Personalized identification of differentially expressed pathways in pediatric sepsis

BINJIE LI[1] and QIYI ZENG[2]

[1]The First Clinical Medical College, Southern Hospital, Southern Medical University, Guangzhou, Guangdong 510515; [2]Department of Pediatric Center, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong 510280, P.R. China

Received November 10, 2015; Accepted November 4, 2016

DOI: 10.3892/mmr.2017.7217

Abstract. Sepsis is a leading killer of children worldwide with numerous differentially expressed genes reported to be associated with sepsis. Identifying core pathways in an individual is important for understanding septic mechanisms and for the future application of custom therapeutic decisions. Samples used in the study were from a control group (n=18) and pediatric sepsis group (n=52). Based on Kauffman's attractor theory, differentially expressed pathways associated with pediatric sepsis were detected as attractors. When the distribution results of attractors are consistent with the distribution of total data assessed using support vector machine, the individualized pathway aberrance score (iPAS) was calculated to distinguish differences. Through attractor and Kyoto Encyclopedia of Genes and Genomes functional analysis, 277 enriched pathways were identified as attractors. There were 81 pathways with P<0.05 and 59 pathways with P<0.01. Distribution outcomes of screened attractors were mostly consistent with the total data demonstrated by the six classifying parameters, which suggested the efficiency of attractors. Cluster analysis of pediatric sepsis using the iPAS method identified seven pathway clusters and four sample clusters. Thus, in the majority pediatric sepsis samples, core pathways can be detected as different from accumulated normal samples. In conclusion, a novel procedure that identified the dysregulated attractors in individuals with pediatric sepsis was constructed. Attractors can be markers to identify pathways involved in pediatric sepsis. iPAS may provide a correlation score for each of the signaling pathways present in an individual patient. This process may improve the personalized interpretation of disease mechanisms and may be useful in the forthcoming era of personalized medicine.

Correspondence to: Dr Qiyi Zeng, Department of Pediatric Center, Zhujiang Hospital, Southern Medical University, 253 Industrial Road, Haizhu, Guangzhou, Guangdong 510280, P.R. China
E-mail: qiyizeng@126.com

Abbreviations: ANS, accumulated normal samples; DEG, differentially expressed genes; FC, fold change; FDR, false discovery rate; FN, false negatives; FP, false positives; iPAS, individualized pathway aberrance score; KEGG, Kyoto Encyclopedia of Genes and Genomes; RMA, robust multi-array average; SD, standard deviation; SVM, support vector machine; TN, true negatives; TP, true positives

Key words: pediatric sepsis, attractors, differentially expressed pathway, support vector machine, individualized pathway aberrance score

Introduction

Sepsis is a leading cause of mortality in children worldwide, with the mortality rate in developing countries reported to be as high as 50% (1). Sepsis is caused by a wide variety of infectious pathogens that can lead to organ dysfunction, shock and mortality (2-4).

Gene expression profiling of diseased human tissues may provide insights into the molecular mechanisms that mediate sepsis and eventually lead to the identification of novel therapeutic targets (5). High-throughput microarray experiments were developed to analyze genetic expression patterns of differentially expressed genes (DEGs). They were sufficient to identify potential target genes, however, they lacked integrality. There are various pathways associated with pediatric sepsis in the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways database. Among them, few reports focused on the core pathways involved. The attractor theory proposed by Kauffman (6) was famous for identifying one or more well-defined model network ensembles with statistical features that matched those of real cells and organisms. Mar *et al* (7) reported that *attract* was a novel approach that could leverage both existing pathway databases and the DEGs between cell phenotypes. It expanded these inferences by identifying new coordinately-regulated gene sets that were relevant to the mechanisms underlying differences that may be associated with the molecular mechanism of diseases. This method was employed in the current study to screen attractor pathways from the vast data of KEGG pathways, in order to narrow down the number of correlated pathways in the individual pathway analysis.

The majority of current pathway analyses have been developed to investigate deregulated pathways between two phenotype groups. The individualized pathway aberrance score (iPAS) method is straightforward to compare the expression profiles of an individual disease and normal cells to identify molecular changes that are specific to the disease (8).

The present study used iPAS to compare one sepsis sample with multiple accumulated normal samples (ANS). This is a biologically intuitive guideline to analyze a single sample that lacks cohort data, which is different from the traditional gene expression analysis. The method involves four steps: Data processing, gene-level statistics, iPAS and a significance test. It captures biological and clinical information in a sensible, valid and useful way, and has been previously used for the analysis of colorectal and lung cancer (8). iPAS was used in the present study to investigate attractors in pediatric septic samples in order to distinguish differences from the control group. A pipeline that identified the dysregulated attractors in individuals of pediatric sepsis was created. Pathways from the KEGG database were set as attractors in order to screen differentially expressed pathways. Following narrowing down the number of correlated pathways using support vector machines (SVM), iPAS was used to distinguish differences in individual pathways.

Materials and methods

Gene expression and data preprocessing. The transcription profile was obtained from EMBI-EBI ArrayExpress (EMBL-EBI, Hinxton, UK) (9). Gene expression profiling of 70 tissues were collected from the E-GEOD-13904 dataset [included five sample types: Sepsis, controls, septic shock, systemic inflammatory response syndrome (SIRS) and SIRS resolved] (10). The samples of the dataset were obtained from 18 control individuals and 52 pediatric patients with sepsis. The array platform used to obtain the profile was A-AFFY-44-Affymetrix GeneChip Human Genome U133 Plus 2.0 (HG-U133_Plus_2; Affymetrix, Inc., Santa Clara, CA, USA).

Data of the gene array chip was read in the R package 'affy' bioconductor (11). The Linear Models for Microarray Data package (12) was then used to preprocess the data. Background adjustment and quantile data normalization were performed by robust multi-array average (13). To protect against outlier probes, a robust procedure, median polish (14), was used to estimate model parameters. The average value of a gene symbol with multiple probes was calculated. P≤0.01 and |log fold change (FC) |≥2 were set as the threshold levels.

Information from gene sets representing human biological pathways human was obtained from the KEGG pathways database (15). Pathways with large number of genes have more complex metabolism. Therefore, pathways with gene set size >100 were filtered. On account of certain differences among data from different platforms, pathways with the number of genes >5 were selected.

Kauffman' attractor analysis. Based on Kauffman's attractor theory (6), *attract* was used to screen differentially expressed pathways associated with pediatric sepsis. To test pathway-level data, gene set enrichment analysis (GSEA)-analysis of variance (ANOVA) was employed as a variance-based implementation of a gene set enrichment algorithm, which was different from other methods in multiple classes (7). Under GSEA-ANOVA, an ANOVA model was fit to each gene where the expression of a gene is modeled by a single factor representing the cell types as distinct levels of this class. From the ANOVA model, the *F*-statistic for gene '*i*' was calculated as follows:

$$F^{(i)} = \frac{MSS_i}{RSS_i}$$

Where MSS_i represents the mean treatment sum of squares, and captures the amount of variation due to the cell type group-specific effects:

$$MSS_i = \frac{1}{K-1}\sum_{k=1}^{k} r_k \left[y_k^{(i)} - y^{(i)}\right]^2$$

and RSS_i represents the residual sum of squares:

$$RSS_i = \frac{1}{N-K}\sum_{k=1}^{K}\sum_{j=1}^{r_j}\left[y_{jk}^{(i)} - y^{(i)}\right]^2$$

where N is the total number of samples, and the overall mean is given by:

$$y^{(i)} = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{1}{r_k}\sum_{j=1}^{r_k} y_{jk}^{(i)}\right)$$

The *F*-statistic captures the strength of association observed in the expression of a gene in pediatric sepsis. Large *F*-statistic values indicate a strong association with sepsis-specific expression changes.

For pathway *P* consisting of g_p genes, the *T*-statistic takes the following form:

$$T_p = \frac{\left[\frac{1}{g_p}\sum_{i=1}^{g_p} F^{(i)}\right] - \left[\frac{1}{G}\sum_{j=1}^{G} F^{(j)}\right]}{\sqrt{\left(\frac{s_p^2}{g_p}\right) + \left(\frac{s_G^2}{G}\right)}}$$

Where *G* represents the total number of genes with a pathway annotation and the sample variances s_p^2 and s_G^2 are defined as:

$$s_p^2 = \frac{1}{g_p - 1}\sum_{j=1}^{g_p}\left(F^{(j)} - \frac{1}{g_p}\sum_{i=1}^{g_p} F^{(i)}\right)^2$$

$$s_G^2 = \frac{1}{G-1}\sum_{j=1}^{G}\left(F^{(j)} - \frac{1}{G}\sum_{i=1}^{G} F^{(i)}\right)^2$$

and the degrees of freedom are specified by the Welch-Satterwhaite equation:

$$v = \frac{\left(\frac{s_p^2}{g_p} + \frac{s_G^2}{G}\right)^2}{\frac{s_p^4}{g_p^2(g_p - 1)} + \frac{s_G^4}{G^2(G-1)}}$$

KEGG had great value in using the data to extend putative functional annotation. Obtained differentially expressed pathways were identified as attractors. They were ranked according to the significance of the differential expression.

Distribution outcomes in SVM. SVM is a supervised learning method used for classification and regression tasks that originated from statistical learning theory (16). Traditionally, accuracy has been used to evaluate classifier performance. This measure is defined as the total number of good classifications over the total number of available examples. Typically, the majority of the classification problems have two classes, positive and negative cases (17). Thus, the classified test points can be divided into four categories that usually are represented in the well-known confusion matrix as follows: True positives (TPs), true negatives (TNs), false positives (FPs) and false negatives (FNs). Sensitivity is equal to the TP rate, i.e., the ratio of TPs that were identified in all positive class samples. Analogously, specificity is the rate of TNs. Sensitivity and specificity describe how well the classifier discriminates the positive and the negative classes, respectively.

SVM was used to test the consistency of distribution results between screened attractors and vast pathways from KEGG. Subsequent analyses using attractors were performed when the distribution outcome of P<0.05 and P<0.01 were consistent with that of the total data.

Individualized analysis. ANS obtained from the Gene Expression Omnibus database of NCBI (www.ncbi.nlm.nih.gov/geo/) (18) referred to the ANS. A total of 18 ANS were collected for identifying individual attractors.

Data preprocessing and gene-level statistics. Expression level was defined using the robust multichip average (13). For individual cases,

$$\mathrm{proj}_{a}q_{k}=\left(\frac{1}{n}\sum_{j=1}^{n}qkj,\ldots,\frac{1}{n}\sum_{j=1}^{n}qkj\right)$$

as quantile normalization was performed following combining the single sepsis data (19). In cases of genes with multiple probes, the gene expression level was summarized by averaging probe-level expression. Individual septic sample gene expression was standardized using the mean and standard deviation.

Pathway-level statistics. The average Z method is a biologically valid modification of current pathway analysis techniques for iPAS, enabling analysis of an individual pediatric sepsis sample pathway aberrance by using ANS (8).

A vector Z= (z_1, z_2, . . . , z_n) denotes the expression status of a pathway where z_i symbolizes the standardized expression value of *i*-th gene, where the number of genes belonging to the pathway is *n*. Gene level statistics of each gene from each sample were calculated as follows:

$$z_i=\frac{g_{Ti}-\mathrm{mean}(g_{nANS})}{\mathrm{stdev}(g_{nANS})},$$

Pathway statistics of every attractor were calculated as follows:

$$iPAS=\frac{\sum_{i}^{n}z_i}{n}$$

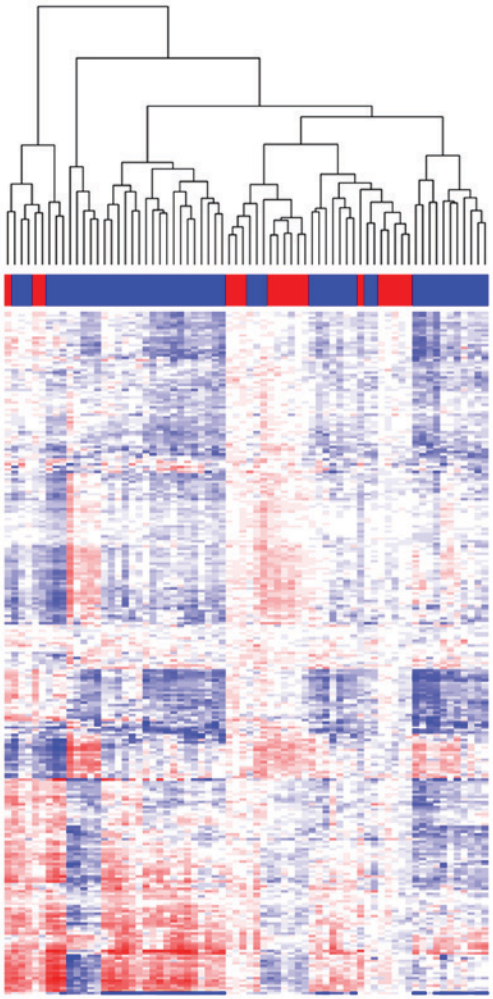


Figure 1. Heat map of individual pathway scores for 277 attractors in pediatric sepsis and control samples. Pathways were clustered with the abscissa and gene chips were clustered with the ordinate. Blue indicates upregulated pathways and red indicates downregulated ones. The expression level is proportional to the brightness of the color.

z_i represents the standardized gene level statistics of *1-i* gene and the number of genes belonging to the pathway is *n*. As the test may induce false positive results, a method (20) was used to adjust the raw P-values into false discovery rate (FDR) to circumvent the problem. FDR<0.05 and |log FC|>1 were used as the cut-off criteria.

Results

Attractor screening analyses. Through *attract* and KEGG functional analysis, 277 enriched pathways were selected as attractors (Fig. 1). They were ranked according to the significance of difference. There were 81 pathways with P<0.05 and 59 pathways with P<0.01 (Fig. 2). Pathways of P<0.01 were broadly classified into two functional themes, signaling pathways (T cell receptor, B cell receptor and nuclear factor-κB signaling pathway) and diseases (tuberculosis, legionellosis and inflammatory bowel disease). The first group of pathways involved in signal transduction is consistent with the hypothesis that the ability to recognize and respond to internal signals drives cell differentiation capacity. The second theme involves growth and metabolism and highlights the fact that pediatric

Table I. Distribution results of screened attractors with P-values tested by Support Vector Machine.

Parameter	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Balanced accuracy	Accuracy
Total attractors	1.00	0.88	0.95	1.00	0.94	96.00
Attractors of P<0.05	1.00	0.88	0.94	0.70	0.86	86.00
Attractors of P<0.01	0.90	0.88	0.95	0.78	0.89	89.00

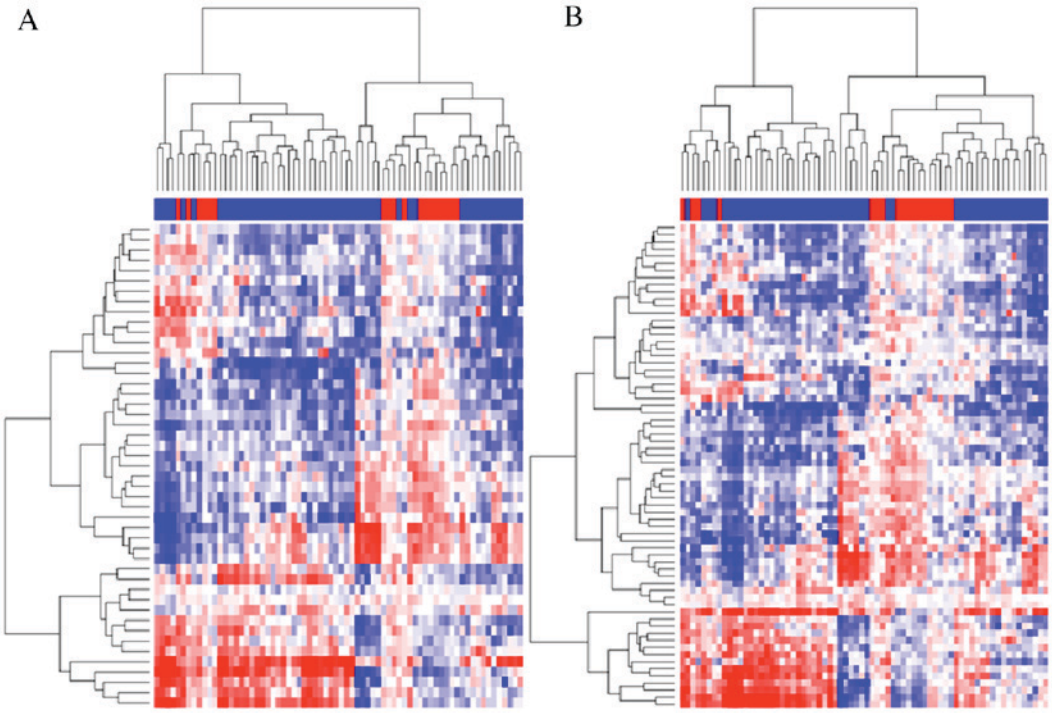


Figure 2. Heat map of significantly different pathways in pediatric sepsis using Kauffman' attractor theory. (A) 59 pathways of P<0.01. (B) 81 pathways of P<0.05. Pathways were clustered with the abscissa and gene chips were clustered with the ordinate. Blue indicates upregulated pathways and red indicates down regulated pathways; the expression level is proportional to the brightness of the color.

patients with sepsis across the differentiation spectrum are expected to have different iPAS scores.

Distribution outcomes in SVM. SVM was used to test the distribution results of screened attractors (Table I). The classifier was used to evaluate the classification efficiency. Distribution outcomes of screened attractors were predominantly consistent with that of total data demonstrated by the six classifying efficiency parameters, which suggested the efficiency of attractors. This demonstrated that screened attractors may replace the vast pathways from KEGG to identify differences in pediatric patients with sepsis using iPAS.

Differentially expressed pathways in individual. Cluster analysis of pediatric sepsis using the iPAS method on Wong's data (10) identified seven pathway clusters (denoted by 1-7 in Fig. 3) and four sample clusters (S1-S4). Sample clusters represent histopathological differentiation status. Cluster S1 represented the differentiation status of pediatric sepsis well and cluster S4 is close to ANS. This assures that unbiased clustering-based iPAS has the required sensitivity to capture clinically relevant associations. Of the pathway clusters presented in Fig. 3, P5, P6 and P7 are markedly different from the control group in pediatric sepsis samples, and may be core pathways involved in the disease. The pathways were antigen processing and presentation, Ras signaling pathway, Fc εRI signaling pathway and others. Pathway cluster P1 and P4 exhibited relatively weak differences when comparing pediatric septic samples and ANS. In conclusion, in the majority of pediatric sepsis sample differences in core pathways are detected, when compared with ANS. Therefore, iPAS analyses may aid the development of personalized pharmacogenetic treatment in pediatric sepsis.

Discussion

Attractor screening analyses. Kauffman' attractor analysis was suggested as an experimental approach to study implications of ensembles of regulatory networks (6). Mar *et al* (7) reported *attract* identified the ribosomal pathway as significant in embryonic stem cells, and it was driven by elevated expression in these cells. The implementation of *attract* was demonstrated to identify the pathways that best describe four example cell types (7).

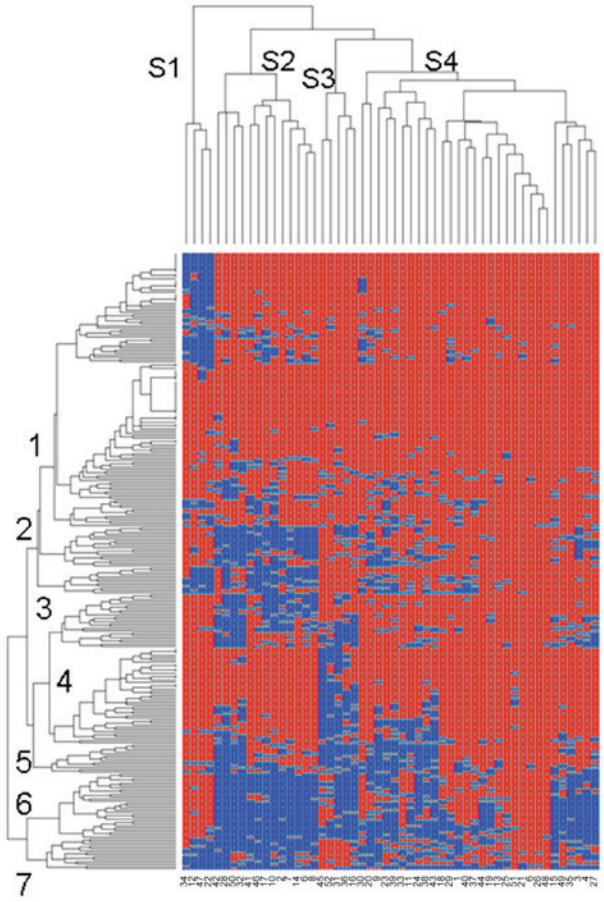


Figure 3. Clustered individualized pathway aberrance score of pediatric sepsis dataset. Pathways (n=277) were clustered with the abscissa and samples (n=52) were clustered with the ordinate. Red indicates the accumulated normal samples and blue indicates there are differences in individual pediatric sepsis samples (P<0.05).

Among 277 attractors, the most significantly different pathway in sepsis samples was osteoclast differentiation (KEGG ID: 04380; P=3.49E-11), where 128 genes were enriched. Osteoclasts, multinucleated cells originating from the hematopoietic monocyte-macrophage lineage, are responsible for bone resorption (21). Osteoclastogenesis is predominantly regulated by signaling pathways activated by receptor activator of nuclear factor-κB and immune receptors, whose ligands are expressed on the surface of osteoblasts.

Various differentially expressed pathways among the 277 attractors regulate coagulation and vascular inflammation that develops in response to pathogens and cytokines. The renin-angiotensin system (KEGG ID: 4614) is reported repair leaky blood vessels and prevent microvascular dysfunction in sepsis (22). The peroxisome proliferator activated receptor signaling pathway (KEGG ID: 3320) is involved in regulation of coagulation caused by sepsis (23). This indicates that these attractors are associated with sepsis, and may be important in the occurrence and development of sepsis.

Distribution outcomes in SVM. SVM is widely used in computational biology due to its high accuracy, powerful ability to deal with high dimensional and large databases, and flexibility in modeling diverse sources of data (24). These characteristics are useful for the analysis performed in the current study. Analysis of clinical data can be problematic, as the available data is usually multi-dimensional and heterogeneous, with a large number of features and different types of data (25).

SVM was used in the current study to screen the efficiency of 277 attractors and detect their distribution outcomes, which were mainly consistent with the pathways from KEGG analysis, with high accuracy of 86.00-96.00. This indicates the classification analysis is helpful to precise diagnosis and prognosis of diseases. Therefore, screened attractors could take the place of vast pathways from KEGG to perform individualize therapy in pediatric sepsis. However, the distribution outcomes of 277 attractors have not been verified in abundance. We suggest that screened attractors ought to be checked in plenty of researches and receive further modification.

Differentially expressed pathways in individual. Pathway analysis has become the first choice for extracting and explaining the underlying biology using high-throughput molecular measurements (26). Identification of differentially expressed pathways in individuals is important for understanding disease mechanisms and for the future application of custom therapeutic decisions. Current pathway analysis methods are not suitable for identifying the pathway aberrance that may occur in an individual sample (8). Therefore, iPAS was used in the present study for the personalized identification of attractors, making use of accumulated normal data.

iPAS using ANS in pediatric sepsis is an important innovation. Ahn *et al* (8) demonstrated that the Average Z method efficiently identified pathway aberrance and further revealed clinical importance. It had the best statistical power when identifying a previously known survival-associated pathway and the best averaged validation rate. As data repositories are developing rapidly, it is expected that ANS data will be available for more and more diseases in the future.

In the present study, the iPAS was used to calculate pathway statistics for every attractor and the Average Z method was selected as a modification of existing pathway analysis methods. There were 77 attractors with P<0.01 following adjustment of the FDR, which were identified as the most significant pathways. The majority of them were clustered in P6 and P7 (Fig. 3). Therefore, these screened attractors were efficient in distinguishing differences in individual pediatric sepsis samples. This may provide a broader insight for personalized medicine (27).

Based on the results of the present study, it is concluded that a novel analysis procedure identified dysregulated attractors in individual pediatric sepsis samples. Attractors can be markers for the identification of pediatric sepsis. We hope the constructed process can help in the personalized interpretation of data and can be efficient in the upcoming era of personalized medicine.

References

1. Oliveira C, Nogueira de Sá FR, Oliveira DS, Gottschald AF, Moura JD, Shibata AR, Troster EJ, Vaz FA and Carcillo JA: Time- and fluid-sensitive resuscitation for hemodynamic support of children in septic shock: Barriers to the implementation of the American College of Critical Care Medicine/Pediatric Advanced Life Support Guidelines in a pediatric intensive care unit in a developing world. Pediatr Emerg Care 24: 810-815, 2008.

2. Wiens MO, Kumbakumba E, Kissoon N, Ansermino JM, Ndamira A and Larson CP: Pediatric sepsis in the developing world: Challenges in defining sepsis and issues in post-discharge mortality. Clin Epidemiol 4: 319-325, 2012.
3. Kissoon N, Carcillo JA, Espinosa V, Argent A, Devictor D, Madden M, Singhi S, van der Voort E and Latour J; Global Sepsis Initiative Vanguard Center Contributors: World federation of pediatric intensive care and critical care societies: Global sepsis initiative. Pediatr Crit Care Med 12: 494-503, 2011.
4. Khilnani P, Singhi S, Lodha R, Santhanam I, Sachdev A, Chugh K, Jaishree M, Ranjit S, Ramachandran B, Ali U, *et al*: Pediatric sepsis guidelines: Summary for resource-limited countries. Indian J Crit Care Med 14: 41-52, 2010.
5. Chen ZH, Kim HP, Ryter SW and Choi AM: Identifying targets for COPD treatment through gene expression analyses. Int J Chron Obstruct Pulmon Dis 3: 359-370, 2008.
6. Kauffman S: A proposal for using the ensemble approach to understand genetic regulatory networks. J Theor Biol 230: 581-590, 2004.
7. Mar JC, Matigian NA, Quackenbush J and Wells CA: attract: A method for identifying core pathways that define cellular phenotypes. PLoS One 6: e25445, 2011.
8. Ahn T, Lee E, Huh N and Park T: Personalized identification of altered pathways in cancer using accumulated normal tissue data. Bioinformatics 30: i422-i429, 2014.
9. Parkinson H, Kapushesky M, Shojatalab M, Abeygunawardena N, Coulson R, Farne A, Holloway E, Kolesnykov N, Lilja P, Lukk M, *et al*: ArrayExpress-a public database of microarray experiments and gene expression profiles. Nucleic Acids Res 35 (Database Issue): D747-D750, 2007.
10. Wong HR, Cvijanovich N, Allen GL, Lin R, Anas N, Meyer K, Freishtat RJ, Monaco M, Odoms K, Sakthivel B, *et al*: Genomic expression profiling across the pediatric systemic inflammatory response syndrome, sepsis, and septic shock spectrum. Crit Care Med 37: 1558-1566, 2009.
11. Gautier L, Cope L, Bolstad BM and Irizarry RA: Affy-analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20: 307-315, 2004.
12. Smyth GK, Gentleman R, Carey V, Huber W, Irizarry RA and Dudoit S: Limma: Linear models for microarray data. In: Bioinformatics and computational biology solutions using R and bioconductor. Springer, New York, pp397-420, 2005.
13. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U and Speed TP: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4: 249-264, 2003.
14. Holder D, Raubertas RF, Pikounis VB, Svetnik V and Soper K: Statistical analysis of high density oligonucleotide arrays: A SAFER approach. In: Proceedings of the ASA Annual Meeting. American Sociological Association, Atlanta, GA, pp1-5, 2001.
15. Kanehisa M and Goto S: KEGG: Kyoto Encyclopedia of genes and genomes. Nucleic Acids Res 28: 27-30, 2000.
16. Vapnik VN: Statistical Learning Theory. Wiley, New York, NY, 1998.
17. Mazurowski MA, Habas PA, Zurada JM, Lo JY, Baker JA and Tourassi GD: Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. Neural Netw 21: 427-436, 2008.
18. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, *et al*: NCBI GEO: Archive for functional genomics data sets-update. Nucleic Acids Res 41 (Database Issue): D991-D995, 2013.
19. Bolstad BM, Irizarry RA, Astrand M and Speed TP: A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19: 185-193, 2003.
20. Benjamini Y and Hochberg Y: Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc 57: 289-300, 1995.
21. Takayanagi H, Sato K, Takaoka A and Taniguchi T: Interplay between interferon and other cytokine systems in bone metabolism. Immunol Rev 208: 181-193, 2005.
22. David S, Kümpers P, van Slyke P and Parikh SM: Mending leaky blood vessels: The angiopoietin-Tie2 pathway in sepsis. J Pharmacol Exp Ther 345: 2-6, 2013.
23. Di Paola R, Esposito E, Mazzon E, Genovese T, Muià C, Crisafulli C, Malleo G, Sessa E, Meli R and Cuzzocrea S: Absence of peroxisome proliferators-activated receptors (PPAR) alpha enhanced the multiple organ failure induced by zymosan. Shock 26: 477-484, 2006.
24. Ben-Hur A, Ong CS, Sonnenburg S, Schölkopf B and Rätsch G: Support vector machines and kernels for computational biology. PLoS Comput Biol 4: e1000173, 2008.
25. Vieira SM, Mendonça LF, Farinha GJ and Sousa JMC: Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients. Appl Soft Comput 13: 3494-3504, 2013.
26. Khatri P, Sirota M and Butte AJ: Ten years of pathway analysis: Current approaches and outstanding challenges. PLoS Comput Biol 8: e1002375, 2012.
27. Slattery ML, Herrick JS, Mullany LE, Gertz J and Wolff RK: Improved survival among colon cancer patients with increased differentially expressed pathways. BMC Med 13: 75, 2015.