Check for updates

OPEN ACCESS

EDITED BY George Grant, Independent Researcher, Aberdeen, United Kingdom

REVIEWED BY Georgia Damoraki, National and Kapodistrian University of Athens, Greece Chih-Hao Fang, Amazon, United States

*CORRESPONDENCE Bangjiang Fang ☑ fangbji@163.com KaiLiang Fan ☑ 18560769418@163.com Chuanqi Wan ☑ wcqxy1995@126.com

^tThese authors have contributed equally to this work

RECEIVED 25 April 2025 ACCEPTED 10 June 2025 PUBLISHED 26 June 2025

CITATION

Lu J, Zhang W, He Y, Jiang M, Liu Z, Zhang J, Zheng L, Zhou B, Luo J, He C, Shan Y, Zhang R, Fan K, Fang B and Wan C (2025) Multi-omics decodes host-specific and environmental microbiome interactions in sepsis. *Front. Microbiol.* 16:1618177. doi: 10.3389/fmicb.2025.1618177

COPYRIGHT

© 2025 Lu, Zhang, He, Jiang, Liu, Zhang, Zheng, Zhou, Luo, He, Shan, Zhang, Fan, Fang and Wan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Multi-omics decodes host-specific and environmental microbiome interactions in sepsis

Jiamin Lu^{1,2†}, Wen Zhang^{1,2†}, Yuzhou He^{3†}, Mei Jiang⁴, Zhankui Liu², Jirong Zhang^{1,2}, Lanzhi Zheng⁵, Bingzhi Zhou³, Jielian Luo^{1,2}, Chenming He^{1,2}, Yunan Shan⁶, Runze Zhang^{1,2}, KaiLiang Fan⁷*, Bangjiang Fang^{1,2}* and Chuanqi Wan^{1,2}*

¹Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China, ²Institute of Acute and Critical Care, Shanghai University of Traditional Chinese Medicine, Shanghai, China, ³The Second Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China, ⁴Tianjin University of Traditional Chinese Medicine, Tianjin, China, ⁵The First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China, ⁶School of Traditional Chinese Medicine, Hubei University of Chinese Medicine, Wuhan, China, ⁷The Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan, China

Sepsis is a life-threatening organ dysfunction caused by a dysregulated host response to infection, and its pathogenesis involves complex interactions between the host and the microbiome. The integration of multi-omics has important value in revealing the mechanism of host-microbiome interaction. It is a key tool for promoting accurate diagnosis and guiding dynamic treatment strategies in sepsis. However, multi-omics data integration faces technical challenges, such as data heterogeneity and platform variability, as well as analytical hurdles, such as the "curse of dimensionality." Fortunately, researchers have developed two integration strategies: data-driven and knowledge-guided approaches, which employ various dimensionality reduction techniques and integration methods to handle multi-omics datasets. This review discusses the applications of multiomics technologies in host-microbiome interactions in sepsis, highlighting their potential in identifying novel diagnostic biomarkers and developing personalized and dynamic treatment strategies. It also summarizes commonly used systems biology resources and computational tools for data integration; the review outlines the challenges in this field and proposes potential directions for future studies.

KEYWORDS

bioinformatics tools, comparative genome analysis, microbiome, multi-omics, sepsis

1 Introduction

Sepsis is a life-threatening disease with a global impact, posing a severe threat to humans and representing one of the significant challenges in global healthcare. It is characterized by high incidence and mortality rates, with data indicating that it causes approximately 5.3 million deaths annually worldwide (Evans et al., 2021; Fleischmann et al., 2016). Although the ability to treat patients with sepsis is improved, the mortality rate of sepsis remains unacceptably high. Therefore, identifying novel pathogenic mechanisms is critical for improving outcomes in sepsis patients. Sepsis can be classified by severity into sepsis and septic shock, with the latter representing the severe stage of sepsis. This study encompasses research related to both sepsis and septic shock (Singer et al., 2016).

Over the past two decades, research on the microbiome has grown exponentially. The human microbiome consists of diverse microorganisms, including bacteria, viruses, fungi, and archaea, which coexist symbiotically with the human host and play a crucial role in maintaining



homeostasis (Koppel and Balskus, 2016). Dysbiosis of the microbial community is both a consequence and a contributing factor in the pathogenesis of sepsis, occupying a critical position in its progression. This imbalance can affect the host's immune response, metabolic processes, and barrier function, thereby influencing the outcome of sepsis (Weersma et al., 2020). Therefore, understanding the complex interactions between the host and microbiome in sepsis is of significant value for identifying novel diagnostic biomarkers and developing personalized and dynamic therapeutic strategies.

Genomic sepsis studies can elucidate associations between host genetic variations and susceptibility (Rayzan et al., 2023). Metagenomics reveals changes in microbial diversity and function (Kalantar et al., 2022). Transcriptomics aids in identifying sepsis biomarkers and infection types (Chen et al., 2024). Proteomics uncovers host immune responses and metabolic remodeling processes (Miao et al., 2021). Metabolomics assists in sepsis diagnosis and therapeutic monitoring (Pandey, 2024). Epigenomics clarifies epigenetic mechanisms by which microbes influence host immunity (Córneo et al., 2021). However, single-omics analyses provide an incomplete perspective. Therefore, integrating multiomics approaches enables a comprehensive and systematic dissection of biological systems. First, this integration helps identify novel biomarkers of host-microbe interactions in sepsis, such as specific microbes and metabolites, thereby enhancing diagnostic accuracy (Khan et al., 2019). Second, it facilitates the development of personalized therapeutic strategies, such as

Abbreviations: CAZy, Carbohydrate-Active enZYmes Database; DR, Dimension Reduction; FMT, Fecal Microbiota Transplantation; GSEA, Gene Set Enrichment Analysis; GSA, Gene set variation analysis; GSVA, gene set variation analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC–MS, Liquid Chromatograph Mass Spectrometer; LDA, Linear Discriminant Analysis; MS, Mass Spectrometry; NMDS, Non-metric multidimensional scaling; PCA, Principal Component Analysis; scRNA-seq, Single-cell RNA sequencing; SNF, Similarity Network Fusion; t-SNE, t-Stochastic Neighbor Embedding; VFDB, Virulence Factor Database; WGCNA, Weighted correlation network analysis; WGS, Whole Genome Sequencing; WES, Whole Exome Sequencing.

modulating gut microbiota and precision drug administration (Isac et al., 2024). Thirdly, multi-omics technologies enable real-time therapeutic efficacy monitoring, predict patient outcomes, and assist clinicians in promptly adjusting treatment regimens (Sun et al., 2023).

The reproducibility of multi-omics data, combined with advanced modeling techniques, facilitates comprehensive and integrative analysis of these complex systems (Santiago-Rodriguez and Hollister, 2021). Public datasets enable the reuse of high-quality data, while machine learning and network analysis tools uncover patterns and relationships within them. Integrative multi-omics approaches, such as multi-omics factor analysis and systems biology modeling, allow for simultaneous analysis across multiple biological levels, revealing correlations and causal relationships between omics layers. We posit that in sepsis research, there is a need for further synthesis and integration of multi-omics approaches and technologies, as well as a holistic perspective host-microbe interrelationships on and interactions.

2 Omics studies of host and microbiome in sepsis

Due to its high incidence and mortality, sepsis remains a priority health concern highlighted by the World Health Organization (WHO) (Rudd et al., 2020). As such, omics technologies have increasingly been applied to explore sepsis's pathogenesis and therapeutic targets. Kurian et al. (2024) demonstrated that research on sepsis utilizing omics technologies has shown a growing trend in publications from the European Union and the United Kingdom, with microbiology being one of the primary research directions. Notably, studies in this field have proliferated since 2000, with 1,608 articles published between 2011 and May 2023. These findings provide new insights into the pathophysiology of sepsis and contribute to rapid diagnosis, targeted therapy, and personalized medicine (Ahmed, 2022).

2.1 Genomics and macrogenomics

Genomics elucidates the relationship between host genetic variations and sepsis susceptibility, therapeutic responses, and clinical outcomes by analyzing the genetic material in host somatic cells. Standard detection techniques include whole genome sequencing (WGS), whole exome sequencing (WES), and single nucleotide polymorphism (SNP) genotyping (Ahmed et al., 2021). Several studies have used WES to reveal the impact of rare immune-deficiency gene variants on sepsis susceptibility in pediatric patients (Rayzan et al., 2023; Borghesi et al., 2020). However, the static nature of genomics limits in-depth exploration of the dynamic course of sepsis. Future research needs to shift toward studying phenotypes and dynamic gene expression.

Metagenomics enables sequencing microbial genomes (archaea, bacteria, viruses, and fungi) present in samples, providing critical insights into microbial diversity and functional potential (Almeida et al., 2019). Common approaches include amplicon-based marker gene analysis (e.g., 16S rRNA gene sequencing) and metagenomic shotgun sequencing. It will be interesting to mention that besides 16S rRNA gene, which is useful for identifying bacterial species on a

broader taxonomic scale, more exact identification requires further genetic techniques, such as Nanopore sequencing, which produces long sequencing reads (Abedini-Nassab, 2017; Ying et al., 2013). This 3rd generation sequencing technology is able to sequence the whole 16S rRNA gene and not just some of its variable regions. Metagenomic sequencing has revealed significant alterations in the fecal microbiota of sepsis patients, such as increased abundances of Bacillota, Bacteroidota, Lactobacillaceae, and opportunistic pathogens such as Klebsiella spp. and Escherichia-Shigella in septic rats (Sun et al., 2020). Functional analyses of phyla such as Bacteroidota and Proteobacteria help elucidate the dysbiosis associated with sepsis (Daliri et al., 2021). Additionally, metagenomic data on the nasal microbiota, including genera such as Staphylococcus spp., Moraxella, and Streptococcus, enhance predictive diagnostic capabilities for patients with lower respiratory tract infections (Li et al., 2022). The 16S rRNA gene sequencing technique provides critical microbiomic evidence for the subtyping and diagnosis of sepsis, facilitating the elucidation of the microbial mechanisms underlying sepsis susceptibility. Liu et al. (2021) employed this technology to demonstrate a strong correlation between gut microbiota composition and sepsis subtype susceptibility, proposing two dysbiosis models: ICU Enterotype I is characterized by a predominance of Bacteroides and unclassified Enterobacteriaceae, with hosts more prone to progressing to septic shock. ICU Enterotype II is dominated by Enterococcus spp., corresponding to hosts predominantly exhibiting a sepsis phenotype without concomitant shock. The application of whole-genome sequencing (WGS) for precision monitoring microbial genomes facilitates accurate clinical diagnosis of infectious microorganisms, enhancing diagnostic accuracy in clinical settings (Huang et al., 2019).

2.2 Transcriptomics and metatranscriptomics

The mRNA offers insights into the pathophysiology of sepsis patients. Whole-blood transcriptomics identifies biologically homogeneous subgroups by revealing differentially expressed genes and detects dynamic changes in sepsis (Saxena et al., 2024). In recent years, transcriptomics-based biomarkers have shown great potential in diagnosing, disease monitoring, and prognosis evaluation of sepsis. Some studies have conducted transcriptomic analyses on whole-blood RNA samples from sepsis patients and combined these with bioinformatics methods to identify nine genes, such as LRG1, ELANE, and TP53, as potential biomarkers for sepsis (Gong et al., 2020). The newly developed IMX-BVN-1 classifier, utilizing 29 preselected host mRNAs, employs a neural network to distinguish between bacterial and viral infections. This approach offers a novel method for rapid diagnosis in sepsis patients (Mayhew et al., 2020). Transcriptomics not only aids in rapidly differentiating infection types but also guides personalized sepsis treatment. Cano-Gamez et al. (2022) constructed a crossplatform transcriptomic reference map using transcriptomic data from different technical platforms. They proposed an immune dysfunction score (SRSq) for sepsis patient stratification, reflecting the degree of immune dysregulation and predicting clinical outcomes, thus providing new directions for early sepsis diagnosis and personalized treatment. Two subclasses of pediatric infectious shock patients were identified through genome-wide expression profiling based on whole blood RNA sequencing, providing information for clinical decision-making in sepsis (Yang et al., 2023). Kalantar et al. (2022) combined host transcriptional profiling with broad-range metagenomic pathogen detection from nucleic acids to develop a novel diagnostic tool for sepsis.

Metatranscriptomics can compensate for the limitations of metagenomics by analyzing microbial transcription levels to reveal their metabolic and functional states (Santiago-Rodriguez et al., 2015). Metagenomics focuses on the genetic potential of microbial communities, while metatranscriptomics provides insights into the actual gene expression and functional activities of microbes under specific conditions (Franzosa et al., 2014). This approach is beneficial for understanding microbial responses to environmental changes and their dynamic metabolic processes. Wang et al. (2023) performed metatranscriptomic sequencing on bronchoalveolar lavage fluid (BALF) from COVID-19 patients, analyzing host transcriptomic profiles, viral, bacterial, and fungal content, as well as virulence factors. They found that SARS-CoV-2, human β -herpesvirus, and phyla such as Proteobacteria and *Bacillota* were highly represented. They revealed a significant correlation between microbial composition and host immune responses.

2.3 Proteomics

Proteomic profiling of host body fluids such as blood and urine, using high-resolution mass spectrometry technologies such as LC-MS/MS to analyze post-translationally modified proteins, can reveal the immune responses and metabolic remodeling processes in sepsis (Duong and Lee, 2023). Proteomic profiling, based on shotgun proteomics and utilizing tandem mass spectrometry (MS) to identify species-specific peptides, distinguishes microbes at the amino acid level. This technique, when applied to analyze clinical samples from sepsis patients, aids in diagnosing sepsis infections (van Houten et al., 2018). However, its clinical application is limited due to technical complexity and high costs. Wang C. et al. (2021) performed a quantitative proteomic analysis of neutrophil proteins in the blood of sepsis patients and analyzed the blood microbiota. They found significant changes in different stages of sepsis, with interactions between microbiota changes and immune cell functional alterations. Bacterial genera were identified as potential predictive biomarkers for sepsis, providing new research and clinical management directions. Liu et al. (2019) performed proteomic analysis of fecal samples from sepsis patients and combined it with 16S rDNA sequencing and metabolomics. They found that sepsis leads to significant changes in gut microbiota protein expression, closely related to immune responses and coagulation function. These findings highlight the potential of gut microbiota proteins as biomarkers for sepsis.

2.4 Metabolomics

Metabolomics aids in sepsis diagnosis by detecting host metabolites (e.g., in blood, urine, and tissues) and microbial metabolites (e.g., SCFAs and amino acids), revealing host-microbe interactions and their impact on health and disease. Commonly used techniques include nuclear magnetic resonance (NMR) and liquid chromatography-mass spectrometry (LC–MS) (Bauermeister et al., 2022). Analysis of gut fecal metabolites can serve as an early non-invasive diagnostic tool for sepsis, particularly for late-onset sepsis in preterm infants caused by Gram-negative bacteria. Specific

metabolic markers, such as ethyl acetate and cyclopentane, are associated with Gram-negative late-onset sepsis. Additionally, octanal has been identified as a unique metabolic marker for late-onset sepsis caused by coagulase-negative staphylococci (Frerichs et al., 2023).

Metabolomics data can be utilized to monitor therapeutic effects, particularly by analyzing specific metabolites correlated with disease outcomes, thereby supporting personalized treatment strategies. For instance, metabolomic analysis in mice demonstrated that gut microbiota-derived short-chain fatty acids (SCFAs), including acetate, propionate, and butyrate, were significantly associated with reduced levels of Lactobacillus and Bifidobacterium in the microbiota of mice with Klebsiella pneumonia induced pneumonic sepsis. These findings highlight the therapeutic potential of SCFAs in sepsis (Wu et al., 2020b). Wang et al. (2024b) found that changes in the gut microbiota and metabolites of sepsis patients are closely related to changes in serum vitamin levels. Vitamin B9 protects the intestinal barrier and reduces the severity of sepsis by altering the type and abundance of gut microbiota and upregulating the concentration of microbiota metabolites, thereby increasing the expression of intestinal barrierrelated genes. Stewart et al. (2017) conducted metabolomic profiling of fecal samples from preterm infants with late-onset sepsis, combined with 16S rRNA sequencing. Their findings suggest that increased prebiotic oligosaccharides and the growth of Bifidobacterium in the gut exert a protective effect. While metabolomics can provide information on the correlation between metabolites and diseases, a single technique cannot trace the origin of metabolites (host, microbial, or co-microbial). Additionally, real-time metabolite technologies in critically ill clinical detection patients remain underdeveloped.

2.5 Epigenomics

In sepsis-related host-microbiota interactions, the inflammatory state promotes the influence of diverse microbes, particularly gut microbiota, on host cellular transcriptional programs and immune cell function. This process occurs through epigenetic mechanisms such as the production of epigenetic substrates and enzymatic regulators, modulation of DNA methylation and histone modifications, and regulation of non-coding RNA. These processes ultimately disrupt immune function and induce organ dysfunction (Woo and Alenghat, 2022; Binnie et al., 2020). In the gut microbiota, Bacteroides fragilis, Clostridium perfringens, and Lactobacillus acidophilus, along with their metabolites, enhance DNA methylation levels in T-lymphocytes, induce the production of various cytokines, mitigate the inflammatory response during sepsis, regulate the balance of immune cells, and maintain intestinal immune homeostasis (Heffernan et al., 2021). SCFAs produced by the gut microbiota induce histone posttranslational modifications (PTMs), thereby influencing gene expression and other cellular processes (Gates et al., 2024). These modifications regulate protein activity and affect cellular signaling events in sepsis, helping to control cytokine storms and immunosuppression during the disease process (Zhang L. et al., 2023).

Some pathogens induce epigenetic changes in the host, promoting inflammation, regulating immune cell function, and modulating responses to microbial infections (Ma et al., 2024). This process provides new insights for preventing sepsis caused by pathogenic microorganisms. In sepsis caused by *Staphylococcus aureus* in mice,

DNMT3A, an essential enzyme for DNA methylation, is suppressed in blood leukocytes *in vivo* and in macrophages and neutrophils *in vitro*. This suppression affects IL-10 regulation, impacting host immune responses. Through a DNA methylation-dependent mechanism, this process influences the host's resistance to MRSA (Mba Medie et al., 2019).

3 Interaction between host immune system and microbiome in sepsis

3.1 Shaping of microbial communities by host immunity

The host immune system is critical in regulating microbial communities, as shown in Figures 1A,B. Under normal conditions, the host maintains intestinal microbiota homeostasis through innate immune mechanisms (e.g., secretion of antimicrobial peptides, complement activation) and adaptive immune mechanisms (e.g., antibody production, immune cell-mediated cytotoxicity) (Potruch et al., 2022). Sepsis-induced robust immune responses disrupt the ecological balance of the gut microbiota. Toll-like receptors (TLRs) recognize pathogen-associated molecular patterns (PAMPs) from microorganisms, triggering intracellular signaling pathways that stimulate macrophages to secrete pro-inflammatory cytokines. This inflammatory response alters the gut microenvironment and affects microbial survival.

Additionally, therapeutic interventions such as antibiotic treatment lead to the massive elimination of susceptible bacteria, allowing drug-resistant or opportunistic pathogens to overgrow, thereby promoting bacterial translocation and the displacement of microbial metabolites (Thazha et al., 2019; Piccioni et al., 2024; Kalra et al., 2022). Broad-spectrum antibiotics reduce microbial diversity

and increase the abundance of *Enterococcus* species, which is directly associated with an increased risk of sepsis (Kang and Thomas, 2021). In infants treated with broad-spectrum antibiotics (e.g., penicillin, gentamicin, or amoxiclav) for suspected early-onset neonatal sepsis, a significant reduction in *Bifidobacterium* abundance and an increase in *Klebsiella* spp. and *Enterococcus* abundance are observed, profoundly impacting the development of the gut microbiota (Reyman et al., 2022).

3.2 Modulation of host immunity by the microbiome

The microbiota plays a dual role in sepsis: it is both a pathogenic factor and a critical modulator of disease progression and host immune responses, as shown in Figure 1C. Microbiota diversity significantly impacts the immunophenotype and mortality of sepsis. In septic mice with high gut microbiota β -diversity, survival rates are improved considerably, accompanied by enhanced CD4+T cell responses. Conversely, reduced diversity impairs immune function, leading to uncontrolled inflammation and increased mortality (Fay et al., 2019). Furthermore, alterations in the gut microbiota before disease onset increase susceptibility to sepsis through multiple mechanisms, such as the expansion of pathogenic gut bacteria, activation of pro-inflammatory immune responses, and reduced production of beneficial microbial metabolites (Adelman et al., 2020).

Microbial dysbiosis can lead to host immune dysregulation during sepsis. The microbiota primarily modulates host immune responses through metabolites such as SCFAs and molecular patterns, thereby influencing the progression of sepsis (Mann et al., 2024). SCFAs (e.g., acetate, propionate, and butyrate) regulate epithelial barrier function, mucosal, and systemic immunity by



FIGURE 1

Host microbiome interactions in sepsis. (A) Under normal conditions, the host relies on innate immunity mechanisms (such as antimicrobial peptide secretion, complement activation, etc.) and adaptive immunity (such as antibody production, immune cell-mediated killing, etc.) to maintain the balance of intestinal microbial community. (B) The strong immune response triggered by sepsis breaks the original ecological balance of intestinal microbial community. (C) Microorganisms influence sepsis development and host immune response.

signaling through G protein-coupled receptors (GPCRs) such as GPR41 and GPR43 or by modulating histone deacetylase activity, thereby alleviating sepsis-associated inflammation (Wang et al., 2024a). Butyrate enhances histone H3 acetylation at the Foxp3 promoter and other conserved non-coding sequences, inducing the differentiation of gut Treg cells and reducing TNF- α and IL-6 levels (Furusawa et al., 2013). Acetate lowers neutrophil apoptosis via the FFAR2 pathway. It downregulates FABP4 through the endoplasmic reticulum stress pathway, modulating neutrophil apoptosis, increasing inflammatory factors in lung epithelial cells, and aggravating sepsis-induced respiratory distress syndrome (Xuan et al., 2024).

3.3 Association of microbiome with sepsis-related organ dysfunction

3.3.1 Gut microbiome and intestinal barrier function

Intestinal microbiota maintains gut barrier integrity through multiple mechanisms, including promoting the expression and stability of tight junction proteins, enhancing mucosal barrier function, and regulating intestinal immune cell homeostasis. Tight junction proteins (e.g., occludin, claudin family proteins) are central components of the mechanical barrier, forming tight intercellular junctions that effectively prevent the translocation of intestinal bacteria, toxins, and other harmful substances across the mucosal barrier into the bloodstream (Serek and Oleksy-Wawrzyniak, 2021).

In sepsis, microbial dysbiosis can compromise intestinal barrier function. On the one hand, the overgrowth of pathogenic bacteria directly erodes intestinal epithelial cells, disrupting tight junction structures. This process leads to increased tight junction proteins claudin-2 and JAM-A and decreased expression of claudin-5 and occludin, significantly increasing intestinal permeability (Yoseph et al., 2016). On the other hand, an imbalance in the gut microbiota impairs the gut's immune regulatory functions, weakening the ability of immune cells to clear pathogens and further exacerbating intestinal barrier damage. Once the intestinal barrier is compromised, intestinal bacteria and their metabolites can translocate in large quantities into the mesenteric lymph nodes, portal venous system, and even directly into the bloodstream, causing systemic infection and serving as a critical trigger for sepsis (Kullberg et al., 2021). For example, D-lactic acid, produced by gut commensal bacteria, is transported via the portal vein to the liver, where it is essential for maintaining the integrity of the intra-vascular firewall mediated by Kupffer cells, enabling the capture and elimination of circulating pathogens (McDonald et al., 2020). Oral supplementation of SCFAs activates G-protein-coupled receptor 43, enhancing macrophage phagocytosis of Klebsiella pneumoniae (Wu et al., 2020a).

3.3.2 Effects of microbial changes in sepsis on other organs

Sepsis disrupts the gut microbiota and causes dysbiosis in the lung microbiota. In patients with sepsis or acute respiratory distress syndrome (ARDS), alveolar microbial diversity decreases, and the lung microbiota network shifts from being dominated by beneficial commensal bacteria (e.g., *Streptococcus salivarius* and *Streptococcus*

oralis) to being dominated by gastrointestinal and periodontal pathogens (Lu et al., 2024). This altered microbial network can induce immunosuppression by promoting the expression of the EGFR gene and suppressing the expression of BST2 and HLA-C genes. The "gutlung axis" modulates pulmonary immune responses through microbial-associated molecular patterns (MAMPs) and metabolites. Oral/pharyngeal bacteria (e.g., *Klebsiella pneumoniae* and *Pseudomonas aeruginosa*) and gut bacteria (e.g., *Enterococcus* and *Klebsiella* spp.) can translocate to the lungs, causing infections and further disrupting systemic immunity (Lee and Banerjee, 2020).

Sepsis-associated neuroinflammation is closely linked to the imbalance of the gut-brain axis. Gut microbiota and SCFA metabolic disorders play a key role in sepsis-associated encephalopathy (SAE). SCFA supplementation can alleviate cognitive impairment and neuroinflammation in SAE mice (Li et al., 2023). Moreover, metformin has therapeutic potential for sepsis-related neuroinflammation by modulating the gut microbiota and metabolites, but its specific metabolic mechanisms require further validation (Zhao et al., 2022).

4 Strategies and methods of integrative omics in sepsis research

4.1 Necessity and challenges of multi-omics data integration

4.1.1 Necessity of a comprehensive understanding of biological systems

Sepsis involves complex interactions between the host and pathogens, with interconnected information from multiple levels, including the microbiome, genome, transcriptome, proteome, and metabolome (Mangioni et al., 2020). For example, changes in the composition and function of the gut microbiota are closely linked to the host's immune response during sepsis. A single-omics approach cannot fully elucidate this intricate relationship. In studies on *Staphylococcus epidermidis* in neonatal sepsis, the pathogenic potential and molecular mechanisms of this bacterium in the sepsis microbiome can only be fully understood by combining the open and diverse nature of its genome with the expression of virulence genes at the transcriptomic level. This integrated approach provides a basis for precision diagnosis and treatment (Joubert et al., 2022).

4.1.2 Technical and analytical challenges of data integration

In clinical sepsis research, the three common goals of using omics are to investigate host responses, develop diagnostic methods, and identify clinically relevant clusters (Schuurman et al., 2021). However, the integration of multi-omics data faces numerous technical barriers. On the one hand, different types of omics data vary in data structure, measurement scale, and noise level (Chen et al., 2023; Wang W. et al., 2021). For example, 16S rRNA sequencing, used to analyze bacterial community structure, produces relatively sparse data. Researchers can consider integrating bacterial, viral, fungal, and protozoan communities to obtain more comprehensive information, but this increases complexity (Zhou et al., 2022). Metagenomic sequencing provides more comprehensive genetic information but generates large volumes of complex data, differing from nucleic acid-based omics data.

Difference	Data-driven methods	Knowledge-guided learning methods
Basic idea	Directly extract patterns and relationships from raw data without relying on prior biological hypotheses.	Utilize existing biological knowledge and databases to simplify data complexity using prior information.
Common techniques/tools	PCA, t-SNE, NMDS, LDA, WGCNA, correlation analysis, network fusion (e.g., SNF)	GSEA, GSVA, pathway annotation based on KEGG/Reactome, specialized databases such as CAZy and VFDB
Advantages	Flexible and unbiased, capable of uncovering novel patterns; well-suited for handling large-scale, high-dimensional data.	Offers strong biological interpretability by integrating data into functional modules, reducing noise, and identifying known key pathways.
Limitations	Requires a high sample size; some patterns discovered may be challenging to interpret.	Database update frequency may affect accuracy depending on existing knowledge and potentially missing novel or unrecorded biological mechanisms.
Integration stage	Commonly used in early to intermediate integration stages (e.g., data concatenation, feature transformation, network construction).	Typically applied in later stages, using functional annotation and pathway aggregation to interpret and validate analysis results.

TABLE 1 Differences between data-methods and knowledge-guided methods.

On the other hand, differences in data acquisition platforms and experimental conditions also complicate integration, as batch effects often interfere with data consistency and comparability. Analytically, the "curse of dimensionality" from high-dimensional data challenges traditional statistical methods, which struggle to handle the complexity of multi-omics data. Models are prone to overfitting, lack generalizability, and fail to apply to new datasets (Gliozzo et al., 2024).

4.2 Multi-omics integration strategy

In recent years, multi-omics research has developed numerous integration strategies to address challenges. These strategies can be broadly categorized into data-driven and knowledge-guided approaches, with a comparison provided in Table 1. Data-driven strategies focus on extracting meaningful patterns and relationships directly from raw omics data. This approach is highly flexible and suitable for complex datasets commonly encountered in multi-omics studies.

However, since extracted correlations may not always align with biological reality, more studies have begun incorporating external biological knowledge for guidance. By leveraging existing biological or functional knowledge (typically stored in multiple databases in network format, as shown in Table 2), knowledge-guided approaches effectively reduce multi-omics data's complexity and enhance the integrated results' biological significance and interpretability (Li W. et al., 2024). For example, in microbiome data integration, many microbial genes can be aggregated into functional modules or pathway levels using functional families from the PATRIC database (Davis et al., 2020) or gene families from the KEGG database (Kanehisa et al., 2023). Additionally, specialized databases focused on microbial genome annotation and comparison, such as IMG/M (Integrated Microbial Genomes & Microbiomes) (Chen et al., 2019) and MicroScope (Vallenet et al., 2020), can also be employed for this purpose. Additionally, for specific functional domains, specialized databases such as CARD (Alcock et al., 2020) for antibiotic resistance genes, VFDB (Zhou et al., 2025) for bacterial virulence factors, and mobileOG-db (Brown et al., 2022) for mobile genetic elements can be integrated. This approach provides a more comprehensive understanding of microbial community functional characteristics and interactions. Furthermore, bioinformatics tools such as PICRUSt (Langille et al., 2013), Piphillin (Narayan et al., 2020), PUMAA (Mitchell et al., 2020), and iVikodak (Nagpal et al., 2018) enable functional prediction based on 16S rRNA data, addressing functional research needs in cases where metagenomic sequencing is unavailable or resource-limited.

Key signaling molecules in the host's transcriptomic and proteomic data can be utilized to construct detailed gene regulatory networks by leveraging gene expression and regulatory information integrated in Harmonizome (Rouillard et al., 2016), as well as RNA-protein interaction data provided by catRAPID (Armaos et al., 2021), RBPDB (Cook et al., 2011), and RNAct (Lang et al., 2019). In practice, data-driven and knowledge-derived DR can complement each other, providing a more comprehensive understanding of the critical interactions between the host and microbes in sepsis.

The integration steps of these two methods are primarily similar, as shown in Figure 2, with the key difference in how prior biological information is incorporated into the analysis (Abdelaziz et al., 2024). Data integration strategies are typically divided into horizontal and vertical integration (Zitnik et al., 2019). Horizontal integration involves studying the same omics across different sample groups, while vertical integration examines multiple omics data on the same samples. Vertical integration is more complex and widely applied in practice; the following discussion in this paper focuses on vertical integration.

4.3 Dimension reduction of multi-omics integration

Dimensionality reduction is often an essential preprocessing step in multi-omics analysis, as simple integration of omics data may lead to the loss of critical features inherent to each omics layer, exacerbating the "curse of dimensionality" in data integration (Downing and Angelopoulos, 2023). While this step is optional, early and intermediate integration strategies typically require dimensionality reduction to enhance their effectiveness. Standard dimensionality reduction techniques are listed in Table 3.

Dimensionality reduction is typically achieved through feature selection and feature extraction. Feature selection identifies the most representative and informative variables from raw data, such as Recursive Feature Elimination (RFE) (Jeon and Oh, 2020), L1/L2 Regularization (Elastic Net) (Xu et al., 2010), and Least Absolute Shrinkage and Selection Operator (LASSO) (D'Angelo et al., 2009). This approach retains features' original physical or biological meaning,

TABLE 2 Representative databases for various types of biological knowledge.

Database name	Full name	Biological knowledge	Source
UniProt-KB	UniProt Knowledgebase (UniProt Consortium, 2025)	Sequence and functional information on proteins	https://www.uniprot.org/
KEGG	Kyoto Encyclopedia of genes and genomes (Kanehisa and Goto, 2000)	Molecular interaction, reaction, and relation networks	https://www.genome.jp/kegg/ pathway.html
Reactome	Reactome pathway database (Croft et al., 2011)	Signaling and metabolic pathways	https://reactome.org/
STRING	Search Tool for the retrieval of interacting genes/ proteins (Szklarczyk et al., 2023)	Protein-protein interaction networks	https://string-db.org/
PathBank	PathBank (Wishart et al., 2024)	Metabolic, signaling, disease, drug, and physiological pathways	https://www.pathbank.org/
Pathway Commons	Pathway Commons (Cerami et al., 2011)	Biological pathway and interactions: biochemical reactions; gene regulatory networks; protein, nucleic acid, small molecule interactions	https://www.pathwaycommons.org/
BioCyc	BioCyc Pathway/Genome Database Collection (Karp et al., 2019)	Metabolic pathways, regulatory networks	https://biocyc.org/
WikiPathways	WikiPathways (Agrawal et al., 2024)	Signaling pathways	https://www.wikipathways.org/
GRNdb	Gene Regulatory Network database (Fang et al., 2021)	Gene regulatory networks among transcription factors and genes	http://www.grndb.com/
BioGRID	Biological General Repository for Interaction Datasets (Oughtred et al., 2021)	Protein and genetic interactions	https://thebiogrid.org/
IID	Integrated Interactions Database (Kotlyar et al., 2022)	Protein-protein interaction	http://ophid.utoronto.ca/iid
Harmonize	Harmonize (Rouillard et al., 2016)	Integrative gene and protein expression data across various tissues and conditions	http://amp.pharm.mssm.edu/ Harmonizome
Cat RAPID	Computational Analysis of Targets of RNA-Protein Interactions and Discovery (Armaos et al., 2021)	RNA-protein interaction prediction	http://service.tartaglialab.com/page/ catrapid_group
RBPDB	RNA-Binding Protein DataBase (Cook et al., 2011)	RNA-binding protein recognition motifs and their target interactions	http://rbpdb.ccbr.utoronto.ca/
RPISeq	RNA-Protein Interaction Sequence-based Predictor (Muppirala et al., 2011)	RNA-protein interactions	http://pridb.gdcb.iastate.edu/RPISeq/
RNAct	RNA - Critical Targets (Lang et al., 2019)	Protein-RNA interactions	http://rnact.crg.eu
StarBase	StarBase (Li et al., 2014)	miRNA-target, RNA-RNA interactions	http://starbase.sysu.edu.cn/
MetaCyc	Metabolic Pathways From all Domains of Life (Caspi et al., 2006)	Metabolic pathways	http://MetaCyc.org/
PHI-base	Pathogen Host Interactions Database (Winnenburg et al., 2006)	fungal and Oomycete pathogenicity genes-host interactions	http://www.phi-base.org/
ConsensusPathDB	ConsensusPathDB (The ConsensusPathDB Interaction Database, 2013)	Integrative database for molecular interactions	http://consensuspathdb.org/
PATRIC	Pathosystems Resource Integration Center (Wattam et al., 2014)	Microbial genomes and associated functional annotations	https://www.patricbrc.org/
IMG/M	The Integrated Microbial Genomes & Microbiomes (Chen et al., 2019)	Multi-Source Microbial Genomes and Metagenomes prediction and functional annotation	https://img.jgi.doe.gov/m/
MicroScope	MicroScope (Vallenet et al., 2020)	Functional annotation of microbial species genes and genomic regions, metabolic network reconstruction, and post-genomic experiments	https://www.genoscope.cns.fr/agc/ microscope

reduces model complexity, minimizes noise, and lowers the risk of overfitting. However, it may overlook complex feature interactions, potentially missing latent joint information.

In contrast, feature extraction constructs a new feature space, such as through Principal Component Analysis (PCA) (Ringnér, 2008), t-stochastic Neighbor Embedding (t-SNE) (Mrowka and Schmauder, 2024), and Variational Autoencoders (Kingma and Welling, 2019). These methods map high-dimensional data into low-dimensional representations, capturing intrinsic data structures and nonlinear relationships more effectively. Although it achieves



dimensionality reduction, multi-omics integration, downstream analysis (classification, regression, clustering), model/method evaluation, interpretability, and explainability. The difference between the Knowledge-Guided Learning Methods and the Data-Driven Methods is that in steps 3 and 5, the former uses a variety of bioinformatics databases related to hosts and microorganisms as a guide to performing data dimensionality reduction, classification, regression, and cluster analysis.

efficient information compression, it often results in reduced interpretability.

Depending on whether they utilize data labels, dimensionality reduction techniques can be categorized as supervised or unsupervised. Unsupervised methods, which eliminate redundant variables based on correlations without considering target variables, are suitable for data exploration and structure revelation. Standard unsupervised techniques include Principal Component Analysis (PCA) (Ringnér, 2008), t-stochastic Neighbor Embedding (t-SNE) (Mrowka and Schmauder, 2024), and Nonnegative Matrix Factorization (NMF) (Devarajan, 2008). In contrast, supervised methods leverage known class information to guide the dimensionality reduction process, making them appropriate for classification tasks. Examples of supervised techniques include Linear Discriminant Analysis (LDA) (Ayesha et al., 2020) and Partial Least Squares Discriminant Analysis (PLS-DA) (Lee et al., 2018).

Data integration can be categorized into single-omics, dual-omics, and multi-omics integration, with the number of datasets to be integrated influencing the choice of dimensionality reduction techniques (see Table 3 for details). In practical applications, PCA and LDA can project multi-omics data from the host and microbiome onto a few principal components, retaining the significant variation in the data. This process allows visualization of key gene, protein, and microbiota changes under sepsis conditions (Jonathan et al., 2020). When integrating multi-omics data from different sources, cPCA can eliminate batch effects between datasets and extract standard features reflecting host-microbiome interactions. PLS helps identify host and microbial metabolites that co-vary during disease progression, providing evidence for potential therapeutic targets. WGCNA constructs co-expression or co-abundance networks, clustering related genes or microbes into modules. This step aids in discovering synergistic changes in host genes and microbes at the network level, revealing key regulatory networks and potential biomarkers in sepsis pathogenesis. SNF (Wang et al., 2014) integrates multi-omics data (e.g., genomics, transcriptomics, proteomics) by constructing and fusing similarity networks, capturing shared patterns across omics layers. UMAP reduces the dimensionality of complex host and

TABLE 3 Common dimension reduction methods.

Method	Learning approach	Туре	Common tools	Advantages	Limitations	Omic datasets
Principal Component Analysis (PCA)	Unsupervised	Feature extraction	base R (R) (Jolliffe and Cadima, 2016); dimRed (R) (Kraemer et al., 2018); mixOmics (R) (Rohart et al., 2017); FactoMineR (R) (Lê et al., 2008); pcaMethods (R) (Stacklies et al., 2007)	Simple, fast, interpretable	Assumes linearity; sensitive to scaling	Single
Independent component analysis (ICA)	Unsupervised	Feature extraction	fastICA (R) (Miettinen et al., 2018); fICA (R) (Miettinen et al., 2018)	Separates mixed signals; extracts non- Gaussian sources	Sensitive to noise; independence assumption may not hold	Single
Multi-dimensional scaling (MDS)	Unsupervised	Feature extraction	base R (R) (Jolliffe and Cadima, 2016); dimRed (R) (Kraemer et al., 2018); MetaboAnalyst	Preserves pairwise distances; effective for visualization	Computationally intensive; sensitive to outliers	Single
Correspondence analysis (CA)	Unsupervised	Feature extraction	vegan (R) (Chen et al., 2024); ade4 (R) (Dray and Dufour, 2024); FactoMineR (R) (Lê et al., 2008); ca (R) (Nenadic and Greenacre, 2007)	Reveals associations in contingency tables	Limited to count/ categorical data; interpretation can be subjective	Single
Multiple Factor Analysis (MFA); Hierarchical Multiple Factor Analysis (HMFA)	Unsupervised	Feature extraction	FactoMineR (R) (Lê et al., 2008)	Integrates multiple datasets; balances influence of each data block	Requires careful preprocessing; interpretation may be complex	Single; multiple
Nonnegative Matrix Factorization (NMF)	Unsupervised	Feature extraction	NMF (R) (Wang and Zhang, 2013)	Produces parts-based, interpretable representation	Non-convex optimization; sensitive to initialization	Single
Linear Discriminant Analysis (LDA)	Supervised	Feature extraction	MASS (R) (Xu et al., 2009); caret (R) (Kuhn et al., 2024)	Maximizes class separability; simple and fast	Assumes normality and equal covariance; limited to linear boundaries	Single
Locally linear embedding (LLE)	Unsupervised	Feature extraction	RDRToolbox (R) (Bartenhagen, n.d.)	Captures local structure; reveals non- linear manifolds	Sensitive to noise and parameter settings; computationally heavy	Single
Consensus PCA (cPCA)	Unsupervised	Feature extraction	mogsa (R) (Meng et al., 2016)	Integrates multiple datasets; robust to dataset-specific noise	Less standardized; higher computational demand	Pair
Canonical Correlation Analysis (CCA)	Unsupervised	Feature extraction	CCA (R) (González and Déjean, 2023); vegan (R) (Oksanen et al., 2025); PMA (R) (Chu et al., 2013); mixOmics (R) (Rohart et al., 2017)	Identifies linear relationships between two variable sets	Assumes linearity; sensitive to noise	Pair
Co-inertia analysis (CIA)	Unsupervised	Feature Extraction	made4 (R) (Culhane et al., 2005)	Highlights common structure between datasets	Requires matched samples; interpretation can be complex	Pair
Partial Least Squares (PLS)	Unsupervised	Feature extraction	pls (R) (Mevik and Wehrens, 2007); caret (R) (Kuhn, 2024)	Handles multicollinearity; effective with high- dimensional predictors	The risk of overfitting requires rigorous cross- validation	Pair

(Continued)

TABLE 3 (Continued)

Method	Learning approach	Туре	Common tools	Advantages	Limitations	Omic datasets
Partial Least Squares Discriminant Analysis (PLS-DA)	Supervised	Feature extraction; Feature selection	mixOmics (R) (Rohart et al., 2017); PLS-DA tool (Open source MATLAB tool) (Zontov et al., 2020)	Combines dimensionality reduction with classification	Overfitting risk; model validation is crucial	Pair
Generalized Procrustes Analysis (GPA)	Unsupervised	Feature extraction	Vegan(R) (Oksanen et al., 2025); FactoMineR (R) (Lê et al., 2008)	Aligns multiple datasets; removes scale/ rotation differences	Sensitive to outliers; may require iterative convergence	Pair
Multiple co-inertia analysis (mCIA)	Unsupervised	Feature Extraction	omicade4 (R) (Meng et al., 2014)	Simultaneously analyzes several datasets.	Computationally intensive; interpretation is complex	Multiple
Tensor Component Analysis (TCA)	Unsupervised	Feature extraction	tensor BSS (R) (Virta et al., 2024); rTensor (R) (Li et al., 2018); tensorr (R) (Rougier, 2012); ThreeWay (R) (Giordani et al., 2014); SDA4D (R) (Gill and Marchini, 2020)	Captures multi- dimensional interactions; parts- based representation	High computational complexity; challenging parameter tuning	Multiple
Weighted Correlation Network Analysis (WGCNA)	Unsupervised	Feature extraction	WGCNA (R) (Zhang et al., 2021a)	Identifies co-expression modules; robust network construction	Sensitive to parameter selection; high computational cost for large networks	Multiple
Two-way orthogonal PLS (O2PLS)	Supervised	Feature extraction	OmicsPLS (R) (Bouhaddani et al., 2018)	Separates shared vs. dataset-specific variation	Model complexity; challenging interpretation	Multiple
Jint and individual variation explained (JIVE)	Unsupervised	Feature extraction	r.jive (R) (Murden et al., 2022)	Distinguishes shared and unique variation; enhances interpretability	Sensitive to noise; proper model selection is critical	Multiple
Generalized CCA (GCCA)	Unsupervised	Feature extraction	RGCCA (R) (Gloaguen et al., 2022); mixOmics (R) (Rohart et al., 2017)	Integrates more than two datasets simultaneously	High computational demand; similar linearity assumptions as CCA	Multiple
Similarity Network Fusion (SNF)	Supervised	Feature extraction	SNFtool (R) (Burton- Pimentel et al., 2021)	Combines multiple similarity networks; robust to noise	Requires careful tuning of similarity and network parameters	Multiple
Uniform Manifold Approximation and Projection (UMAP)	Unsupervised	Feature extraction	Umap (R) (Yang et al., 2021)	Preserves both local and global structures, faster computation compared to t-SNE	Sensitive to initial conditions, it may not capture all global structures.	Single;Multiple

microbiome multi-omics data for visualization, enabling researchers to intuitively observe data distribution and clustering patterns. Extended CCA integrates multi-omics datasets (e.g., microbiome and metabolome) to capture disease-related multi-omics modules, uncovering significant correlations and highly predictive modules across omics layers (Muller et al., 2024).

4.4 Data integration methods and challenges

Due to the problems of high dimensionality, population heterogeneity, and complex data association, it is challenging to utilize the multi-omics data of sepsis effectively. In order to meet these challenges, Iperi et al. (2025) used BiomiX to solve the bottleneck of high-throughput omics data analysis, which can efficiently and integratively analyze multi-omics data from different queues. Moreover, multi-omics factor analysis (MOFA) is commonly used for solving multi-omics data integration (Guo et al., 2025). Besides, direct linkage prioritizes longitudinal multi-omics profiling (e.g., concurrent genomic, metabolomic, and microbiome measurements within identical sepsis cohorts) to reconstruct patient-specific interactomes using advanced integration algorithms (MOFA+/mixOmics). This approach systematically reveals latent covariance structures and time-resolved biological associations. Indirect linkage addresses non-paired datasets through cross-cohort meta-analytic frameworks, statistically harmonizing heterogeneous sepsis cohorts to detect subtle trans-study differential signals.

Three procedures can be employed to address population heterogeneity to mitigate population bias from different data sources. Initially, during the preprocessing phase, the ComBat-harmony algorithm can be applied to eliminate batch effects. Subsequently, a causal inference framework (such as do-calculus) can be introduced in the modeling phase to distinguish confounding factors from accurate biological signals (Dibaeinia et al., 2025). Additionally, transfer learning can bridge the gap between general databases and sepsis-specific data, which is conducive to better handling population heterogeneity across different data sources (Liu et al., 2024). Finally, in the validation phase, when the sample size of specific subgroups is limited, a digital twin cohort can be developed to verify the performance and robustness of the model. Although these methods are reasonable and feasible in theory, adjustments and optimizations are necessary during implementation according to the data and research objectives.

In the context of cross-species (human and animal) data integration and analysis, although the use of phyloP scores effectively identifies conserved molecular features across species (Sullivan et al., 2023), thereby reducing the impact of interspecies differences on research outcomes, we recommend that when collecting data, efforts should be made to ensure similarity between human and animal samples in terms of experimental conditions and processing methods. Data should not be simply combined without consideration; suitable data for aggregation should be carefully selected. Following data cleaning and imputation, the data should undergo standardization and batch effect correction to eliminate systematic differences introduced by different experimental batches or platforms.

Addressing the sparsity issue in omics data, employing deep learning methods such as OmiEmbed (a cross-modal generation approach based on Variational Autoencoders, VAEs) to generate partially missing data represents an advanced strategy (Zhang et al., 2021b). This method can predict missing values based on existing data to address partial missingness in omics datasets. However, when facing comprehensive omics sparsity, leveraging existing biological knowledge to augment data analysis emerges as an effective alternative. Utilizing BioBERT to extract knowledge from literature and constructing virtual omics layers with this knowledge can help mitigate the issue of data sparsity (Lee et al., 2020). Given that deep learning methods and sophisticated statistical models may demand substantial computational resources, the computational cost must be considered when applying these approaches. Moreover, when imputing and integrating data, it is crucial to ensure that the results are biologically meaningful, which can be confirmed through biological validation to verify the reliability of the outcomes.

4.5 Types of data integration

Following dimensionality reduction, further analysis and integration of multi-omics data are required. Various integration strategies are available, as proposed by Picard et al. (2021), into five types: early (concatenation-based), mixed (transformation-based), late (model-based), intermediate and hierarchical. Early integration can be selected if each dataset has been preprocessed according to its omics type; this approach involves directly concatenating samples and assembling the resulting matrix as input for machine-learning models. While this method is straightforward, it increases data complexity; therefore, various strategies have been developed to transform or map datasets to facilitate integration. Mixed integration involves independently transforming or mapping each omics dataset, whereas intermediate integration constructs a joint low-dimensional representation across omics. Late integration analyzes each omics dataset separately and aggregates predictions after model training. Hierarchical integration leverages known regulatory relationships between omics, as defined by the central dogma of molecular biology, to integrate datasets.

Numerous tools and platforms based on diverse integration methods have emerged to integrate multi-omics data and investigate host-microbe interactions in sepsis. For example, multi-omics factor analysis (MOFA) and MOFA+ are dimensionality reduction techniques and intermediate integration-based integration tools. These tools take abundance matrices of microbial communities (e.g., bacteria, fungi, viruses) as input and learn low-dimensional representations of samples along with corresponding feature-loading matrices (Argelaguet et al., 2020; Losert et al., 2024). Haak et al. (2021) utilized MOFA to integrate gut microbiota data from sepsis patients and healthy volunteers, including bacterial (16S rRNA), fungal (internal transcribed spacer 1 (ITS1) rRNA), viral (viral metagenomic next-generation sequencing) components and revealed that the proliferation of aerobic pathogens, bacteriophages, and opportunistic yeasts disrupts anaerobic environments, thereby contributing to sepsis pathogenesis.

MintTea (Muller et al., 2024) is an intermediate integration method to identify disease-associated microbial modules from multi-omics data. It elucidates the mechanistic roles of the microbiome in disease and provides evidence for generating system-level, multi-dimensional hypotheses of microbiome-disease interactions. MaAsLin 2 (Mallick et al., 2021) is a tool for multivariable association analysis in microbial community multi-omics studies. It is classified as a late integration method but incorporates features of mixed integration. It is designed to explore associations between microbial community features and complex metadata, such as human health outcomes, diet, and environmental conditions. BZINB-iMMPath (Lin et al., 2023) is a late integration method that constructs metabolite-species and speciesspecies correlation networks. It identifies species modules through similarity-based clustering and facilitates the joint modeling and analysis of microbiome and metabolome data. Zhang S. et al. (2023) introduced a Bayesian modeling method for integrating sparse multivariate count data in microbiome studies. This method leverages joint sparsity to capture feature interactions and supports robust structural estimation in small-sample datasets. Hypergraph-induced orthogonal Nonnegative Matrix Factorization (HONMF) (Ma et al., 2023) is an unsupervised learning framework and an intermediate integration method for microbiome multi-omics data. By integrating three groups of latent variables, it preserves the high-order geometric structure of the original data. It supports sample clustering, data visualization, feature selection, and cross-domain association analysis (e.g., bacteria-virus, fungi-virus interactions). Multimodal Functional Deep Learning for Multi-Omics Data (MFDL) (Zhou et al., 2024) is a deep learning-based intermediate integration method. It integrates various omics data by fitting them into a shared dimensionality-reduction hidden layer at the input level. It enables learning complex relationships between multi-omics data and phenotypes through multi-layer training. EMPress (Cantrell et al., 2021) is an open-source, interactive visualization tool for multi-omics data and employs a late integration approach. It effectively links community-level sample groupings with feature-level structures, supporting exploratory analysis of complex multi-omics datasets.

4.6 Guidelines for clinical sample collection in sepsis

One crucial aspect is that researchers need to minimize sample heterogeneity when collecting samples. Factors to be considered include chemotherapy, active or passive immunotherapy, antibiotic medication use, lifestyle, etc., which can disrupt the microbial balance and consequently damage the network within the bacterial community and its relationship with the host.

The sepsis-specific biobank was designed, and different sampling time windows, core omics layer, and extended omics layer were recommended for different stages of sepsis. During the early recognition phase (0-6 h), it is suggested to collect metabolomics (plasma) and microbiome (fecal) samples within the first hour of initial presentation in the emergency department, along with singlecell transcriptomics (peripheral blood mononuclear cells, PBMC) analysis. Metabolomics and microbiome samples are relatively inexpensive and easy to collect, with high clinical feasibility and costeffectiveness. During the progression phase (6-72 h), dynamic sampling of proteomics (serum) and lipidomics (plasma) every 12 h is recommended. Although these samples are easy to collect, they require substantial resources and incur higher costs. Spatial transcriptomics (tissue biopsy) is more challenging to sample but offers deeper analytical insights. In the recovery or sequelae phase, follow-up visits are recommended at 30, 90, and 180 days postdischarge to collect circulating cell-free DNA (cfDNA) and exosome omics samples and perform longitudinal analysis of gut metagenomics. This stratified sampling strategy facilitates comprehensive acquisition of the multi-omics features of sepsis, providing critical evidence for early disease recognition, progression monitoring, and prognostic assessment.

In order to reduce costs, some intensive care centers are equipped with core devices such as mass spectrometers and sequencers to meet the real-time decision-making needs for septic shock. An automated diagnostic platform is constructed and can be operated by grassroots medical staff. At the same time, data analysis is automatically completed by cloud-based learning models, thereby achieving intensive and efficient use of resources.

5 Application of integrative omics in the diagnosis and treatment of sepsis

5.1 Discovery of diagnostic markers

Since infection and host-pathogen interactions constitute a complex nonlinear system, no single biomarker can accurately diagnose sepsis (Pierrakos et al., 2020). Multi-omics integration offers new avenues for identifying diagnostic markers of sepsis. Metagenomic sequencing and 16S rRNA analyses reveal altered gut microbiota structure and function in sepsis patients. Specific microbes (e.g., *Enterococcus* abundance) and metabolites (e.g., SCFA levels) are linked

to sepsis progression and hold potential as diagnostic markers. Integrating host omics data, such as inflammatory gene expression in the host transcriptome and specific protein markers in the proteome, with microbiome data enhances diagnostic accuracy and specificity (Malik et al., 2024).

In a multicenter prospective study (van Houten et al., 2018), researchers collected blood, nasal swabs, and fecal samples from sepsis and non-infected individuals. Multi-omics analyses included host RNA and protein biomarker detection, nasal and gut microbiota analysis, host genomics, and bacterial proteomics. A multiparameter model was developed to distinguish between bacterial and viral infection etiologies. This study provided a diagnostic tool for differentiating bacterial and viral infections and also offered data support for bacterial subtype classification, providing a more accurate basis for sepsis diagnosis. Ma et al. (2022) integrated host transcriptomics, microbiomics, and metabolomics data and found that in cecal sepsis caused by methicillin-resistant Staphylococcus aureus, CYP1A1 deficiency improved gut barrier function and reduced the accumulation of the harmful metabolite cadaverine, whose level was positively correlated with the clinical SOFA score, and could offer new biomarkers for early sepsis diagnosis and treatment. Using the BZINB model, an integrative multi-omics model, disease-related modules in the microbiome and metabolome were identified. Specific correlations between certain metabolites and microbes were found in healthy and orally diseased patients, providing new insights for developing disease diagnostic biomarkers (Lin et al., 2023).

5.2 Development of personalized treatment strategies

Integrating multi-omics technologies enables in-depth analysis of patients' genomics, transcriptomics, proteomics, metabolomics, and gut microbiomes. This approach precisely characterizes each patient's unique pathophysiological profiles, immune status, microbial composition, and interactions, thereby enabling the gradual realization of personalized medicine in sepsis therapy.

If a deficiency of beneficial gut microbiota or an overgrowth of pathogenic bacteria is detected, interventions such as probiotic supplementation, prebiotic administration, or fecal microbiota transplantation (FMT) can be employed (Lisko et al., 2017). For instance, FMT has been shown to restore and increase butyrate levels the IRF3/NF-κB signaling pathway, counteract the via immunosuppressive effects of sepsis pathogens, and improve sepsisinduced muscle atrophy (Kim et al., 2020). By integrating host genomic and transcriptomic data, it is possible to predict patient responses to medications, select more appropriate antimicrobial agents, prevent antibiotic misuse, and achieve precision medicine. Additionally, dynamic monitoring based on omics data can facilitate timely adjustments in therapeutic strategies, thereby enhancing treatment outcomes and improving patient prognosis. Yang et al. (2024) utilized a two-sample bidirectional Mendelian randomization analysis, integrating data from genome-wide association studies (GWAS), eQTL datasets, single-cell transcriptomics, and large-scale RNA sequencing. By combining insights into gut microbiota regulatory mechanisms with drug databases, they precisely evaluated the association between gut microbiota and sepsis, identified potential genes and targets, and deeply explored and validated potential therapeutic agents for sepsis. Mu et al. (2023) used clinical isolates of four sepsis-causing bacterial strains and integrated multi-omics data, including genomics, transcriptomics, proteomics, and metabolomics, to investigate the responses of sepsis pathogens in a human serum environment. Their findings provide critical evidence for the development of therapeutic targets for sepsis. Li D. et al. (2024) conducted bulk RNA sequencing, single-cell transcriptomic analysis, metabolomics, and 16S rDNA sequencing of gut microbiota using peritoneal lavage fluid (PLF) cells in mice. They discovered that rhamnose derived from the gut or therapies targeting SLC12A4 may enhance macrophage phagocytosis during sepsis, offering a novel potential direction for clinical sepsis treatment.

5.3 Therapeutic effect monitoring and prognosis evaluation

In the treatment of sepsis, multi-omics technologies play a critical role. By integrating and analyzing multi-omics data, including the microbiome, host transcriptome, proteome, and metabolome, these technologies enable real-time and comprehensive monitoring of therapeutic efficacy, provide deep insights into disease progression, predict patient outcomes, and offer a scientific basis for formulating and adjusting treatment plans with precision. This approach ultimately contributes to improving clinical outcomes for patients with sepsis.

In treating neonatal sepsis, monitoring the expression changes of Staphylococcus epidermidis virulence genes and the dynamics of transcriptional regulatory networks, combined with the unique host environmental factors of neonates, comprehensively evaluates the role of treatment in controlling infection and improving prognosis. Providing key information for optimizing subsequent treatment decisions and, more precisely, grasping the progress and direction of sepsis treatment (Joubert et al., 2022). Using the central security database (HoPOIT database) to standardize, integrate, and dynamically monitor multi-omics data of patients (such as host biomarkers, pathogen characteristics, and drug resistance information), the temporal dynamics of host-pathogen interactions are analyzed, revealing the changing trends of biomarkers during treatment (van Houten et al., 2018). Sepsis patients are categorized into different enterotypes by integrating gut microbiota 16S rRNA analysis and metabolomics. It is found that patients with enterotype E3 have the most severe conditions. The OTU773 of Bacteroidota and OTU822 from Rikenellaceae are significantly positively correlated with ICU length of stay. 5-Hydroxyindoleacetylglycine is positively correlated with the APACHE II score, and three compounds negatively correlate with ICU length of stay. Long et al. (2023) utilized metagenomics and metabolomics to delineate the dynamic changes in gut microbiota and their metabolites in sepsis patients at different stages of ICU admission. They observed that in sepsis patients, gut microbiota diversity, the relative abundance of Firmicutes, and SCFA levels were significantly reduced, whereas the relative abundance of Proteobacteria and primary bile acid levels markedly increased with prolonged hospitalization. Among the differential microbiota and metabolites, the relative abundance of Klebsiella and the concentrations of butyrate and taurocholic acid exhibited strong correlations with sepsis patient prognosis, providing direction for prognostic evaluation and the adjustment of treatment plans based on microbiota profiles. These findings indicate that alterations in gut microbiota and metabolites are associated with sepsis's progression and clinical outcomes, providing a basis for early prediction of clinical outcomes and exploration of new therapies (Sun et al., 2023).

6 Current challenges and future directions

In the field of sepsis, multi-omics integration has achieved some progress in exploring host-microbiota interactions, but many challenges remain, and there is still considerable room for development. From the perspective of omics technology development, novel technologies such as single-cell and spatial omics are continuously emerging and evolving. Single-cell RNA sequencing (scRNA-seq) (Cheng et al., 2023) and single-cell nascent RNA sequencing (scGRO-seq) (Mahat et al., 2024) can reveal cellular heterogeneity at the single-cell level and provide deeper insights into cellular functions and characteristics. Spatial omics, on the other hand, allow for the investigation of cellular microenvironments and intercellular interactions while preserving the spatial structure of tissues (Wang J. et al., 2024). Currently, Janosevic et al. (2021) used scRNA-seq to detect changes in various renal cell populations in septic mice during the disease process and combined spatial omics sequencing to provide a spatiotemporal dynamic map of septic kidneys at the cellular and molecular levels. However, there are currently no studies on the relationship between microbiota and host in sepsis at the single-cell and spatial levels. Integrating emerging omics technologies to locate the distribution of microbiota within the host precisely and their impact on surrounding tissues and cells, as well as deeply exploring the spatial interactions between host and microbiota in sepsis, holds great potential.

From the perspective of the disease process, most existing multiomics studies are focused on static analysis, that is, testing samples at specific time points. However, sepsis is a dynamically evolving pathological process, with microbial community structures, host gene expression, and metabolite levels all fluctuating continuously over time. The lack of dynamic monitoring limits our understanding of critical turning points and intervention timing during the disease process, making it difficult to provide precise time-window guidance for precision treatment. Dynamic modeling approaches should be adopted to study bacterial populations from a dynamic perspective and integrate data from multiple time points or conditions.

Author contributions

JLu: Writing – original draft. WZ: Formal analysis, Writing – original draft. YH: Formal analysis, Writing – review & editing. MJ: Formal analysis, Writing – review & editing. ZL: Investigation, Methodology, Writing – review & editing. JZ: Conceptualization, Writing – review & editing. LZ: Conceptualization, Writing – review & editing. BZ: Conceptualization, Writing – review & editing. JLuo: Data curation, Formal analysis, Writing – review & editing. CH: Conceptualization, Formal analysis, Writing – review & editing. YS: Data curation, Formal analysis, Writing – review & editing. RZ: Conceptualization, Writing – review & editing. KF: Writing – review & editing. BF: Writing – review & editing. CW: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported by the National Administration of Traditional Chinese Medicine (NATCM) High-Level Key Discipline of TCM (zyyzdxk-2023067); the National TCM Dominant Specialty Construction Project (2024YSZKZZYX006); the National Natural Science Foundation of China (82374350); the third round of "the academic honor scheme" of Shanghai University of Traditional Chinese Medicine (No. 38 (2024) of the TCM Administration); the National Natural Science Foundation of China (82204890); the Young Elite Scientists Sponsorship Program by China Association for Science and Technology (2023QNRC001).

References

Abdelaziz, E. H., Ismail, R., Mabrouk, M. S., and Amin, E. (2024). Multi-omics data integration and analysis pipeline for precision medicine: systematic review. *Comput. Biol. Chem.* 113:108254. doi: 10.1016/j.compbiolchem.2024.108254

Abedini-Nassab, R. (2017). Nanotechnology and Nanopore sequencing. Recent Pat. Nanotechnol. 11, 34–41. doi: 10.2174/1872210510666160602152913

Adelman, M. W., Woodworth, M. H., Langelier, C., Busch, L. M., Kempker, J. A., Kraft, C. S., et al. (2020). The gut microbiome's role in the development, maintenance, and outcomes of sepsis. *Crit. Care* 24:278. doi: 10.1186/s13054-020-02989-1

Agrawal, A., Balcı, H., Hanspers, K., Coort, S. L., Martens, M., Slenter, D. N., et al. (2024). WikiPathways 2024: next generation pathway database. *Nucleic Acids Res.* 52, D679–D689. doi: 10.1093/nar/gkad960

Ahmed, Z. (2022). Multi-omics strategies for personalized and predictive medicine: past, current, and future translational opportunities. *Emerg. Top. Life Sci.* 6, 215–225. doi: 10.1042/ETLS20210244

Ahmed, Z., Renart, E. G., and Zeeshan, S. (2021). Genomics pipelines to investigate susceptibility in whole genome and exome sequenced data for variant discovery, annotation, prediction and genotyping. *PeerJ* 9:e11724. doi: 10.7717/peerj.11724

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., et al. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 48, D517–D525. doi: 10.1093/nar/gkz935

Almeida, A., Mitchell, A. L., Boland, M., Forster, S. C., Gloor, G. B., Tarkowska, A., et al. (2019). A new genomic blueprint of the human gut microbiota. *Nature* 568, 499–504. doi: 10.1038/s41586-019-0965-1

Argelaguet, R., Arnol, D., Bredikhin, D., Deloro, Y., Velten, B., Marioni, J. C., et al. (2020). MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol.* 21:111. doi: 10.1186/s13059-020-02015-1

Armaos, A., Colantoni, A., Proietti, G., Rupert, J., and Tartaglia, G. G. (2021). catRAPID omics v2.0: going deeper and wider in the prediction of protein-RNA interactions. *Nucleic Acids Res.* 49, W72–W79. doi: 10.1093/nar/gkab393

Ayesha, S., Hanif, M. K., and Talib, R. (2020). Overview and comparative study of dimensionality reduction techniques for high dimensional data. *Inf. Fusion* 59, 44–58. doi: 10.1016/j.inffus.2020.01.005

Bartenhagen, C. RDRToolbox. (n.d) Available online at: http://bioconductor.org/packages/RDRToolbox/ (Accessed March 25, 2025).

Bauermeister, A., Mannochio-Russo, H., Costa-Lotufo, L. V., Jarmusch, A. K., and Dorrestein, P. C. (2022). Mass spectrometry-based metabolomics in microbiome investigations. *Nat. Rev. Microbiol.* 20, 143–160. doi: 10.1038/s41579-021-00621-9

Binnie, A., Tsang, J. L. Y., Hu, P., Carrasqueiro, G., Castelo-Branco, P., and dos Santos, C. C. (2020). Epigenetics of Sepsis. *Crit. Care Med.* 48, 745–756. doi: 10.1097/CCM.00000000004247

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Borghesi, A., Trück, J., Asgari, S., Sancho-Shimizu, V., Agyeman, P. K. A., Bellos, E., et al. (2020). Whole-exome sequencing for the identification of rare variants in primary immunodeficiency genes in children with sepsis: a prospective, population-based cohort study. *Clin. Infect. Dis.* 71:e614. doi: 10.1093/cid/ciaa290

Bouhaddani, S. E., Uh, H.-W., Jongbloed, G., Hayward, C., Klarić, L., Kiełbasa, S. M., et al. (2018). Integrating omics datasets with the OmicsPLS package. *BMC Bioinformatics* 19:371. doi: 10.1186/s12859-018-2371-3

Brown, C. L., Mullet, J., Hindi, F., Stoll, J. E., Gupta, S., Choi, M., et al. (2022). mobileOG-Db: a manually curated database of protein families mediating the life cycle of bacterial Mobile genetic elements. *Appl. Environ. Microbiol.* 88:e0099122. doi: 10.1128/aem.00991-22

Burton-Pimentel, K. J., Pimentel, G., Hughes, M., Michielsen, C. C., Fatima, A., Vionnet, N., et al. (2021). Discriminating dietary responses by combining transcriptomics and metabolomics data in nutrition intervention studies. *Mol. Nutr. Food Res.* 65:e2000647. doi: 10.1002/mnfr.202000647

Cano-Gamez, E., Burnham, K. L., Goh, C., Allcock, A., Malick, Z. H., Overend, L., et al. (2022). An immune dysfunction score for stratification of patients with acute infection based on whole-blood gene expression. *Sci. Transl. Med.* 14:eabq4433. doi: 10.1126/scitranslmed.abq4433

Cantrell, K., Fedarko, M. W., Rahman, G., McDonald, D., Yang, Y., Zaw, T., et al. (2021). Empress enables tree-guided, interactive, and exploratory analyses of multi-omic data sets. *mSystems* 6:20. doi: 10.1128/mSystems.01216-20

Caspi, R., Foerster, H., Fulcher, C. A., Hopkinson, R., Ingraham, J., Kaipa, P., et al. (2006). MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Res.* 34, D511–D516. doi: 10.1093/nar/gkj128

Cerami, E. G., Gross, B. E., Demir, E., Rodchenkov, I., Babur, O., Anwar, N., et al. (2011). Pathway commons, a web resource for biological pathway data. *Nucleic Acids Res.* 39, D685–D690. doi: 10.1093/nar/gkq1039

Chen, I.-M. A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., et al. (2019). IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res.* 47, D666–D677. doi: 10.1093/nar/gky901

Chen, W., Guo, W., Li, Y., and Chen, M. (2024). Integrative analysis of metabolomics and transcriptomics to uncover biomarkers in Sepsis. *Sci. Rep.* 14:9676. doi: 10.1038/s41598-024-59400-0

Chen, C., Wang, J., Pan, D., Wang, X., Xu, Y., Yan, J., et al. (2023). Applications of multi-omics analysis in human diseases. *MedComm* 4:e315. doi: 10.1002/mco2.315

Cheng, C., Chen, W., Jin, H., and Chen, X. (2023). A review of single-cell RNA-Seq annotation, integration, and cell-cell communication. *Cells* 12:1970. doi: 10.3390/cells12151970

Chu, D., Liao, L.-Z., Ng, M. K., and Zhang, X. (2013). Sparse canonical correlation analysis: new formulation and algorithm. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 3050–3065. doi: 10.1109/TPAMI.2013.104

Cook, K. B., Kazan, H., Zuberi, K., Morris, Q., and Hughes, T. R. (2011). RBPDB: a database of RNA-binding specificities. *Nucleic Acids Res.* 39, D301–D308. doi: 10.1093/nar/gkq1069

Córneo, E. d. S., Michels, M., and Dal-Pizzol, F. (2021). Sepsis, immunosuppression and the role of epigenetic mechanisms. *Expert Rev. Clin. Immunol.* 17, 169–176. doi: 10.1080/1744666X.2021.1875820

Croft, D., O'Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L., et al. (2011). Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* 39, D691–D697. doi: 10.1093/nar/gkq1018

Culhane, A. C., Thioulouse, J., Perrière, G., and Higgins, D. G. (2005). MADE4: an R package for multivariate analysis of gene expression data. *Bioinformatics* 21, 2789–2790. doi: 10.1093/bioinformatics/bti394

D'Angelo, G. M., Rao, D., and Gu, C. C. (2009). Combining least absolute shrinkage and selection operator (LASSO) and principal-components analysis for detection of gene-gene interactions in genome-wide association studies. *BMC Proc.* 3, S7–S62. doi: 10.1186/1753-6561-3-S7-S62

Daliri, E. B.-M., Ofosu, F. K., Chelliah, R., Lee, B. H., and Oh, D.-H. (2021). Challenges and perspective in integrated multi-omics in gut microbiota studies. *Biomol. Ther.* 11:300. doi: 10.3390/biom11020300

Davis, J. J., Wattam, A. R., Aziz, R. K., Brettin, T., Butler, R., Butler, R. M., et al. (2020). The PATRIC bioinformatics resource center: expanding data and analysis capabilities. *Nucleic Acids Res.* 48, D606–D612. doi: 10.1093/nar/gkz943

Devarajan, K. (2008). Nonnegative matrix factorization: an analytical and interpretive tool in computational biology. *PLoS Comput. Biol.* 4:e1000029. doi: 10.1371/journal.pcbi.1000029

Dibaeinia, P., Ojha, A., and Sinha, S. (2025). Interpretable AI for inference of causal molecular relationships from omics data. *Sci. Adv.* 11:eadk0837. doi: 10.1126/sciadv.adk0837

Downing, T., and Angelopoulos, N. (2023). A primer on correlation-based dimension reduction methods for multi-omics analysis. J. R. Soc. Interface 20:20230344. doi: 10.1098/rsif.2023.0344

Dray, S., and Dufour, A.-B. (2024). The ade4 package: implementing the duality diagram for ecologists. *ResearchGate*. doi: 10.18637/jss.v022.i04

Duong, V.-A., and Lee, H. (2023). Bottom-up proteomics: advancements in sample preparation. *Int. J. Mol. Sci.* 24:5350. doi: 10.3390/ijms24065350

Evans, L., Rhodes, A., Alhazzani, W., Antonelli, M., Coopersmith, C. M., French, C., et al. (2021). Surviving Sepsis campaign: international guidelines for Management of Sepsis and Septic Shock 2021. *Intensive Care Med.* 47, 1181–1247. doi: 10.1007/s00134-021-06506-y

Fang, L., Li, Y., Ma, L., Xu, Q., Tan, F., and Chen, G. (2021). GRNdb: decoding the gene regulatory networks in diverse human and mouse conditions. *Nucleic Acids Res.* 49, D97–D103. doi: 10.1093/nar/gkaa995

Fay, K. T., Klingensmith, N. J., Chen, C.-W., Zhang, W., Sun, Y., Morrow, K. N., et al. (2019). The gut microbiome alters immunophenotype and survival from sepsis. *FASEB J.* 33, 11258–11269. doi: 10.1096/fj.201802188R

Fleischmann, C., Scherag, A., Adhikari, N. K. J., Hartog, C. S., Tsaganos, T., Schlattmann, P., et al. (2016). Assessment of global incidence and mortality of hospitaltreated sepsis. Current estimates and limitations. *Am. J. Respir. Crit. Care Med.* 193, 259–272. doi: 10.1164/rccm.201504-0781OC

Franzosa, E. A., Morgan, X. C., Segata, N., Waldron, L., Reyes, J., Earl, A. M., et al. (2014). Relating the Metatranscriptome and metagenome of the human gut. *Proc. Natl. Acad. Sci.* 111, E2329–E2338. doi: 10.1073/pnas.1319284111

Frerichs, N. M., El Hassani, E. M., Deianova, N., van Weissenbruch, M. M., van Kaam, A. H., Vijlbrief, D. C., et al. (2023). Fecal volatile metabolomics predict gramnegative late-onset Sepsis in preterm infants: a Nationwide case-control study. *Microorganisms* 11:572. doi: 10.3390/microorganisms11030572

Furusawa, Y., Obata, Y., Fukuda, S., Endo, T. A., Nakato, G., Takahashi, D., et al. (2013). Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* 504, 446–450. doi: 10.1038/nature12721

Gates, L. A., Reis, B. S., Lund, P. J., Paul, M. R., Leboeuf, M., Djomo, A. M., et al. (2024). Histone Butyrylation in the mouse intestine is mediated by the microbiota and associated with regulation of gene expression. *Nat. Metab.* 6, 697–707. doi: 10.1038/s42255-024-00992-2

Gill, C., and Marchini, J. Four-dimensional sparse Bayesian tensor decomposition for gene expression data (2020). bioRxiv [Preprint]. Available at: https://www.biorxiv.org/ content/10.1101/2020.11.30.403907v1.article-info (Accessed March 15, 2025).

Giordani, P., Kiers, H. A. L., and Ferraro, M. A. D. (2014). Three-way component analysis using the R package ThreeWay. *J. Stat. Softw.* 57, 1–23. doi: 10.18637/jss.v057.i07

Gliozzo, J., Soto-Gomez, M., Guarino, V., Bonometti, A., Cabri, A., Cavalleri, E., et al. (2024). Intrinsic-dimension analysis for guiding dimensionality reduction and data fusion in multi-omics data processing. *Artif. Intell. Med.* 160:103049. doi: 10.1016/j.artmed.2024.103049

Gloaguen, A., Philippe, C., Frouin, V., Gennari, G., Dehaene-Lambertz, G., Le Brusquet, L., et al. (2022). Multiway generalized canonical correlation analysis. *Biostatistics* 23, 240–256. doi: 10.1093/biostatistics/kxaa010 Gong, F.-C., Ji, R., Wang, Y.-M., Yang, Z.-T., Chen, Y., Mao, E.-Q., et al. (2020). Identification of potential biomarkers and immune features of sepsis using bioinformatics analysis. *Mediat. Inflamm.* 2020;3432587. doi: 10.1155/2020/3432587

González, I., and Déjean, S. (2023). CCA: Canonical Correlation Analysis [R package]. Version 1.2.2. Available online at: https://CRAN.R-project.org/package=CCA (Accessed March 25, 2025).

Guo, Y., Zou, H., Alam, M. S., and Luo, S. (2025). Integrative multi-omics and multivariate longitudinal data analysis for dynamic Risk estimation in Alzheimer's Disease. *Stat. Med.* 44:e70105. doi: 10.1002/sim.70105

Haak, B. W., Argelaguet, R., Kinsella, C. M., Kullberg, R. F. J., Lankelma, J. M., Deijs, M., et al. (2021). Integrative transkingdom analysis of the gut microbiome in antibiotic perturbation and critical illness. *mSystems* 6:48. doi: 10.1128/msystems.01148-20

Heffernan, I. M., McGeary, J. E., Chung, C.-S., Ayala, A., and Heffernan, D. S. (2021). Unmasking unique immune altering aspects of the microbiome as a tool to correct Sepsis-induced immune dysfunction. *Surg. Infect.* 22, 400–408. doi: 10.1089/sur.2020. 233

Huang, W., Wang, G., Yin, C., Chen, D., Dhand, A., Chanza, M., et al. (2019). Optimizing a whole-genome sequencing data processing pipeline for precision surveillance of health care-associated infections. *Microorganisms* 7:388. doi: 10.3390/microorganisms7100388

Iperi, C., Fernández-Ochoa, Á., Barturen, G., Pers, J.-O., Foulquier, N., Bettacchioli, E., et al. (2025). BiomiX, a user-friendly bioinformatic tool for democratized analysis and integration of multiomics data. *BMC Bioinformatics* 26:8. doi: 10.1186/s12859-024-06022-y

Isac, S., Isac, T., Tanasescu, M. D., Pavel, B., Andreescu, C. V., Badea, A.-G., et al. (2024). The omics complexity in Sepsis: the limits of the personalized medicine approach. *J. Pers. Med.* 14:225. doi: 10.3390/jpm14030225

Janosevic, D., Myslinski, J., McCarthy, T. W., Zollman, A., Syed, F., Xuei, X., et al. (2021). The orchestrated cellular and molecular responses of the kidney to endotoxin define a precise sepsis timeline. *eLife* 10:e62270. doi: 10.7554/eLife.62270

Jeon, H., and Oh, S. (2020). Hybrid-recursive feature elimination for efficient feature selection. *Appl. Sci.* 10:3211. doi: 10.3390/app10093211

Jolliffe, I. T., and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philos. Transact. A Math. Phys. Eng. Sci.* 374:20150202. doi: 10.1098/rsta.2015.0202

Jonathan, K. J. T., Ong, G., Prasetyaningsih, F. A., Amandito, R., Rohsiswatmo, R., and Malik, A. (2020). Clinical characteristics influence cultivable-Bacteria composition in the meconium of Indonesian neonates. *Heliyon* 6:e05576. doi: 10.1016/j.heliyon.2020.e05576

Joubert, I. A., Otto, M., Strunk, T., and Currie, A. J. (2022). Look who's talking: host and pathogen drivers of *Staphylococcus Epidermidis* virulence in neonatal Sepsis. *Int. J. Mol. Sci.* 23:860. doi: 10.3390/ijms23020860

Kalantar, K. L., Neyton, L., Abdelghany, M., Mick, E., Jauregui, A., Caldera, S., et al. (2022). Integrated host-microbe plasma metagenomics for Sepsis diagnosis in a prospective cohort of critically ill adults. *Nat. Microbiol.* 7, 1805–1816. doi: 10.1038/s41564-022-01237-2

Kalra, S. J. S., Shankar, H., Mansoori, N., and Gupta, D. L. (2022). Antibiotic-resistant bacteria originating from the gut may modulate the mucosal immune response during sepsis and septic shock. *Drug Target Insights* 16, 81–87. doi: 10.33393/dti.2022.2520

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., and Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 51, D587–D592. doi: 10.1093/nar/gkac963

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Kang, H., and Thomas, R. M. (2021). Bacteria and Sepsis: microbiome to the rescue? J. Clin. Med. 10:3578. doi: 10.3390/jcm10163578

Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., et al. (2019). The BioCyc collection of microbial genomes and metabolic pathways. *Brief. Bioinform.* 20, 1085–1093. doi: 10.1093/bib/bbx085

Khan, M. M., Ernst, O., Manes, N. P., Oyler, B. L., Fraser, I. D. C., Goodlett, D. R., et al. (2019). Multi-omics strategies uncover host-pathogen interactions. *ACS Infect. Dis.* 5, 493–505. doi: 10.1021/acsinfecdis.9b00080

Kim, S. M., DeFazio, J. R., Hyoju, S. K., Sangani, K., Keskey, R., Krezalek, M. A., et al. (2020). Fecal microbiota transplant rescues mice from human pathogen mediated Sepsis by restoring systemic immunity. *Nat. Commun.* 11:2354. doi: 10.1038/s41467-020-15545-w

Kingma, D. P., and Welling, M. (2019). An introduction to variational autoencoders. *Foundat. Trends* 12, 307–392. doi: 10.1561/2200000056

Koppel, N., and Balskus, E. P. (2016). Exploring and understanding the biochemical diversity of the human microbiota. *Cell Chem. Biol.* 23, 18–30. doi: 10.1016/j.chembiol.2015.12.008

Kotlyar, M., Pastrello, C., Ahmed, Z., Chee, J., Varyova, Z., and Jurisica, I. (2022). IID 2021: towards context-specific protein interaction analyses by increased coverage, enhanced annotation and enrichment analysis. *Nucleic Acids Res.* 50, D640–D647. doi: 10.1093/nar/gkab1034

Kraemer, G., Reichstein, M., and Mahecha, D. (2018). dimRed and coRanking---Unifying Dimensionality Reduction in R. *R J.* 10, 342–358.

Kuhn, M. (2024). Classification and regression training [R package caret version 7.0–1] Available online at: https://CRAN.R-project.org/package=caret (Accessed March 25, 2025).

Kuhn, M., Wing, J., Weston, S., and Williams, A. (2024). Caret: classification and regression training version 6.0–94 from CRAN. Available online at: https://rdrr.io/cran/caret/ (Accessed March 25, 2025).

Kullberg, R. F. J., Wiersinga, W. J., and Haak, B. W. (2021). Gut microbiota and Sepsis: from pathogenesis to novel treatments. *Curr. Opin. Gastroenterol.* 37, 578–585. doi: 10.1097/MOG.00000000000781

Kurian, J. R., Evangelatos, N., Kamath, A., and Brand, H. (2024). Sepsis research using omics technology in the European Union and the United Kingdom: maps, trends, and future implications. *OMICS* 28, 357–366. doi: 10.1089/omi.2024.0089

Lang, B., Armaos, A., and Tartaglia, G. G. (2019). RNAct: protein-RNA interaction predictions for model organisms with supporting experimental data. *Nucleic Acids Res.* 47, D601–D606. doi: 10.1093/nar/gky967

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* 31, 814–821. doi: 10.1038/nbt.2676

Lê, S., Josse, J., and Husson, F. (2008). Factominer: an R package for multivariate analysis. J. Stat. Softw. 25, 1–18. doi: 10.18637/jss.v025.i01

Lee, J., and Banerjee, D. (2020). Metabolomics and the microbiome as biomarkers in Sepsis. *Crit. Care Clin.* 36, 105–113. doi: 10.1016/j.ccc.2019.08.008

Lee, L. C., Liong, C.-Y., and Jemain, A. A. (2018). Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst* 143, 3526–3539. doi: 10.1039/C8AN00599K

Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., et al. (2020). BioBERT: a pretrained biomedical language representation model for biomedical text mining. *Bioinformatics* 36, 1234–1240. doi: 10.1093/bioinformatics/btz682

Li, W., Ballard, J., Zhao, Y., and Long, Q. (2024). Knowledge-guided learning methods for integrative analysis of multi-omics data. *Comput. Struct. Biotechnol. J.* 23, 1945–1950. doi: 10.1016/j.csbj.2024.04.053

Li, J., Bien, J., and Wells, M. T. (2018). rTensor: an R package for multidimensional array (tensor) unfolding, multiplication, and decomposition. *J. Stat. Softw.* 87, 1–31. doi: 10.18637/jss.v087.i10

Li, J.-H., Liu, S., Zhou, H., Qu, L.-H., and Yang, J.-H. (2014). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 42, D92–D97. doi: 10.1093/nar/gkt1248

Li, Y., van Houten, C. B., Boers, S. A., Jansen, R., Cohen, A., Engelhard, D., et al. (2022). The diagnostic value of nasal microbiota and clinical parameters in a multiparametric prediction model to differentiate bacterial versus viral infections in lower respiratory tract infections. *PLoS One* 17:e0267140. doi: 10.1371/journal.pone.0267140

Li, D., Wei, R., Zhang, X., Gong, S., Wan, M., Wang, F., et al. (2024). Gut commensal metabolite Rhamnose promotes macrophages phagocytosis by activating SLC12A4 and protects against Sepsis in mice. *Acta Pharm. Sin. B* 14, 3068–3085. doi: 10.1016/j.apsb.2024.03.025

Li, Z., Zhang, F., Sun, M., Liu, J., Zhao, L., Liu, S., et al. (2023). The modulatory effects of gut microbes and metabolites on blood-brain barrier integrity and brain function in Sepsis-associated encephalopathy. *PeerJ* 11:e15122. doi: 10.7717/peerj.15122

Lin, B. M., Cho, H., Liu, C., Roach, J., Ribeiro, A. A., Divaris, K., et al. (2023). BZINB model-based pathway analysis and module identification facilitates integration of microbiome and metabolome data. *Microorganisms* 11:766. doi: 10.3390/microorganisms11030766

Lisko, D. J., Johnston, G. P., and Johnston, C. G. (2017). Effects of dietary yogurt on the healthy human gastrointestinal (GI) microbiome. *Microorganisms* 5:6. doi: 10.3390/microorganisms5010006

Liu, W., Cheng, M., Li, J., Zhang, P., Fan, H., Hu, Q., et al. (2021). Classification of the gut microbiota of patients in intensive care units during development of sepsis and septic shock. *Genomics Proteomics Bioinformatics* 18, 696–707. doi: 10.1016/j.gpb.2020. 06.011

Liu, Z., Li, N., Fang, H., Chen, X., Guo, Y., Gong, S., et al. (2019). Enteric Dysbiosis is associated with Sepsis in patients. *FASEB J.* 33, 12299–12310. doi: 10.1096/fj.201900398RR

Liu, Y., Li, D., Zhang, X., Xia, S., Qu, Y., Ling, X., et al. (2024). A protein sequencebased deep transfer learning framework for identifying human proteome-wide Deubiquitinase-substrate interactions. *Nat. Commun.* 15:4519. doi: 10.1038/s41467-024-48446-3

Long, X., Mu, S., Zhang, J., Xiang, H., Wei, W., Sun, J., et al. (2023). Global signatures of the microbiome and metabolome during hospitalization of septic patients. *Shock* 59, 716–724. doi: 10.1097/SHK.00000000002117

Losert, C., Pekayvaz, K., Knottenberg, V., Nicolai, L., Stark, K., and Heinig, M. (2024). Application of unsupervised multi-omic factor analysis to uncover patterns of variation and molecular processes linked to cardiovascular disease. J. Vis. Exp. e66659. doi: 10.3791/66659

Lu, F., Huang, T., Chen, R., and Yin, H. (2024). Multi-omics analysis reveals the interplay between pulmonary microbiome and host in immunocompromised patients with Sepsis-induced acute lung injury. *Microbiol. Spectr.* 12:e0142424. doi: 10.1128/spectrum.01424-24

Ma, X., Jin, H., Chu, X., Dai, W., Tang, W., Zhu, J., et al. (2022). The host CYP1A1microbiota metabolic Axis promotes gut barrier disruption in methicillin-resistant *Staphylococcus Aureus*-induced abdominal Sepsis. *Front. Microbiol.* 13:802409. doi: 10.3389/fmicb.2022.802409

Ma, Y., Liu, L., Ma, Y., and Zhang, S. H. O. N. M. F. (2023). Integration analysis of multi-omics microbiome data via matrix factorization and hypergraph. *Bioinforma. Oxf. Engl.* 39:335. doi: 10.1093/bioinformatics/btad335

Ma, Y., Zhao, Y., and Zhang, X. (2024). Factors affecting neutrophil functions during Sepsis: human microbiome and epigenetics. *J. Leukoc. Biol.* 116, 672–688. doi: 10.1093/jleuko/qiae107

Mahat, D. B., Tippens, N. D., Martin-Rufino, J. D., Waterton, S. K., Fu, J., Blatt, S. E., et al. (2024). Single-cell nascent RNA sequencing unveils coordinated global transcription. *Nature* 631, 216–223. doi: 10.1038/s41586-024-07517-7

Malik, A., Kamli, M. R., Sabir, J. S. M., Rather, I. A., Phan, L. T., Kim, C.-B., et al. (2024). APLpred: a machine learning-based tool for accurate prediction and characterization of asparagine peptide Lyases using sequence-derived optimal features. *Methods San Diego Calif* 229, 133–146. doi: 10.1016/j.ymeth.2024.05.014

Mallick, H., Rahnavard, A., McIver, L. J., Ma, S., Zhang, Y., Nguyen, L. H., et al. (2021). Multivariable association discovery in population-scale Meta-omics studies. *PLoS Comput. Biol.* 17:e1009442. doi: 10.1371/journal.pcbi.1009442

Mangioni, D., Peri, A. M., Rossolini, G. M., Viaggi, B., Perno, C. F., Gori, A., et al. (2020). Toward rapid Sepsis diagnosis and patient stratification: what's new from microbiology and omics science. *J. Infect. Dis.* 221, 1039–1047. doi: 10.1093/infdis/jiz585

Mann, E. R., Lam, Y. K., and Uhlig, H. H. (2024). Short-chain fatty acids: linking diet, the microbiome and immunity. *Nat. Rev. Immunol.* 24:14. doi: 10.1038/s41577-024-01014-8

Mayhew, M. B., Buturovic, L., Luethy, R., Midic, U., Moore, A. R., Roque, J. A., et al. (2020). A generalizable 29-mRNA neural-network classifier for acute bacterial and viral infections. *Nat. Commun.* 11:1177. doi: 10.1038/s41467-020-14975-w

Mba Medie, F., Sharma-Kuinkel, B. K., Ruffin, F., Chan, L. C., Rossetti, M., Chang, Y.-L., et al. (2019). Genetic variation of DNA methyltransferase-3A contributes to protection against persistent MRSA bacteremia in patients. *Proc. Natl. Acad. Sci. USA* 116, 20087–20096. doi: 10.1073/pnas.1909849116

McDonald, B., Zucoloto, A. Z., Yu, I.-L., Burkhard, R., Brown, K., Geuking, M. B., et al. (2020). Programing of an intravascular immune firewall by the gut microbiota protects against pathogen dissemination during infection. *Cell Host Microbe* 28, 660–668.e4. doi: 10.1016/j.chom.2020.07.014

Meng, C., Helm, D., Frejno, M., and Kuster, B. (2016). moCluster: identifying joint patterns across multiple omics data sets. *J. Proteome Res.* 15, 755–765. doi: 10.1021/acs.jproteome.5b00824

Meng, C., Kuster, B., Culhane, A. C., and Gholami, A. M. (2014). A multivariate approach to the integration of multi-omics datasets. *BMC Bioinformatics* 15:162. doi: 10.1186/1471-2105-15-162

Mevik, B.-H., and Wehrens, R. (2007). The Pls package: principal component and partial least squares regression in R. J. Stat. Softw. 18, 1–23. doi: 10.18637/jss.v018.i02

Miao, H., Chen, S., and Ding, R. (2021). Evaluation of the molecular mechanisms of Sepsis using proteomics. *Front. Immunol.* 12:733537. doi: 10.3389/fimmu.2021.733537

Miettinen, J., Nordhausen, K., and Taskinen, S. (2018). fICA: FastICA algorithms and their improved variants. *R J.* 10:148. doi: 10.32614/RJ-2018-046

Mitchell, K., Ronas, J., Dao, C., Freise, A. C., Mangul, S., Shapiro, C., et al. (2020). PUMAA: a platform for accessible microbiome analysis in the undergraduate classroom. *Front. Microbiol.* 11:584699. doi: 10.3389/fmicb.2020.584699

Mrowka, R., and Schmauder, R. (2024). How to visualize high-dimensional data. Acta Physiol. 240:e14219. doi: 10.1111/apha.14219

Mu, A., Klare, W. P., Baines, S. L., Ignatius Pang, C. N., Guérillot, R., Harbison-Price, N., et al. (2023). Integrative omics identifies conserved and pathogen-specific responses of Sepsis-causing Bacteria. *Nat. Commun.* 14:1530. doi: 10.1038/s41467-023-37200-w

Muller, E., Shiryan, I., and Borenstein, E. (2024). Multi-Omic integration of microbiome data for identifying Disease-associated modules. *Nat. Commun.* 15:2621. doi: 10.1038/s41467-024-46888-3

Muppirala, U. K., Honavar, V. G., and Dobbs, D. (2011). Predicting RNA-protein interactions using only sequence information. *BMC Bioinformatics* 12:489. doi: 10.1186/1471-2105-12-489

Murden, R. J., Zhang, Z., Guo, Y., and Risk, B. B. (2022). Interpretive JIVE: connections with CCA and an application to brain connectivity. *Front. Neurosci.* 16:969510. doi: 10.3389/fnins.2022.969510

Nagpal, S., Haque, M. M., Singh, R., and Mande, S. S. (2018). iVikodak-A platform and standard workflow for inferring, analyzing, comparing, and visualizing the

functional potential of microbial communities. Front. Microbiol. 9:3336. doi: 10.3389/fmicb.2018.03336

Narayan, N. R., Weinmaier, T., Laserna-Mendieta, E. J., Claesson, M. J., Shanahan, F., Dabbagh, K., et al. (2020). Piphillin predicts metagenomic composition and dynamics from DADA2-corrected 16S rDNA sequences. *BMC Genomics* 21:56. doi: 10.1186/s12864-019-6427-1

Nenadic, O., and Greenacre, M. (2007). Correspondence analysis in R, with two-and three-dimensional graphics: the ca package. *J. Stat. Softw.* 20, 1–13. doi: 10.18637/jss.v020.i03

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., et al. (2025). Vegan: community ecology package [R package]. Version 2.7-1. Available online at: https://CRAN.R-project.org/package=vegan (Accessed March 25, 2025).

Oughtred, R., Rust, J., Chang, C., Breitkreutz, B.-J., Stark, C., Willems, A., et al. (2021). The BioGRID database: a comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.* 30, 187–200. doi: 10.1002/pro. 3978

Pandey, S. (2024). Sepsis, management & advances in metabolomics. *Nano* 8, 270–284. doi: 10.7150/ntno.94071

Picard, M., Scott-Boyer, M.-P., Bodein, A., Périn, O., and Droit, A. (2021). Integration strategies of multi-omics data for machine learning analysis. *Comput. Struct. Biotechnol. J.* 19, 3735–3746. doi: 10.1016/j.csbj.2021.06.030

Piccioni, A., Spagnuolo, F., Candelli, M., Voza, A., Covino, M., Gasbarrini, A., et al. (2024). The gut microbiome in Sepsis: from Dysbiosis to personalized therapy. *J. Clin. Med.* 13:6082. doi: 10.3390/jcm13206082

Pierrakos, C., Velissaris, D., Bisdorff, M., Marshall, J. C., and Vincent, J.-L. (2020). Biomarkers of sepsis: time for a reappraisal. *Crit. Care* 24:287. doi: 10.1186/s13054-020-02993-5

Potruch, A., Schwartz, A., and Ilan, Y. (2022). The role of bacterial translocation in Sepsis: a new target for therapy. *Ther. Adv. Gastroenterol.* 15:17562848221094214. doi: 10.1177/17562848221094214

Rayzan, E., Mirbeyk, M., Pezeshki, P. S., Mohammadpour, M., Yaghmaie, B., Hassani, S. A., et al. (2023). Whole-exome sequencing to identify undiagnosed primary immunodeficiency disorders in children with community-acquired sepsis, admitted in the pediatric intensive care unit. *Pediatr. Allergy Immunol.* 34:e14066. doi: 10.1111/pai.14066

Reyman, M., van Houten, M. A., Watson, R. L., Chu, M. L. J. N., Arp, K., de Waal, W. J., et al. (2022). Effects of early-life antibiotics on the developing infant gut microbiome and Resistome: a randomized trial. *Nat. Commun.* 13:893. doi: 10.1038/s41467-022-28525-z

Ringnér, M. (2008). What is principal component analysis? Nat. Biotechnol. 26, 303-304. doi: 10.1038/nbt0308-303

Rohart, F., Gautier, B., Singh, A., and Cao, K.-A. L. (2017). mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* 13:e1005752. doi: 10.1371/journal.pcbi.1005752

Rougier, J. (2025). Tensor: tensor product of arrays [R package]. Version 1.5.1. Available at: https://CRAN.R-project. org/package=tensor (Accessed March 25, 2025).

Rouillard, A. D., Gundersen, G. W., Fernandez, N. F., Wang, Z., Monteiro, C. D., McDermott, M. G., et al. (2016). The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database* 2016:baw100. doi: 10.1093/database/baw100

Rudd, K. E., Johnson, S. C., Agesa, K. M., Shackelford, K. A., Tsoi, D., Kievlan, D. R., et al. (2020). Global, regional, and National Sepsis Incidence and mortality, 1990–2017: analysis for the global burden of Disease study. *Lancet* 395, 200–211. doi: 10.1016/S0140-6736(19)32989-7

Santiago-Rodriguez, T. M., and Hollister, E. B. (2021). Multi 'omic data integration: a review of concepts, considerations, and approaches. *Semin. Perinatol.* 45:151456. doi: 10.1016/j.semperi.2021.151456

Santiago-Rodriguez, T. M., Naidu, M., Abeles, S. R., Boehm, T. K., Ly, M., and Pride, D. T. (2015). Transcriptome analysis of bacteriophage communities in periodontal health and Disease. *BMC Genomics* 16:549. doi: 10.1186/s12864-015-1781-0

Saxena, J., Das, S., Kumar, A., Sharma, A., Sharma, L., Kaushik, S., et al. (2024). Biomarkers in Sepsis. *Clin. Chim. Acta* 562:119891. doi: 10.1016/j.cca.2024.119891

Schuurman, A. R., Reijnders, T. D. Y., Kullberg, R. F. J., Butler, J. M., Van Der Poll, T., and Wiersinga, W. J. (2021). Sepsis: deriving biological meaning and clinical applications from high-dimensional data. *Intensive Care Med. Exp.* 9:27. doi: 10.1186/s40635-021-00383-x

Serek, P., and Oleksy-Wawrzyniak, M. (2021). The effect of bacterial infections, probiotics and zonulin on intestinal barrier integrity. *Int. J. Mol. Sci.* 22:11359. doi: 10.3390/ijms222111359

Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., et al. (2016). The third international consensus definitions for Sepsis and septic shock (Sepsis-3). *JAMA* 315, 801–810. doi: 10.1001/jama.2016.0287

Stacklies, W., Redestig, H., Scholz, M., Walther, D., and Selbig, J. (2007). pcaMethods—a Bioconductor package providing PCA methods for incomplete data. *Bioinformatics* 23, 1164–1167. doi: 10.1093/bioinformatics/btm069 Stewart, C. J., Embleton, N. D., Marrs, E. C. L., Smith, D. P., Fofanova, T., Nelson, A., et al. (2017). Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset Sepsis and healthy controls. *Microbiome* 5:75. doi: 10.1186/s40168-017-0295-1

Sullivan, P. F., Meadows, J. R. S., Gazal, S., Phan, B. N., Li, X., Genereux, D. P., et al. (2023). Leveraging Base-pair mammalian constraint to understand genetic variation and human Disease. *Science* 380:eabn2937. doi: 10.1126/science.abn2937

Sun, J., Ding, X., Liu, S., Duan, X., Liang, H., and Sun, T. (2020). Adipose-derived mesenchymal stem cells attenuate acute lung injury and improve the gut microbiota in septic rats. *Stem Cell Res Ther* 11:384. doi: 10.1186/s13287-020-01902-5

Sun, S., Wang, D., Dong, D., Xu, L., Xie, M., Wang, Y., et al. (2023). Altered intestinal microbiome and metabolome correspond to the clinical outcome of Sepsis. *Crit. Care* 27:127. doi: 10.1186/s13054-023-04412-x

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., et al. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 51, D638–D646. doi: 10.1093/nar/gkac1000

Thazha, S. K., Chandrean, S. P., and Akbar, Z. (2019). The role of antibiotics in mucosal colonization and invasion with regards to gut microbes. *Int. J. Health Sci. Res.* 9, 230–240.

The ConsensusPathDB Interaction Database (2013) Update - PubMed Available online at: https://pubmed.ncbi.nlm.nih.gov/23143270/ (Accessed March 15, 2025).

UniProt Consortium (2025). UniProt: the universal protein knowledgebase in 2025. Nucleic Acids Res. 53, D609–D617. doi: 10.1093/nar/gkae1010

Vallenet, D., Calteau, A., Dubois, M., Amours, P., Bazin, A., Beuvin, M., et al. (2020). MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, Pangenomic and metabolic comparative analysis. *Nucleic Acids Res.* 48, D579–D589. doi: 10.1093/nar/gkz926

van Houten, C. B., Oved, K., Eden, E., Cohen, A., Engelhard, D., Boers, S., et al. (2018). Observational multi-Centre, prospective study to characterize novel pathogen-and host-related factors in hospitalized patients with lower respiratory tract infections and/or Sepsis - the "TAILORED-treatment" study. *BMC Infect. Dis.* 18:377. doi: 10.1186/s12879-018-3300-9

Virta, J., Koesner, C. L., Li, B., Nordhausen, K., Oja, H., and Radojicic, U. (2024). TensorBSS: Blind source separation methods for tensor-valued observations [R package]. Version 0.3.9. Available online at: https://CRAN.R-project.org/ package=tensorBSS (Accessed March 25, 2025).

Wang, L., Cao, J.-B., Xia, B.-B., Li, Y.-J., Zhang, X., Mo, G.-X., et al. (2023). Metatranscriptome of human lung microbial communities in a cohort of mechanically ventilated COVID-19 omicron patients. *Signal Transduct. Target. Ther.* 8, 432–412. doi: 10.1038/s41392-023-01684-1

Wang, Y., Deng, H., Xiao, L., and Pan, Y. (2024a). *Escherichia Coli* Nissle 1917 protects against Sepsis-induced intestinal damage by regulating the SCFA/GPRs signaling pathway. *Microorganisms* 12:1622. doi: 10.3390/microorganisms12081622

Wang, Y., Feng, S., Shi, H., Lu, Y., Zhang, J., Zhang, W., et al. (2024b). Analysis of alterations in serum vitamins and correlations with gut microbiome, microbial metabolomics in patients with sepsis. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 1237:124101. doi: 10.1016/j.jchromb.2024.124101

Wang, C., Li, Q., Tang, C., Zhao, X., He, Q., Tang, X., et al. (2021). Characterization of the blood and neutrophil-specific microbiomes and exploration of potential bacterial biomarkers for Sepsis in surgical patients. *Immun. Inflamm. Dis.* 9, 1343–1357. doi: 10.1002/iid3.483

Wang, B., Mezlini, A. M., Demir, F., Fiume, M., Tu, Z., Brudno, M., et al. (2014). Similarity network fusion for aggregating data types on a genomic scale. *Nat. Methods* 11, 333–337. doi: 10.1038/nmeth.2810

Wang, J., Ye, F., Chai, H., Jiang, Y., Wang, T., Ran, X., et al. (2024). Advances and applications in single-cell and spatial genomics. *Sci. China Life Sci.* 68:70. doi: 10.1007/s11427-024-2770-x

Wang, Y.-X., and Zhang, Y.-J. (2013). Nonnegative matrix factorization: a comprehensive review. *IEEE Trans. Knowl. Data Eng.* 25, 1336–1353. doi: 10.1109/TKDE.2012.51

Wang, W., Zhang, X., and Dai, D.-Q. (2021). DeFusion: a Denoised network regularization framework for multi-omics integration. *Brief. Bioinform.* 22:57. doi: 10.1093/bib/bbab057

Wattam, A. R., Abraham, D., Dalay, O., Disz, T. L., Driscoll, T., Gabbard, J. L., et al. (2014). PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res.* 42, D581–D591. doi: 10.1093/nar/gkt1099

Weersma, R. K., Zhernakova, A., and Fu, J. (2020). Interaction between drugs and the gut microbiome. *Gut* 69, 1510–1519. doi: 10.1136/gutjnl-2019-320204

Winnenburg, R., Baldwin, T. K., Urban, M., Rawlings, C., Köhler, J., and Hammond-Kosack, K. E. (2006). PHI-base: a new database for pathogen host interactions. *Nucleic Acids Res.* 34, D459–D464. doi: 10.1093/nar/gkj047

Wishart, D. S., Kruger, R., Sivakumaran, A., Harford, K., Sanford, S., Doshi, R., et al. (2024). PathBank 2.0-the pathway database for model organism metabolomics. *Nucleic Acids Res.* 52, D654–D662. doi: 10.1093/nar/gkad1041

Woo, V., and Alenghat, T. (2022). Epigenetic regulation by gut microbiota. *Gut Microbes* 14:2022407. doi: 10.1080/19490976.2021.2022407

Wu, T., Li, H., Su, C., Xu, F., Yang, G., Sun, K., et al. (2020a). Microbiota-derived shortchain fatty acids promote LAMTOR2-mediated immune responses in macrophages. *mSystems* 5:e00587. doi: 10.1128/mSystems.00587-20

Wu, T., Xu, F., Su, C., Li, H., Lv, N., Liu, Y., et al. (2020b). Alterations in the gut microbiome and Cecal metabolome during *Klebsiella Pneumoniae*-induced Pneumosepsis. *Front. Immunol.* 11:1331. doi: 10.3389/fimmu.2020.01331

Xu, P., Brock, G. N., and Parrish, R. S. (2009). Modified linear discriminant analysis approaches for classification of high-dimensional microarray data. *Comput. Stat. Data Anal.* 53, 1674–1687. doi: 10.1016/j.csda.2008.02.005

Xu, Z., Zhang, H., Wang, Y., Chang, X., and Liang, Y. (2010). L1/2 regularization. Sci. China Inf. Sci. 53, 1159–1169. doi: 10.1007/s11432-010-0090-0

Xuan, W., Wu, X., Zheng, L., Jia, H., Zhang, X., Zhang, X., et al. (2024). Gut microbiota-derived acetic acids promoted Sepsis-induced acute respiratory distress syndrome by delaying neutrophil apoptosis through FABP4. *Cell. Mol. Life Sci.* 81:438. doi: 10.1007/s00018-024-05474-y

Yang, S., Guo, J., Kong, Z., Deng, M., Da, J., Lin, X., et al. (2024). Causal effects of gut microbiota on Sepsis and Sepsis-related death: insights from genome-wide Mendelian randomization, single-cell RNA, bulk RNA sequencing, and network pharmacology. *J. Transl. Med.* 22:10. doi: 10.1186/s12967-023-04835-8

Yang, Y., Sun, H., Zhang, Y., Zhang, T., Gong, J., Wei, Y., et al. (2021). Dimensionality reduction by UMAP reinforces sample heterogeneity analysis in bulk transcriptomic data. *Cell Rep.* 36:109442. doi: 10.1016/j.celrep.2021.109442

Yang, J. O., Zinter, M. S., Pellegrini, M., Wong, M. Y., Gala, K., Markovic, D., et al. (2023). Whole blood transcriptomics identifies subclasses of pediatric septic shock. *Crit. Care* 27:486. doi: 10.1186/s13054-023-04689-y

Ying, Y.-L., Zhang, J., Gao, R., and Long, Y.-T. (2013). Nanopore-based sequencing and detection of nucleic acids. *Angew. Chem. Int. Ed. Engl.* 52, 13154–13161. doi: 10.1002/anie.201303529

Yoseph, B. P., Klingensmith, N. J., Liang, Z., Breed, E. R., Burd, E. M., Mittal, R., et al. (2016). Mechanisms of intestinal barrier dysfunction in Sepsis. *Shock* 46, 52–59. doi: 10.1097/SHK.00000000000565

Zhang, X., Cui, Y., Ding, X., Liu, S., Han, B., Duan, X., et al. (2021a). Analysis of mRNA-IncRNA and mRNA-IncRNA-pathway co-expression networks based on WGCNA in developing pediatric Sepsis. *Bioengineered* 12, 1457–1470. doi: 10.1080/21655979.2021.1908029

Zhang, S., Shen, Y., Chen, I. A., and Lee, J. (2023). Bayesian modeling of interaction between features in sparse multivariate count data with application to microbiome study. *Ann. Appl. Stat.* 17:690. doi: 10.1214/22-AOAS1690

Zhang, L., Shi, X., Qiu, H., Liu, S., Yang, T., Li, X., et al. (2023). Protein modification by Short-chain fatty acid metabolites in Sepsis: a comprehensive review. *Front. Immunol.* 14:834. doi: 10.3389/fimmu.2023.1171834

Zhang, X., Xing, Y., Sun, K., and Guo, Y. (2021b). OmiEmbed: a unified multi-task deep learning framework for multi-omics data. *Cancers* 13:3047. doi: 10.3390/cancers1312 3047

Zhao, H., Lyu, Y., Zhai, R., Sun, G., and Ding, X. (2022). Metformin mitigates Sepsisrelated Neuroinflammation via modulating gut microbiota and metabolites. *Front. Immunol.* 13:797312. doi: 10.3389/fimmu.2022.797312

Zhou, Y., Geng, P., Zhang, S., Xiao, F., Cai, G., Chen, L., et al. (2024). Multimodal functional deep learning for multiomics data. *Brief. Bioinform.* 25:448. doi: 10.1093/bib/bbae448

Zhou, Y., Liu, M., and Yang, J. (2022). Recovering metagenome-assembled genomes from shotgun metagenomic sequencing data: methods, applications, challenges, and opportunities. *Microbiol. Res.* 260:127023. doi: 10.1016/j.micres.2022.127023

Zhou, S., Liu, B., Zheng, D., Chen, L., and Yang, J. (2025). VFDB 2025: an integrated resource for exploring anti-virulence compounds. *Nucleic Acids Res.* 53, D871–D877. doi: 10.1093/nar/gkae968

Zitnik, M., Nguyen, F., Wang, B., Leskovec, J., Goldenberg, A., and Hoffman, M. M. (2019). Machine learning for integrating data in biology and medicine: principles, practice, and opportunities. *Inf. Fusion* 50, 71–91. doi: 10.1016/j.inffus.2018.09.012

Zontov, Y. V., Rodionova, O. Y., Kucheryavskiy, S. V., and Pomerantsev, A. L. (2020). PLS-DA – a MATLAB GUI tool for hard and soft approaches to partial least squares discriminant analysis. *Chemom. Intell. Lab. Syst.* 203:104064. doi: 10.1016/j.chemolab.2020.104064