Cytokine and molecular networks in sepsis cases: a network biology approach

Dong Wook Jekarl1,4,a, Kyung Soo Kim2,a, Seungok Lee1,4, Myungshin Kim3,4, Yonggoo Kim3,4

1 Department of Laboratory Medicine, Incheon St. Mary’s Hospital, College of Medicine, The Catholic University of Korea
2 Department of Thoracic and Cardiovascular Surgery, Seoul St. Mary’s Hospital, College of Medicine, The Catholic University of Korea
3 Department of Laboratory Medicine, Seoul St. Mary’s Hospital, College of Medicine, The Catholic University of Korea
4 Laboratory for Development and Evaluation Center, Seoul St. Mary’s Hospital, College of Medicine, The Catholic University of Korea

Corresponding author: Y. Kim, Department of Laboratory Medicine, Seoul St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, Republic of Korea
<yonggoo@catholic.ac.kr>

ABSTRACT. Background: Sepsis is a life-threatening condition of organ dysfunction caused by a dysregulated host immune response to infection. We performed network analysis of cytokine molecules and compared network structures between a systematic inflammatory response syndrome (SIRS) or normal control (NC) group and a sepsis group. Results: We recruited SIRS (n = 33) and sepsis (n = 89) patients from electronic medical records (EMR) according to whether data on PCT, CRP, interleukin (IL)-1β, IL-2, IL-4, IL-5, IL-6, IL-9, IL-10, IL-12p70, IL-13, IL-17, IL-22, TNF-α, and IFN-γ levels were available. From the public GEO dataset, GSE66099, GSE9960, GSE95233, GSE57065 were downloaded. Genes corresponding to 15 molecules were extracted from an expression array. A correlation matrix was formed for the 15 molecules and statistically significant molecular pairs were used as pairs for network analysis of coexpression. The number of molecular or gene expression pairs significantly correlated among the SIRS or control and sepsis groups are as follows for datasets: EMR, 15 and 15; GEO66099-1, 13 and 15; GEO9960, 13 and 11; GSE95233, 13 and 8; GSE66099-2, 15 and 14; GSE57065, 14 and 13, respectively. Network analysis revealed that network diameter, number of nodes and shortest path were equal to or lower in the sepsis group. Conclusions: The coexpression network in sepsis patients was relatively small sized and had lower shortest paths compared with the SIRS group or healthy control group. Cytokines with one degree (k = 1) are increased in sepsis group compared with SIRS or healthy control group. IL-9 and IL-2 were not included in network of sepsis group indicating that these cytokines showed no correlation with other cytokines. These data might imply that cytokines tend to be dysregulated in the sepsis group compared to that of SIRS or normal control groups.

Key words: sepsis, cytokine molecules, network analysis, network topology, GEO dataset

BACKGROUND

Sepsis is defined as documented infection with early activation of both pro and anti-inflammatory responses with organ dysfunction [1-3]. The host responses are mediated by immune-related molecules, including cytokines and molecules associated with pathophysiology [4-6]. Immunological function or molecules from sepsis patients are studied to elucidate the pathogenesis of sepsis [6-10]. The cytokine molecules such as interleukin (IL)-1β, IL-2, IL-4, IL-5, IL-6, IL-9, IL-10, IL-12p70, IL-13, IL-17, IL-22, TNF-α, IFN-γ, procalcitonin (PCT), and C-reactive protein (CRP) or expressed gene levels have been studied among SIRS and sepsis patients [11-14]. Various pro-inflammatory cytokines such as IL-1, IL-2, IL-6, IL-8, TNF-α, and anti-inflammatory cytokines such as IL-10, IL-1RA in serum are studied [5]. Cytokines are thought to be in a biphasic state during sepsis, resulting in an initial hyper-inflammatory state or cytokine storm state to a late hypo-inflammatory state or immune paralysis state [4, 7]. In another aspect, pro- and anti-inflammatory cytokines both occur during the course of sepsis and sepsis-related cytokines vary at the transcription level [15]. Cytokine analysis has mainly focused on individual cytokine characteristics between SIRS or sepsis groups [13, 14]. On the other hand, gene expression profiles dealt with pathway and differential gene expression between SIRS and sepsis groups [16, 17]. Cytokine measurement and gene expression data expanded our understanding of SIRS and sepsis pathogenesis, respectively. An equilibrium of cytokines between healthy and diseased states is rarely attributed to single molecules, but rather to groups of complex inter-related molecules. Because causalities among the cytokines are complex and the concentration in serum
has wide range of variation, a network approach might enhance our understanding of equilibrium of cytokines in sepsis [18, 19].

A biological network can be composed of clustered molecules with physical interaction and non-physical interactions. A few cases of the networks of molecules with physical interactions include protein versus protein interactions, transcription factor versus DNA interactions, miRNA versus RNA interactions, Drug versus target interactions, and molecular pathway interaction. A few cases of networks of non-physical interactions are coexpression, genetic interactions, functional networks, gene versus phenotype, miRNA versus disease, drug versus disease, disease networks, and phoneme networks [18, 19]. A conceptual network of gene coexpression data might reveal the complex network of cytokine in SIRS and sepsis patients, and could contribute in elucidating pathogenesis.

Therefore, in this study, we analyzed molecules and cytokines using network analysis for SIRS or normal controls and sepsis patients. Protein-based datasets and supporting gene expression signatures from public datasets were used. PCT, CRP, and the following cytokines were analyzed: interleukin (IL)-1β, IL-2, IL-4, IL-5, IL-6, IL-9, IL-10, IL-12, IL-13, IL-17, TNF-α, and IFN-γ.

**METHODS**

**Patients in the cohort**

This was a retrospective study and our protocol was approved by the Institutional Review Board of Incheon St. Mary’s Hospital. Thirteen cytokines, procalcitonin and C-reactive protein (CRP) from SIRS (n=33) and sepsis (n=89) patients were extracted from hospital’s electric medical records (EMR) between January and December 2013, and data was partially presented [14]. In brief, patient ≥18 years old who were admitted to emergency department with SIRS were selected. Among these, those with documented bacterial infection were regarded as being septic. Patients were excluded from the study if they showed any evidence of an immunocompromised state (e.g., malignancy) or of a viral infection. To match the SIRS cases, 33 sepsis cases were randomly selected. To support and validate the EMR dataset, genotypic data from a public microarray database was utilized for network analysis. Gene expression experiments with blood for SIRS or control and sepsis were roughly the same as in our study design. The public datasets were downloaded from the Gene Expression Omnibus (GEO) website (http://www.ncbi.nlm.nih.gov/geo/). Databases were searched using the following common terminology: sepsis, human, and GPL570 (Affymetrix Human Genome U133 Plus 2.0 Array or hgu133plus2.db). This platform was selected, because it allowed studying 47,000 transcripts from 38,000 genes. Approximately 40 datasets were found from which we selected the GSE66099, GSE9960, GSE95233 and GSE57065 datasets, because their study designs were similar to ours [15-17, 20-22]. GSE66099 is composed of several unique datasets, including GSE4607, GSE8121, GSE9692, GSE13904, GSE26378, and GSE26440. As healthy control case and SIRS case are included within GSE66099, GSE66099 was divided into GSE66099-1 with SIRS case and GSE66099-2 with healthy control cases. For each dataset, randomized selection of sepsis cases without replacement was performed to match the number of SIRS or healthy control cases. Matched number of pairs were selected for SIRS or healthy control cases to sepsis cases as follows: EMR datasets, 33 pairs; GSE66099-1, 30 pairs; GSE9960, 16 pairs; GSE95233, 22 pairs; GSE66099-2, 48 pairs; and GSE57065, 25 pairs, respectively. The GSE datasets were designed to compare SIRS or healthy control group with sepsis. The baseline characteristics of the datasets are listed in Table 1.

**Sepsis diagnosis**

Sepsis diagnosis was based on sepsis criteria in all the recruited datasets [1]. The patients were classified as having SIRS if they met two or more of the following criteria: WBC count ≥12.0 × 10⁹/L or ≤4.0 × 10⁹/L, or indicating more than 10% immature cells; body temperature ≥38°C or ≤36°C; respiratory rate ≥20 breaths per minute; or heart rate ≥90 beats per minute. The SIRS patients were classified as septic if they had documented bacterial infection [1].

**Bioinformatics**

All analyses were performed using R software and several appropriate packages [23], including Bioconductor,

---

Table 1  
Baseline characteristics of recruited datasets

<table>
<thead>
<tr>
<th>Source</th>
<th>EMR dataset</th>
<th>GSE66099-1 dataset</th>
<th>GSE9960 dataset</th>
<th>GSE95233 dataset</th>
<th>GSE66099-2 dataset</th>
<th>GSE57065 dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient type</td>
<td>Adult</td>
<td>Pediatric Patient</td>
<td>Adult</td>
<td>Pediatric Patient</td>
<td>Adult</td>
<td>Pediatric Patient</td>
</tr>
<tr>
<td>Age (year)</td>
<td>≥18</td>
<td>≤10</td>
<td>≥1.5</td>
<td>≥18</td>
<td>≤10</td>
<td>≥18</td>
</tr>
<tr>
<td>SIRS</td>
<td>33</td>
<td>30</td>
<td>16</td>
<td>(-)</td>
<td>22</td>
<td>48</td>
</tr>
<tr>
<td>Control</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>22</td>
<td>48</td>
<td>25</td>
</tr>
<tr>
<td>Total sepsis</td>
<td>90</td>
<td>198</td>
<td>54</td>
<td>51</td>
<td>198</td>
<td>28</td>
</tr>
<tr>
<td>Randomized selection among sepsis case</td>
<td>33</td>
<td>30</td>
<td>16</td>
<td>22</td>
<td>48</td>
<td>25</td>
</tr>
<tr>
<td>Reference</td>
<td>[14]</td>
<td>[16]</td>
<td>[20]</td>
<td>[21]</td>
<td>[16]</td>
<td>[22]</td>
</tr>
</tbody>
</table>

SIRS: systemic inflammatory response syndrome; (-) not available
Biobase, affy, GEOquery, limma, hgu133plus2.db, annaffy, and annotate [24-27]. The data were downloaded and the data matrix of gene expression values was formatted using the GEOquery package [25]. The raw microarray data (CEL files) were normalized using the Robust Multichip Averaging (RMA) method, and log2 transformation was performed [26-28]. The data were annotated with corresponding gene names using the R packages annaffy, hgu133plus2.db, and annotate [29-31]. Each dataset was collapsed from approximately 54675 rows to 22486 rows by gene names and highest values among redundant gene expression values were selected. Interleukin (IL)-1, IL-2, IL-4, IL-5, IL-6, IL-9, IL-10, IL-12, IL-13, IL-17, IL-22, TNF, IFNG, PCT (CALCA in gene expression data), and CRP, which was also of analytical interest, were parsed among them.

**Network topology analysis**

Based on Spearman’s correlations for each set of molecules, network analysis was performed for molecules with significant correlation coefficients. Each molecule was regarded as a node and number of correlation pairs between the molecules are regarded as links (edges) [32, 33]. Network analyses were performed, the results were plotted, and the topologic parameters from the network were calculated using Cytoscape and Network Analyzer, which is a plugin software for Cytoscape [33-35].

The degree measurement is defined as the number of links (edges) connected to a node (molecule), and a node with a relatively high degree implies a possible hub. The clustering coefficient is defined as the number of links among associated nodes, connected each other, which all allow for a connection to node i. Network density is defined as the ratio of links in the network to the maximum possible number of links, which is proportionate to the number of edges.

The length of a path is the number of unique edges that form it. Distance is defined as the shortest path length between nodes i and j within the network. Network diameter is defined as the maximum length of the shortest path. The characteristic path length is the average shortest path length and provides expected distance between nodes i and j. The average number of neighbors is the number of same identical nodes connected by nodes i and j. The normalized version of the average number of neighbors is the network density. Network heterogeneity is defined as the variance in the number of links divided by the number of mean links. The average number of neighbors is a parameter that defines the average connectivity of the nodes. The topological coefficient is a relative measure of the extent to which a node shares a neighbor with another node. Network centralization is close to 1 when the network resembles a star figure, while uniformly connected networks have a centralization value close to 0.

Stress centrality of a node is the number of shortest paths passing through the node. Betweenness centrality of node k is defined as the shortest path between nodes i, and j that pass through node k, divided by the shortest path between nodes i and j, and it reflects the amount of control that this node exerts over other nodes. Closeness centrality reflects how close on average. Eccentricity centrality reflects maximum distance from a node [30-32].

To compare network parameters between control or SIRS and sepsis, Z scores were calculated for respective parameters. The parameters of control or SIRS group were used as control for calculation. Absolute Z score of 1.644 or greater was regarded as statistically significant (p ≤ 0.05).

**RESULTS**

**Levels of molecules and correlation analysis of molecules**

The EMR dataset revealed that most of the cytokines were elevated in the sepsis group with statistical significance except IL-2 and IL-9 compared with the SIRS group (table 1). On the contrary, IFN was decreased in the sepsis group with statistical significance compared with the SIRS group. Gene expression analyses in the GEO dataset revealed that most of the gene expression revealed statistical significance for the tested genes. GSE66099-1, GSE9960, GSE66099-2 and GSE57065 revealed higher expression of genes in the sepsis group compared with the SIRS or control group. On the contrary, GSE9960, GSE95233 revealed lower expression of genes in the sepsis group compared with the SIRS or control group. Expression matrix via Spearman’s method was performed for each dataset. The number of molecular or gene expression pairs significantly correlated among the SIRS or control and sepsis groups are as follows for datasets: EMR, 15 and 15; GEO66099-1, 13 and 15; GEO9960, 13 and 11; GSE95233, 13 and 8; GSE66099-2, 15 and 14; GSE57065, 14 and 13, respectively (Supplementary Table S1-S24).

**Network analysis of molecules**

Based on significant correlation coefficients, networks were plotted for pairs of molecules with significant correlations. Based on this co-expression matrix, topologic parameters were calculated (table 2). The EMR dataset and GEO datasets with SIRS and sepsis are plotted (figure 1) and GEO datasets with control and sepsis are plotted (figure 2), respectively. All the networks from the datasets revealed that network diameter was equal to or lower in sepsis compared with SIRS or control. Except for GSE9960, characteristic path length was lower in the sepsis group. Except for GSE66099-1, shortest path and number of nodes were equal to or lower in sepsis compared with SIRS or control. The Z score calculated for shortest path length revealed that SIRS (Z = -1.787) and sepsis group (Z = -2.671) of GSE95233 revealed statistical significance and sepsis group (z = -2.028) of GSE57065 revealed statistical significance. The characteristic path length revealed that the sepsis group (z = -2.359) of GSE66099-2 revealed statistical significance.

Analysis of each node revealed that in the EMR dataset, TNF had the highest degree (k = 10) in the SIRS group and IL-1 (k = 11), IL-17 and TNF had highest degree (k = 10) in the sepsis group. In the GSE66099-1 dataset, IL-9 had highest degree (k = 7) in the SIRS group and IL-6 (k = 10) had highest degree in the sepsis group. In the GSE9960 dataset, IL-17A had highest degree (k = 8) in the SIRS group and IL-22 had highest degree in the sepsis group. In the GSE57065 dataset, CALCA had highest degree (k = 7) in the control group and CRP (k = 4) had highest degree in the sepsis group. In the GSE66099-2 dataset,
Table 2

Cytokine protein or gene expression levels in datasets

<table>
<thead>
<tr>
<th></th>
<th>EMR</th>
<th>GSE66099-1</th>
<th>GSE960</th>
<th>GSE95233</th>
<th>GSE66099-2</th>
<th>GSE57065</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SIRS</td>
<td>Sepsis</td>
<td>P</td>
<td>SIRS</td>
<td>Sepsis</td>
<td>P</td>
</tr>
<tr>
<td>IL–12</td>
<td>1.966</td>
<td>7.306</td>
<td>0.021</td>
<td>1.680</td>
<td>1.724</td>
<td>1.979</td>
</tr>
<tr>
<td>IL–17</td>
<td>9.109</td>
<td>30.547</td>
<td>0.012</td>
<td>2.034</td>
<td>2.114</td>
<td>2.243</td>
</tr>
<tr>
<td>IL–10</td>
<td>32.884</td>
<td>1380.413</td>
<td>&lt;0.001</td>
<td>2.231</td>
<td>2.493</td>
<td>2.261</td>
</tr>
<tr>
<td>IL–13</td>
<td>19.905</td>
<td>44.423</td>
<td>&lt;0.001</td>
<td>2.521</td>
<td>2.536</td>
<td>2.503</td>
</tr>
<tr>
<td>IL–22</td>
<td>173.734</td>
<td>344.843</td>
<td>0.0076</td>
<td>1.793</td>
<td>1.960</td>
<td>2.104</td>
</tr>
<tr>
<td>IL–2</td>
<td>18.107</td>
<td>46.613</td>
<td>0.033</td>
<td>1.457</td>
<td>1.639</td>
<td>1.839</td>
</tr>
<tr>
<td>IL–4</td>
<td>11.052</td>
<td>57.521</td>
<td>0.007</td>
<td>1.817</td>
<td>1.845</td>
<td>2.026</td>
</tr>
<tr>
<td>IL–9</td>
<td>21.938</td>
<td>40.978</td>
<td></td>
<td>1.567</td>
<td>1.608</td>
<td>1.892</td>
</tr>
<tr>
<td>IL–1</td>
<td>28.601</td>
<td>89.072</td>
<td>&lt;0.001</td>
<td>3.086</td>
<td>3.191</td>
<td>3.131</td>
</tr>
<tr>
<td>IL–5</td>
<td>10.803</td>
<td>31.837</td>
<td></td>
<td>1.612</td>
<td>1.555</td>
<td>1.871</td>
</tr>
<tr>
<td>IL–6</td>
<td>385.793</td>
<td>979.991</td>
<td></td>
<td>2.257</td>
<td>2.289</td>
<td>2.294</td>
</tr>
<tr>
<td>IFN</td>
<td>543.449</td>
<td>237.206</td>
<td>&lt;0.001</td>
<td>2.233</td>
<td>2.082</td>
<td>2.142</td>
</tr>
<tr>
<td>TNF</td>
<td>21.142</td>
<td>58.484</td>
<td>&lt;0.001</td>
<td>2.604</td>
<td>2.679</td>
<td>2.760</td>
</tr>
<tr>
<td>CRP</td>
<td>50.932</td>
<td>125.763</td>
<td>&lt;0.001</td>
<td>1.941</td>
<td>1.986</td>
<td>2.055</td>
</tr>
<tr>
<td>PCT</td>
<td>1.726</td>
<td>23.248</td>
<td>&lt;0.001</td>
<td>2.112</td>
<td>2.129</td>
<td>2.166</td>
</tr>
</tbody>
</table>

Units for cytokines are as follows: CRP, mg/L; PCT, ng/mL; other, pg/mL. The gene levels are normalized and log2 transformed expression data. PCT should be regarded as CALCA for gene expression datasets.
CRP and IL-17A had highest degree (k = 8) in the control group and IL-4 and CALCA (k = 10) had highest degree in the sepsis group. In the GSE57065 dataset, IL-17A had highest degree (k = 7) in the control group as well as in the sepsis group (k = 5). Other topologic parameters of each node are listed in S25-36 Tables. Among cytokine genes, IL-9 was not included in sepsis network analysis for the GSE9960, GSE95233 and GSE57065 datasets, indicating that IL-9 had no correlation with other cytokines within the sepsis group. IL-2 was also not included in sepsis network analysis for GSE95233 and GSE66099-2.

Altogether, network diameter, number of nodes, characteristic path length and shortest path were mostly lower in the sepsis group than in the SIRS or normal control group. The sepsis network reflects a relatively small sized network compared with the SIRS group, which was a relatively larger network.

DISCUSSION

The definition of sepsis has become more accurate based on pathophysiology but there are still controversies. The common factors included in the definitions from SEPSIS-1 to SEPSIS-3 are microbial infection and a dysregulated host immune response [1-3]. The definition of a dysregulated immune response still requires elucidation, and previous research showed that immune function might become upregulated and then down regulated as infection progresses to sepsis [3, 36-38]. Features of immune dysfunction in sepsis include defective antigen presentation, defective T and B cell mediated immunity, defective NK cells; dysregulated regulatory T cell, activation of PD-1, defective immunoglobulin level, alteration in neutrophils, abnormal cytokine levels, complement consumption, and defective bacterial killing [9, 36].
Studying cytokine molecules might be a way to better understand immune dysregulation in sepsis. A previous cytokine molecule study in SIRS and sepsis patients analyzed single cytokine levels [13, 14], and the interrelations among cytokines have rarely been analyzed at the protein level. Because of the complexity of cytokine, network analysis might better reflect different aspects of their dynamics that cannot be gleaned via single cytokine analysis. Analysis of single cytokines revealed variance between the chosen datasets (table 2). Except for IFN, cytokines are increased in the sepsis group compared with the SIRS group in the EMR dataset [14]. On the contrary, gene expression patterns were heterogeneous among datasets. Single cytokine analysis revealed variance among the cytokines, which was consistent with a previous study that found variation in cytokine levels during transcription [15]. These characteristics might be derived from the diverse nature of sepsis, which includes a heterogeneous subset of patients with different biological and clinical characteristics, including age, medication, comorbidities, causative organism, organism with different antibiotic resistance, and patients with different immune status. As interrelation of the cytokine is complex, analysis of multiple cytokines by network analysis might reflect different aspects, which could not be considered by a single cytokine analysis. Network analysis was performed based

Figure 2
The network analysis result of the Control (healthy) group and sepsis group. (A) GSE95233 dataset, (B) GSE66099-2 dataset, (C) GSE57065 dataset.
### Table 3

<table>
<thead>
<tr>
<th></th>
<th>EMR</th>
<th>GSE66099-1</th>
<th>GSE9960</th>
<th>GSE95233</th>
<th>GSE66099-2</th>
<th>GSE57065</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SIRS</strong></td>
<td></td>
<td>Control</td>
<td>Sepsis</td>
<td>Control</td>
<td>Sepsis</td>
<td>Control</td>
</tr>
<tr>
<td><strong>Sebes</strong></td>
<td></td>
<td>Control</td>
<td>Sepsis</td>
<td>Control</td>
<td>Sepsis</td>
<td>Control</td>
</tr>
</tbody>
</table>

- **Clustering Coefficient**
  - EMR: 0.741, 0.716, 0.416, 0.349
  - GSE66099-1: 0.549, 0.416
  - GSE9960: 0.303
  - GSE95233: 0.347
  - GSE66099-2: 0.292
  - GSE57065: 0.273

- **Connected components**
  - EMR: 1, 1, 1, 1
  - GSE66099-1: 1, 1, 1, 1
  - GSE9960: 1, 2, 1
  - GSE95233: 1, 2
  - GSE66099-2: 1
  - GSE57065: 1

- **Network diameter**
  - EMR: 4, 3, 6, 4, 4, 4
  - GSE66099-1: 3
  - GSE9960: 4
  - GSE95233: 3
  - GSE66099-2: 2
  - GSE57065: 3

- **Network centralization**
  - EMR: 0.319, 0.302, 0.356, 0.386, 0.381, 0.321, 0.437, 0.308
  - GSE66099-1: 0.378
  - GSE9960: 0.386
  - GSE95233: 0.381
  - GSE66099-2: 0.386
  - GSE57065: 0.385

- **Shortest path**
  - EMR: 210, 210, 156, 210, 210, 142, 210, 210
  - GSE66099-1: 210, 210, 156, 210
  - GSE9960: 210, 210, 100, 210
  - GSE95233: 210, 210, 100, 210
  - GSE66099-2: 210, 210, 100, 210
  - GSE57065: 210, 210, 100, 210

- **Character path length**
  - EMR: 2, 1, 1, 2, 0.65, 0.85, 2, 0.5
  - GSE66099-1: 1, 1, 1, 1, 1, 1, 1, 1
  - GSE9960: 1, 1, 1, 1, 1, 1, 1, 1
  - GSE95233: 1, 1, 1, 1, 1, 1, 1, 1
  - GSE66099-2: 1, 1, 1, 1, 1, 1, 1, 1
  - GSE57065: 1, 1, 1, 1, 1, 1, 1, 1

- **Avg. number of neighbors**
  - EMR: 6.133, 7.333, 3.385, 4.533, 4, 2.909, 3.077, 2
  - GSE66099-1: 4, 2
  - GSE9960: 4
  - GSE95233: 4
  - GSE66099-2: 4
  - GSE57065: 4

- **Network density**
  - EMR: 0.48, 0.5, 0.27, 0.3, 0.4, 0.5, 0.6, 0.3
  - GSE66099-1: 0.51, 0.5
  - GSE9960: 0.51, 0.5
  - GSE95233: 0.51, 0.5
  - GSE66099-2: 0.51, 0.5
  - GSE57065: 0.51, 0.5

- **Network heterogeneity**
  - EMR: 0.43, 0.43, 0.43, 0.43, 0.43, 0.43, 0.43, 0.43
  - GSE66099-1: 0.44, 0.44
  - GSE9960: 0.44, 0.44
  - GSE95233: 0.44, 0.44
  - GSE66099-2: 0.44, 0.44
  - GSE57065: 0.44, 0.44

Network analysis of molecules revealed some common topological parameters. The sepsis group yielded a network with lower diameter, number of nodes and shortest path than the network for the SIRS or normal control group. Roughly, the sepsis group retained a small sized network compared with the SIRS or normal control group. Unlike single cytokine level or gene expression, network analysis revealed common features. Network analysis revealed that cytokines with only one degree (k = 1) are more frequent in the sepsis group compared to the normal or healthy control group. In addition, IL-9 was not included in sepsis network analysis for GSE9960, GSE95233 and GSE57065 datasets, indicating that IL-9 had no correlation with other cytokines within the sepsis group. IL-2 was also not included in sepsis network analysis for GSE95233 and GSE66099-2. These data indicate that cytokines in the sepsis group tend to be unregulated or not correlated or dysregulated or segregated. Cytokine molecules in the sepsis patients are thought to become uncontrolled and dysregulated compared with the SIRS or normal control group. The equilibrium of cytokines state in sepsis is not pro- or anti-inflammatory state but in the dysregulated state losing correlation or harmony with other cytokines. The dysregulation of host immune state might be defined using topologic parameters, such as small diameter, small number of nodes, characteristic path length or shortest path among SIRS and sepsis. In addition, changing levels of cytokines should be also considered.

Path length is reported to be related to function or modularity [39, 40]. As network evolves, complex function and specialized function are executed within modules that are functionally separable subnetworks. Such modules tend to increase path length or enlarge network diameter or size because it increases minimal path length between modules due to increased nodes [39, 40]. In this data, shortest path length, characteristic path length was equal to or shorter in the sepsis group, which implies that function of cytokine network is decreased or lost, which is in line with clinical outcomes. Nodes with highest degree could be regarded as a hub node, which has tendency to link to nodes with fewer edges or interaction partners rather than to other hubs [41]. Removal of hub nodes has great impact on the network which could lead to system failure. In this data, we could not find such common hubs, but IL-17A gene had the highest degree in 3 data set of SIRS or control group as well as in 2 data sets of sepsis group. This result implies that IL-17 might be a house keeping cytokine in cytokine network.

A limitation of this study is that only five of public datasets with the same platform were used.Datasets with other platforms could be integrated into the study to improve samples size and data quantity, at the cost of decreasing homogeneity of platform. The recruited data are heterogenous in that one was based on protein and others are gene expression data. The age was also heterogenous, which could affect equilibrium of cytokine in normal and sepsis state. The 15 molecules we analyzed are relatively few molecules considering the vast cytokine network of the immune system. Increasing the number of cytokines studied by conducting a phenotypic study could expand our understandings of sepsis.
CONCLUSIONS

The coexpression network in sepsis patients was relatively small sized and had lower shortest path compared with the SIRS group or healthy control group. Cytokines with one degree (k = 1) are increased in the sepsis group compared with the SIRS or healthy control group. IL-9 and IL-2 were not included in network of sepsis group indicating that these cytokines showed no correlation with other cytokines. These data might imply that cytokines tend to be dysregulated in the sepsis group compared to that of the SIRS or normal control groups.

Disclosure. Acknowledgments. We thank and cherish the memory of former associate professor KiYoung Lee (Laboratory of Network Biology and Medicine, Department of Medical Informatics, Ajou University, School of Medicine) who inspired us to write this article. We thank MiJeong Kim (Department of Anesthesiology, St. Mary’s Spine & Joint Hospital) for supporting the hardware used for our bioinformatics analyses. Conflict of interest: The authors declare that they have no competing interest.

SUPPLEMENTARY DATA

Supplementary data associated with this article can be found, in the online version, at doi:10.1684/ecn.2018.0414.

REFERENCES


