

REVIEW

# Bench-to-bedside review: Future novel diagnostics for sepsis – a systems biology approach

Simon Skibsted<sup>\*1,2</sup>, Manoj K Bhasin<sup>3</sup>, William C Aird<sup>2</sup> and Nathan I Shapiro<sup>1,2</sup>

## Abstract

The early, accurate diagnosis and risk stratification of sepsis remains an important challenge in the critically ill. Since traditional biomarker strategies have not yielded a gold standard marker for sepsis, focus is shifting towards novel strategies that improve assessment capabilities. The combination of technological advancements and information generated through the human genome project positions systems biology at the forefront of biomarker discovery. While previously available, developments in the technologies focusing on DNA, gene expression, gene regulatory mechanisms, protein and metabolite discovery have made these tools more feasible to implement and less costly, and they have taken on an enhanced capacity such that they are ripe for utilization as tools to advance our knowledge and clinical research. Medicine is in a genome-level era that can leverage the assessment of thousands of molecular signals beyond simply measuring selected circulating proteins. *Genomics* is the study of the entire complement of genetic material of an individual. *Epigenetics* is the regulation of gene activity by reversible modifications of the DNA. *Transcriptomics* is the quantification of the relative levels of messenger RNA for a large number of genes in specific cells or tissues to measure differences in the expression levels of different genes, and the utilization of patterns of differential gene expression to characterize different biological states of a tissue. *Proteomics* is the large-scale study of proteins. *Metabolomics* is the study of the small molecule profiles that are the terminal downstream products of the genome and consists of the total complement of all low-molecular-weight molecules that cellular processes leave behind. Taken together, these individual fields of study may be linked during a systems biology approach. There remains a valuable opportunity to deploy these technologies further in human research. The techniques described in this paper not only have the potential to increase the spectrum of diagnostic and prognostic biomarkers in sepsis, but they may also enable the discovery of new disease pathways. This may in turn lead us to improved therapeutic targets. The objective of this paper is to provide an overview and basic framework for clinicians and clinical researchers to better understand the 'omics technologies' to enhance further use of these valuable tools.

## Introduction

The early, accurate diagnosis and risk stratification of sepsis remains an important challenge in the critically ill. Despite significant improvements in clinical care, sepsis continues to be a lethal and expensive condition with mortality rates approaching 20 to 30% [1-4]. While clinicians typically incorporate elements of history, physical examination, laboratory and radiographic testing, no single accepted biomarker, combination of biomarkers or clinical prediction rule is used to aid in diagnosis and

risk stratification [5,6]. Since traditional biomarker strategies whereby one measures the concentration of circulating proteins have not yielded a definitive biomarker or set of biomarkers for sepsis, focus is shifting towards strategies that improve assessment capabilities. The overall objective is to facilitate early and appropriate therapeutic intervention, improve triage decisions, provide a means to follow response to therapy, establish new therapeutic targets, and/or provide ways to identify patients amenable to tailored therapies.

Technological advancements, along with the information generated through the human genome project, have positioned systems biology at the forefront of biomarker discovery. This has facilitated approaches that may not only yield an improved insight into complicated sepsis pathophysiology, but may also identify unexplored pathways [7,8]. Over the past decade, technologies focusing on DNA, gene expression, gene regulatory mechanisms,

\*Correspondence: sskibste@bidmc.harvard.edu

<sup>1</sup>Department of Emergency Medicine and Center for Vascular Biology Research, Beth Israel Deaconess Medical Center and Harvard Medical School, 330 Brookline Avenue, Boston, MA 02215, USA

Full list of author information is available at the end of the article

protein and metabolite discovery have been introduced. Taken together, a systems biology approach links these individual fields of study. Systems biology refers to the integration and analysis of complex datasets derived from multiple facets of the body's signaling and response pathways (that is, genomics, transcriptomics, proteomics, and metabolomics). The science of biomarkers discovery has substantially evolved, with focused fields of study on each part of the lifecycle of biologic signaling and response. While the 'omics technologies' have been available in varying capacities for well over a decade (reviewed in [9-18]), advances in technology are continually increasing the feasibility and accessibility, while decreasing the costs. The objective of this paper is to provide the reader with an overview and understanding of these approaches and techniques that are at the forefront of sepsis research.

### Genomics

Genomics (Figure 1, target 1) is the study of the entire complement of genetic material of an individual. In 2003, after 13 years, the international human genome project completed its task of sequencing the 3 billion bp of human genomic DNA, estimating a total of ~28,000 to 34,000 genes. This laid the foundation for functional genomics, genomics medicine, bioinformatics and systems biology to investigate functions and regulatory mechanisms. The sequencing of the human genome has moved systems biology into the forefront of biomarker discovery. In sepsis, genomics focuses primarily on genomic variation analysis [19].

### Genetic variation analysis

#### Background

The notion that genetics plays an important role in sepsis is not new. In 1988 Sorensen and colleagues conducted a study of adoptees in Denmark, focusing on death from all causes. The study found that if a biological parent died of infection before the age of 50 years, the child had a 5.8 relative risk of also dying from infection [20]. In fact, there is a much higher heritability of death due to infection than due to cancer or heart disease. This suggests that the genetics of sepsis is an important factor in determining outcome and makes genomics research particularly interesting.

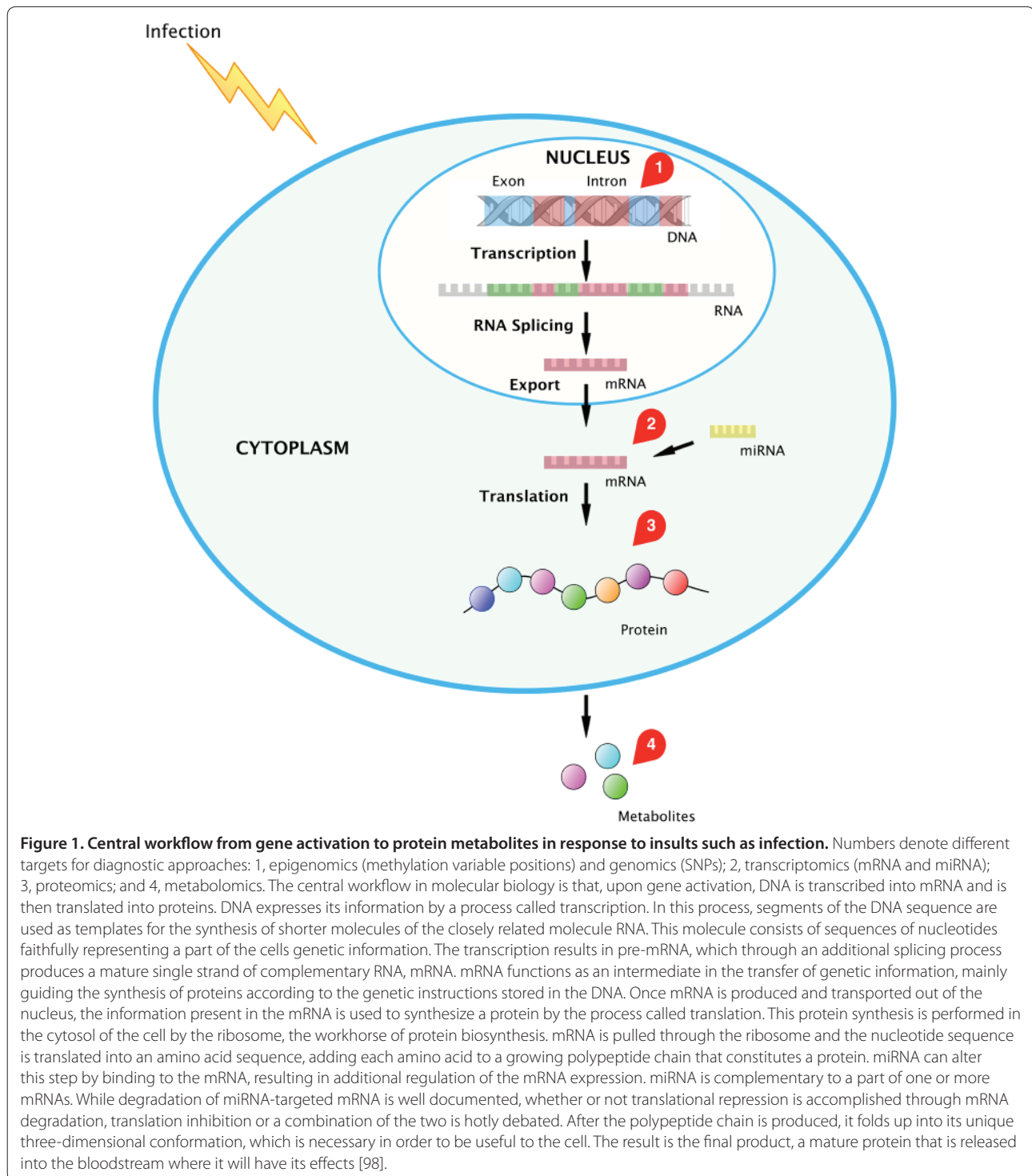
The primary approach for studying human genetics in relation to disease is to analyze genetic variations called single nucleotide polymorphisms (SNPs). A SNP results from a single base mutation in the DNA sequence and has a frequency of 1% or more in the population. Large-scale SNP discovery projects such as the HapMap have identified that a comparison of two chromosomes between any two individuals will generally reveal, on average, 5 to 10 million common SNPs across the genome

[21,22]. The large number of SNPs identified has led to the development of an SNP database – the dbSNP, established by the National Center for Biotechnology Information [23]. In February 2013 the dbSNP contained 53 million unique human SNPs. The UCSC Genome browser lists results from genome-wide association studies (GWAS) and has developed an integrated map of genetic variation from 1,092 human genomes as of March 2013 [24,25].

The effort to find genetic variants that are responsible for susceptibility to sepsis has been primarily gene association studies in case-control or cohort studies. In 2010, a large case-control study including over 8,000 subjects found a genetic association between five SNPs in the gene coding for IL-2 with increased susceptibility to bacteremia, malaria and tuberculosis [26]. The overall risk of one of these infectious diseases was increased to 81% in persons with four or more IL-2-specific SNPs. Such studies often rely on already known genetic variations from the HapMap database. However, due to the ever increasing dbSNP database and the fact that genotyping technology has become less expensive, focus has shifted towards GWAS [27].

GWAS simultaneously probe all segments of the genome for evidence of association between a known SNP and disease by comparing diseased and nondiseased populations to identify SNPs that are more prevalent in the diseased state. GWAS have thus far uncovered >800 SNP associations for more than 150 disease and other traits [28]. For these hypothesis-generating studies, DNA microarrays or whole genome sequencing are often used because high throughput is feasible, and *a priori* selection of specific SNPs is not required. DNA microarrays identify pathways and associations in an extremely efficient fashion (Figure 2). DNA microarrays are more comprehensive than PCR methods because PCR uses primers that the investigator needs to specify before the analysis is run. Furthermore, PCR is only capable of analyzing hundreds of SNPs as compared with millions when using DNA microarrays.

GWAS provide the researcher with information about whether a SNP is present or not. GWAS do not offer any information about gene regulation or disease progression. The ChIP-on-chip technique combines chromatin immunoprecipitation (the ChIP) with DNA microarray technology (the chip). Chromatin immunoprecipitation refers to the technique of precipitating a protein antigen from a solution using an antibody that specifically binds to that particular protein. This technique is used to investigate interactions between proteins and DNA. The proteins are generally those operating in the context of chromatin. The overall goal of ChIP-on-chip is to localize protein-binding sites that may help identify functional elements in the genome.

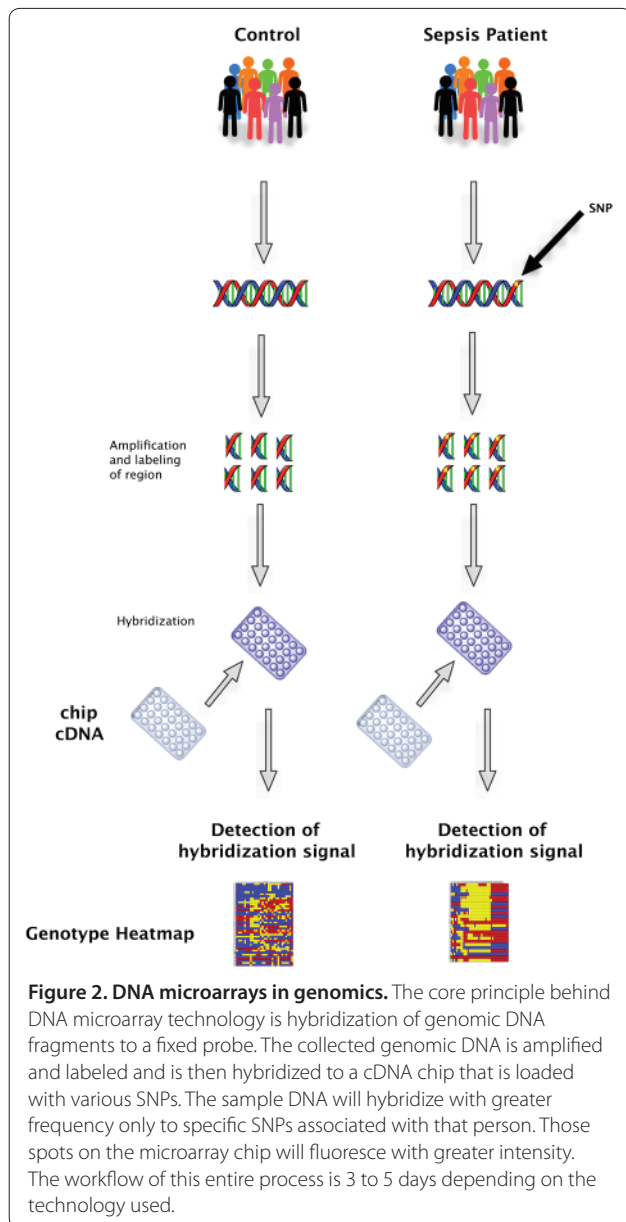


**Figure 1. Central workflow from gene activation to protein metabolites in response to insults such as infection.** Numbers denote different targets for diagnostic approaches: 1, epigenomics (methylation variable positions) and genomics (SNPs); 2, transcriptomics (mRNA and miRNA); 3, proteomics; and 4, metabolomics. The central workflow in molecular biology is that, upon gene activation, DNA is transcribed into mRNA and is then translated into proteins. DNA expresses its information by a process called transcription. In this process, segments of the DNA sequence are used as templates for the synthesis of shorter molecules of the closely related molecule RNA. This molecule consists of sequences of nucleotides faithfully representing a part of the cells genetic information. The transcription results in pre-mRNA, which through an additional splicing process produces a mature single strand of complementary RNA, mRNA. mRNA functions as an intermediate in the transfer of genetic information, mainly guiding the synthesis of proteins according to the genetic instructions stored in the DNA. Once mRNA is produced and transported out of the nucleus, the information present in the mRNA is used to synthesize a protein by the process called translation. This protein synthesis is performed in the cytosol of the cell by the ribosome, the workhorse of protein biosynthesis. mRNA is pulled through the ribosome and the nucleotide sequence is translated into an amino acid sequence, adding each amino acid to a growing polypeptide chain that constitutes a protein. miRNA can alter this step by binding to the mRNA, resulting in additional regulation of the mRNA expression. miRNA is complementary to a part of one or more mRNAs. While degradation of miRNA-targeted mRNA is well documented, whether or not translational repression is accomplished through mRNA degradation, translation inhibition or a combination of the two is hotly debated. After the polypeptide chain is produced, it folds up into its unique three-dimensional conformation, which is necessary in order to be useful to the cell. The result is the final product, a mature protein that is released into the bloodstream where it will have its effects [98].

#### **Clinical utility of genetic screening for infection**

Several SNPs in genes coding for the innate immune system have been identified as playing a role in pathophysiology and outcomes in severe sepsis or septic shock [29-36]. One study found that when administering activated protein C (Xigris™; Eli Lilly and Company,

Indianapolis, IN, USA) a polymorphism at position -1,641 AA in the promoter region coding for protein C was associated with decreased survival, increased organ dysfunction and increased systemic inflammation in severe sepsis [37]. Even though Xigris™ has been withdrawn by the manufacturer, this theragnostic



**Figure 2. DNA microarrays in genomics.** The core principle behind DNA microarray technology is hybridization of genomic DNA fragments to a fixed probe. The collected genomic DNA is amplified and labeled and is then hybridized to a cDNA chip that is loaded with various SNPs. The sample DNA will hybridize with greater frequency only to specific SNPs associated with that person. Those spots on the microarray chip will fluoresce with greater intensity. The workflow of this entire process is 3 to 5 days depending on the technology used.

approach – where patients who are more likely to respond to a given therapy are targeted – is certainly a likely future path forward. There are a number of other examples of the importance of SNPs in infection [31-33,38]. For example, in a 176-patient human trial in sepsis, a specific SNP at the -308 position in the TNF $\alpha$  promoter region was associated with elevated TNF $\alpha$  expression *in vitro* and *in vivo*, and conferred to a 1.5-fold to fourfold increased mortality in studies of septic shock. The association between gene polymorphisms and mortality awaits large-scale validation but there is a strong support for the inclusion of genotyping when designing sepsis trials [19,39,40].

### Advantage

The primary attraction of genomic approaches for gene identification is that one may study the genome to determine how genetic predisposition influences disease acquisition and response. The GWAS approach can elucidate the molecular basis for disease without any prior understanding of the biology underlying the disease. This unbiased genome level approach is particularly interesting in a heterogeneous syndrome such as sepsis where the underlying pathophysiology is poorly understood.

### Limitations

In sepsis research, the focus is on the elements that cause disease and alter phenotypes through alteration of molecular function. There are challenges to identifying such functional variants. Earlier studies have primarily evaluated SNPs in regulatory or coding regions, which alter expression of a gene or produce an altered protein structure that may be dysfunctional [41]. This approach may be oversimplified and ignores the fact that other SNPs might alter outcome – SNPs that we do not yet know the exact function of. Furthermore, our understanding of how SNPs disrupt molecular function is poorly understood [42]. Without a complete understanding of the regulation of transcription of a gene, an association study strategy based on the functional plausibility of single SNPs may overlook polymorphisms essential for the expression of genes [43]. Indeed, it seems that the major obstacle is an information gap, not simply a technology gap.

### Epigenetics

Gene activity is regulated by a variety of mechanisms, known as epigenetics. These mechanisms rely on reversible modifications of DNA (most commonly DNA methylation and histone modification) that affect gene expression without altering the DNA sequence. The distribution of these modifications may not only be specific to a particular organism or a particular tissue, but may also mark specific disease states. The epigenome (Figure 1, target 1) is the distribution of epigenetic regulation. The epigenome is not static, like the genome, but changes in response to environmental changes, and plays a fundamental role in gene expression following environmental and extracellular stimuli.

### Background

DNA methylation is the biochemical process involving addition of a methyl group to the 5-position of the cytosine pyrimidine ring or the number-6 nitrogen of the adenine purine ring. The changes in DNA methylation are often associated with chromosome instability and gene repression. Histone modification is another

well-studied epigenetic regulation mechanism. The function of histones is to package and order the DNA into structural units by wrapping the DNA around core histones. Huge catalogues of histone modifications have been described but the functional meaning of these are not yet fully understood. Acetylation of the histones probably opens up the DNA and facilitates transcription. Specific deacetylases reverse this by closing the DNA, making it more condensed and promoting gene repression. Methylation of the histones may also activate or repress the DNA. Only recently have epigenomic profiling technologies reached the stage at which large-scale studies are becoming feasible. A variety of both array-based and sequencing-based methods are available, with the choice based on balancing coverage, resolution, accuracy, specificity, throughput and costs [44].

DNA methylation is detected by the use of bisulfite treatment of DNA. This treatment changes unmethylated cytosines to uracil but leaves methylated cytosines unchanged. When looking at epigenetic changes in the genome, the primary focus is on the CpG islands, which are >200 bp stretches of DNA that have a significantly higher frequency of the nucleotides cytosine and guanine. These islands have been found in approximately 40% of promoters of human mammalian genes [45]. Usually the CpG islands occur near the transcription site of genes and are involved in the transcriptional regulation. Differential methylation hybridization allows the simultaneous determination of the methylation levels of a large number of CpG island loci [46]. The genome wide study of histone modifications is achieved by HPLC (a technique for separating DNA or protein molecules by molecular weight and conformation) and high-performance capillary electrophoresis (a separation technique that uses narrow-bore fused-silica capillaries to separate a complex mixture of chemical compounds) [47].

#### **Clinical utility**

There are no major studies of epigenetic modifications in sepsis of which we are aware. Several histone modifications are demonstrated to differentially regulate subsets of lipopolysaccharide-induced genes. A phosphorylation of a specific histone (histone 3 at serine 10) may have a gene-specific role in NF- $\kappa$ B recruitment [48]. NF- $\kappa$ B is a transcriptional regulatory factor and a central participant in modulating the expression of many of the immunoregulatory mediators involved in sepsis. After lipopolysaccharide stimulation, the genes encoding for several cytokines, including IL-6, undergo phosphorylation at their promoters. This facilitates NF- $\kappa$ B recruitment and gene induction. DNA methylation is a common epigenetic signaling tool that cells use to silence genes and thereby regulate gene expression [49].

#### **Advantages**

Despite the success of the GWAS, there is still a substantial proportion of causality that remains unexplained. Increasingly evident is that the epigenome is highly dynamic and consists of a complex interplay of genetic and environmental factors [50]. A method to uncover this interplay is the epigenomics equivalent of GWAS – epigenome-wide association studies [51]. For DNA methylation, technology is now available that is directly comparable in resolution and throughput with the GWAS chips [51]. The epigenetic equivalent of a SNP is DNA methylation at a single site, known as a methylation variable position. Based on simulations assuming conservative methylation odds ratios, epigenome-wide association studies should be able to detect associations in fewer samples compared with GWAS [51]. For non-malignant, common complex diseases such as diabetes or autoimmunity, the investigation of the epigenetic component is only beginning.

#### **Limitations**

Disease-associated epigenetic variation can be tissue or cell specific. All tissues are composed of multiple cell types (blood contains more than 50 specific cell types). If the disease-associated variation is restricted to a specific cell type, then assessing the wrong tissue type (for example, the easily accessible leukocyte from circulating blood) will miss a target [51]. Furthermore, there is no epigenomic equivalent of the HapMap project, which helped to elucidate some of the genetic variation in the human genome. Since no database exists, we are not yet capable of making any statements about the frequency of an epigenetic regulation since we do not know the level of normal epigenetic variation that exists in human populations [51]. Lastly, since epigenetic variation can be causal of disease or can arise as a consequence of disease, it can be difficult to conclusively distinguish between disease-driving or passenger epigenetic variants, thus making reverse causation a concern [51].

#### **Transcriptomics**

Transcriptomics (Figure 1, target 2) is the quantification of messenger RNA levels for a large number of genes in specific cells or tissues to measure differences in the expression levels of different genes and the utilization of patterns of differential gene expression to characterize different biological states of a tissue. Unlike the genome, which is mostly similar for a given cell line, the transcriptome responds constantly to external environmental conditions and internal conditions, such as sepsis. The study of transcriptomics, also referred to as expression profiling, examines the expression level of mRNAs or miRNAs in a given cell population. The transcriptome is thus indicative of gene activity and regulation. In



humans, nearly every cell contains the same genome, and thus the same genes. However, not every gene is transcriptionally active in every cell; different cells show different patterns of gene expression [52]. In context, the transcriptome is seen as a precursor for the entire set of proteins expressed by the genome – the proteome.

### **Gene expression profiling**

#### **Background**

In 2001, a novel molecular approach using microarrays to monitor genome-wide changes in relative mRNA abundance in the host response to infection was described [53]. This was the beginning of genome-wide transcriptomics as an investigational tool to study sepsis. DNA microarrays are a commonly used technique to profile gene expression as they allow for genome-wide assessments of changes in gene expression by surveying expression patterns for tens of thousands of genes in a single experiment (Figure 3) [54]. Since its introduction, microarray technology has been applied to sepsis by several investigators (reviewed in [9]); however, further work is needed to advance our understanding and to increase the scope of implementation in research.

Another approach to provide a quantitative view of the expression of selected genes is multi-gene transcriptional profiling, which quantifies mRNA copy numbers. [55] Compared with DNA microarrays that are limited in their accuracy and reproducibility, multi-gene transcriptional profiling uses real-time PCR – a method widely regarded as the gold standard for nucleic acid quantification [56,57]. Compared with DNA microarrays, the results are quantitative and, if real-time PCR is employed, the turnaround time is short. Real-time PCR lacks the discovery breadth of DNA microarrays since it cannot be used for a genome-wide scan, but it has the ability to rapidly and quantitatively measure hundreds of genes and could allow for targeted screening for multiple biomarkers.

The demand for low-cost sequencing has driven the development of next-generation sequencing technologies, such as RNA-seq. This technique does not require the sequence information in order to detect and evaluate transcripts, and has deep coverage and base-scale resolution. To our knowledge, however, this technique has not yet been applied to sepsis research. For many, the DNA microarray approach is still the method of choice due to lower costs and availability.

#### **Clinical utility**

Alterations in transcript/gene abundance in cells such as white blood cells and endothelial cells that affect cytokine synthesis, cytokine receptor expression, protein synthesis regulation and apoptosis have been reported in patients with severe sepsis [58]. One whole blood gene expression analysis revealed over 500 unique genes that were

differentially expressed comparing pre-septic patients (patients with systemic inflammatory response syndrome, who developed clinical sepsis during the study) and uninfected patients with systemic inflammatory response syndrome [59]. In addition to discriminating inflammation from sepsis, gene expression profiling also has been widely implemented to identify predictive biomarkers. A recent systematic review by Tang and colleagues found that a total of 12 cohorts consisting of 784 individuals has been investigated using genome-wide expression data [60]. Even though the studies had consistent results in terms of activation of signal transduction cascades and pathogen recognition receptors, the studies had highly variable changes in inflammation-related genes. In a genome-wide survey of mRNA expression in 38 patients with septic shock, a set of 28 genes that discriminated between survivors and nonsurvivors was identified. These genes were upregulated between 31 and 714% [61]. In 2010, a transcriptional-based stratification strategy for pediatric septic shock was published [62]. This strategy was based on 100 gene signatures and gene expression mosaics, and provides proof of the concept for the use of gene expression data in a clinical setting. The results from the first gene expression profiling studies are promising and are hypothesis generating; however, they await further larger scale studies in more generalizable populations.

#### **Advantages**

The use of transcriptomics in sepsis has enabled the discovery of specific and sensitive transcriptional signatures consistent with activation of pathogen recognition receptors in the human cell. Associated alterations in signal transduction pathways in sepsis have the potential to increase the knowledge of the pathophysiology of sepsis. Several gene expression patterns have been associated with the early diagnosis of sepsis, and this could be exploited to direct early interventions.

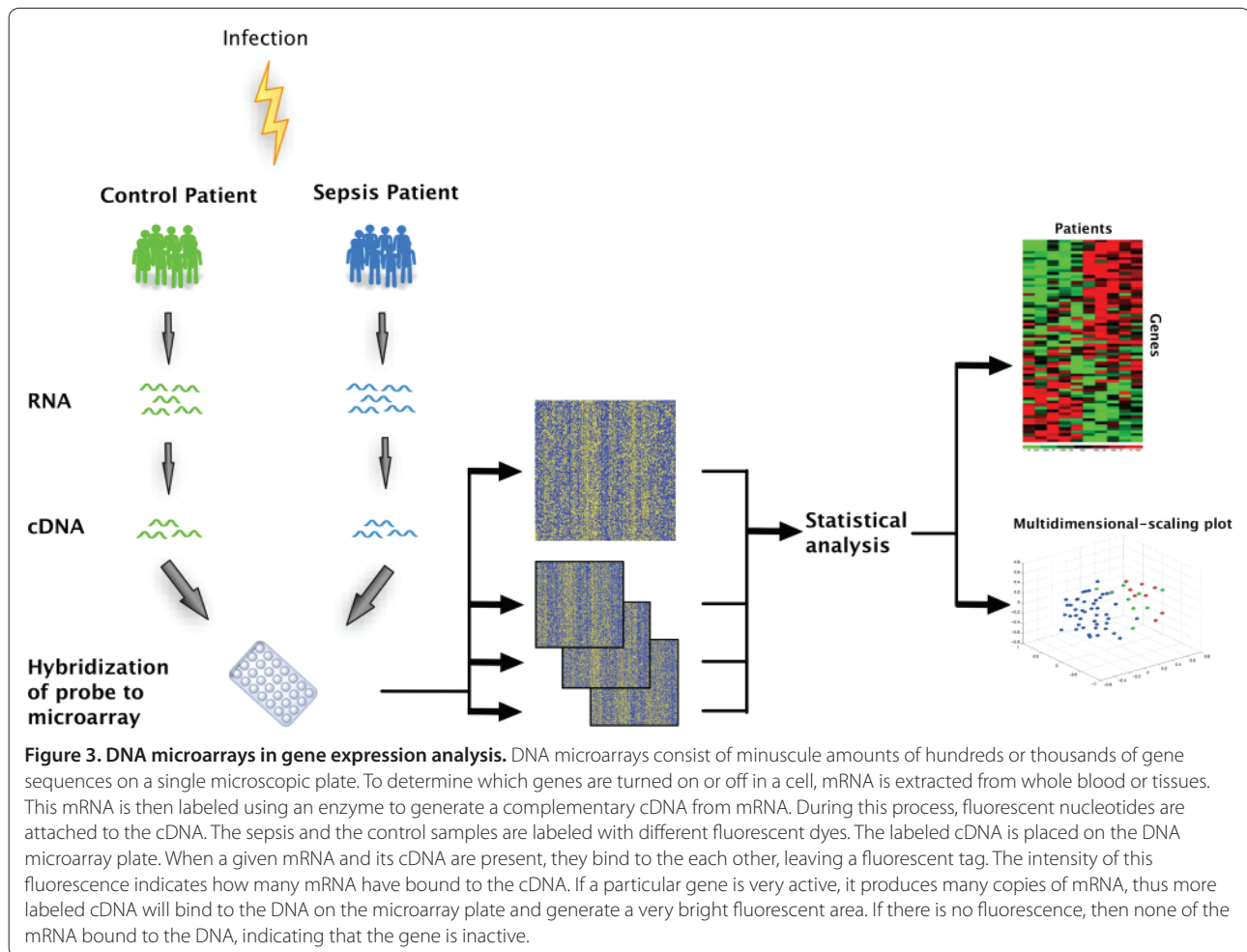
#### **Limitations**

Genome-level transcriptional studies have found highly variable changes in the transcriptional profiles of genes associated with inflammation. Indeed, there is a lack of consistent patterns in the expression of sepsis markers [63]. There are several possible reasons for this – for example, the studies have typically assessed gene expression changes in circulating leukocytes, and gene expression changes in resident leukocytes in local tissue may be different. Furthermore most of the studies have not reported leukocyte differential, which indeed is important due to the variability.

#### **microRNAs**

##### **Background**

MicroRNA (miRNA) are small ~22 nucleotide-long non-coding RNAs that regulate gene expression at the level of



RNA processing, RNA stability and translation (Figure 1, target 2). The effects of miRNAs on gene expression and control are generally inhibitory, and the corresponding regulatory mechanisms are therefore collectively termed RNA silencing. They are thought to regulate expression of protein-coding genes by direct interaction with and degradation of mRNA or by inhibition of protein translation [64]. Although estimated to represent around 2% of the genome, miRNA gene products are proposed to regulate as many as 92% of the genes in humans [65].

Recent studies also reveal that miRNAs may function as mediators of cell-to-cell communication. There is thus a possibility that miRNAs are taken up by distant cells to regulate gene expression. Since miRNA is involved in numerous cellular processes, including cell proliferation, differentiation and apoptosis, it is proposed that levels of specific miRNAs could serve as novel biomarkers of disease. In fact, the regulation of miRNA production may provide the human cell with a fast-acting response to environmental changes, such as an infection. Since miRNA acts on many different mRNAs simultaneously,

the miRNA regulation is widespread. Additionally, the interaction between miRNA and mRNA may have important biologic implications (for example, mRNA may be present but miRNA may regulate its activity) – such that to properly interpret an mRNA signal, the miRNA interaction must be considered. miRNAs hold a particularly appeal in the clinical setting because they are very stable in both plasma and serum [66,67]. Approximately ~20,000 miRNAs are thus far identified and registered in the miRBase – a database that acts as an archive of miRNA sequences and annotations.

#### **Clinical utility**

Owing to the regulatory role of miRNA on gene expression, it is not surprising that miRNA expression levels are altered in human pathological conditions, due to the changes in the transcriptional or post-transcriptional regulation and miRNA expression. Indeed, data suggest that investigating miRNA expression has potential for the identification of new and early diagnostic as well as prognostic and clinical markers [68]. Studies in animals

and humans have found that miRNAs are differentially expressed in many types of immune cells and that miRNAs have potentially critical functions in the immune system [69-71]. As an example, *in vitro* profiling of the human leukocyte response to endotoxemia indicated that five miRNAs consistently responded to lipopolysaccharide infusion, four of which were downregulated (miR-146b, miR-150, miR-342 and let-7g) and one of which was upregulated (miR-143) [71]. In another prospective clinical study enrolling 17 sepsis patients and 32 healthy controls, genome-level profiling by microarray in leukocytes identified that miR-150 was significantly downregulated in sepsis patients. Furthermore, the levels of miR-150 correlated with the Sequential Organ Failure Assessment scores as a measure of disease severity.

#### **Advantages**

Extracellular miRNAs are remarkably stable in the bloodstream. This makes probing easy, thus having the potential to serve as novel biomarkers in sepsis. Furthermore, miRNA sequences are evolutionarily conserved and are often tissue or pathology specific [72]. This suggests that miRNA functions might play an important role in regulating networks. Advances in technology platforms for miRNA detection such as microarrays and next-generation sequencing have allowed for the simultaneous interrogation of the complete small noncoding RNA repertoire. Finally, the interaction between miRNA (coding and noncoding) and mRNA may drive functions such that interpreting mRNA in the absence of miRNA data would be a flawed approach.

#### **Limitations**

Precisely how miRNAs regulate the expression of protein-coding genes is not completely understood, and the underlying mechanisms remain an important question that will impact on our understanding of gene regulation and its alteration in disease.

#### **Proteomics**

Proteomics (Figure 1, target 3) is the large-scale discovery of proteins. Proteomics confirms the presence of the protein and provides a direct measure of the quantity present.

#### **Background**

Compared with traditional protein biomarker technologies, proteomics uses more discovery enabling techniques such as mass spectrometry. The proteome will vary with time and distinct requirements, or stresses, that a cell or organism undergoes. Proteomics is considered the next step in the study of biological systems downstream from genomics and transcriptomics. mRNA expression levels

do not necessarily correlate with protein content [73,74]. This noncorrelation is partly due to the fact that not all mRNA is translated into proteins and the amount of protein produced for a given amount of mRNA depends on the gene from which it is transcribed.

#### **Clinical utility**

Proteomic methods are divided into expressional proteomics and functional proteomics. Expressional proteomics is the cataloging of the expression of all proteins present in cells, tissues or organisms [75]. In biomedical application, this comparative approach is usually employed to identify proteins that are upregulated or downregulated in a disease-specific manner for use as diagnostic markers. Expressional proteomics analyzes proteins that undergo a specific change after a given stimulus – such as severe sepsis [76]. As an example, a prospective cohort study of liver transplant patients assessed which plasma protein peaks were associated with postoperative sepsis. The study found that a combination of five proteins provided material for useful diagnostic biomarkers [77]. A total of 31 patients developed sepsis postoperatively and found an area under the curve of (0.72, 95% confidence interval = 0.57 to 0.85), which was similar to procalcitonin (0.68, 95% confidence interval = 0.53 to 0.82). Another study of 18 patients with sepsis found differential protein expression in survivors versus nonsurvivors [78]. These plasma proteins included both known cytokines as well as a group of proteins with unknown functions [78,79].

Functional proteomics is a fundamentally and strategically different approach. Functional proteomics is an emerging research area that focuses on the elucidation of biological functions of unknown proteins and the definition of cellular mechanisms at the molecular level. Due to the number of genome sequencing projects, there is an exponential growth in the number of protein sequences whose function is still unknown. One obstacle in biology is to identify those proteins that participate in specific biological processes and to assign a function to each.

#### **Advantages**

Plasma is not only the primary clinical specimen but also represents the largest version of the human proteome present in any sample. Proteomics have several advantages over genomics and transcriptomics. Genetic markers reveal only the genotype and hence do not reveal anything about the regulation of biological processes in response to disease that is expressed at the mRNA or protein levels. While mRNA does reveal regulation, it is routinely obtained from blood lymphocytes and is not well correlated with protein expression [80]. Compared with traditional protein biomarker technologies, proteomics have the major advantage of being able to drastically



increase the number of proteins detected. Proteomics is much less restrictive than ELISA and multiplex technologies in that, theoretically, an unlimited number of proteins can be analyzed simultaneously.

### Limitations

Every presently known plasma proteomic method still only samples a relatively small fraction of the proteome that mostly consists of the relatively highly expressed proteins [81,82]. Presently used proteomics methods mainly sample classical plasma proteins in the range of  $\mu\text{g/ml}$  to  $\text{mg/ml}$ , thereby excluding messengers and proteins leaking from specific diseased tissue leakage products [83]. The abundance of different proteins in blood varies by more than 10 orders of magnitude [79]. In fact, attempts to conduct a large-scale characterization of the human plasma proteome had been disappointing. The Human Proteome Organization has estimated that only 10% of the core plasma proteome (estimated to contain at least 10,000 proteins [84]) is being effectively sampled with current approaches. To identify a peptide, it must be detected and sequenced. Due to the overwhelming presence of peptides derived from the most abundant proteins, there is a significant suppression of lower abundance analytes that mask signals of less abundant species with similar chemical properties. This limits the amount of sample that can be loaded for mass spectrometry. Furthermore the currently used mass spectrometers have a limited working dynamic range that typically spans only three orders of magnitude within a single mass spectrum.

### Metabolomics

Metabolomics (Figure 1, target 4) is the study of the small-molecule end products of cellular processes that are the terminal downstream products of the genome, and consists of the total complement of all low-molecular-weight molecules that cellular processes leave behind [85].

### Background

Metabolomics may provide a viable supplement to genomics, transcriptomics and proteomics, to which it is intimately coupled. Metabolomics provides information furthest downstream from genomics; the key concept is that changes in the genome, transcriptome or proteome are reflected in the metabolome as alterations of metabolite concentration. In recent years, metabolomics has been revolutionized. Significant advances in computational and small-molecule detection tools allow the measurement of complex metabolic profiles in biological fluids [86].

Metabolomics strategies are divided into two distinct approaches, untargeted and targeted. Untargeted metabolomics is the comprehensive analysis of all measurable

analytes in a sample and offers the opportunity for novel biomarker discovery. Targeted metabolomics measures defined groups of chemically characterized and biochemically annotated metabolites. The most common techniques are high-resolution NMR spectrometry and mass spectrometry. NMR spectrometry exploits the behavior of molecules when placed in a magnetic field, allowing the identification of different nuclei based on their resonant frequency. Spectrometry is limited in their ability to identify more than a few small molecules. NMR is limited by its relatively insensitivity to very small amounts of molecules. Recent advances have been made by coupling NMR and mass spectrometry to the quantitative measurement of small-molecule metabolites in patient samples. This quantitative metabolomics approach makes it possible to associate changes in multiple metabolites to the diagnosis or characterization of disease processes. Since metabolomics has the capacity of identifying thousands of small molecules, it greatly improves our ability to characterize patterns of metabolites correlating with disease. Differences in metabolites may be predictive of disease severity, and changes over time may be useful in characterizing therapeutic response, disease progression or clinical outcome [87].

### Clinical utility

Metabolomics is positioned at a key point in the interpretation of any biological system because of its role as the downstream end product. Sepsis is a disease with significant disruption in biochemical homeostasis, and initial differences in metabolites may be predictive of disease severity and changes over time may be useful in characterizing therapeutic response. Recently, metabolomics has been applied in research in sepsis-induced acute lung injury [88]. When comparing 13 sepsis-induced acute lung injury patients with six healthy controls, the study found that distinct metabolites – including glutathione, adenosine, phosphatidylserine and sphingomylin – differed between the two groups. Not only did this pilot study demonstrate the feasibility of plasma H-NMR quantitative metabolomics, but it also justifies the continued study of this approach. Indeed, larger scale studies are needed to verify the potential of metabolomics in sepsis.

### Advantages

One of the most important advantages of metabolomics is the fact that the metabolome is relatively small compared with the other compartments. Around 5,000 unique molecules are estimated to be present. Metabolites are furthermore sensitive to biological perturbations and respond rapidly. Precise measurements are possible with available technologies. Lastly, new metabolite biomarkers may translate well to existing clinical chemistry laboratory technologies.

### Limitations

Untargeted metabolomics strategies are extremely time-consuming. Furthermore, there are difficulties in identifying and characterizing unknown small molecules and there tends to be a bias towards detection of highly abundant molecules [85]. At present, even the combination of a wide range of analytical tools allows us to see only a portion of the total metabolite complement of the cell. Furthermore, the physical and chemical properties of metabolites are highly divergent. This divergence means there is no single extraction process that does not incur substantial loss to some of the metabolites, let alone a single analytical platform that can measure all of the metabolites. A totally comprehensive approach is therefore lacking.

### Computational analysis in the omics setting

To make sense of the vast amounts of data generated by the omics technologies, analytical methodologies and tools are key requirements. A crucial step in the discipline of computational analysis is the process of automatically searching large volumes of data for patterns. We have entered the so-called  $p \gg n$  paradigm where the number of independent samples is substantially smaller than the number of variables (for example, the number of genes in an expression profile) [89]. In classical research settings a few prespecified null hypotheses are evaluated, whereas we are now simultaneously testing thousands of hypotheses.

The process of bioinformatics analysis can be divided into: data processing and quality control analysis; statistical data analysis; biological functions and pathways analysis; and data modeling in a system-wide context [90]. Data preprocessing and annotation involves transformation of raw machine data into readable and normalized data. Quality control assessment is a crucial first step in successful data analysis. Before any comparisons are performed, one must check that there were no problems with sample processing, and that samples are of sufficient quality to be included. After quality control, normalization is the next step, which is a transformation of signal values so that different sample results become comparable. For example the normalization for microarray data includes background correction, normalization of signal and summarization of signal values of probe sets for a transcript. Some of the well-known approaches include total intensity normalization [91], rank invariant methods [92] and locally weighted linear regression [93].

After normalization, genomics or proteomics variables from control and diseased groups are compared using various statistical models ( $P$  value, analysis of variation, signal-to-noise ratio, correlation) to identify variables that are specially associated with disease condition. To

reduce false positive results, multiple test-corrected statistical methods are employed. Typically these analyses yield a long list of variables (for example, gene, proteins, metabolites) that are significantly altered in the disease condition and require further pathways and functional enrichment analyses to understand the biological mechanism. Currently, a large number of commercial and academic software packages (for example, Ingenuity Systems, Cytoscape, GeneGO, Partek) are available for this purpose. These software packages integrate proteins/genes into biological pathways based on scientific literature by using natural language processors and expert human curation [94,95]. These analyses help in understanding the biological effects of genome-level variables induced in disease as well as yielding candidate pathways for therapeutic intervention. Furthermore, systems-level modeling of crosstalk or interaction among gene/proteins that are altered in disease is routinely explored to obtain a coherent systems-level view of the underlying biology. This modeling assists in generating the scale-free literature-driven networks to determine key regulatory nodes of the network that are essential for the stability of the network. Disruption of key regulatory nodes is considered to provide the most effective way to break a pathophysiological network, thus providing a potential method to design gene/protein-based effective therapies. In summary, high-level bioinformatics analysis will help to identify the key molecules associated with disease from thousands of molecules measured in genome-level assays.

### The development of diagnostics

The discovery of new biomarker targets is merely the first step in the comprehensive approach to developing new diagnostics in sepsis. After the biomarker candidate discovery, a derivation study is required in order to maximize the area under the curve and to choose thresholds that can optimize the sensitivity and specificity. In this derivation step, the targeted biomarker must perform superiorly compared with past biomarkers. Then follows a validation study, which typically requires the measurement of thousands of patient samples. Within this phase the analytical evaluation of the selected biomarkers is assessed. These include accuracy and predictability. Lastly, once the clinical evidence of a biomarker has been demonstrated, companies will determine whether the marker is worth pursuing from a technical, medical, financial and legal standpoint.

### Future directions

We so far lack a definitive gold standard biomarker that distinguishes sepsis from nonsepsis, or that reliably predicts outcome. The current literature is filled with numerous single-protein, or occasionally multi-protein,

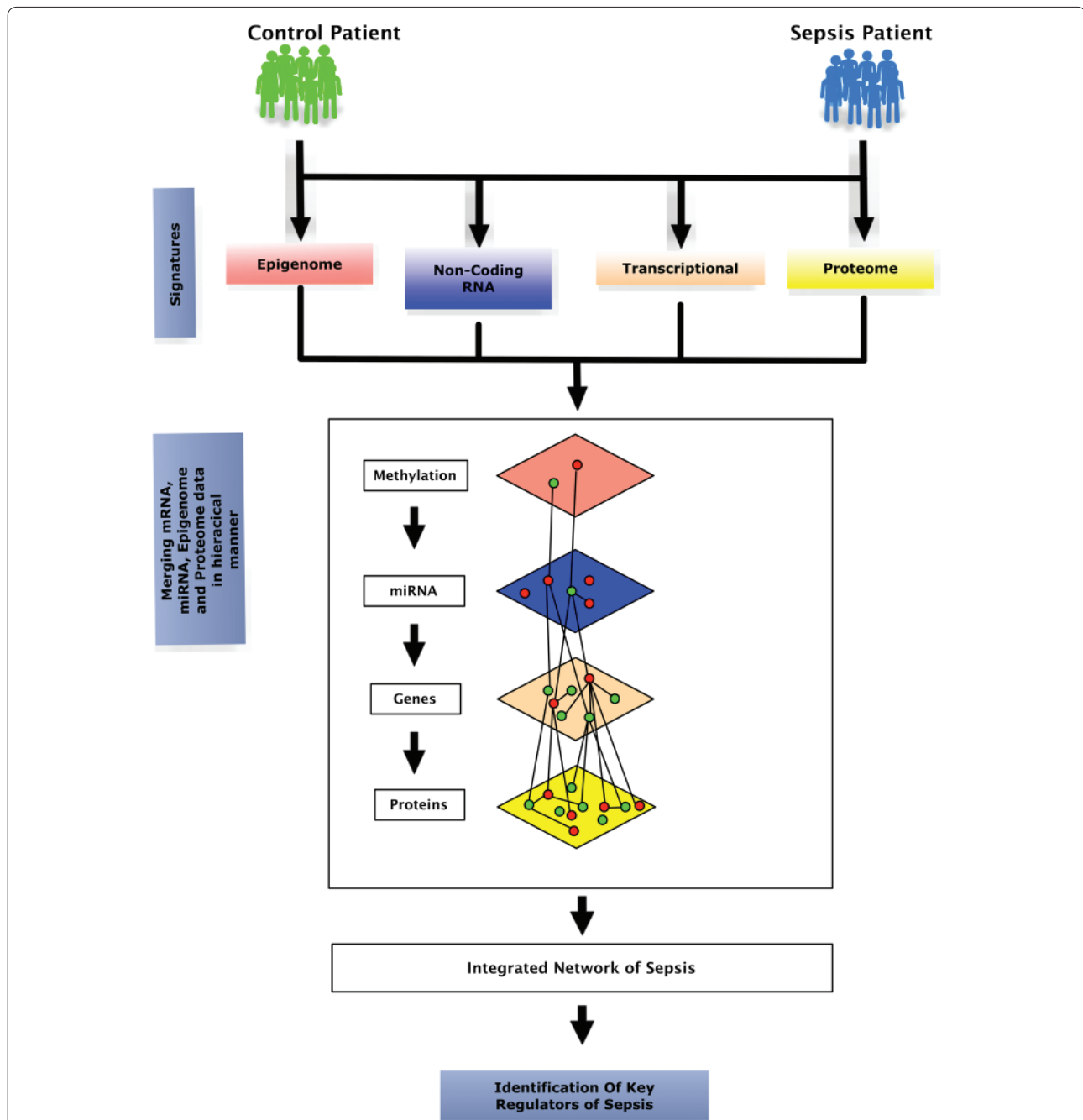
**Table 1. Overview of omics technologies: summary of strengths, limitations and clinical utility for each technology**

Omics	Strengths	Limitations	Clinical utility
<b>Genomics</b>			
SNP	<ul style="list-style-type: none"> <li>• Unbiased approach when using GWAS</li> <li>• Cost-effective large-scale genetic screening</li> <li>• Well-established analysis tools</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to find functional and structural gene variants</li> <li>• Only regulatory or coding regions are included</li> </ul>	<ul style="list-style-type: none"> <li>• Theragnostic approach</li> <li>• Risk stratification</li> </ul>
Epigenetics	<ul style="list-style-type: none"> <li>• Unbiased approach when using epigenome-wide association studies</li> <li>• Can elucidate the interplay between genetic and environmental factors</li> </ul>	<ul style="list-style-type: none"> <li>• Tissue-specific alterations</li> <li>• Different composition of cell types during sepsis</li> <li>• Frequency of epigenetic changes not known</li> <li>• Reverse causation</li> </ul>	<ul style="list-style-type: none"> <li>• Epigenetic signatures for sepsis diagnosis and/or prognosis</li> <li>• Prediction of therapeutic response</li> </ul>
<b>Transcriptomics</b>			
Expression profiling	<ul style="list-style-type: none"> <li>• Can generate global view transcriptome alterations</li> <li>• Provide good coverage of genome</li> <li>• Can elucidate alterations in signal transduction pathways during sepsis</li> </ul>	<ul style="list-style-type: none"> <li>• Tissue-specific expression of genes</li> <li>• Fails to measure low-expression genes with good sensitivity</li> </ul>	<ul style="list-style-type: none"> <li>• mRNA expression signatures for sepsis diagnosis and/or prognosis</li> <li>• Prediction of therapeutic response</li> </ul>
High-throughput gene sequencing (for example, RNA-seq)	<ul style="list-style-type: none"> <li>• Comprehensive sequence information</li> <li>• Unbiased approach</li> <li>• Estimates abundance of genes in term of copies</li> </ul>	<ul style="list-style-type: none"> <li>• Tissue-specific expression of genes</li> </ul>	<ul style="list-style-type: none"> <li>• No clinical utility</li> </ul>
miRNA	<ul style="list-style-type: none"> <li>• Stable in blood</li> <li>• Suggestive evidence that miRNAs play an important role in regulation of networks</li> <li>• The inclusion of miRNA when interpreting mRNA expression</li> </ul>	<ul style="list-style-type: none"> <li>• Functions not completely understood</li> </ul>	<ul style="list-style-type: none"> <li>• Novel diagnostic and/or prognostic biomarkers in sepsis.</li> <li>• Necessary for correctly interpretation of gene expression</li> </ul>
<b>Proteomics</b>			
	<ul style="list-style-type: none"> <li>• Provides global or unbiased alteration</li> <li>• Highly sensitive</li> <li>• No need for antibody-based technologies for measuring proteins</li> </ul>	<ul style="list-style-type: none"> <li>• Needs large amount of preprocessing or fractions</li> <li>• Current instruments unable to measure all proteins from complex biological fluids</li> <li>• Inefficient quantification of low expression proteins</li> </ul>	<ul style="list-style-type: none"> <li>• Novel diagnostic and/or prognostic biomarkers in sepsis</li> <li>• Prediction of therapeutic response</li> </ul>
<b>Metabolomics</b>			
	<ul style="list-style-type: none"> <li>• Relatively few targets</li> <li>• Good translation to existing laboratory technology</li> </ul>	<ul style="list-style-type: none"> <li>• Difficulty in identifying small molecules</li> <li>• Diverse physical and chemical properties and thus no single extraction tool</li> </ul>	<ul style="list-style-type: none"> <li>• Novel diagnostic and/or prognostic biomarkers in sepsis</li> <li>• Prediction of therapeutic response</li> <li>• Disease progression</li> </ul>

GWAS, genome-wide association studies.

markers in various stages of preclinical, translational and clinical investigation. However, the results have been somewhat disappointing, peak in their diagnostic accuracy or fail to validate. As this paper describes, there are upstream and downstream techniques that may find new and better targets. These techniques have both the potential to increase the spectrum of diagnostic and

prognostic biomarkers in sepsis, but they also have the potential to lead to the discovery of new disease pathways (Table 1). This may in turn lead us to improved targets for therapeutics. The incorporation of omics into the clinic has had successes in other fields. For example, expression signatures based on multigene sets are now used clinically for breast cancer prediction [96]. The use of



**Figure 4. An integrated analysis.** Integrated analysis of multidimensional genomics, epigenomics and proteomics data to capture the interaction between genetics, gene expression and regulatory RNA as well as proteomics. The analysis will enable identification of critical pathways or biological processes that drive the perturbation across multiple genome-level spaces, and thus are critical for disease pathophysiology.

GWAS found specific SNPs that can predict virologic response rates following specific treatment for hepatitis C [97].

It is our hope that this paper provides the reader with a basic understanding of the molecular biology and concepts across the spectrum of the omics technologies. Clinical utility and application in sepsis may lead to a

paradigm shift in diagnosis, management and our understanding of sepsis. The traditional flow of genetic information is from epigenome, genome and transcriptome to proteome and metabolome, but most studies focus on one space, thereby ignoring changes in other spaces. The biology of human disease is complex; we must therefore submit that a multidimensional view involving the input



from each genomic space is required to develop a true understanding. Studying the interaction and crosstalk of genomic information exchange between the epigenomic, genomic and proteomic space may assist in identifying core pathways that are continuously dysregulated, starting from epigenome to proteome (Figure 4).

### Additional file 1

**Additional file 1.** A document presenting a glossary of terms. Please refer to Glossary for common terms used in genomics, epigenetics, transcriptomics, proteomics, and metabolomics.

### Abbreviations

bp, base pairs; ELISA, enzyme-linked immunosorbent assay; GWAS, genome-wide association studies; HPLC, high-performance liquid chromatography; IL, interleukin; miRNA, microRNA; NF, nuclear factor; NMR, nuclear magnetic resonance; PCR, polymerase chain reaction; SNP, single nucleotide polymorphism; TNF, tumor necrosis factor.

### Acknowledgements

SS is funded by a research grant from Aarhus University, Denmark. NIS is funded in part by National Institutes of Health grants HL091757, GM076659, and SR01HL093234-02. WCA is supported by National Institutes of Health grants HL091757 and GM088184.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Department of Emergency Medicine and Center for Vascular Biology Research, Beth Israel Deaconess Medical Center and Harvard Medical School, 330 Brookline Avenue, Boston, MA 02215, USA. <sup>2</sup>Center for Vascular Biology Research, Beth Israel Deaconess Medical Center, 99 Brookline Street, Boston, MA 02215, USA. <sup>3</sup>Beth Israel Deaconess Medical Center Genomics and Core, 99 Brookline Avenue, Boston, MA 02115, USA.

Published: 4 October 2013

### References

1. Friedman G, Silva E, Vincent JL: **Has the mortality of septic shock changed with time.** *Crit Care Med* 1998, **26**:2078-2086.
2. Wheeler AP, Bernard GR: **Treating patients with severe sepsis.** *N Engl J Med* 1999, **340**:207-214.
3. Sands KE, Bates DW, Lanken PN, Graman PS, Hibberd PL, Kahn KL, Parsonnet J, Panzer R, Orav EJ, Snyderman DR, Black E, Schwartz JS, Moore R, Johnson BL Jr, Platt R; Academic Medical Center Consortium Sepsis Project Working Group: **Epidemiology of sepsis syndrome in 8 academic medical centers.** *JAMA* 1997, **278**:234-240.
4. Rangel-Frausto MS, Pittet D, Costigan M, Hwang T, Davis CS, Wenzel RP: **The natural history of the systemic inflammatory response syndrome (SIRS). A prospective study.** *JAMA* 1995, **273**:117-123.
5. Vincent J-L, Abraham E, Annane D, Bernard G, Rivers E, van den Berghe G: **Reducing mortality in sepsis: new directions.** *Crit Care* 2002, **6**(Suppl 3):S1-S18.
6. Regnault V, Levy B: **Recombinant activated protein C in sepsis: endothelium protection or endothelium therapy?** *Crit Care* 2007, **11**:103.
7. Hood L, Heath JR, Phelps ME, Lin B: **Systems biology and new technologies enable predictive and preventative medicine.** *Science* 2004, **306**:640-643.
8. Jain KK: **Role of oncoproteomics in the personalized management of cancer.** *Expert Rev Proteomics* 2004, **1**:49-55.
9. Wong HR: **Clinical review: sepsis and septic shock – the potential of gene arrays.** *Crit Care* 2012, **16**:204.
10. Wong HR: **Genetics and genomics in pediatric septic shock.** *Crit Care Med* 2012, **40**:1618-1626.
11. LaRosa SP, Opal SM: **Biomarkers: the future.** *Crit Care Clin* 2011, **27**:407-419.
12. Feezor RJ, Cheng A, Paddock HN, Baker HV, Moldawer LL: **Functional genomics and gene expression profiling in sepsis: beyond class prediction.** *Clin Infect Dis* 2005, **41**(Suppl 7):S427-S435.
13. Cobb JP, O'Keefe GE: **Injury research in the genomic era.** *Lancet* 2004, **363**:2076-2083.
14. Nguyen A, Yaffe MB: **Proteomics and systems biology approaches to signal transduction in sepsis.** *Crit Care Med* 2003, **31**(1 Suppl):S1-S6.
15. Cornell TT, Wynn J, Shanley TP, Wheeler DS, Wong HR: **Mechanisms and regulation of the gene-expression response to sepsis.** *Pediatrics* 2010, **125**:1248-1258.
16. Namath A, Patterson AJ: **Genetic polymorphisms in sepsis.** *Crit Care Clin* 2009, **25**:835-856, x.
17. Prucha M, Ruryk A, Boriss H, Moller E, Zazula R, Herold I, Claus RA, Reinhart KA, Deigner P, Russwurm S: **Expression profiling: toward an application in sepsis diagnostics.** *Shock* 2004, **22**:29-33.
18. Tang BM, McLean AS, Dawes IW, Huang SJ, Lin RC: **The use of gene-expression profiling to identify candidate genes in human sepsis.** *Am J Respir Crit Care Med* 2007, **176**:676-684.
19. Holmes CL, Russell JA, Walley KR: **Genetic polymorphisms in sepsis and septic shock: role in prognosis and potential for therapy.** *Chest* 2003, **124**:1103-1115.
20. Sorensen TI, Nielsen GG, Andersen PK, Teasdale TW: **Genetic and environmental influences on premature death in adult adoptees.** *N Engl J Med* 1988, **318**:727-732.
21. International HapMap Consortium: **A haplotype map of the human genome.** *Nature* 2005, **437**:1299-1320.
22. Altshuler D, Pollara VJ, Cowles CR, Van Etten WJ, Baldwin J, Linton L, Lander ES: **An SNP map of the human genome generated by reduced representation shotgun sequencing.** *Nature* 2000, **407**:513-516.
23. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K: **dbSNP: the NCBI database of genetic variation.** *Nucleic Acids Res* 2001, **29**:308-311.
24. UCSC Genome Bioinformatics [http://genome.ucsc.edu]
25. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D: **The human genome browser at UCSC.** *Genome Res* 2002, **12**:996-1006.
26. Khor CC, Vannberg FO, Chapman SJ, Guo H, Wong SH, Walley AJ, Vukcevic D, Rautanen A, Mills TC, Chang KC, Kam KM, Crampin AC, Ngwira B, Leung CC, Tam CM, Chan CY, Sung JJ, Yew WW, Toh KY, Tay SK, Kwiatkowski D, Lienhardt C, Hien TT, Day NP, Peshu N, Marsh K, Maitland K, Scott JA, Williams TN, Berkley JA, Floyd S, Tang NL, Fine PE, Goh DL, Hill AV: **CISH and susceptibility to infectious diseases.** *N Engl J Med* 2010, **362**:2092-2101.
27. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K: **A comprehensive review of genetic association studies.** *Genet Med* 2002, **4**:45-61.
28. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA: **Potential etiologic and functional implications of genome-wide association loci for human diseases and traits.** *Proc Natl Acad Sci USA* 2009, **106**:9362-9367.
29. Lorenz E, Mira JP, Frees KL, Schwartz DA: **Relevance of mutations in the TLR4 receptor in patients with gram-negative septic shock.** *Arch Intern Med* 2002, **162**:1028-1032.
30. Gibot S, Cariou A, Drouet L, Rossignol M, Ripoll L: **Association between a genomic polymorphism within the CD14 locus and septic shock susceptibility and mortality rate.** *Crit Care Med* 2002, **30**:969-973.
31. Mira JP, Cariou A, Grall F, Delclaux C, Losser MR, Heshmati F, Cheval C, Monchi M, Teboul JL, Riché F, Leleu G, Arbibe L, Mignon A, Delpech M, Dhainaut JF: **Association of TNF2, a TNF-alpha promoter polymorphism, with septic shock susceptibility and mortality: a multicenter study.** *JAMA* 1999, **282**:561-568.
32. Tang GJ, Huang SL, Yien HW, Chen WS, Chi CW, Wu CW, Lui WY, Chiu JH, Lee TY: **Tumor necrosis factor gene polymorphism and septic shock in surgical infection.** *Crit Care Med* 2000, **28**:2733-2736.
33. Stuber F, Petersen M, Bokelmann F, Schade U: **A genomic polymorphism within the tumor necrosis factor locus influences plasma tumor necrosis factor-alpha concentrations and outcome of patients with severe sepsis.** *Crit Care Med* 1996, **24**:381-384.
34. Majetschak M, Flohe S, Obertacke U, Schroder J, Staubach K, Nast-Kolb D, Schade FU, Stuber F: **Relation of a TNF gene polymorphism to severe sepsis in trauma patients.** *Annals Surg* 1999, **230**:207-214.
35. Fang XM, Schroder S, Hoeft A, Stuber F: **Comparison of two polymorphisms of the interleukin-1 gene family: interleukin-1 receptor antagonist polymorphism contributes to susceptibility to severe sepsis.** *Crit Care Med*

- 1999, **27**:1330-1334.
36. Medzhitov R, Janeway C, Jr: **Innate immunity.** *N Engl J Med* 2000, **343**:338-344.
37. Walley KR, Russell JA: **Protein C –1641 AA is associated with decreased survival and more organ dysfunction in severe sepsis.** *Crit Care Med* 2007, **35**:12-17.
38. Poltorak A, He X, Smirnova I, Liu MY, Van Huffel C, Du X, Birdwell D, Alejos E, Silva M, Galanos C, Freudenberg M, Ricciardi-Castagnoli P, Layton B, Beutler B: **Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene.** *Science* 1998, **282**:2085-2088.
39. Lin MT, Albertson TE: **Genomic polymorphisms in sepsis.** *Crit Care Med* 2004, **32**:569-579.
40. Angus DC, Burgner D, Wunderink R, Mira JP, Gerlach H, Wiedermann CJ, Vincent JL: **The PIRO concept: P is for predisposition.** *Crit Care* 2003, **7**:248-251.
41. Risch N, Merikangas K: **The future of genetic studies of complex human diseases.** *Science* 1996, **273**:1516-1517.
42. Mooney SD, Krishnan VG, Evani US: **Bioinformatic tools for identifying disease gene and SNP candidates.** In *Genetic Variation; Methods and Protocols, Methods in Molecular Biology*. Edited by Barnes MR, Breen G. New York: Springer; 2010.
43. Sutherland AM, Russell JA: **Issues with polymorphism analysis in sepsis.** *Clin Infect Dis* 2005, **41**(Suppl 7):S396-S402.
44. Beck S: **Taking the measure of the methylome.** *Nat Biotechnol* 2010, **28**:1026-1028.
45. Fatemi M, Pao MM, Jeong S, Gal-Yam EN, Egger G, Weisenberger DJ, Jones PA: **Footprinting of mammalian promoters: use of a CpG DNA methyltransferase revealing nucleosome positions at a single molecule level.** *Nucleic Acids Res* 2005, **33**:e176.
46. Huang TH, Perry MR, Laux DE: **Methylation profiling of CpG islands in human breast cancer cells.** *Hum Mol Genet* 1999, **8**:459-470.
47. Esteller M: **Cancer epigenomics: DNA methylomes and histone-modification maps.** *Nat Rev Genet* 2007, **8**:286-298.
48. Saccani S, Pantano S, Natoli G: **p38-Dependent marking of inflammatory genes for increased NF-kappa B recruitment.** *Nat Immunol* 2002, **3**:69-75.
49. Phillips T: **The role of methylation in gene expression.** *Nat Educ* 2008, **1**:1.
50. Bernstein BE, Meissner A, Lander ES: **The mammalian epigenome.** *Cell* 2007, **128**:669-681.
51. Rakyan VK, Down TA, Balding DJ, Beck S: **Epigenome-wide association studies for common human diseases.** *Nat Rev Genet* 2011, **12**:529-541.
52. Adams J: **Transcriptome: connecting the genome to gene function.** *Nat Educ* 2008, **1**:1.
53. Huang Q, Liu D, Majewski P, Schulte LC, Korn JM, Young RA, Lander ES, Hacohen N: **The plasticity of dendritic cell responses to pathogens and their components.** *Science* 2001, **294**:870-875.
54. Brown PO, Botstein D: **Exploring the new world of the genome with DNA microarrays.** *Nat Genet* 1999, **21**(1 Suppl):33-37.
55. Shih S-C, Robinson GS, Perruzzi CA, Calvo A, Desai K, Green JE, Ali IU, Smith LEH, Senger DR: **Molecular profiling of angiogenesis markers.** *Am J Pathol* 2002, **161**:35-41.
56. Bassett DE, Jr, Eisen MB, Boguski MS: **Gene expression informatics – it's all in your mine.** *Nat Genet* 1999, **21**(1 Suppl):51-55.
57. Ding C, Cantor CR: **Quantitative analysis of nucleic acids – the last few years of progress.** *J Biochem Mol Biol* 2004, **37**:1-10.
58. Tang BM, McLean AS, Dawes IW, Huang SJ, Lin RC: **Gene-expression profiling of peripheral blood mononuclear cells in sepsis.** *Crit Care Med* 2009, **37**:882-888.
59. Johnson SB, Lissauer M, Bochicchio GV, Moore R, Cross AS, Scalea TM: **Gene expression profiles differentiate between sterile SIRS and early sepsis.** *Annals Surg* 2007, **245**:611-621.
60. Tang BM, Huang SJ, McLean AS: **Genome-wide transcription profiling of human sepsis: a systematic review.** *Crit Care* 2010, **14**:R237.
61. Pachot A, Lepape A, Vey S, Bienvenu J, Mouglin B, Monneret G: **Systemic transcriptional analysis in survivor and non-survivor septic shock patients: a preliminary study.** *Immunol Lett* 2006, **106**:63-71.
62. Wong HR, Wheeler DS, Tegtmeyer K, Poynter SE, Kaplan JM, Chima RS, Stalets E, Basu RK, Doughty LA: **Toward a clinically feasible gene expression-based subclassification strategy for septic shock: proof of concept.** *Crit Care Med* 2010, **38**:1955-1961.
63. Cobb JP, Hayden DL, Schoenfeld DA: **Novel diagnostics for sepsis: a decade of promise for gene expression profiling.** *Crit Care Med* 2011, **39**:2579-2581.
64. Anglicheau D, Sharma VK, Ding R, Hummel A, Snopkowski C, Dadhania D, Seshan SV, Suthanthiran M: **MicroRNA expression profiles predictive of human renal allograft status.** *Proc Natl Acad Sci U S A* 2009, **106**:5330-5335.
65. Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B, Rigoutsos I: **A pattern-based method for the identification of microRNA binding sites and their corresponding heteroduplexes.** *Cell* 2006, **126**:1203-1217.
66. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Brian KC, Allen A, Lin DW, Urban N, Drescher CW, Knudsen BS, Stirewalt DL, Gentleman R, Vessella RL, Nelson PS, Martin DB, Tewari M: **Circulating microRNAs as stable blood-based markers for cancer detection.** *Proc Natl Acad Sci U S A* 2008, **105**:10513-10518.
67. Kroh EM, Parkin RK, Mitchell PS, Tewari M: **Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR).** *Methods* 2010, **50**:298-301.
68. Fan J-B: *Next-generation microRNA Expression Profiling Technology: Methods and Protocols.* New York: Springer; 2012.
69. Tili E, Michaille JJ, Cimino A, Costinean S, Dumitru CD, Adair B, Fabbri M, Alder H, Liu CG, Calin GA, Croce CM: **Modulation of miR-155 and miR-125b levels following lipopolysaccharide/TNF-alpha stimulation and their possible roles in regulating the response to endotoxin shock.** *J Immunol* 2007, **179**:5082-5089.
70. Vasilescu C, Rossi S, Shimizu M, Tudor S, Veronese A, Ferracin M, Nicoloso MS, Barbarotto E, Popa M, Stanculea O, Fernandez MH, Tulbure D, Bueso-Ramos CE, Negrini M, Calin GA: **MicroRNA fingerprints identify miR-150 as a plasma prognostic marker in patients with sepsis.** *PLoS One* 2009, **4**:e7405.
71. Schmidt WM, Spiel AO, Jilma B, Wolzt M, Muller M: **In vivo profile of the human leukocyte microRNA response to endotoxemia.** *Biochem Biophys Res Commun* 2009, **380**:437-441.
72. Creemers EE, Tijssen AJ, Pinto YM: **Circulating microRNAs: novel biomarkers and extracellular communicators in cardiovascular disease?** *Circ Res* 2012, **110**:483-495.
73. Rogers S, Girolami M, Kolch W, Waters KM, Liu T, Thrall B, Wiley HS: **Investigating the correspondence between transcriptomic and proteomic expression profiles using coupled cluster models.** *Bioinformatics* 2008, **24**:2894-2900.
74. Dhingra V, Gupta M, Andacht T, Fu ZF: **New frontiers in proteomics research: a perspective.** *Int J Pharm* 2005, **299**:1-18.
75. Kocher T, Superti-Furga G: **Mass spectrometry-based functional proteomics: from molecular machines to protein networks.** *Nat Methods* 2007, **4**:807-815.
76. MD SPL, MD SMO: **Biomarkers: the future.** *Crit Care Clin* 2011, **27**:407-419.
77. Paugam-Burtz C, Albuquerque M, Baron G, Bert F, Voitot H, Delefosse D, Dondero F, Sommacale D, Francoz C, Hanna N, Belghiti J, Ravaut P, Bedossa P, Mantz J, Paradis V: **Plasma proteome to look for diagnostic biomarkers of early bacterial sepsis after liver transplantation: a preliminary study.** *Anesthesiology* 2010, **112**:926-935.
78. Kalenka A, Feldmann RE, Jr, Otero K, Maurer MH, Waschke KF, Fiedler F: **Changes in the serum proteome of patients with sepsis and septic shock.** *Anesth Analg* 2006, **103**:1522-1526.
79. Service RF: **Proteomics. Will biomarkers take off at last?** *Science* 2008, **321**:1760.
80. Anderson L, Seilhamer J: **A comparison of selected mRNA and protein abundances in human liver.** *Electrophoresis* 1997, **18**:533-537.
81. Schiess R, Wollscheid B, Aebersold R: **Targeted proteomic strategy for clinical biomarker discovery.** *Mol Oncol* 2009, **3**:33-44.
82. States DJ, Omenn GS, Blackwell TW, Fermin D, Eng J, Speicher DW, Hanash SM: **Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study.** *Nat Biotechnol* 2006, **24**:333-338.
83. Putnam FW: **The trace components of plasma: an overview.** *Prog Clin Biol Res* 1976, **5**:1-24.
84. Anderson NL, Anderson NG: **The human plasma proteome: history, character, and diagnostic prospects.** *Mol Cell Proteomics* 2002, **1**:845-867.
85. Roberts LD, Souza AL, Gerszten RE, Clish CB: **Targeted metabolomics.** *Curr Protoc Mol Biol* 2012, **Chapter 30**:Unit 30.2.1-24.
86. Lacy P: **Metabolomics of sepsis-induced acute lung injury: a new approach for biomarkers.** *Am J Physiol Lung Cell Mol Physiol* 2011, **300**:L1-L3.
87. Serkova NJ, Standiford TJ, Stringer KA: **The emerging field of quantitative blood metabolomics for biomarker discovery in critical illnesses.** *Am J Respir Crit Care Med* 2011, **184**:647-655.
88. Stringer KA, Serkova NJ, Karnovsky A, Guire K, Paine R, 3rd, Standiford TJ:

- Metabolic consequences of sepsis-induced acute lung injury revealed by plasma (1)H-nuclear magnetic resonance quantitative metabolomics and computational analysis. *Am J Physiol Lung Cell Mol Physiol* 2011, **300**:L4-L11.
89. Dunkler D, Sanchez-Cabo F, Heinze G: **Statistical analysis principles for omics data.** *Methods Mol Biol* 2011, **719**:113-131.
90. Schneider MV, Orchard S: **Omics technologies, data and bioinformatics principles.** *Methods Mol Biol* 2011, **719**:3-30.
91. Quackenbush J: **Microarray data normalization and transformation.** *Nat Genet* 2002, **32**(Suppl):496-501.
92. Tseng GC, Oh MK, Rohlin L, Liao JC, Wong WH: **Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects.** *Nucleic Acids Res* 2001, **29**:2549-2557.
93. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP: **Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation.** *Nucleic Acids Res* 2002, **30**:e15.
94. Panda S, Antoch MP, Miller BH, Su AI, Schook AB, Straume M, Schultz PG, Kay SA, Takahashi JS, Hogenesch JB: **Coordinated transcription of key pathways in the mouse by the circadian clock.** *Cell* 2002, **109**:307-320.
95. Jarvis RM, Goodacre R: **Genetic algorithm optimization for pre-processing and variable selection of spectroscopic data.** *Bioinformatics* 2005, **21**:860-868.
96. Oakman C, Santarpia L, Di Leo A: **Breast cancer assessment tools and optimizing adjuvant therapy.** *Nat Rev Clin Oncol* 2010, **7**:725-732.
97. Clark PJ, Thompson AJ, McHutchison JG: **IL28B genomic-based treatment paradigms for patients with chronic hepatitis C infection: the future of personalized HCV therapies.** *Am J Gastroenterol* 2011, **106**:38-45.
98. Alberts B: *Molecular Biology of the Cell*. 5th edition. New York: Garland Science; 2008.

doi:10.1186/cc12693

**Cite this article as:** Skibsted S, *et al.*: **Bench-to-bedside review: Future novel diagnostics for sepsis – a systems biology approach.** *Critical Care* 2013, **17**:231.