Contribution of m6A subtype classification on heterogeneity of sepsis

Shi Zhang, Feng Liu, Zongsheng Wu, Jianfeng Xie, Yi Yang, Haibo Qiu

Department of Critical Care Medicine, Zhongda Hospital, School of Medicine, Southeast University, Nanjing 210009, China

Contributions: (I) Conception and design: S Zhang; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: S Zhang, Z Wu, F Liu, Y Yang; (V) Data analysis and interpretation: S Zhang, Z Wu, F Liu, Y Yang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Haibo Qiu, MD, PhD. Department of Critical Care Medicine, Zhongda Hospital, School of Medicine, Southeast University, Nanjing 210009, China. Email: haiboq2000@163.com.

Background: Sepsis is a highly heterogeneous syndrome with diverse immune status and varied bioprocesses among individuals. The heterogeneity of sepsis could be associated with N6-methyladenosine (m6A) RNA methylation, due to m6A as a common and reversible posttranscriptional RNA modification involved in the regulation of whole bioprocesses. Therefore, we aim to identify m6A induced molecular subtypes of sepsis and furthermore explore the probable mechanism.

Methods: Gene expression datasets with 479 consecutive patients admitted for sepsis to the intensive care unit (ICU) in the Amsterdam Academic Medical Center were included in present study at first. Secondly, twelve m6A methylation regulatory genes were determined via systematic review in published researches. Furthermore, we utilized unsupervised clustering (consensus k means clustering) to identify m6A induced molecular subtypes in sepsis based on m6A prognostic molecular, and assess the association of these subtypes with clinical traits and survival outcomes. Moreover, the probable mechanism and regulatory relationship of m6A in sepsis was also explored through Gene Set Enrichment Analysis (GSEA), Weighted gene co-expression network analysis (WGCNA), Gene Ontology (GO) analysis and Co-expression analysis.

Results: Three m6A subtypes with different outcome were identified in sepsis cohort through unsupervised clustering on m6A prognostic molecular, designated Cluster 1/2/3 (log-rank P=0.004). The best outcome was found for patients classified as having cluster 3, and at 28 days, 21 of 144 people with cluster 3 had died [hazard ratio (HR) vs. all other clusters 5.42 (95% CI: 0.359–0.819); P=0.011], compared with 57 of 224 people with cluster 1 (HR 0.579, 95% CI: 0.364–0.920; P=0.037), and 36 of 112 people with cluster 2 (HR 0.477, 95% CI: 0.272–0.833; P=0.003). For exploration of the relationship between m6A subtypes and immunity, the GSEA found that patients in cluster 1 suffered from hyper-activated immunocompetent status; patients in cluster 2 indicated immunosuppressive status; and patients in cluster 3 showed the moderate immune activity (P<0.05). Co-expression analysis furthermore identified 82 immune molecules and 40 autophagy-related molecules could be regulated by prognostic m6A RNA methylation regulators (P<0.05) and correlation coefficient >0.6. In addition, WGCNA and GO analysis indicated that autophagy was significantly and widely activated in patients with cluster 3 (P<0.05).

Conclusions: According to the heterogeneity in m6A methylation regulatory genes, three distinct subtypes in sepsis were identified with different RNA epigenetics, immune status, biological processes and outcomes, which initially uncovered that heterogeneity of sepsis may be largely caused by m6A RNA methylation.

Keywords: Sepsis; heterogeneity; m6A RNA methylation; immunity; autophagy

Submitted Dec 07, 2019. Accepted for publication Feb 05, 2020.
doi: 10.21037/atm.2020.03.07
View this article at: http://dx.doi.org/10.21037/atm.2020.03.07
Introduction

Sepsis is defined as a highly heterogeneous syndrome that is associated with a dysregulated systemic inflammatory host response to infection and causes organs dysfunction (1,2). It is generally believed that the failure of proposed sepsis therapies is due to the heterogeneity of sepsis patients and the inability to accurately classify sepsis patients at the molecular biology level. The identification of distinct subtypes of sepsis could adopt appropriate treatment regimens for patients with different subtypes, which can improve survival in sepsis (3,4).

Previous studies attempted to identify diverse subtypes through clinical features or biomarkers. For example, Seymour et al. classified sepsis patients to four derived phenotypes via 29 clinical features (temperature, mean arterial pressure, fluid resuscitation response and central venous oxygen saturation, etc.). Using different genes between sepsis and health controls, researchers identified four subphenotypes of sepsis, among them. However, these classification studies of sepsis have not uncovered the mechanism of different subtypes. N6-methyladenosine (m6A), a chemical modification for RNA regulation, accounts for more than 60% of all RNA epigenetics (5,6). As m6A is the commonest posttranscriptional RNA modification involved in the regulation of whole bioprocesses (including immunity, metabolism, proliferation and apoptosis), the heterogeneity in sepsis may associate with m6A RNA regulations. In addition, since m6A emerged as a reversible bioprocess involved in many various disease progressions, potential therapeutic targets of m6A were widely investigated (7-11). Previous studies have demonstrated that m6A regulators were closely associated with tumor heterogeneity (12,13). However, details of m6A RNA regulatory factors mediating sepsis heterogeneity remained largely unknown. Fortunately, the public database provides a large number of sepsis-related transcriptome data, while the use of RNA omics and RNA-seq are becoming cheaper and more extensive, helping us to perform integrative analysis on m6A RNA methylation and sepsis molecules subtypes.

In this scenario, we systematically analyzed the expression of widely reported m6A RNA methylation regulators in sepsis based on public database—Gene Expression Omnibus (GEO database) (5-16). Firstly, we sought to screen prognostic m6A RNA methylation regulators in sepsis through univariate Cox’s proportional-hazards regression. Furthermore, we classified sepsis into diverse m6A endotypes through unsupervised clustering according to prognostic m6A molecules. In addition, we attempted to elaborate the reasons of different prognosis in m6A induced subtypes via enrichment analysis.

Methods

Data sources and study selection

The public database—GEO database was searched for all expression microarrays that matched terms of sepsis. The datasets collected from clinical studies investigating sepsis in adults using peripheral blood within 48 hours after intensive care unit admission (ICU) were included. Exclusion criteria were as follows: (I) datasets that utilized endotoxin or lipopolysaccharide infusion as vitro or animal models for sepsis; (II) clinical-gene expression microarray derived from sorted cells; (III) without the outcome of 28 day mortality or survival curves. The flow-process diagrams for screening datasets were shown in Figure S1.

Collecting m6A RNA methylation regulators via systematic review

A systematic literature review of all of the pertinent English language studies was undertaken in the PubMed databases from inception through December 28, 2019. A manual search of the selected articles and relevant review articles was performed by two reviews (F Liu and ZS Wu) independently. Only m6A RNA methylation regulators gene which confirmed by animal or cell experiments will eligible to present analysis.

Data preprocessing

All datasets were downloaded as txt files, and outputs from mRNA array were normal-exponential background corrected and then between-arrays quantile normalized using limma R package. For compatibility with microarray study, expression was normalized using a weighted linear regression, and the estimated precision weights of each observation were then multiplied with the corresponding log2 to yield final gene expression values.

Screening prognosis-associated m6A RNA methylation regulators in sepsis

In order to determine prognostic m6A RNA methylation
regulators, the univariate Cox’s proportional-hazards regression with Bonferroni correction for multiple comparisons was utilized, using survival R package, with cut-off value as P<0.05.

**Explore m6A induced molecular subtypes of sepsis**

The consensus k means clustering was utilized to perform consistent clustering and selecting of m6A subtypes based on prognostic m6A RNA methylation regulators expression profiles. The clustering was performed using 100 iterations, with each iteration containing 80% of samples. The optimal cluster number was determined by cumulative distribution function (CDF) curves of the consensus score, clear separation of the consensus matrix heatmaps, characteristics of the consensus cumulative distribution function plots, and adequate pair wise-consensus values between cluster members.

**Assessment for immune heterogeneity among subtypes**

Kaplan-Meier (KM) curves would be performed to evaluate prognosis of various m6A Subtypes, with cut-off value as P<0.05. Differential expression analysis using moderated t-tests would be utilized to assess the expression distribution of twelve m6A RNA methylation regulators in different subtypes.

We furthermore evaluated the variation of immune status among subtypes. Gene Set Enrichment Analysis (GSEA) and cell type identification by estimating relative subset of known RNA transcripts (CIBERSORT) were utilized to evaluate proportions of immune cells. Briefly, GSEA calculates separate enrichment scores (ES) for each pairing of a sample and gene set (17) via GSVA and GSEABase R package. Each GSEA ES represents the degree to which the genes in a particular gene set are coordinately up- or downregulated within a sample. CIBERSORT algorithm could calculate proportions of human immune cells according to RNA matrix (https://cibersort.stanford.edu/) (18). Immune cells included immune enhancing cells (Th1 cells, T cells CD4 activated, NK cells activated and B cells activated) and immune suppressive cells (Th2 cells and Treg cells). Differential expression analysis using moderated t-tests would be utilized to assess the expression distribution of pro-inflammatory cytokines and enrichment of immune cells among subgroups, P<0.05 was recognized as significant results. Pro-inflammatory cytokines included interleukin-1β (IL-1β), interleukin-6 (IL-6) and Tumor necrosis factor (TNF).

**Assessing the heterogeneity of biological function among subtypes**

There are two steps to implement this analysis. On the one hand, we attempted to find gene-sets which significantly correlated to m6A subtypes through Weighted gene co-expression network analysis (WGCNA), using WGCNA R package to determine co-expressed genes using all expressed genes in microarrays (19,20). Analysis setting included bi-weight miscorrelation (corType = ‘bicor’) to account for outliers, sign of correlations between neighbors (TOM type and networkType = ‘signed’), and a more sensitive module detection parameter (deepSplit = 3). Module eigengene (ME) was calculated as the first principal component of gene expression for the module and inter-relatedness of each module by eigengene network clustering (Figure S2).

MEs were compared with m6A subtype information using Spearman’s correlation corrected for subtypes, and P values were adjusted for multiple comparisons by false discovery rate. Genes from modules which highly associated with m6A subtypes (the maximum correlation coefficient and P<0.05) were selected for further Gene Ontology (GO) analysis.

On the other, GO analysis was performed to elaborate the functions of selected gene-sets using org.Hs.eg.db, clusterProfiler and enrichplot R packages, with cut-off value as P<0.05.

Co-expression analysis was performed to furthermore explore the possible downstream molecules regulated by m6A RNA methylation, with cut-off value as P<0.05, Pearson correlation coefficient >0.6.

**Software and versions**

Rx64 3.6.1 was conducted to process data, analyze data and plot diagrams; Cytoscape 3.6.1 was performed to plot network diagrams.

**Results**

**Characteristics of Datasets and Patients**

After search strategy and inclusive criteria, one mRNA dataset from University Medical Center in Utrecht and Academic Medical Center in Amsterdam is finally enrolled in current study. A total of 479 patients in this dataset were
all older than 18 years and were diagnosed with sepsis. Septic shock ratio is 34.8%, and more details are shown in Table 1.

Twelve m6A RNA methylation regulators were collected via systematic review

We first collated a list of sixteen m6A RNA methylation regulators from published literature (5-16), and then we restricted the list to genes with available RNA expression data in the GEO datasets. This yielded a total of 12 m6A RNA methylation regulators. These m6A RNA regulators included m6A writers, m6A readers and m6A erasers. m6A writers included KIAA1429, methyltransferase like 3 (METTL3), methyltransferase like 14 (METTL14), RNA binding motif protein 15 (RBM15) and WT1 associated protein (WTAP). m6A readers included HNRNPC, YTH domain containing 1 (YTHDC1), YTH domain containing 2 (YTHDC2), YTH N6-methyladenosine RNA binding protein 1 (YTHDF1) and YTH N6-methyladenosine RNA binding protein 2 (YTHDF2). m6A erasers included alkB homolog 5, RNA demethylase (ALKBH5) and FTO alpha-ketoglutarate dependent dioxygenase (FTO), shown in Table 2.

The prognosis associated m6A RNA methylation regulators in sepsis

According to the univariate Cox regression analyses, ALKBH5, HNRNPC, KIAA1429, WTAP and YTHDF2 are significantly correlated with 28-day cumulative mortality (P<0.05). ALKBH5 and WTAP are risky genes with Hazard Ratio (HR) >1, and HNRNPC, KIAA1429 and YTHDF2 are protective genes with HR <1 (Figure 1). The correlation results among m6A RNA methylation regulators were shown in Figure S2.

m6A molecular subtypes in sepsis

Based on the expression similarity of m6A RNA methylation regulators, k =2 or 3 seemed to be an adequate selection with clustering stability increasing from k = 2 to 9 in cohort (Figure 2 and Figure S3). KM suggested significantly different 28-day cumulative mortality among three subtypes (P=0.004). However, 28-day cumulative mortality did not show an obviously statistic difference between two subtypes (P=0.092). In consideration of KM curves and clustering stability, patients with sepsis were finally divided into three subtypes (cluster 1, cluster 2 and cluster 3).

In the sepsis cohort, the best outcome was found for patients classified as having a cluster 3 , and at 28 days, 21 (14.6%) of 144 people with a cluster 3 had died (HR vs. all other clusters 5.42, 95% CI: 0.359–0.819; P=0.011), compared with 57 (25%) of 224 people with a cluster 1 (HR 0.579, 95% CI: 0.364–0.920; P=0.037), and 36 (23%) of 112 people with a cluster 2 (HR 0.477, 95% CI: 0.272–0.833; P=0.003).

Immune status heterogeneity among m6A subtypes

Figure 3A showed a distinct expression pattern in the m6A...
regulators profiles of each subtype (P<0.05). After GSEA and CIBERSORT, the obviously heterogeneity of immune status among subtypes were identified in Figure 3B,C. This heterogeneity of immune status included significant difference of enrichment scores in immune cells (Figure 3B), and obvious discrepancy of expression in pro-inflammatory cytokines (Figure 3C). In cluster 1 (moderate prognosis), immune enhancing cells and pro-inflammatory cytokines were significantly up-regulated compared with other subtypes, companied with immune suppressive cell down-regulated simultaneously (P<0.05). The distribution of immune cells and pro-inflammatory cytokines in cluster 2 (the worst prognosis) showed diametrically opposed to those in cluster 1 (P<0.05). The enrichment score of immune cells and expression of pro-inflammatory cytokines in cluster 3 (the best prognosis) were in the middle of cluster 1 and cluster 2.

**Heterogeneity of other biological process among subtypes**

WGCNA identified 24 modules in the sepsis cohort (Figures S4,S5). Cluster 1 positively correlated with Seventeen module (including 1,860 genes), with correlation coefficient as 0.73 and P value as 8×10^{-51}. Cluster 2 markedly positively associated with One module (including 3,423 genes), with correlation coefficient as 0.68 and P value as 8×10^{-70}. Cluster-3 significantly positively related to Twenty-one (including 251 genes), with correlation coefficient as 0.61 and P value as 8×10^{-51}. Genes from these strongly relevant modules defined as the eigengenes of each subtypes (Figure 4A).

In order to furthermore explore module biological function, GO enrichment analyses were conducted and the results were shown in Figure 4B,C,D. The biological functions of Seventeen module (correlated with cluster 1) were mainly enriched in RNA processing and vesicle transport; One module (correlated with cluster 2) were mainly related to nucleotide repair and extracellular matrix; Twenty-one (correlated with cluster 3) were mainly associated with autophagy.

Co-expression analysis furthermore identified 82 immune molecules and 40 autophagy-related molecules could be regulated by prognostic m6A RNA methylation regulators, P<0.05 and correlation coefficient >0.6 (Figure S6). The immune molecules and autophagy-related molecules were referenced to immunology database (https://www.
Figure 2  Identification of optimal sepsis subtypes based on the five prognostic m6A RNA methylation regulators. (A) The cumulative distribution function (CDF) curves is the integral of probability density function, which can completely describe the probability distribution of a real random variable, and established using consensus clustering approach. CDF curves of consensus scores based on different subtype number (k =2, 3, 4, 5, 6, 7, 8, 9) and the corresponding color are represented. (B) The CDF Delta area curve of all samples when k =2 or 3. (C) Survival curves of two subtypes (when k =2). Kaplan-Meier (KM) curves did not showed an obviously statistic difference between two subtypes, with P value as 0.092. (D) Survival curves of three subtypes (when k =3). Kaplan-Meier (KM) curves indicated a significantly statistic difference among three subtypes, with P value as 0.004.

Discussion

Sepsis is a clinically common syndrome with high heterogeneity. Heterogeneity in patients with sepsis is currently considered to be an important cause of treatment failure. Therefore, it is very important to find a regulatory analysis that affects the heterogeneity of concentration and then guide the treatment of sepsis. Current research indicates that the expression of RNA methylation, regulators of epigenetic, is closely related to the heterogeneity and prognosis of sepsis. Three m6A molecular subtypes were identified in sepsis, cluster 1/2/3, by consensus clustering based on the expression of most aberrant m6A regulators. Three m6A subtypes exhibited significantly different RNA epigenetics, immune status, biological processes and outcomes. We further investigated that the moderate immune activated status and potential autophagy mechanisms could benefit septic patients, which were regulated by RNA methylation. To our knowledge, this was the first study on the transcriptome-wide mapping of m6A regulators which focuses on investigating the landscapes and functions of the reversible RNA modifications in sepsis.

YTHDF2, HNRNPC, KIAA1429, WTAP, and ALKBH5 were proved to be more important in prognosis of sepsis. The major mechanism by which m6A exerts its effects is by recruiting m6A-binding proteins (m6A readers) to target RNAs. Subsequently, a methyltransferase complex
within the nuclear speckle, m6A writers, installs m6A modification on targeted RNAs via the methyl groups of S-adenosylmethionine (SAM) transferase. Demethylases (erasers) were found that removes methyl groups from m6A, indicating that m6A is a dynamically reversible RNA modification. The previous studies identified YTHDF2 and HNRNPC as m6A readers, KIAA1429 and WTAP as m6A writers and ALKBH5 as m6A eraser (4-14). The current study further uncovered the prognostic roles of these five m6A regulators in sepsis. The future studies could

### Figure 3

The heterogeneity of m6A regulators distribution and difference of host response among subtypes. (A) The distribution of twelve m6A RNA methylation regulators among subtypes. (B) The heterogeneity of immune cells among subtypes. Enrichment score of every patient was calculated by single sample Gene Set Enrichment Analysis. (C) The expression of pro-inflammatory cytokines among subtypes.

* P<0.05; ** P<0.01; *** P<0.001; ns, P>0.05.
Figure 4 The distinction of mainly biological function among subtypes. (A) Weighted gene co-expression network analysis (WGCNA) in the three sepsis subtypes. The 24 modules were validated and are designated by the following numbers: “One”, “Two”, “Three”, “Four”, “Five”, “Six”, “Seven”, “Eight”, “Nine”, “Ten”, “Eleven”, “Twelve”, “Thirteen”, “Fourteen”, “Fifteen”, “Sixteen”, “Seventeen”, “Eighteen”, “Nineteen”, “Twenty”, “Twenty-one”, “Twenty-two”, “Twenty-three”, and “Twenty-four”. Heatmap showing the correlation between feature vectors of 24 modules and three subtypes. (B) Gene Ontology (GO) enrichment analysis on genes from Seventeen module. (C) GO enrichment analysis on genes from One module. (D) GO enrichment analysis on genes from Twenty-one module.
be conducted to sequentially verify and research these m6A elaborate mechanisms in sepsis.

The current study provided an insight that immune heterogeneity of sepsis was largely associated with m6A RNA methylation. Patients with same syndrome (sepsis) were classified into finer taxa and had differential prognosis and distinct immune status. Patients in cluster 2 with worst prognosis suffered from immunosuppression; patients in cluster 1 with worse prognosis suffered from hyper-activated immunocompetent status; and patients in cluster 3 (the best prognosis) had the moderate immune activity. The extent of hyper-activated and hypo-activated immune response varies among individuals in sepsis, which result in immune heterogeneities (3,4). It is well known that hyper-activated immunocompetent and Immunosuppression are both detrimental status in the progression of sepsis (5,7), especially immunosuppression. However, for a given patient, it is too hard to characterize the immune status due to complex and unknown mechanisms of immune heterogeneity in sepsis. As immune functions were widely regulated by m6A RNA methylation (9,10), three m6A subtypes enhanced the comprehension of molecular characteristics and subgroup-specific immune status in highly heterogeneous syndrome (sepsis). These results provided a better understanding of the predicting clinical outcomes and valuable research targets in sepsis.

The molecular characterization of the transcriptome of subtypes subsequently indicated that autophagy mechanisms, regulated by RNA methylation, could play an important role in improvement of outcomes in sepsis, which were consistent with previous studies of autophagy in sepsis. Autophagy is the regulated process cells use to recycle nonessential, redundant, or inefficient components and is an adaptive response during times of stress (21). In addition to its function in enabling the cell to gain vital nutrients in times of stress, autophagy can also be involved in elimination of intracellular microorganisms, tumor suppression, and antigen presentation (22). Therefore, previous studies also confirmed that autophagy-related processes were properly activated and accelerated during sepsis due to the sepsis induced mitochondrial injury (22). Besides, autophagy was widely recognized as protective biological processes in sepsis (21-28). Unfortunately, these autophagy-related studies did not show upstream regulatory mechanisms. As autophagy seemed to be not reversible once triggered, exploration for regulatory factors of autophagy would be beneficial for develop potential therapeutic targets in sepsis. Combined with current analysis and previous study, outcomes of septic patients in cluster 3 were obviously improved largely depending on autophagy, which were quite possibly regulated by m6A RNA methylation.

As m6A RNA methylation is generally known a reversible and regulatable process, chemicals targeting m6A methylation would be explored as a new therapeutic method for sepsis therapy. The present study provided possible regulatory relationships between m6A regulators with immune and autophagy-related molecules.

There are several limitations in the present study. First, as a retrospective study of primarily publically available data, we were not able to download more demographics and clinical features such as severity, complications, individual treatment of each patient for detailed and longitudinal analyses. In addition, for the limitation of bioinformatics, the further studies and experiments should be conducted to sequentially verify and research elaborate RNA methylation mechanisms screened by our analyses.

**Conclusions**

We identified three highly heterogenous m6A subtypes in sepsis, with significantly different RNA epigenetics, immune status, biological processes and outcomes, which could be intervention targets for improvement of therapeutic system in sepsis, with validation experiments to be warranted to assess these further.

**Acknowledgments**

We would like to thank all of the microarray data contributors of this study and all of the patients and volunteers who participated in this study.

**Funding:** The study was supported in part by the National Natural Science Foundation of China (81571847, 81930058), the projects of Jiangsu Provincial Medical Key Discipline (ZDXKA2016025) and Jiangsu Provincial Special Program of Medical Science (BE2018743, BE2019749).

**Footnote**

**Conflicts of Interest:** The authors have no conflicts of interest to declare.

**Ethical Statement:** The authors are accountable for all
aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The present study was not required ethics approval since it is a secondary bioinformatics analysis based on public database (GEO database).

References


Supplementary

Figure S1 The flow-process diagram for screening datasets.
Figure S2 The correlation among 12 m6A RNA methylation regulators. Number in the plots represents Pearson correlation coefficient. × represents P>0.05.
Figure S3 Consensus heatmap. Consensus heatmaps for k = 2–9 are shown. A relative stable partitioning of the samples is seen at k = 3.
**Figure S4** Analysis of network topology for various soft thresholding powers of weighted gene co-expression network analysis.

**Figure S5** Hierarchical cluster tree displaying 24 modules of co-expressed genes through weighted gene co-expression network analysis.
Figure S6 Co-expression network of prognostic m6A RNA methylation regulators with immune and autophagy related genes. Two related genes are linked (P<0.05 and correlation coefficient more than 0.6). ◇ represent m6A RNA methylation regulators; △ represent immune molecules; blue links represent positive correlation and green links represent negative correlation.