OJ, Verstrepen KJ. Timescales of Genetic and Epigenetic Inheritance. *Cell* 2007; **128**: 655–68.
- 60 Paschos K, Allday MJ. Epigenetic reprogramming of host genes in viral and microbial pathogenesis. *Trends Microbiol* 2010; **18**: 439–47.
- 61 Hamon MA, Cossart P. Histone Modifications and Chromatin Remodeling during Bacterial Infections. *Cell Host Microbe* 2008; **4**: 100–9.
- 62 Marazzi I, Ho JSY, Kim J, *et al.* Suppression of the antiviral response by an influenza histone mimic. *Nature* 2012; **483**: 428–33.
- 63 Vayssier-Taussat M, Albina E, Citti C, *et al.* Shifting the paradigm from pathogens to pathobiome: new concepts in the light of meta-omics. *Front Cell Infect Microbiol* 2014; **4**: 29.
- 64 Peltola VT, Murti KG, McCullers J a. Influenza virus neuraminidase contributes to secondary bacterial pneumonia. *J Infect Dis* 2005; **192**: 249–57.
- 65 Thaïss CA, Zmora N, Levy M, Elinav E. The microbiome and innate immunity. *Nature* 2016; **535**: 65–74.
- 66 Baldridge MT, Nice TJ, McCune BT, *et al.* Commensal microbes and interferon-determine persistence of enteric murine norovirus infection. *Science* 2015; **347**: 266–9.
- 67 Ichinohe T, Pang IK, Kumamoto Y, *et al.* Microbiota regulates immune defense against respiratory tract influenza A virus infection. *Proc Natl Acad Sci U S A* 2011; **108**: 5354–9.
- 68 Thompson JA, Oliveira R, Djukovic A, Ubeda C, Xavier KB. Manipulation of the quorum sensing signal AI-2 affects the antibiotic-treated gut microbiota. *Cell Rep* 2015; **10**: 1861–71.
- 69 Peterson J, Garges S, Giovanni M, *et al.* The NIH Human Microbiome Project. *Genome Res* 2009; **19**: 2317–23.

- 70 Dickson RP. The microbiome and critical illness. *Lancet Respir Med* 2016; **4**: 59–72.
- 71 Wu H-J, Ivanov II, Darce J, *et al.* Gut-Residing Segmented Filamentous Bacteria Drive Autoimmune Arthritis via T Helper 17 Cells. *Immunity* 2010; **32**: 815–27.
- 72 Rogers GB, Zain NMM, Bruce KD, *et al.* A novel microbiota stratification system predicts future exacerbations in bronchiectasis. *Ann Am Thorac Soc* 2014; **11**: 496–503.
- 73 Le Chatelier E, Nielsen T, Qin J, *et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature* 2013; **500**: 541–6.
- 74 Zhang X, Zhang D, Jia H, *et al.* The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nat Med* 2015; **21**: 895–905.
- 75 Grice EA, Segre JA. The human microbiome: our second genome. *Annu Rev Genomics Hum Genet* 2012; **13**: 151–70.
- 76 MARS Consortium. Molecular Diagnosis and Risk Stratification of Sepsis. <http://www.ctmm.nl/en/downloads/projects-downloads/ctmm-mars-output-report> (accessed Dec 12, 2016).
- 77 Scicluna BP, Klein Klouwenberg PMC, Van Vught LA, *et al.* A molecular biomarker to diagnose community-acquired pneumonia on intensive care unit admission. *Am J Respir Crit Care Med* 2015; **192**: 826–35.
- 78 Tendl KA, Schulz SMF, Mechtler TP, *et al.* DNA methylation pattern of CALCA in preterm neonates with bacterial sepsis as a putative epigenetic biomarker. *Epigenetics* 2013; **8**: 1261–7.
- 79 Huang AM, Newton D, Kunapuli A, *et al.* Impact of rapid organism identification via matrix-assisted laser desorption/ionization time-of-flight combined with antimicrobial stewardship team intervention in adult patients with bacteremia and candidemia. *Clin Infect Dis* 2013; **57**: 1237–45.
- 80 Tang BMP, McLean AS, Dawes IW, Huang SJ, Lin RCY. Gene-expression profiling of peripheral blood mononuclear cells in sepsis. *Crit Care Med* 2009; **37**: 882–8.

- 81 Pankla R, Buddhisa S, Berry M, *et al.* Genomic transcriptional profiling identifies a candidate blood biomarker signature for the diagnosis of septicemic melioidosis. *Genome Biol* 2009; **10**: R127.
- 82 Wynn JL, Cvijanovich NZ, Allen GL, *et al.* The influence of developmental age on the early transcriptomic response of children with septic shock. *Mol Med* 2011; **17**: 1146–56.
- 83 Ahn SH, Tsalik EL, Cyr DD, *et al.* Gene Expression-Based Classifiers Identify *Staphylococcus aureus* Infection in Mice and Humans. *PLoS One* 2013; **8**. DOI:10.1371/journal.pone.0048979.
- 84 Parnell GP, McLean AS, Booth DR, *et al.* A distinct influenza infection signature in the blood transcriptome of patients with severe community-acquired pneumonia. *Crit Care* 2012; **16**: R157.
- 85 Cernada M, Serna E, Bauerl C, Collado MC, Pérez-Martínez G, Vento M. Genome-Wide Expression Profiles in Very Low Birth Weight Infants With Neonatal Sepsis. *Pediatrics* 2014; **133**: e1203-1211.
- 86 Smith CL, Dickinson P, Forster T, *et al.* Identification of a human neonatal immune-metabolic network associated with bacterial infection. *Nat Commun* 2014; **5**: 4649.

## Figure legends

**Figure 1:** A complex interaction of factors involving the host, microbiome, pathogen, and environment results in a heterogeneous sepsis disease phenotype, which varies between individuals during the dynamic time course of their illness. This phenotype can be observed at both a whole organism and molecular level (see example biomarkers)<sup>18,52–54,57,77,78</sup>. During the course of an individual's illness, a variable host response is observed over different timescales, e.g. Patient B's illness peaks on day 3, leading to recovery by day 5. The optimal therapy for an individual at a given time needs to reflect the composite phenotype (arising through the interaction of host, microbiome, pathogen, and environment) and requires a precision medicine approach.

*FAIM3* = Fas apoptotic inhibitory molecule 3. *miR* = micro RNA. *PLAC8* = placenta-specific 8. *SOFA* = Sequential Organ Failure Assessment.

**Figure 2:** Unique and common biological pathways are activated by different pathogens, resulting in pathogen-specific -omic signatures. The number of differentially upregulated and downregulated genes are provided for illustrative purposes and were obtained from microarray data comparing peripheral blood gene expression in sepsis patients with *E. coli* and *Candida* positive blood cultures.<sup>16</sup>

*E. coli* = *Escherichia coli*

## Figures and tables

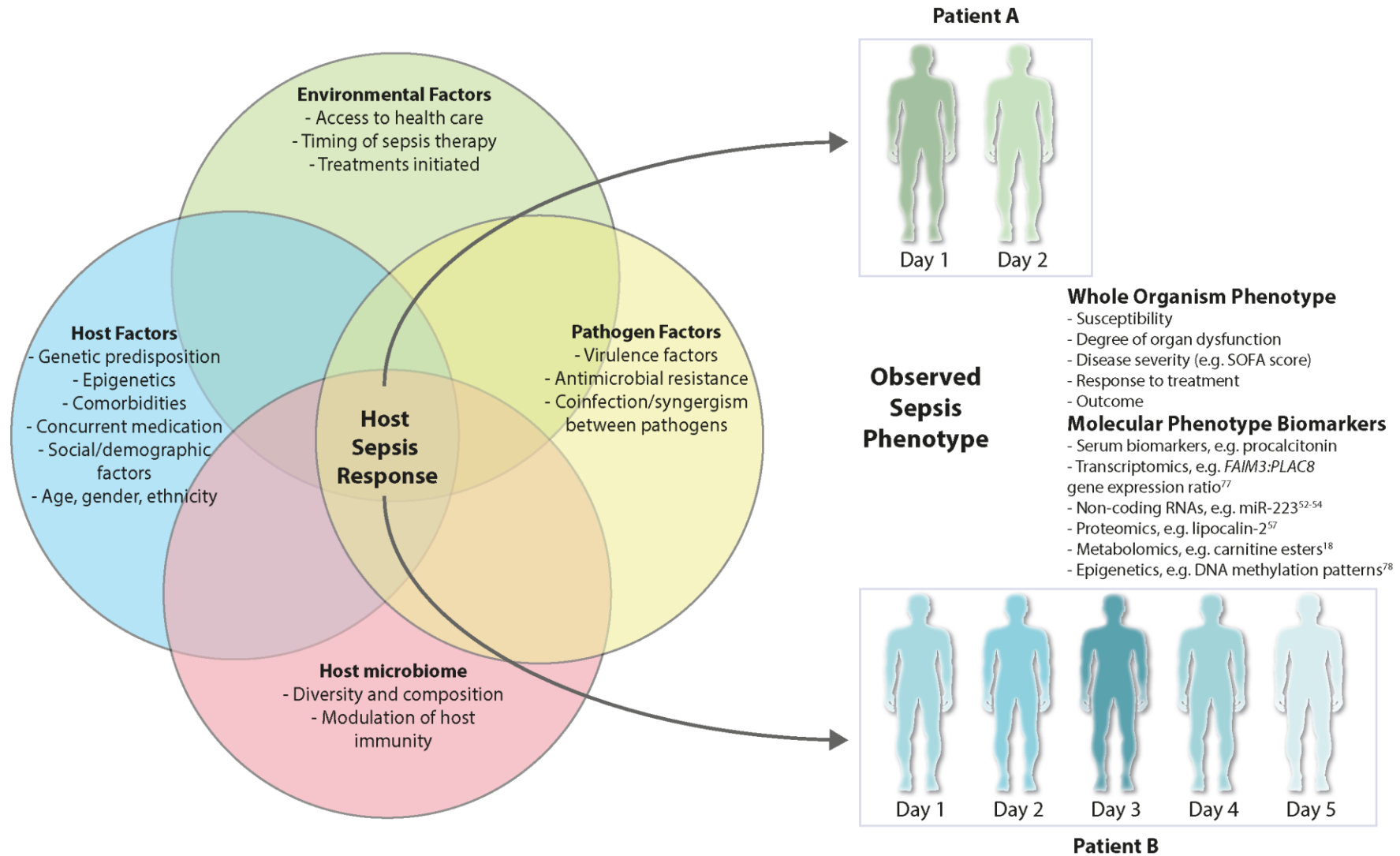
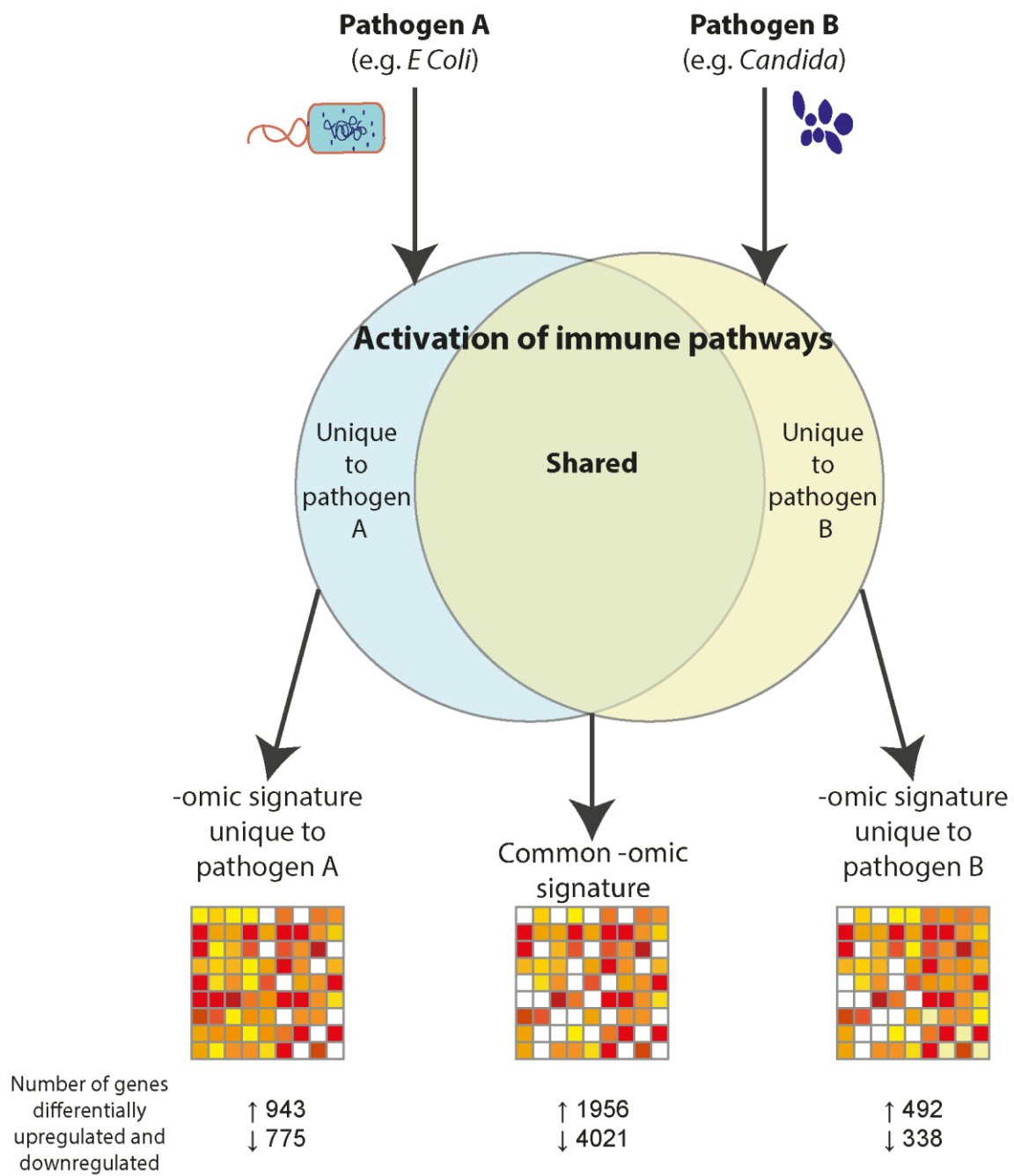


Figure 1





**Figure 2**

**Table 1: New techniques in clinical microbiology**

Technique	Applications and example platforms	Sample pre-processing	Benefits	Challenges
Multiplex PCR	<ul style="list-style-type: none"> <li>Single wells with multiple probe sets: Roche LightCycler SeptiFast, Curetis Unyvero P55</li> <li>Parallel singleplex reactions in microwells: bioMérieux FilmArray</li> <li>Microarray: Luminex Verigene</li> <li>Liquid-array: Luminex xTAG</li> </ul>	<p>Minimal (SeptiFast, Unyvero, FilmArray Respiratory)</p> <p>Direct from positive blood culture (FilmArray BCID, Verigene)</p> <p>Nucleic acid extraction, PCR (xTAG)</p>	<p>High sensitivity</p> <p>Simultaneous detection of multiple pathogens</p> <p>Can provide antimicrobial susceptibility data</p> <p>Potential for automation</p> <p>Point-of-care platforms with rapid turnaround time available</p>	<p>Low specificity</p> <p>Detection limited to pathogens in pre-specified probe panel</p> <p>Differentiating viable from non-viable pathogens</p> <p>Reduced accuracy in polymicrobial specimens</p>
Mass spectrometry	MALDI-TOF MS: bioMérieux VITEK	Pre-culture (e.g. positive blood culture)	<p>Studies show clinical benefit in ICU setting<sup>79</sup></p> <p>Provision of antimicrobial susceptibility data if combined with PCR, e.g. Cepheid GeneXpert MRSA</p>	<p>Reduced accuracy in polymicrobial specimens</p> <p>Limited detection if low pathogen load</p> <p>Bacterial/fungal species only</p> <p>Specimen pre-culture required</p>
	ESI-MS: Abbott Iridica	Nucleic acid extraction and PCR	<p>Strain typing and antimicrobial susceptibility data possible from SNP detection</p> <p>More sensitive than MALDI-TOF for low pathogen loads</p> <p>No pre-culture necessary</p>	<p>Reduced accuracy in polymicrobial specimens</p> <p>Pre-processing stages reduces turnaround time</p> <p>More suited to batch processing of multiple samples</p> <p>Limited by PCR primer design and reference library</p>
Nucleic acid sequencing	High-throughput sequencing: Illumina MiSeq and HiSeq	Nucleic acid extraction and sequencing library preparation	<p>High-resolution pathogen sequence information</p> <p>Unbiased simultaneous detection of multiple organisms from a single sample</p> <p>Detection of non-culturable/fastidious organisms</p> <p>Highly sensitive</p>	<p>Human nucleic acid far exceeds that of microorganisms</p> <p>Expensive, complex sample pre-processing</p> <p>Computational challenges with data analysis</p> <p>Interpretation of positive results given high contamination rates and complex microbiome</p>
	Long-read sequencing: Oxford Nanopore MinION		<p>Less demanding sample pre-processing than Illumina platforms</p> <p>Platform size and cloud-based computing enable use in resource-poor settings</p>	<p>High error rate in repetitive regions</p>

ESI-MS = electrospray ionisation-mass spectrometry. MALDI-TOF MS = matrix-assisted laser desorption ionisation-time of flight mass spectrometry.

**Table 2: Summary of representative clinical studies comparing the host transcriptomic response in sepsis of different microbiological aetiologies**

	Cell type	Accession no; Platform	Clinical setting	Patient group	Control group	Microbiological aetiology	Infection source	Main findings
Tang et al (2008) <sup>42</sup>	Neutrophils	GSE6535; GPL4274	Adult ICU	Sepsis (n=55)	ICU patients, no sepsis (n=17)	Gram-positive (n=18) Gram-negative (n=25) Polymicrobial (n=12)	Lung, abdomen, urinary tract, CNS, other	94 genes differentially expressed between sepsis and non-sepsis patients. Gram-positive, Gram-negative, and mixed sepsis share a similar transcriptomic profile.
Tang et al (2009) <sup>80</sup>	PBMCs	GSE9960; GPL570	Adult ICU	Sepsis (n=46)	ICU patients, non-infective SIRS (n=24)	Gram-positive (n=17) Gram-negative (n=19)	Lung, abdomen, urinary tract, other	138-gene set differentiates sepsis from SIRS. Gram-positive and Gram-negative sepsis share similar transcriptomic profile.
Pankla et al (2009) <sup>81</sup>	Whole blood	GSE13015; GPL6106, GPL6947	Adult inpatient	Sepsis with positive blood culture (n=63)	Heterogeneous: healthy volunteers and outpatients (n=29)	<i>B. pseudomallei</i> (n=32) Other bacterial/fungal (n=31)	Lung, other unidentified or unspecified	37-gene set differentiates sepsis from <i>B. pseudomallei</i> and sepsis from other causes.
Wynn et al (2011) <sup>82</sup>	Whole blood	GSE26440, GSE26378; GPL570	Paediatric ICU	Sepsis (n=180)	Patients from ambulatory department without infective or inflammatory pathology (n=53)	Gram-positive (n=47) Gram-negative (n=46)	Unspecified	Only 11 genes differentially expressed between Gram-positive and Gram-negative sepsis.
Ahn et al (2012) <sup>83</sup>	Whole blood	GSE33341; GPI571	Adult inpatient	Sepsis with positive blood culture (n=51)	Healthy volunteers (n=43)	<i>S. aureus</i> (n=32) <i>E. coli</i> (n=19)	Lung, urinary tract, endocarditis, skin, catheter, bone, CNS, unknown	Human-derived (2-factor) and murine-derived (4-factor) classifier sets differentiate <i>S. aureus</i> bacteraemia from healthy controls or <i>E. coli</i> bacteraemia.
Parnell et al (2012) <sup>84</sup>	Whole blood	GSE40012; GPL6947	Adult ICU	CAP with ≥ SIRS criteria (n=27)	Healthy volunteers (n=18) ICU patients with non-infected SIRS (n=12)	Influenza A H1N1 (n=8) Bacterial (n=16) Influenza/bacterial coinfection (n=3)	Lung	29-gene influenza-specific signature identified. Small number of differentially expressed genes in bacterial group but no specific biological pathways represented.
Cernada et al (2014) <sup>85</sup>	Whole blood	E-MTAB-4785; A-AFFY-141	Neonatal ICU	Very low birth weight with bacterial sepsis (n=17)	Very low birth weight with no infection (n=19)	Gram-positive (n=13) Gram-negative (n=4)	Unspecified	Unsupervised principal component analysis of gene expression identified 3 clusters corresponding to Gram-positive, Gram-negative and non-infected aetiologies.
Smith et al (2014) <sup>86</sup>	Whole blood	GSE25504; GPL13667, GPL6947	Neonatal unit	Sepsis (n=47)	Inpatients without infective pathology (n=45)	Bacterial (n=43) Rhinovirus or CMV (n=4)	Lung, CNS, urinary tract, abdominal, soft tissue, unspecified	52-gene-classifier derived which discriminated bacterial infection from viral/no infection in an independent population.
Cheng et al (2016) <sup>16</sup>	Whole blood	GSE65682; GPL13667	Adult ICU	Sepsis with positive blood culture (n=46)	Healthy volunteers (n=42)	<i>E. coli</i> bacteraemia (n=33) <i>Candida</i> fungaemia (n=13)	Unspecified	Majority of differentially expressed genes common to both pathologies but a unique transcriptional response specific to <i>E. Coli</i> bacteraemia and <i>Candida</i> fungaemia also seen.
Tsalik et al (2016) <sup>6</sup>	Whole blood	GSE63990; GPL571	Adult emergency department	ARI with ≥ SIRS criteria (n=185)	Healthy volunteers (n=44) Non-infected SIRS patients (n=88)	Bacterial (n=70) Viral (n=115)	Lung	Gene set classifying bacterial ARI (71 probes mapping to 69 genes), viral ARI (33 genes) or non-infectious illness (26 genes) identified.

ARI = acute respiratory illness. *B. pseudomallei* = *Burkholderia pseudomallei*. CNS = central nervous system. *E. coli* = *Escherichia coli*. PBMC = peripheral blood mononuclear cell. *S. aureus* = *Staphylococcus aureus*. *S. pneumoniae* = *Streptococcus pneumoniae*. SIRS = systemic inflammatory response syndrome.

**Panel 1: Translational goals for sepsis –omics research**

Diagnosis

- Early recognition of patients at risk for sepsis
- Differentiation from non-infective aetiologies
- Rapid, accurate microbiological diagnosis

Treatment

- Precision therapy for specific sepsis sub-phenotypes
- Identification of novel therapeutic targets
- Targeted, effective antimicrobial therapy
- Biomarker development for definition of disease sub-phenotypes and monitoring response to therapy

Prognosis

- Identification of patients at risk of specific organ failure or poor outcome