Genome-Wide Expression Profiles in Very Low Birth Weight Infants With Neonatal Sepsis

What’s Known on This Subject: Rapid and reliable tools for the diagnosis of neonatal sepsis are still unavailable. No single biomarker studied has yielded conclusive results. Genome-wide expression profiles (GWEPs) have been successfully determined for the diagnosis of sepsis in pediatric and adult populations.

What This Study Adds: GWEPs are described for the first time in very low birth weight infants with proven bacterial sepsis. Our results suggest that GWEPs could be used for early discrimination of septic newborn versus nonseptic infants.

Abstract

Background: Bacterial sepsis is associated with high morbidity and mortality in preterm infants. However, diagnosis of sepsis and identification of the causative agent remains challenging. Our aim was to determine genome-wide expression profiles of very low birth weight (VLBW) infants with and without bacterial sepsis and assess differences.

Methods: This was a prospective observational double-cohort study conducted in VLBW (<1500 g) infants with culture-positive bacterial sepsis and non-septic matched controls. Blood samples were collected as soon as clinical signs of sepsis were identified and before antibiotics were initiated. Total RNA was processed for genome-wide expression analysis using Affymetrix gene arrays.

Results: During a 19-month period, 17 septic VLBW infants and 19 matched controls were enrolled. First, a three-dimensional unsupervised principal component analysis based on the entire genome (28,000 transcripts) identified 3 clusters of patients based on gene expression patterns: Gram-positive sepsis, Gram-negative sepsis, and noninfected control infants. Furthermore, these groups were confirmed by using analysis of variance, which identified a transcriptional signature of 554 genes. These genes had a significantly different expression among the groups. Of the 554 identified genes, 66 belonged to the tumor necrosis factor and 56 to cytokine signaling. The most significantly overexpressed pathways in septic neonates related with innate immune and inflammatory responses and were validated by real-time reverse transcription polymerase chain reaction.

Conclusions: Our preliminary results suggest that genome-wide expression profiles discriminate septic from nonseptic VLBW infants early in the neonatal period. Further studies are needed to confirm these findings. Pediatrics 2014;133:1–9
Bacterial sepsis in the neonatal period remains a relevant cause of morbidity and mortality in preterm infants. In very low birth weight (VLBW) infants, the incidence of early-onset sepsis varies from 1.5% to 1.9%, whereas late-onset sepsis rates are estimated to be ∼20% with mortality of ∼18%. Moreover, neonatal infections in preterm infants are associated with a higher risk of cerebral palsy at age 5 years.

Diagnosis of sepsis and identification of the causative bacteria still represent great challenges. Hence, clinical signs are nonspecific, and results of blood cultures are frequently delayed or yield inconclusive results. Factors such as low bacteremia, small blood inoculation volumes, and previous administration of antibiotics to the mother during labor or to the infant after birth pose additional difficulties.

A variety of diagnostic biomarkers have been used to identify neonates early in the course of the sepsis or to monitor response to therapy. However, no single biomarker has yield conclusive results. Factors such as low bacteremia, small blood inoculation volumes, and previous administration of antibiotics to the mother during labor or to the infant after birth pose additional difficulties.

The purpose of this study was to compare gene expression profiles in VLBW infants with confirmed bacterial sepsis and with paired nonseptic controls.

METHODS

Study Design and Patient Samples

This is a prospective, double-cohort, single-center, observational study. All VLBW infants (birth weight <1500 g) admitted to the University and Polytechnic Hospital La Fe from April 2011 to September 2012 were screened. Eligible patients were those with risk factors and/or clinical signs of sepsis. Risk factors were (1) maternal chorioamnionitis; (2) mothers with incompletely treated group B streptococcus infection (≥2 doses of ampicillin before labor); (3) mothers nontested for group B streptococcus with premature rupture of membranes for >18 hours, maternal fever during labor, or premature labor incompletely treated with antibiotics; (4) neonates with indwelling devices (<24 hours before symptoms) or surgery 72 hours before onset of symptoms. Sepsis was considered when ≥3 of the following were present: temperature instability (rectal temperature ≥38°C or ≤36°C); respiratory symptoms (distress, apnea, or cyanosis); cardiovascular symptoms including hypotension (blood pressure <5th percentile for age), tachycardia (heart rate >180/minute), bradycardia (heart rate <100/minute), or poor perfusion; neurological symptoms (clinical or electrical seizures, hypotonia, or lethargy); or gastrointestinal symptoms (vomiting, poor feeding or feeding intolerance, or abdominal distension).

Noninfected controls were recruited from hospitalized neonates without clinical signs of infection and paired 1:1 with cases based on gestational age, birth weight, gender, ethnicity, type of delivery, antenatal steroids, age, and clinical status. Exclusion criteria were chromosomal abnormalities, major congenital malformations, profound resuscitation, parental history of immunodeficiency, or congenital infections.

A database (Access 2010, Microsoft, Redmond, WA) was designed for the study. Data were retrieved and transferred to the database daily and crosschecked weekly. The local institutional review board approved the study, and all patients had informed consent signed by their parents or representatives.

Diagnosis of Sepsis

Blood cultures (BacT/Alert PF; BioMérieux, Durham, NC) were performed in patients with suspected sepsis before starting antibiotics. When a microorganism was isolated from blood and clinical signs and risk factors were present concomitantly, the patient was assigned to the sepsis-positive culture group. Two positive blood cultures were required for the diagnosis of coagulase-negative staphylococcal (CONS) sepsis. Cultures obtained from other sites were not considered for the diagnosis of sepsis.

RNA Isolation and Microarray Hybridization

Venous blood (0.5 mL) was obtained from eligible patients before starting antibiotics and from matched controls, mixed with 1 mL of RNA stabilizing solution (Tempus Blood RNA tubes, Applied Biosystems, Foster City, CA), and stored at −20°C until further processing. Briefly, total RNA was isolated from each sample using a MagMAX RNA isolation kit (Ambion/Applied Biosystems, 51 stems) according to the manufacturer’s instructions. RNA integrity was assessed by using the 2100 Bioanalyzer (Agilent, Palo Alto, CA), and stored at −20°C until further processing. Briefly, total RNA was isolated from each sample using a MagMAX RNA isolation kit (Ambion/Applied Biosystems, 51 stems) according to the manufacturer’s instructions. RNA integrity was assessed by using the 2100 Bioanalyzer (Agilent, Palo Alto, CA). Hybridization was performed if the RNA integrity number was ≥7. Microarray experiments were conducted according to the manufacturer’s instructions (Instituto Clínico Valencia; University of Medicine and Dentistry of the State of New York, Cooperstown, NY).
Clinical and Demographic Analysis

Clinical Data, Microarray and Statistical analysis

Clinical and Demographic Analysis

Categorical variables were compared using $\chi^2$ or Fisher’s exact test (2-tailed). Continuous variables were expressed as mean ± SD or medians with interquartile range depending on data distribution. Two-tailed Student’s t or Mann-Whitney U tests and analysis of variance (ANOVA) or Kruskal-Wallis were used to compare 2 or >2 groups as appropriate. Kolmogorov-Smirnov analysis was performed to test the normal distribution of the data. Data analysis was performed by using SPSS version 17.0 (SPSS Inc, Chicago, IL). Significance was considered for a $P \leq .05$.

Microarray

Analyses were performed by using 1 patient sample per Genechip. Files were captured using an Affymetrix GeneChip Scanner 3000 7G at 570-nm wavelength excitation. Data (.CEL files) were analyzed and statistically filtered by using Partek Genomic Suite 6.6 software (Partek Inc, St Louis, MO). Data were normalized with the robust multiple-array average (RMA) algorithm for gene array on core meta probe sets. One-way ANOVA was performed with the Partek Genomics Suite across all samples.

Data were analyzed by principal components analysis (PCA) to determine the significant sources of variability in the data sets. PCA reduces the complexity of high-dimensional data and simplifies the task of identifying patterns and sources of variability in a large data sets in a tridimensional fashion. The distance between any pair of points is related to the similarity between the 2 samples in high-dimensional space (in this case, each variable corresponded to a one-dimensional space). Samples in the plot close to each other are similar in a large number of variables, whereas samples far apart differ in a large number of variables. Next, significant differentially expressed genes between neonates with bacterial sepsis and controls were identified using a model ANOVA with a fold discovery rate (FDR) <0.05. Significant genes derived from the ANOVA analyses were visualized according to their expression levels in an unsupervised hierarchical clustering. The area under the receiver operator curve was calculated for the most relevant genes from a statistical and biological perspective. Finally, the selected differentially expressed genes were imported into Pathway Studio version 9 (Ariadne Genomics software, Elsevier Inc, Rockville, MD) to identify the molecular function and biological processes between bacterial sepsis and controls and to determine main master regulators.

Real Time Polymerase Chain Reaction Validation

The top significant differentially expressed transcripts were validated by using a single real-time reverse-transcription polymerase chain reaction (RT-PCR) using TaqMan Gene Expression Assays probes (Applied Biosystems, Foster City, CA), according to the manufacturer’s instructions.

RESULTS

Gene Expression Patterns in Bacterial Sepsis Compared With Controls

First, we analyzed all groups of patients simultaneously to identify gene signatures that helped to identify VLBW with sepsis based on transcriptional profiles. For this purpose, we performed a three-dimensional unsupervised PCA mean centering and scaling based on all genome (28 000 well-annotated genes) of 17 neonates with sepsis and 19 controls (Fig 1). PCA identified 3 main clusters according to gene expression: Gram-positive sepsis, Gram-negative sepsis, and controls. All patients with sepsis were clearly separated by PCA; however, 6 nonseptic controls were mixed with the Gram-positive and Gram-negative patients. Nevertheless GWEPs showed a good diagnostic capacity and were able to discriminate septic patients from nonseptic controls with an overall sensitivity of 100%, specificity of 68%, positive predictive value of 74%, and negative predictive value of 100%. One of the controls developed Enterobacter cloacae sepsis 1 week after enrollment; another control had a septic episode 1 month before sample collection, and a third had a CONS sepsis 2 weeks before enrollment. Thereafter, we applied statistical group comparisons using 1-way ANOVA between positive blood culture. Thus, 17 neonates with bacterial sepsis and 19 controls were suitable for further analysis. No significant demographic and clinical differences were observed between cases and controls (Table 1). A single pathogen was identified by blood culture in neonates with culture-positive sepsis: CON staphylococcus: $n = 10$ (59%); Escherichia coli: $n = 3$ (17%); Enterococcus faecalis: $n = 2$ (12%); Staphylococcus aureus: $n = 1$ (6%); and Morganella morganii: $n = 1$ (6%).
To understand the biological significance of the 554 differentially regulated genes we used Pathway Studio (Ariadne Genomics). The most representative biological processes identified were related to innate immunity (P-value 4.46 × 10^{-15}) and inflammatory response (P-value 1.24 × 10^{-7}). Table 2 shows the top 5 pathways based on P values and main master regulators. Overall, the tumor necrosis factor (TNF-α) network represented the most enriched network and contained the greater number of genes and the most significant ones in terms of P value. The area under the receiver operator curve was calculated for the most relevant genes from both statistical and biological perspective. These genes were validated by RT-PCR. The referred genes were: Matrix-metalloproteinase 8 (TNF8), CD177, CD 40 ligand, CCAAT enhancer binding protein β (CEBPB), and TNF-α. Of these, MMP8 and CD177 showed the greatest area under curve. MMP8: 0.91 (95% confidence interval: 0.79–1.0). CD177: 0.89 (95% confidence interval: 0.79–0.99) and were therefore considered the most relevant biomarkers.

**Correlation Between Gene Expression Profiles and RT-PCR**

The most significantly expressed genes from a statistical and biological perspective MMP8, CD177, CD 40 ligand, CEBPB, and TNF-α were validated by RT-PCR with P values ranging from <0.05 to .004 compared with nonsupercritical controls (Fig 3).

### DISCUSSION

In the current study, VLBW infants with culture-proven sepsis displayed gene expression profiles that were different from nonsupercritical controls. Moreover, patients with culture-proven sepsis overexpressed TNF-α network included differentially expressed cytokines that play an important role in sepsis. To our knowledge, this is the first study that has used GWEPs to evaluate the usefulness of gene expression profiles for the diagnosis of sepsis in VLBW infants. Moreover, GWEP were performed in only 0.5 mL of whole blood, which enhances applicability for the newborn period.

Neonatal sepsis is still a major cause of mortality and morbidity. Newborn and especially VLBW infants have a functionally immature immune system, which influences clinical presentation and response to sepsis. As a consequence, the clinical response to infection may often be subtle until it unexpectedly turns into a life-threatening situation. To reach a prompt and accurate diagnosis is rendered difficult. Remarkably, ancillary laboratory tests often exhibit ambiguous and variable results, and blood cultures have a low yield. Under these circumstances, reaching a prompt and accurate diagnosis is rendered difficult.
inconclusive results. Thus, in recent updated reviews, the need for additional "disease-specific" biomarkers has been emphasized. Microarray-based expression profiling provides the opportunity to gain a broader "picture" of such complex and heterogeneous diseases such as sepsis with the advantage of decreasing the investigator's bias and increase discovery capability as all genes are potentially interrogated. We found a significant overexpression of genes related to innate instead of acquired immune response corresponding to an immature response. Hence, the important role of anti-inflammatory master regulators supports an immature-type of response. Expression targets of interleukin (IL)-10 or IL-4 showed a significant overexpression although the specific genes for neither IL-4 nor IL-10 were present in the list of 554 significant genes. The immature immune system of VLBW infants undoubtedly influences the transcriptomic response to infection and may partially explain the differences observed in our cohort of preterm infants with sepsis compared with studies performed in adults and older children. Thus, the most significant biological processes expressed in septic neonates were related to innate immunity and inflammatory response. Specifically, the T-cell receptor and cytokines pathways were the most differentially expressed. This is in agreement with Wong's findings in a cohort of pediatric patients with sepsis. The most significantly overexpressed gene in VLBW neonates with sepsis compared with controls was CD177, also known as human granulocyte alloantigen NB1 or HNA-2a, which is heterogeneously expressed on neutrophils of 88% to 97% of healthy individuals. This gene has granulocyte-specific antigen function and is upregulated upon neutrophil activation by peptides derived from bacterial protein degradation, internalization upon cross-linking, and activation of the respiratory burst after antigen-antibody binding, suggesting a receptor function. However, little is known about CD177 function. Antibodies to CD177 can cause a variety of disorders, including alloimmune neonatal neutropenia, autoimmune neutropenia, immune neutropenia after bone marrow transplantation, drug-induced immune neutropenia, transfusion-related lung injury, essential thrombocythemia, and myeloproliferative diseases. Although it was first described as a neutrophil-specific antigen linked to neonatal alloimmune neutropenia, to our knowledge, this is the first time that this gene is reported as a biomarker for sepsis in neonates. In agreement with previous studies in children with septic shock, MMP8 was also one of the most expressed genes in VLBW infants with bacterial sepsis. Furthermore, in a murine model of sepsis, either genetic or pharmacologic inhibition of MMP8 activity.
significantly increased survival rates.\textsuperscript{37} MMPs play an important role in the remodeling of a number of tissues. These enzymes are the only ones capable of degrading the components of the extracellular matrix. MMP8, also known as collagenase-2, has as common substrates collagen I and III, \(\alpha_2\) macroglobulin, \(\alpha_1\)-proteinase inhibitor, Clq, fibrinogen, and substance P. It has been shown that MMPs are upregulated during infections, especially by Gram-negative bacteria, through LPS stimulation, suggesting a relevant role in the pathogenesis of endotoxemia. Remarkably, broad-spectrum MMP inhibitors confer protection against septic shock in both animal and in vitro models and may constitute an important therapeutic option in the future.\textsuperscript{38} A recent study using GWEPs in children with septic shock showed that gene response was age dependent.\textsuperscript{39} Hence, term neonates with sepsis showed lower expression of B-cell receptor, triggering receptor expressed on myeloid (TREM) cells, and nuclear factor (NF)-\(\kappa B\)-related genes compared with older age groups with septic shock. In our study, T-cell receptor signaling was also underexpressed, whereas NF-\(\kappa B\) was overexpressed. These differences could be partially explained by the lower gestational age of our patients and by the fact that we compared VLBW infants with sepsis versus noninfected VLBW controls, whereas Ng et al used other developmental-age groups with septic shock.\textsuperscript{29}

Our data also revealed that host immune response assessed by using gene expression patterns may be different in children infected with Gram-positive versus Gram-negative bacteria. This is in agreement with studies by Ramilo et al,\textsuperscript{18} Feezor et al,\textsuperscript{39} and Yu et al.\textsuperscript{40} In contrast, Tang et al\textsuperscript{41} did not find significant differences in critically ill adults. Nevertheless, these data need to be interpreted with caution because of the limited number of children with Gram-negative sepsis.

Our study has a number of limitations. The patients included represent a heterogeneous group of preterm infants exhibiting different clinical symptoms, microbial pathogens, and concomitant comorbidities intrinsic to prematurity. This concurrence of factors can explain the overlap shown in cluster 2. It should be underscored that controls were not healthy neonates but were VLBW infants admitted to the NICU. However, they were carefully selected based on clinical and

### TABLE 2

<table>
<thead>
<tr>
<th>Sepsis Versus Control Neonates</th>
<th>(P)</th>
<th>No. of Genes</th>
<th>Pathways</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell surface receptor linked signaling</td>
<td>8.07E-06</td>
<td>13</td>
<td>Cell surface receptor linked signaling</td>
</tr>
<tr>
<td>T-cell receptor signaling</td>
<td>1.08E-05</td>
<td>8</td>
<td>T-cell receptor signaling</td>
</tr>
<tr>
<td>Interferon (\gamma)-mediated signaling</td>
<td>2.32E-05</td>
<td>7</td>
<td>Interferon (\gamma)-mediated signaling</td>
</tr>
<tr>
<td>Negative regulation of insulin-receptor signaling</td>
<td>5.21E-05</td>
<td>4</td>
<td>Negative regulation of insulin-receptor signaling</td>
</tr>
<tr>
<td>Cytokine-mediated signaling</td>
<td>5.73E-05</td>
<td>11</td>
<td>Cytokine-mediated signaling</td>
</tr>
<tr>
<td>TNFRSF5 (\rightarrow) STAT signaling</td>
<td>0.0005</td>
<td>3</td>
<td>TNFRSF5 (\rightarrow) STAT signaling</td>
</tr>
<tr>
<td>Toll-like receptor 3 signaling</td>
<td>0.0009</td>
<td>5</td>
<td>Toll-like receptor 3 signaling</td>
</tr>
<tr>
<td>MYD88-independent Toll-like receptor signaling</td>
<td>0.001</td>
<td>5</td>
<td>MYD88-independent Toll-like receptor signaling</td>
</tr>
<tr>
<td>Toll-like receptor 1 signaling</td>
<td>0.001</td>
<td>5</td>
<td>Toll-like receptor 1 signaling</td>
</tr>
<tr>
<td>Negative regulation of smoothened signaling</td>
<td>0.001</td>
<td>3</td>
<td>Negative regulation of smoothened signaling</td>
</tr>
</tbody>
</table>

### Master regulators

- TNF
- Cytokine
- Nuclear factor \(\kappa\) B light chain enhancer of activated B cells
- Specificity protein 1
- IL-1 \(\beta\)
- Jun/Fos
- Phosphatidylinositol 3-kinase
- IL-1 family
- Interferon \(\gamma\)
- Mitogen-activated protein-kinases

STAT, signal transduction and activation of transcription; TNFRSF5 is the name of a member of the TNF superfamily gene.

\textsuperscript{*} The number of genes that belong to each gene list that are involved in each signaling or network.
Demographic characteristics comparable to patients with sepsis, emphasizing that differences in gene expression were due to infection and not to other confounding conditions.

Blood samples for expression analyses were drawn as soon as clinical signs were observed and before starting antibiotics to avoid changes in the immune response due to medication. Some authors have advocated the use of peripheral blood mononuclear cells to better account for the differences in white blood cell populations. Nevertheless, whole-blood gene expression has the potential to provide a comprehensive picture without missing biologically relevant expression signatures from cells that are not included in PBMCs, such as neutrophils, which play an important role in sepsis. RNA was extracted from whole blood, which provides a rapid stabilization of nucleic acids on collection, facilitates its use on clinical studies, and has been widely validated with the potential for fast-turnaround diagnosis. The lack of a validation cohort is a limitation that needs to be addressed in future studies. However, we used strict analytical tools and strategies such as FDR to correct for all multiple comparisons. In the current study, GWEPs were adequately performed in infected and noninfected preterm infants with only 0.5 mL of whole blood. This is of practical relevance because obtaining large blood volumes in neonates is especially difficult. Although the limited number of patients in this study precludes generalization of our results, we consider this study a starting point to perform strongly powered, prospective collaborative studies in the neonatal population.

CONCLUSIONS

Our study provides initial insight into the applicability of GWEP for the diagnosis of bacterial sepsis in preterm infants using a translational approach. Future studies are warranted to confirm these results and to identify specific candidate genes for the diagnosis and management of neonatal sepsis.

ACKNOWLEDGMENTS

We thank Octavio Ramilo, MD and Asunción Mejías, MD (Department of Pediatrics, Division of Infectious Diseases, Center for Vaccines and Immunity, Research Institute at Nationwide Children’s Hospital, Ohio State University College of Medicine, Columbus, OH) for their support and help. Their advice has been extremely useful in improving this article.

FIGURE 3

RT-PCR validation. Boxplots represent mean and error standard of mRNA expression (2-ΔΔCt) in controls and patients with bacterial sepsis for the following genes: CD177, MMP8, CD40 ligand, TNF, and CCAAT enhancer binding protein β (CEBPB). P value was calculated using Student’s t test.
REFERENCES


38. Vanlaere I, Libert C. Matrix metalloproteinases as drug targets in infections caused by gram-

(Continued from first page)
Genome-Wide Expression Profiles in Very Low Birth Weight Infants With Neonatal Sepsis
María Cernada, Eva Serna, Christine Bauert, María Carmen Collado, Gaspar Pérez-Martínez and Máximo Vento

Pediatrics; originally published online April 7, 2014;
DOI: 10.1542/peds.2013-2552

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/early/2014/04/02/peds.2013-2552