Data in Brief

Gene expression profiling of mononuclear cells from patients with sepsis secondary to community-acquired pneumonia

Patricia Severino [a,*], Eliezer Silva [b], Giovana Lotici Baggio-Zappia [c], Milena Karina Colo Brunialti [c], Laura Alejandra Nucci [c], Otelo Rigato Junior [d], Ismael Dale Cotrim Guerreiro da Silva [e], Flavia Ribeiro Machado [f], Reinaldo Salomao [c]

[a] Center for Experimental Research, Instituto Israelita de Ensino e Pesquisa, Hospital Israelita Albert Einstein, Sao Paulo, Brazil
[b] Intensive Care Unit, Hospital Israelita Albert Einstein, Sao Paulo, Brazil
[c] Division of Infectious Diseases, Escola Paulista de Medicina, Universidade Federal de Sao Paulo, Sao Paulo, Brazil
[d] Intensive Care Unit, Hospital Sirio Libanes, Sao Paulo, Brazil
[e] Department of Gynecology, Escola Paulista de Medicina, Universidade Federal de Sao Paulo, Sao Paulo, Brazil
[f] Intensive Care Unit, Hospital Sao Paulo, Universidade Federal de Sao Paulo, Sao Paulo, Brazil

ARTICLE INFO

ABSTRACT

Mechanisms governing the inflammatory response during sepsis involve crosstalk between diverse signaling pathways, but current knowledge provides an incomplete picture of the syndrome. Microarray-based expression profiling is a powerful approach for the investigation of complex clinical conditions such as sepsis. In this study, we investigated whole-genome expression profiles in mononuclear cells from septic patients admitted in intensive care units with community-acquired pneumonia. Blood samples were collected at the time of sepsis diagnosis and seven days later since we aimed to evaluate the role of biological processes or genes possibly involved in patient recovery. Here we provide a detailed description of the study design, including clinical information, experimental methods and procedures regarding data analysis. Metadata corresponding to microarray results deposited in the database Gene Expression Omnibus (GEO) under the accession number GSE48080 are also described in this report. Our dataset allows the identification of genes possibly associated with host defense to infection as well as gene expression patterns associated with patient outcome.

Specifications	
Organism/cell line/tissue	Homo sapiens
Sex	Male and female
Sequencer or array type	Agilent Whole Human Genome Microarray 4 × 44K arrays G4112F
Data format	Background-corrected signals
Experimental factors	Samples were peripheral blood mononuclear cells from patients with sepsis secondary to community-acquired pneumonia
Experimental features	Gene expression profiling comparing peripheral blood mononuclear cells from healthy volunteers and septic patients. Samples from septic patients were collected at two time points: D0 [within 48 h of the first occurrence of organ dysfunction indicative of severe sepsis or septic shock] and D7 [seven days after the first sample was collected]. Results compared healthy individuals with patients, and survivor and non-survivor septic patients.

(continued)

Specifications	
Consent	Written informed consent was obtained from all participants or, when necessary, from relatives before enrollment in the study protocol.
Sample source location	Intensive care units from Sao Paulo Hospital, Albert Einstein and Sirio Libanes Hospital, Sao Paulo, Brazil

Direct link to deposited data

Deposited data can be found at http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE48080.

Materials and methods

Experimental design

We investigated whole-genome gene expression profiles of peripheral blood mononuclear cells from survivors and non-survivors of sepsis

* Corresponding author.
E-mail address: patricia.severino@einstein.br (P. Severino).

as well as from healthy volunteers. Blood samples were collected at the time of sepsis diagnosis and seven days later, allowing us to evaluate the role of biological processes or genes that may be involved in patient recovery. Aiming to at least partially circumvent the heterogeneity of septic patient populations, we included only patients admitted with sepsis caused by community-acquired pneumonia. However, patients varied in SOFA and APACHEII scores.

Patients and healthy volunteers

A total of ten severe sepsis or septic shock patients and three healthy volunteers were included in this study. Five patients were admitted with severe sepsis and five with septic shock (Table 1). Sepsis is a systemic inflammatory response secondary to an infection and its mechanisms have been shown to be complex and dynamic [1,2]. Severe sepsis is sepsis complicated by acute organ dysfunction, and septic shock includes either hypotension that is refractory to fluid resuscitation or hyperlactatemia [3]. Severe sepsis may occur as a result of community-acquired and health care-associated infections but these particular conditions may require specific treatment approaches and clinical or molecular markers may differ [4]. Thus, as a means to circumvent variability in the results, all patients included in this study were adults with community-acquired, radiologically confirmed, pneumonia (CAP) as the primary source of infection. Reportedly, the most frequent etiological agent in CAP is *Streptococcus pneumonia* [5,6] but we have not had access to the etiological agent for all our patients and, therefore, could not include this data in the analysis. Patients were divided into two groups: patients that either succumbed to or survived their sepsis episode. Their age ranged from 25 to 92 years. Patients were enrolled within 48 h of the first occurrence of organ dysfunction indicative of severe sepsis. APACHE II scores (Acute Physiology and Chronic Health Evaluation II) [7] ranged from 7 to 23, and SOFA scores (Sequential Organ Failure Assessment score) [8] ranged from 2 to 11 at enrollment. Both systems classify severity of the disease and are usually applied within 24 h of admission of the patient in the ICU. We also included as a variable the administration of vasopressors (i.e. drugs that induce vasoconstriction, elevating mean arterial pressure when septic shock persists despite fluid resuscitation) since we assumed this could influence our gene expression results. Healthy controls, two females and one male, were 36, 58 and 84 years old, and were included in the study so we could evaluate gene expression patterns characterizing septic patients. Exclusion criteria were: individuals under 18 years old, patients with immunosuppressive therapy, AIDS or end stage chronic illness, and patients who had been submitted to any kind of experimental therapy. This study was approved by the ethics committees of the participating hospitals (Sao Paulo Hospital 1477/06, Albert Einstein Hospital 07/549, and Sirio Libanes Hospital 2006/27) and written

Table 1
Demographic data and outcomes from severe sepsis and septic shock patients.

Patients	Age [years]	Status at admission	Vasopressors	APACHE II	SOFA	Outcome
P107	82	Septic shock	Yes	23	6	Alive
P140	47	Severe sepsis	Yes	16	11	Alive
P143	71	Severe sepsis	No	10	3	Alive
P146	57	Septic shock	Yes	14	10	Dead
P217	83	Severe sepsis	No	14	4	Dead
P227	92	Septic shock	Yes	12	10	Dead
P229	32	Severe sepsis	No	7	7	Alive
P239	84	Septic shock	Yes	19	9	Dead
P254	58	Septic shock	Yes	23	9	Dead
P260	25	Severe sepsis	No	15	2	Alive

Acute Physiology and Chronic Health Evaluation II; SOFA: Sequential Organ Failure Assessment score; Vasopressors: drugs that induce vasoconstriction.

informed consent was obtained from all participants or, when necessary, from relatives before enrollment in the study protocol.

Isolation of peripheral mononuclear cells and RNA extraction

Fifty milliliters of blood was collected in sodium heparin-treated tubes (BD Biosciences, Franklin Lakes, NJ, USA) from healthy volunteers and septic patients. Samples from septic patients were collected within 48 h of the first occurrence of organ dysfunction indicative of severe sepsis or septic shock (D0) and seven days after the first sample was collected (D7). Peripheral blood mononuclear cells were obtained using the Ficoll gradient method (Ficoll-Paque PLUS, GE Healthcare Life Sciences, Uppsala, Sweden). Cells were frozen in fetal bovine serum (Gibco, Life Technologies, USA) with 10% dimethyl sulfoxide (Calbiochem, La Jolla, CA, USA) and stored in liquid nitrogen until required for RNA extraction. The standard cell concentration was 1×10^7 cells/mL. Total RNA was isolated from peripheral mononuclear cells using Illustra RNAspin Mini Kit (GE Healthcare Life Sciences). The quality and concentration of the RNA were determined using an RNA Nano Chip Kit and a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). RNA purity (ratio of absorbance at 260 nm and 280 nm) was accessed using the NanoVue spectrophotometer (GE Healthcare Life Sciences). Only samples presenting suitable RNA integrity number (RIN ≥ 7) and RNA purity ($A_{260/280}$ ~ 2) were used for gene expression profiling.

Gene expression profiling using DNA microarrays

Microarray analysis was performed using Agilent Whole Human Genome Microarray 4 × 44K arrays (design ID 014850). This array is manufactured using the 60-mer SurePrint Agilent technology and its format targets 19,596 Entrez Gene RNAs. For the gene list, annotation and probe sequences one should visit eArray application at http://www.genomics.agilent.com. Labeling of probes with cyanine 3 dye (Cy3), hybridization and washing procedures followed, strictly, the manufacturer's protocol (One Color Quick Amp Labeling Kit, Agilent Technologies). The arrays were scanned using a GenePix 4000B Microarray Scanner (Molecular Devices) using default parameters for Agilent 44K microarrays. Initial data analysis was performed using the Agilent Feature Extraction software (version 9.5). This software places the microarray grids, determines feature intensities and flags outlier pixels. The quality of the microarray data was assessed using the positive controls and RNA spike-ins. The gProcessedSignal (i.e. end result of standard Agilent Feature Extraction normalization and background correction procedures) from each array was loaded into the Partek Genomics Suite (v6.6), normalized between arrays using quantile normalization, and log transformed for further analyses. Principal Components Analysis (PCA) was used as an exploratory tool to identify major effects influencing data. One could say that the PCA plot reveals the structure of the data in a way that best explains its variance, allowing us to visualize the multivariate dataset [i.e. our gene expression values per sample] as a lower-dimensional picture. The final result can be described as a projection of the data, viewed from its most informative viewpoint. In our study this approach allowed for the identification of gene expression patterns associated with sepsis, when comparisons were made between septic patients and healthy volunteers, and with patient outcome when we compared survivors and non-survivors.

For subsequent statistical analysis we used the ANOVA implementation of Partek. The ANOVA model was defined by the experimental design and included variations due to volunteer group (sepsis, control), day of sample collection (D0, D7) and survival status (survivor, non-survivor). Since RNA extraction was not performed in the same day, we also used the analysis of variance to remove possible batch effects. Gene expression patterns related to biological function were assessed using Gene Ontology (GO) term enrichment analysis and KEGG pathway mapping through DAVID Bioinformatics Resources 6.7

(http://david.abcc.ncifcrf.gov). For the identification of differentially expressed genes we took in consideration the characteristics of each experiment for determining a cut-off and, therefore, the cut-off varied depending on the studied group. The raw microarray data (background-corrected signal) can be assessed at Gene Expression Omnibus (GEO accession GSE48080).

Competing interests

The authors declare that there are no competing interests.

Funding

References

[1] R.C. Bone, R.A. Balk, F.B. Cerra, R.P. Dellinger, A.M. Fein, W.A. Knaus, R.M. Schein, W.J. Sibbald, Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. Chest 101 (1992) 1644–1655.

[2] R. Salomao, M.K.C. Brunialti, M.M. Rapozo, G.L. Baggio-Zappia, C. Galanos C, M. Freudenberg, Bacterial sensing, cell signaling, and modulation of the immune response during sepsis. Shock 38 (2012) 227–242.

[3] D.C. Angus, T. van der Poll, Severe sepsis and septic shock. N. Engl. J. Med. 369 (2013) 840–851.

[4] B. Prieto, F.V. Alvarez-Menéndez, Diagnosis of infection in critical care. Adv. Clin. Chem. 65 (2014) 43–67.

[5] I. Köksal, T. Ozlü, O. Bayraktar, G. Ylmaz, Y. Bülbül, F. Oztuna, R. Caylan, K. Aydın, N. Sucu, TUCAP Study Group, Etiological agents of community-acquired pneumonia in adult patients in Turkey; a multicentric, cross-sectional study. Tuberk Toraks 58 (2010) 119–127.

[6] C. Jokinen, L. Heiskanen, H. Juvonen, S. Kallinen, M. Kleemola, M. Koskela, M. Leinonen, P.R. Rönnberg, P. Saikku, M. Stén, A. Tarkiainen, H. Tukiainen, K. Pyörälä, P.H. Mäkelä, Microbial etiology of community-acquired pneumonia in the adult population of 4 municipalities in eastern Finland. Clin. Infect. Dis. 32 (2001) 1141–1154.

[7] W.A. Knaus, E.A. Draper, D.P. Wagner, J.E. Zimmerman, APACHE II: a severity of disease classification system. Crit. Care Med. 13 (1985) 818–829.

[8] J.L. Vincent, A. de Mendonça, F. Cantraine, R. Moreno, J. Takala, P.M. Suter, C.L. Sprung, F. Colardyn, S. Blecher, Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Crit. Care Med. 26 (1998) 1793–1800.