# Molecular mechanisms of sepsis

Title: Protein-protein interaction network andfunctional module analysis to reveal the mechanism ofsepsis inpolytrauma patients

## Highlights:

1. We explored the molecular basis of sepsis induced by polytrauma using PPI network.
2. A total of 342 DEGs including 110 up- and 232 down-regulated genes were obtained.
3. *TRAF3* was related with the innate immune responses in sepsis.
4. *ITGB3* was the key gene involved in coagulation dysregulation in sepsis.
5. *CASP6* and *RASA1* played key roles in the cell apoptosis mechanism of sepsis.

## Abstract

*Objective* Sepsis represents thesystemic inflammatory response to microbial infection. The pathogenesis of sepsis remains unclear. In this study, we aimed to explore the molecular mechanism ofsepsis inpolytrauma patients.

*Methods* Thedifferentially expressed genes (DEGs) between the polytrauma patients with and without sepsis were identified by analyzing theGSE12624 microarray data using the limma package of R. The protein-protein interaction (PPI) network was extracted from the human PPI datasets by using MATLAB. The functional modules in the PPI network were identified by the MCODE network clustering algorithm. The KEGG pathway analysis was performed in each module. The phylogenetic tree was constructed using phylogeny inference package (PHYLIP).

*Result* Total of 342DEGs including 110 up- and 232 down-regulated genes were obtained. The PPI network identified several hub genes which had more interactions with others, such as *TRAF3* , *ITGB3* , *CASP6* and *RASA1* . Further phylogenetic analysis indicated the high conservation of these hub genes. In the module analysis, four significant modules were identified. All the genes (COL1A2, FN1, ITGA2B, ITGB3 and CD36) in module 2 were enriched in extracellular matrix (ECM)-receptor interaction pathway. In module 4, *CASP6* and *CASP3* were enriched in apoptosis pathway.

*Conclusion* We predicted genes such as *TRAF3* , *ITGB3, CASP6* and *RASA1* which were closely associated with sepsis induced by polytrauma *.* Among them, *ITGB3* may play key role in the coagulation dysregulation of polytrauma patients with sepsis, and *CASP6* and *RASA1* may be the key genes in the cell apoptosis mechanism of sepsis.

KeywordsSepsis, DEGs, GO, PPI network, phylogenetic tree

# Introduction

Polytrauma is a syndrome of multiple injuries exceeding a defined severity with sequential systemic reactions that can lead to dysfunction or failure of remote organs and vital systems, which have not themselves been directly injured [1]. Sepsis, as one of the complications of polytrauma [2], is the systemic inflammatory response to microbial infection that often leads to increasing susceptibility to secondary infections, multiorgan failure, and death [3]. A sixteen years clinical study indicated that 10. 2% of polytrauma patients infected sepsis during their hospital course [4]. Polytrauma is a major cause of morbidity and mortality in global and sepsis (3. 1-17%) is one

of the predominant causes of late death in polytrauma patients [5]. The disease severity is increasing according to the order of sepsis, severe sepsis and septic shock in the systemic inflammatory response syndrome (SIRS) [6]. Mortality has been reported to be as high as 45. 6% for patients with severe sepsis or septic shock [7].

Based on the pathogenesis of sepsis, many therapies have been applied in the clinical practice such as antimicrobial therapy [8, 9] and hemodynamic support and adjunctive therapy [10, 11]. Currently, the Surviving Sepsis Campaign (SSC) has attempted to increase the awareness and establish the practice guidelines to improve the recognition and treatment for the patients with sepsis [12, 13].

At present, there are four approved mechanisms in the pathogenesis of sepsis [14]. The first one is dysregulated coagulation. Sepsis patients frequently manifest disseminated intravascular coagulation (DIC) with consumption of platelets and prolongation of clotting times [15]. The second one is inflammatory response. The inflammatory response is an important and central component of sepsis because the elements of response drive the physiological alterations that manifest as the SIRS [16]. Third, many cellular aspects become dysfunctional in sepsis which behave either excessive activation or depressed function [17]. The last one is metabolic alterations. It was reported that endogenous glucose production was markedly increased in the patients [18]. However, the specific molecular mechanisms of them remain entirely unclear. In this study, the differentially expressed genes (DEGs) between the polytrauma patients with sepsis and without sepsis were identified. Gene ontology (GO) analysis, protein-protein interaction (PPI)

network and phylogenetic tree construction were performed to explore the molecular basis of sepsis induced by polytrauma.

# Materials and methods

### *Microarray data*

The gene expression profile ofGSE12624based on the CodeLink UniSet Human I Bioarray platform (GE Healthcare/Amersham Biosciences) was downloaded fromNational Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database (http://www. ncbi. nlm. nih. gov/geo/). The dataset available in this analysis contained 70 samples including 34 polytrauma patients with sepsis and36 polytrauma patients without sepsis.

### *Data preprocessing and DEGs screening*

For themicroarray data, Robust Multichip Average (RMA) in the Affy package of R was used tocompute normalized expression measures from the raw expression values. Probe annotation was obtained by using theBioconductorpackage. Thelimma package was used to identify the DEGs withp-value <0. 05 and | logFC|> 1 [19].

### *GO enrichment analysis of DEGs*

GO analysis was performedusingthe DAVID online tool (http://david. abcc. ncifcrf. gov/) [20]. For GO enrichment of DEGs, we selected GOTERM_BP_FAT, GOTERM_CC_FAT and GOTERM_MF_FATas the gene set categories. A p-value of less than 0. 05 was set as the cut-off criterion.

### *PPI network construction*

Thehuman PPI datasets with 108477 interacting protein pairs were downloaded from PINA2 (http://cbg. garvan. unsw. edu. au/pina/interactome. stat. do) at December 26, 2013. The PPI networks of the DEGs in sepsis were extracted from the human PPI datasets by usingMATLAB [21]. The proteins in the network served as nodes and the degree of a node corresponded to the number of interactions with other proteins [22]. The protein with high degree was considered as the hub node.

### *Identification of functional modules in PPI network*

PPI network visualization and network parameters evaluation were performed by using Cytoscape software. The modules were identified by theMCODE(a cytoscape plug-in) network clustering algorithm with the default parameters [23]. Themodule withscore larger than 2 was considered as significant. KEGG pathway analysis of each module was performed by applying the DAVID annotation tool.

### *Phylogenetic tree construction*

In this study, we constructed the phylogenetic tree based on the nucleotide sequences to investigate thesequenceconservation of the DEGs whosedegree were large than 30. The BLAST program is used to search for homologous sequences of these DEGs. The DNA sequence of these DEGsand their homologous genes in FASTA format were obtained from the nucleic acid database in NCBI (http://www. ncbi. nlm. nih. gov/nuccore). The phylogenetic tree was constructed by using phylogeny inference package (PHYLIP) with

the default parameters [24]. The gene conservation was estimated by the distance from gene to the phylogenetic tree root.

# Result

### *DEGs between the patients with and without sepsis*

After statistical analysis of the microarray data, a total of342 DEGs were screened out. Among them, 110 were down-regulated and 232 were up-regulated insepsis. The top 20 significantly up- and down-regulated DEGs are shown in Table 1.

### *GO enrichment analysis*

The 342 DEGs were significantly enriched into 95GOterms including 81biological processes terms, 10 cellular component terms and 4 molecular function terms. The top 10 GO biological processes termswere mainly related to the purine base (purine base biosynthetic process, purine base metabolic process, purine nucleoside monophosphate biosynthetic process and purine ribonucleoside monophosphate biosynthetic process), nucleobase (nucleobase metabolic process and nucleobase biosynthetic process) and regulation of protein modification (regulation of protein modification process and positive regulation of protein modification process). The 10 significantly enriched GO termsof cellular component included four lumen related terms (organelle lumen, membrane-enclosed lumen, intracellular organelle lumen and nuclear lumen), two membrane related terms (extrinsic to membrane and plasma membrane part) andfour other cellular component terms (peroxisome, microbody, nuclear body and Golgi apparatus). For molecular function, four significant GO terms were enriched finally. They were acyl-CoA

binding, sons of mothers against decapentaplegic homologue (SMAD) binding, aryl hydrocarbon receptor binding and potassium channel inhibitor activity (Table 2).

### *PPI network of DEGs*

A PPI network consisting of 225 DEGs and 1048 non-DEGs is shown in Fig. 1. This network included 1145 gene nodes and 1273 interactions. The connectivity degree of each node in this PPI network was calculated and the results of top 20 nodes are listed in Table 3. Among them, the genes *CRK* (encoding CDC2 related protein kina), *RASA1* (encoding RAS p21 protein activator 1), *TRAF3* (encoding tumour-necrosis-factor receptor associated factor 3), *ZHX1* (encode zinc-fingers and homeoboxes), *ITGB3* (encoding integrin β3), *RPA1* (encoding replication protein A1), *JAK3* (encoding Janus kinases 3), and *CASP6* (encoding caspase-6) with the degree over 30 were selected as the hub genes.

### *Module analysis of PPI network*

A total of 7 modules were constructed by usingMCODE plug-in. After excluding themodules with the score less than 2, 4 significant modules were considered as functional ones associated with sepsis (Table 4). According to the Fig. 2, the numbers of nodes and edges were similar in each model. The detailed results of KEGG pathway analysis for each module are provided in Table 5. For module 1, no pathway was enriched in the KEGG pathway analysis. For module 2, a total of 14 significant enriched pathways were identified. Among them, all the genes in this module were enriched in the pathway of extracellular matrix (ECM)-receptor interaction. In addition,

except *CD36* (encoding glycoprotein IV), the other four genes ( *ITGB3 and ITGA2B* encoding integrin αIIbβ3 *, COL1A2* encoding the α 2 chain of type 1 collagen and *FN1* encoding fibrinogen 1) were enriched in the focal adhesion and phosphatidylinositol 3-kinase (PI3K)-Akt signaling pathway. There were three significant enriched pathways in module 3. *HIF1A* (encoding hypoxia inducible factor-1) *, ARNT* (encoding arylhydrocarbon receptor nuclear translocator) and *ARNT2* (encoding arylhydrocarbon receptor nuclear translocator 2) were enriched in the pathway of renal cell carcinoma and pathways in cancer. *HIF1A* and *ARNT* were enriched in the pathway of Hypoxia-inducible factor 1 (HIF-1) signaling. For the module 4, five significant pathways were found. Among them, *CASP3* (encoding caspase 3) and *BIRC5* (encoding baculoviral IAP repeat–containing 5 and also called survivin) were enriched in the pathway of colorectal cancer, hepatitis B and pathways in cancer. *CASP6* and *CASP3* were enriched in apoptosis pathway. *CASP3* and *RASA1* were enriched in the mitogen-activated protein kinase (MAPK) signaling pathway (Table 5).

### *Phylogenetic tree analysis*

Based on the result of PPI network analysis, the selected hub genes were chosen to construct the phylogenetic tree. The phylogenetic tree of ZHX1 was unable to be constructed, as only three homologous sequences were searched out. The phylogenetic trees of the other seven hub genes were constructed by the DEGs and their top nine significant homologous genes. The results showed that *CRK* , *RASA1* , *TRAF3* , *ITGB3* , *RPA1* and *CASP6* werethe genes thatwere closer to tree roots indicating that the conservation

of these genes was high during evolution. However, the conservation of *JAK3* was low because of appearing in the late period of evolution (Fig. 3).

## Discussion

Currently, sepsis remains a serious clinical problem. The four approved mechanisms of sepsis were dysregulated coagulation, inflammatory response, and cellular dysfunctional and metabolic alterations. However, the specific molecular mechanisms are still incompletely understood. For better understanding the pathogenesis, we identified and analyzed the DEGs between the patients with and without sepsis. As a result, a total of 342 DEGs including 110 up-regulated genes and 232 down-regulated genes were found. These genes were significantly enriched in GO terms including purine base biosynthetic process, regulation of protein modification process and peroxisome. Among them, the process of purine base biosynthesis is the most significantly enriched process. It was reported that de novo purine biosynthesis was essential for infectivity, growth and virulence of many bacteria in mammals [25]. The pathogenesis of sepsis was related with the bacterial infection [26]. Therefore, the purine base biosynthesis process may associate with sepsis based on the tissue response to bacterial infection. For the regulation of protein modification, Wu et al. reported that the alterations in the phosphorylation of myofibrillar proteins and the Ca2+ sensitivity of myofibrillar ATPase might contribute to alter cardiac function during the progression of sepsis [27]. The cardiac dysfunction was the clinical characteristic in severe sepsis and septic shock [28]. Thus, the phosphorylation of myofibrillar proteins may be related with the sepsis-induced cardiac dysfunction.

Furthermore, we mapped the DEGs to the PPI network and identified high conserved hub genes. Among them, the high conservation of CRK, RASA1, TRAF3, ITGB3, RPA1 and CASP6 were proved by the phylogenetic tree analysis. They may be the crucial genes in the pathogenesis of sepsis. For TRAF3, it is a member of the TNF receptor (TNFR) associated factor (TRAF) protein family [29]. This protein participates in the activation of the innate immune response [30]. In the PPI network, TBK1 (encoding TANK-binding kinase 1) was a non-DEG interacted with TRAF3. It was reported that TIR domain-containing adaptor-inducing IFN-β (TRIF) could interact with noncanonical IKKs (IKKÏµ and TBK-1) and IKKι (also called IKKÏµ) through TRAF3 in the Toll-like receptors (TLR) signaling pathway [31]. The innate immune system constitutes the first line of defense by rapidly detecting invading pathogens through the TLR [32] and is a danger signal in systemic inflammatory response syndrome and sepsis [33]. Thus, TRAF3 may be the mediator of innate immune responses in sepsis induced by polytrauma.

We also performed the modular analysis of the PPI network and four functional modules were identified. Based on the result of the KEGG pathway analysis of each module, we found that the pathways in module 2 and 4 were more related with sepsis. The ECM-receptor interaction pathway was the most significant pathway in module 2 and all the genes of this module were enriched in this pathway. Fibronectin and collagen are the components of ECM [34]. Integrin family are the receptors transducing signals from the ECM [35]. Among them, integrin αIIbβ3 is the platelet integrin promoting theaggregation of platelets [36-38]. Moreover, it was reported that collagen type I could induce the aggregation of platelet [39]. Integrin αIIbβ3 is one of

the platelet collagen receptors in platelets [40]. It was reported that platelet-specific elements initiated at the cytoplasmic domains of integrin αIIbβ3, which was a signal that leaded to conformational changes within the extracellular domains of integrin and expression of the fibrinogen receptor, then the simultaneous occupancy on adjacent platelets of receptors with dimeric fibrinogen molecules leaded to platelet aggregation [41]. In addition, CD36 is spatially associated with the αIIbβ3 integrin on the surface of platelets [42]. Thus, we speculated that the binding of collagen type I and αIIbβ3 might need the participation of CD36, and then conformational changes within the extracellular domains of integrin and the binding between fibrinogen and fibrinogen receptor could lead to platelet aggregation. Disseminated platelet aggregation is one of the characteristics of the DIC in sepsis [43, 44]. The up-regulated expression of ITGB3 in sepsis may lead to the disseminated platelet aggregation. Hence, we concluded that the coagulation dysregulation in the polytrauma patients with sepsis may be associated with the increase of disseminated platelet aggregation caused by the up-regulated expression of ITGB3. Thus, ITGB3 may play key roles in the coagulation dysregulation of the polytrauma patients with sepsis.

Hub nodes CASP6 and RASA1 were predicted to be closely interacted with each other in module 4. Besides, CASP3, TOP1, BIRC5 and AURKB (Aurora B kinase) were also included in module 4. Among them, CASP6 and CASP3 were enriched in apoptosis pathway. It was reported that CASP6 may be associated with the cell apoptosis in sepsis [45] and blocking caspases might have some beneficial effects in decreasing cell apoptosis in sepsis [46]. Thus, we further confirmed that the up-regulated expression of CASP6 may

promote cell apoptosis in sepsis. Besides, TOP1 is cleaved late during cell apoptosis by CASP6 and CASP3 [47]. The TOP1 cleavage complexes contribute to cell apoptosis [48]. Therefore, the increase of these complexes induced by the up-regulated CASP6 can promote the cell apoptosis in sepsis. Moreover, full-length TOP1 could induce DNA cleavage by single-strand breaks which is the signal of cell apoptosis [49, 50]. Therefore, the exaggerated gene expression of TOP1 in our study might contribute to cell apoptosis in sepsis. In addition, it was reported that CASP3 could modulate a given set of proteins to generate, depending on the intensity of the input signals, opposite outcomes (survival vs death) through differential processing of RASA1 [51]. Some articles reported that low CASP3 activity leaded to the cleavage of the RASA1 protein into an amino-terminal fragment [52, 53]. RASA1 bound BIRC5 is a bifunctional protein complex that can suppress cell apoptosis and regulated cell division, so as to generate anti-apoptotic signals [54]. AURKB exists in a complex with BIRC5 [55]. Considering the up-regulated expression of RASA1 and AURKB, we speculated that there may be a switch mechanism of CASP3-RASA1 in cell apoptosis and BIRC5 and AURKB might play roles in the anti-apoptosis mechanism of RASA1. In summary, CASP6 and RASA1 are the key genes in the pathogenesis of sepsis induced by polytrauma.

## Conclusion

In this study, we obtained four key genes related with pathogenesis of sepsisinduced by polytrauma. Among them, TRAF3 was related with the innate immune responses in sepsis, *ITGB3* may play key role in the coagulation dysregulation of the polytrauma patients with sepsis and *CASP6*

and *RASA1* were associated with the mechanism of cell apoptosisin sepsis. For further investigating the association of these hub nodes with sepsis and verifying the role of the interactions among the genes in the pathogenesis of sepsis, more studies are required in the future.