Association of proteome and metabolome signatures with severity in patients with community-acquired pneumonia

Manuela Gesell Salazar a,1, Sophie Neugebauer b,1, Tim Kacprowski b,c, Stephan Michalik k, Peter Ahnert d, Petra Creutz e, Maciej Rosolowski d, PROGRESS Study Group, Markus Lößler d, Michael Bauer f, Norbert Suttrop e, Michael Kiehnert f,1, Uwe Völker f,1

a Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, University Medicine Greifswald, Felix-Hausdorff-Str. 8, 17475 Greifswald, Germany
b Jena University Hospital, Institute of Clinical Chemistry and Laboratory Diagnostics and Integrated Biobank Jena (IBBJ), Am Klinikum 1, 07740 Jena, Germany
c Junior Research Group on Computational Systems Medicine, Chair of Experimental Bioinformatics, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising-Weihenstephan, Germany
d University of Leipzig, Institute for Medical Informatics, Statistics and Epidemiology (IMISE), Härtelstr. 16-18, 04107, Leipzig, Germany
e Department of Infectious Disease and Respiratory Medicine, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Campus Virchowklinikum, Augustenburgerplatz 1, 13353 Berlin, Germany
f Department of Anesthesiology and Intensive Care Medicine, Jena University Hospital, Am Klinikum 1, 07747 Jena, Germany

ARTICLE INFO

Keywords:
Community acquired pneumonia
Plasma
Serum
Proteomics
Lipid metabolism
LC-MS/MS

ABSTRACT

A combined OMICS screening approach of human plasma and serum was used to characterize protein and metabolome signatures displaying association to severity of Community-acquired pneumonia (CAP). 240 serum and BD P100 EDTA plasma samples from patients diagnosed with CAP, collected during the day of enrolment to the hospital, were analyzed by a metabolomic and proteomic approach, respectively. Disease severity of CAP patients was stratified using the Sequential Organ Failure Assessment (SOFA) score. Quantitative proteome and metabolome data, derived by LC-MS/MS, were associated to SOFA and specific parameters of SOFA using linear regression models adjusted for age, BMI, sex, smoking and technical variables.

Both proteome and metabolome profiling revealed remarkable strong changes in plasma and serum composition in relation to severity of CAP. Proteins and metabolites displaying SOFA associated levels are involved in immune response, particularly in processes of lipid metabolism. Proteins, which show an association to SOFA score, are involved in acute phase response, coagulation, complement activation and inflammation. Many of these metabolites and proteins displayed not only associations to SOFA, but also to parameters of SOFA score, which likely reflect the strong influence of lung-, liver-, kidney- and heart-dysfunction on the metabolome and proteome patterns.

Significance: Community-acquired pneumonia is the most frequent infection disease with high morbidity and mortality. So far, only few studies focused on the identification of proteins or metabolites associated to severity of CAP, often based on smaller sample sets. A screening for new diagnostic markers requires extensive sample collections in combination with high quality clinical data. To characterize the proteomic and metabolomics pattern associated to severity of CAP we performed a combined metabolomics and proteomics approach of serum and plasma sample from a multi-center clinical study focused on patients with CAP, requiring hospitalization. The results of this association study of omics data to the SOFA score enable not only an interpretation of changes in molecular patterns with severity of CAP but also an assignment of altered molecules to dysfunctions of respiratory, renal, coagulation, cardiovascular systems as well as liver.

https://doi.org/10.1016/j.jprot.2019.103627
Received 19 September 2019; Received in revised form 29 November 2019; Accepted 22 December 2019
Available online 30 December 2019
1874-3919/ © 2020 Elsevier B.V. All rights reserved.
1. Introduction

Community-acquired pneumonia (CAP) is a considerable cause of morbidity and mortality even in developed countries [1–4]. By definition, CAP is an acute lung infection which emerges in patients without former hospitalization. Incidence rate of CAP in Germany is 9.7 cases per 1000 persons per year and patients with CAP have a high in-hospital mortality of approximately 13% [5]. Furthermore, CAP, most commonly caused by an infection with Streptococcus pneumoniae, is one of the most frequent reasons for the development of sepsis [6]. Assessment of severity is essential for individual treatment of patients with CAP. Unfortunately, specific clinical markers for the determination of the severity of the disease have not been identified yet. However, sensitive and specific parameters are required to stratify disease risk, severity, as well as prognosis. In the multi-center PROGRESS study, time series data of > 1500 patients were used to evaluate different scoring systems for their ability to estimate CAP severity [7] revealing the Sequential Organ Failure Assessment (SOFA) score to be by far the best scoring system to detect patients in a severe state of CAP [8].

Omics profiling provides an alternative approach for identification of biomarkers for diagnosis, risk stratification, and prognosis of different diseases. One focus of these efforts is the screening for new diagnostic markers for the early detection of high-risk populations, and the individual monitoring and control of therapy. However, such efforts require extensive sample collections which are complemented by high quality phenotypic and clinical data. In the recent past, such extensive sample collections have been established but the number of proteomic and metabolomic studies that have been conducted in relation to CAP and the development of sepsis are still limited [9,10].

In the present study, we performed a comparative analysis of plasma and serum samples of a comprehensively phenotyped patient cohort of CAP patients using a combination of proteomic and metabolomic techniques to survey the metabolite and protein profiles of adults with different severity levels of CAP.

2. Methods

2.1. Study design

Samples were obtained from a prospective multi-centric longitudinal observational study (PROGRESS) of patients hospitalized after confirmation of a CAP [7]. Out of > 1500 patients we selected a subset of 240 patients with comprehensive phenotypic documentation. The selected cases represent the whole observed spectrum of CAP severity across SOFA scores (Supplemental Fig. S3), developing an uncomplicated or severe CAP and extending to CAP up to septic shock requiring specific treatment on ICU or resulting in death of the patient. For the current metabolome and proteome analyses, 240 serum and BD P100 EDTA plasma samples of patients were collected during the day of enrollment to the hospital. Patients were stratified with respect to the sequential organ failure assessment score (SOFA), clinical and microbiological characterization, along with information on respiratory support and medication.

2.2. Serum metabolome analysis

Metabolite concentrations were determined in serum according to the manufacturer’s protocol using the Absolute IDQ p180 kit (Biocrates Life Science AG, Innsbruck, Austria), as described previously [11].

2.3. Plasma proteome analysis

To deplete six highly abundant proteins in plasma, a multiaffinity chromatography using a Multiple Affinity Removal Column Human-6 (MARS-6, Agilent Technologies, CA, USA) was performed on a Proteomelab PF2D system (Beckman Coulter, Brea, CA, USA) in accordance with the manufacturer’s protocol. Proteins of the non-bound fraction were precipitated with trichloroacetic acid (final concentration 10%). Protein concentrations were determined using a Bradford protein assay [12] via a Varioskan Flash multimode reader (Thermo Fisher Scientific Inc., Waltham, MA, USA) in a 96-well-plate format after resuspension of protein pellets in 100 μl 8 M urea/2 M thiourea. Prior to mass spectrometric analysis, 11 μg of protein were diluted with ammonium bicarbonate (20 mM) to a final concentration of < 1 M urea, reduced with dithiothreitol (2.5 mM final concentration, 60 min at 60 °C) and alkylated with iodoacetamid (10 mM final concentration, 30 min at 37 °C in the dark) and proteolytically digested using trypsin over night at 37 °C (Promega, Madison, WI, USA; ratio trypsin to protein 1:24 w/w). Purification of peptide solutions was performed on OASIS PRIME HLB plates (Waters, Milford, MA, USA) via solid phase extraction. Eighteen to twenty tryptic peptide mixtures of the BD P100 EDTA plasma samples were analyzed by LC-MS/MS in batches on a reverse phase HPLC chromatography (Acquity UPLC system, Waters, Milford, MA, USA) coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) as described before [13]. To monitor technical variation and performance of instruments, a plasma pool consisting of 10 healthy subjects was analyzed with each batch using the same workflow. To constantly monitor the reproducibility of LC-MS/MS analysis, initially a plasma pool was prepared and digested before the start of the analysis of any patient samples and stored as lyophilized aliquots. Aliquots of this pre-made global standard were reconstituted in 0.1% acetic acid/2% acetonitrile and measured as the first sample of each batch (global standard). To monitor differences in the whole sample processing workflow, an aliquot of the same plasma pool was digested, desalted and analyzed with patient samples of each individual batch as the second sample of each batch (batch standard).

Preprocessing of LC-MS/MS data was performed using the Refiner MS software version 7.6.6 (GenData, Basel, Switzerland) with the following steps: chemical noise removal, retention time alignment across all samples, feature extraction and isotope group clustering. Using Mascot Server (rel. 2.4), data was searched against a Uniprot/Swissprot database limited to human entries (version 2016.06), search parameters are shown in Supplemental Table S1. For relative quantification, peptides identified as unique were considered with the rank of 1 by setting the FDR to 1% for annotated peptides after identification. The three most intense peptides for every protein were summed to the protein intensity.

2.4. Association analyses of metabolites and proteins to SOFA and selected parameters of the SOFA Score (sub-SOFA scores)

Due to the large number of samples, sample preparation and LC-MS/MS analyses were performed in batches. Maintenance of the mass spectrometer was necessary between the measurement of different batches, which may trigger higher technical variance in protein intensities. To monitor the performance of HPLC and mass spectrometer and adjust for technical variation if necessary, fully processed aliquots of a plasma pool consisting of plasma samples of 10 healthy subjects was repeatedly analyzed with each batch (global standard). A second aliquot of the same plasma pool was fully processed with each individual batch to monitor technical variance during the entire workflow (batch standard).

A Principal Component analysis (PCA) (Supplemental Fig. S1) illustrates the observed effects. In general, the batch and global standards clustered closely together, indicating good reproducibility even in this larger samples series. To minimize the impact of technical factors onto quantification, a median-median-normalization of the data and an adjustment for technical variance was performed. The technical variation was captured in principle component 2 for the proteome data, and in principle components 6 and 7 for the metabolite data. Hence, for analysis of the correlation of protein levels with SOFA and parameters of
SOFA, technically adjusted data corresponding to the residuals of the linear models ‘protein ~ PC2’ and ‘metabolite ~ PC6 + PC7’, respectively, and calculated with the lm() function of the ‘base’/‘stats’ package in R v3.5.3 were used. The PCA itself has been calculated in Python using the scikit-learn module v0.21.0.

Only metabolites and proteins present in at least 40% of all samples and proteins with a unique cluster ID were used for any further analyses. Association analyses were conducted using linear regression models. Within the PROGRESS study, it was decided that molecular analyses should be adjusted for confounders which were associated with at least one of the relevant outcome measures (e.g. SOFA and parameters used to construct the SOFA) in the entire study population in multivariate analysis. This identified age, BMI, sex, and smoking as confounders. Therefore, in the linear regression model proteome and metabolome data were adjusted for age, body mass index (BMI), sex, and smoking (years). Supplemental Fig. S2 shows the final completely adjusted protein data as residual plot of the linear model ‘protein ~ PC2 + age + BMI + sex + smoking’. Linear regression models were calculated in R using the ‘stats’/‘base’ package version 3.5.3 [14] and corresponding figures were generated using the tidyverse (version 1.2.1) [15], pheatmap (version 1.0.12) [16] packages or the Python package ‘seaborn’ v0.9.0 [https://seaborn.pydata.org].

Adjustment for technical variation was done by including principal component (PC) two or six and seven of the respective normalized data into the regression models for proteome and metabolome data, respectively. These PCs were very strongly associated with analysis batches and order of measurement albeit only very weakly with the SOFA score. Hence, the final models were of the forms ‘protein ~ SOFA + age + sex + BMI + smoking + PC2’ and ‘metabolite ~ SOFA + age + sex + BMI + smoking + PC6 + PC7’ for protein and metabolite level associations, respectively. Significant associations of proteins and metabolites with SOFA or parameters of SOFA were considered after applying Benjamini-Hochberg multiple testing correction if \(q \leq 0.05\).

2.5. Ingenuity pathway analysis

Ingenuity pathway analysis (Content version: 33559992, Ingenuity Systems, Redwood City, CA, USA) was performed to interpret proteomics data based on GO annotations [17].

3. Results

3.1. Characteristics of study cohort

For metabolome and proteome analyses, samples of 101 female and 139 male patients who were recruited at 40 different PROGRESS study sites were used. Age of patients ranged from 18 to 94 years (mean of 63.1 years) and BMI from 16.9 to 52.9 kg/m\(^2\) (mean of 27.1 kg/m\(^2\)). Patients were hospitalized due to confirmed CAP [7]. SOFA was obtained to estimate the severity of disease which is based on the laboratory variables displayed in Table 1 indicating the failure of the following six organ systems: respiratory system, neurological system, cardiovascular system, renal system, liver and coagulation [18]. SOFA scores ranged from 0 to 19 with a mean of 3.8. According to the sepsis definition [18,19], 38 of our 240 studied patients did not have sepsis, i.e. they had a SOFA score below 2 (Supplemental Fig. S3). In conclusion, we consider sepsis as SOFA ≥ 2, assuming that SOFA = 0 before the current episode of CAP. Most patients in our study displayed symptoms of sepsis of varying severity caused by CAP.

In correlation analyses of phenotypes, using a rank-based Kendall’s tau approach, correlations between SOFA and the six clinical parameters of SOFA oxygenation index, bilirubin, creatinine, mean arterial pressure (MAP), Glasgow Coma Scale (GCS) and thrombocytes were found, as would be expected due to the definition of the SOFA score. Furthermore, an association of SOFA to death of patients (\(T\) (tau correlation coefficient) = 0.32, \(p\) (p-value of Kendall’s tau rank correlation test) = 1.73E8) could be shown for the study cohort (Supplemental Fig. S4). The analyses also indicate a weak correlation of SOFA score to C reactive protein (CRP, \(T = 0.2, p = 1.22E5\)) and a significant correlation of SOFA score to procalcitonin (PCT, \(T = 0.39, p = 2.7E17\)). Both acute phase proteins show a clear increase during inflammatory processes [20] and are markers of systemic inflammation and bacterial infection, respectively.

3.2. Associations of plasma protein levels with SOFA and individual SOFA parameters

In total, LC-MS/MS analysis revealed 2570 peptides which could be assigned to 492 proteins. The number of identified proteins varied from 288 to 447 proteins among the different subjects, 253 proteins (51%) were identified in all analyzed samples. For association analysis, 367 proteins present in at least 40% of all samples and with unique cluster ID were considered for further analyses, 184 proteins were quantified with more than one unique peptide. Among the 367 proteins, 247 proteins could be characterized as secreted proteins, 96 as leakage proteins and 10 as components of immunoglobulin complexes (Fig. 1), according to GO annotation. The remaining 10 proteins could not be assigned to any of these groups.

Using linear regression, associations to the SOFA score and six variables of SOFA were investigated. Protein intensities were adjusted for technical variation, age, BMI, sex and smoking (years). After applying Benjamini-Hochberg correction (threshold \(q \leq 0.05\)), a large number of plasma proteins (n = 126) was still associated with the SOFA score (Supplemental Table S2). In the linear regression model for association to SOFA, the lowest \(q\)-values were recorded for pancreatic ribonuclease (RNase1, positive association to SOFA, \(q = 2.33E-08\)), transmembrane protease serine 6 (TMS6, negative association to SOFA, \(q = 5.55E-08\)) and hyaluronan-binding protein (HABP2, negative association to SOFA, \(q = 6.42E-07\)).

Ingenuity pathway analysis of SOFA associated proteins indicated acute phase response signaling and LXR/RXR, FXR/RXR and PPARα/PPXR activation as top 3 canonical pathways, processes which are involved in regulation of lipid metabolism and inflammation.

The results of this linear regression analysis for associations of protein levels with SOFA score are represented in Fig. 1 and Supplemental Table S3. In Fig. 1, the y axis shows the slope of regression and proteins were grouped according to their biological function. Benjamini-Hochberg multiple testing correction was applied and statistically significant data for the association analysis (\(q\)-value ≤ 0.05) are highlighted. In total, 26 SOFA associated proteins were annotated as immune response, including complement component C8 beta chain (C8B, positive association to SOFA), histidin-rich glycoprotein (HRG, negative association to SOFA) and promyelocytic leukemia protein (PML, positive association to SOFA) with lowest \(q\)-values (Supplemental Fig. S5). In detail, 10 proteins associated to SOFA score (threshold \(q \leq 0.05\)) were annotated as acute phase response, 15 as complement activation and 13 as defense response (Supplemental Fig. S6). Regarding the acute phase response, proteins with the lowest \(q\)-values and a negative association to SOFA include alpha-2-HS-glycoprotein (FETUA) and serum amyloidA-4 protein (SAA4). Alpha-1-antichymotrypsin (AECT), alpha-1-antitrypsin (A1AT), C-reactive protein (CRP), lipopolysaccharide-binding protein (LBPR) were positively associated with SOFA (Supplemental Fig. S6). The above-mentioned proteins HRG, CRP, LBP and PML are also involved in defense response mechanisms (Supplemental Fig. S6). Proteins with very low \(q\)-values, like complement component C8 (C8B), mannann-binding lectin serine protease 1 and 2 (MASP1, MASP2, both with positive associations to SOFA), complement 3 (C3, negative association), vitronectin (VTNC, positive association), should be mentioned to be involved in complement activation (Supplemental Fig. S6). In total, 21 SOFA associated proteins were annotated as inflammatory proteins (Supplemental Fig. S6).
SS), for instance pigment epithelium-derived factor (PEDF, positive association), apolipoprotein A-1 (APOA1, negative association), AACT, CRP, LBP.

In general, nine apolipoproteins could be found in the linear regression model to be associated with SOFA (Fig. 2). Higher abundances of APOB and APOE and lower abundances of APOA1, APOA2, APOD, APOF, APOM, serum amyloid A proteins (SAA2, SAA4) were associated with higher SOFA score values. Furthermore, lipoprotein associated proteins like transthyretin (TTHY) and serotransferrin (TRFE) showed decreased abundances with higher SOFA score values.

Regarding the specific SOFA parameters, only small numbers of associated proteins were obtained in the linear regression analysis for mean arterial pressure (MAP, \( n = 2 \)), Glasgow Coma Scale (GCS, \( n = 1 \)) and thrombocytes (\( n = 7 \)), which are indicators for failure of the

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
<th>Mean + SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>18–94</td>
<td>63.1 ± 16.3</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>16.9–52.9</td>
<td>27.1 ± 6.2</td>
</tr>
<tr>
<td>Smoking (years)</td>
<td>0–60</td>
<td>18.4 ± 17.9</td>
</tr>
<tr>
<td>Sex</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Death at all</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>Death 28 days</td>
<td>10</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 1
Clinical characteristics of CAP patients of a subset (\( n = 240 \)) of the prospective multi-centric longitudinal observational study (PROGRESS). Data are presented as range and mean ± standard deviation.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Category (organ system)</th>
<th>range</th>
<th>Mean + SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOFA</td>
<td></td>
<td>0–19</td>
<td>3.8 ± 2.9</td>
</tr>
<tr>
<td>Bilirubin (mol/l)</td>
<td>Liver</td>
<td>3.4–71</td>
<td>15.6 ± 12.7</td>
</tr>
<tr>
<td>Creatinine (mol/l)</td>
<td>Renal</td>
<td>47–1034.3</td>
<td>120.9 ± 98.3</td>
</tr>
<tr>
<td>Oxygen index (kPa)</td>
<td>Respiratory</td>
<td>6.2–92.06</td>
<td>33.4 ± 14.2</td>
</tr>
<tr>
<td>MAP (kPa)</td>
<td>Cardiovascular</td>
<td>5.1–16.0</td>
<td>10.8 ± 2.0</td>
</tr>
<tr>
<td>Glasgow Coma Scale</td>
<td>Neurological</td>
<td>3–15</td>
<td>14.7 ± 1.6</td>
</tr>
<tr>
<td>Thrombocytes (( 10^9/l ))</td>
<td>Coagulation</td>
<td>2.0E+05–7.0E+05</td>
<td>2.3E+05 ± 1.0E+05</td>
</tr>
<tr>
<td>CRP (mg/l)</td>
<td></td>
<td>2.6–641.1</td>
<td>190.8 ± 111.1</td>
</tr>
<tr>
<td>PCT (mg/l)</td>
<td></td>
<td>0.02–195.2</td>
<td>7.6 ± 22.1</td>
</tr>
</tbody>
</table>

Fig. 1. Association of plasma proteins to SOFA score. Proteins are plotted showing the slope of linear regression model on the y axis and grouped by biological function according to GO categories. GO location of proteins is displayed by symbols. Opaque symbols represent proteins with no statistical significant association to SOFA after multiple test correction (threshold: \( q\)-value \( ≤ 0.05 \)).
cardiovascular, neurological and coagulation system, respectively (Supplemental Tables S2-S3). A much higher number of proteins was associated to oxygenation index ($n = 87$), creatinine ($n = 78$) and bilirubin levels ($n = 73$), indicating their strong influence on the plasma proteome due to dysfunction of respiratory, renal organ system and liver (Supplemental Table S2). For association with oxygenation index, the lowest $q$-values were observed for proteins involved in immune response like CO8B, integrin alpha-L protein (ITAL), PML, LBP or in inflammatory processes like for instance ITAL, A1AT, AACT (all negatively associated to oxygenation index, Supplemental Fig. S5). Furthermore, proteins like pancreatic ribonuclease (RNAS1, negative association) and glutathione peroxidase 3 (GPX3, positive association), were found to be altered in relation to plasma creatinine levels. Supplemental Figs. S5 to S8 show associations of proteins to SOFA and to each specific variable of SOFA. The lowest $q$-values were observed for two proteins positively associated to bilirubin levels: LBP and pigment epithelium-derived factor (PEDF); and for three proteins which showed a negative association to bilirubin levels: TMP86, plasminogen (PLMN) and clusterin (CLUS).

### 3.3. Association of serum metabolite levels with SOFA and individual SOFA parameters

183 out of 188 metabolites included in the Absolute IDQ p180 kit (Biocrates Life Science AG, Innsbruck, Austria), including 40 acylcarnitines, 21 amino acids, 16 biogenic amines, 14 lysophosphatidylcholines (LPCs), 76 phosphatidylcholines (PCs), 15 sphingolipids, 1 sugar, could be detected. Analyses of the associations between metabolites and SOFA or parameters of SOFA subscores were performed using linear regression models, metabolomics data were adjusted by technical variance, age, BMI, sex and smoking (years).

Metabolite level associations to the SOFA score according to a linear regression model are shown in Fig. 3 and Supplemental Table S4. In Fig. 3, the $y$-axis shows the slope of regression and metabolites are grouped by compound classes defined by the Absolute IDQ p180 kit (Biocrates Life Science AG, Innsbruck, Austria). Significant associations were detected for 108 metabolites ($q \leq 0.05$, statistically significant results are highlighted in Fig. 3) including 19 acylcarnitines, 5 amino acids, 7 biogenic amines, 13 LPCs, 51 PCs and 13 sphingolipids.

Short- (C2-C8) and medium-chain acylcarnitines (C6-C10), as classified by Ikeda [21], showed increased concentrations with higher SOFA score values, while long-chain acylcarnitines (C14-C22) decreased (Supplemental Fig. S9). In contrast to methionine, the amino acids aspartic acid, glutamic acid, glycine, serine and tryptophan decreased with increasing SOFA score (Supplemental Fig. S10). Biogenic amines creatinine, kynurenine, putrescine, symmetric dimethylarginine (SDMA) as well as total dimethylarginine increased and serotonin and taurine levels decreased with higher SOFA score values. In general, lower concentrations of LPCs, PCs and sphingolipids were associated with a higher SOFA score with the exception of the phosphatidylcholines PCaaC32:0, PCaaC32:1, PCaaC34:1 and PCaeC36:0.

The above-mentioned six physiological SOFA parameters describing respiratory-, cardiovascular-, liver- and renal-function as well as coagulation and neurological status were significantly associated with 79, 78, 42, 29, 21 and 14 metabolites, respectively (Supplemental Table S2). Disease-related changes in acylcarnitine concentrations are mostly correlated with an increased creatinine level and, thus, with renal dysfunction (Supplemental Fig. S9). With decreasing oxygenation ratio and MAP, higher concentrations of short- and medium-chain acylcarnitines and lower concentrations of long-chain acylcarnitines could be detected.

Amino acids were related to all parameters included in the SOFA score (Supplemental Fig. S10). Low concentrations of aspartic acid were associated with low oxygenation ratios, MAP, platelet count and high bilirubin levels. Poor oxygenation and liver dysfunction were associated with decreased glutamic acid concentrations. Reduced platelet counts were associated with low histamine levels and a poor neurological status with high isoleucine and methionine concentrations. Decreased serine concentrations were related to low MAP and liver dysfunction and low tryptophan levels with poor oxygenation, low MAP and renal dysfunction. Alterations in biogenic amine concentrations were related with renal dysfunction and failure in coagulation. Of course, the endogenous marker for renal function SDMA was positively correlated with creatinine levels. Higher kynurenine concentrations were associated with low MAP, platelet counts and high creatinine levels. Poor oxygenation and
renal dysfunction were related with increased putrescine concentrations. Low spermidine and spermine concentrations were associated with poor coagulation.

Lower concentrations of LPCs were strongly associated with a decreased oxygenation ratio, MAP, platelet count and increased bilirubin levels (Supplemental Fig. S9). PCs and sphingolipids were strongly correlated with the respiratory function, the cardiovascular system and the liver function. Decreased PC and sphingolipid levels were related with poor oxygenation, low MAP and high concentrations of bilirubin (Supplemental Fig. S9).

4. Discussion

From a global perspective, CAP is one of the leading acute infectious diseases causing sepsis and death. The necessary hospitalizations lead to a considerable economic burden [1-4]. We used the phenotypic information of a well characterized cohort of 240 CAP patients to perform association studies of complex proteomics and metabolomics data to SOFA. A previous study has shown that the SOFA score can support the prediction of survival in patients with CAP associated sepsis [22]. Only few studies attempted to screen for markers of pneumonia on a proteomics [10] or metabolomics level [23,24]. A comparative study of a comprehensive cohort of CAP patients using a combination of both techniques has not been reported, so far. In relation to SOFA score we have observed tremendous changes in proteome as well as in metabolome profiles. The association analyses of the current study enable a precise specification of metabolome and proteome signatures not only in relation to severity of the disease but also to the extent of specific organ dysfunctions. All observed changes are related to CAP since all patients enrolled in PROGRESS are confirmed CAP patients. However, we cannot proclaim that observed changes can only be observed in CAP and not in septic patients with a different underlying etiology, like peritoneal sepsis.

In our linear regression models, e.g., two interesting proteins (RNAS1, HABP2) with strong association to SOFA (q < 5E07) could be identified. RNAS1 showed positive association to severity of CAP, negative association to oxygenation index and positive association to creatinine levels. This indicates dysfunction of the respiratory and renal system. The so-called pancreatic ribonuclease, which can be detected in body fluids like plasma and urine as well as in different tissue types, plays an important role in innate immune response due to its ability to induce the maturation of dendritic cells [25]. The other protein, extracellular serine protease HABP2, was negatively associated to severity of pneumonia and positively associated to oxygenation index. HABP2 was described to be related to lung diseases like LPS-induced lung injury, ventilator-induced lung injury (VILI), acute respiratory distress syndrome (ARDS) and pulmonary vascular leakiness (summarized in [26]). Previous studies have established that HABP2 can contribute to an increase of the vascular permeability [27].

The analysis of the GO annotation of the SOFA-associated proteins revealed immune response related processes, particularly acute phase response, coagulation, complement activation and inflammation as most important categories. Functional assignment of SOFA-associated proteins could be confirmed by ingenuity pathway analysis, which

![Fig. 3. Association of metabolites to SOFA score.](image)

Metabolites were plotted showing the slope of linear regression on the y axis and grouped by metabolic pathways with central biological relevance. Opaque symbols represent metabolites with no statistical significant association to SOFA after multiple testing correction (threshold: q-value ≤ 0.05).
showed an enrichment in canonical pathways like acute phase response signaling and LXR/RXR, FXR/RXR and PPARα/PXRα activation.

Our findings in the Ingenuity Pathway Analyses are consistent with formerly reported proteomics results. Cao et al. [9] described LXR/RXR activation and acute phase response signaling as most represented pathways for altered proteins between CAP patients with and without severe sepsis. LXR/RXR activation processes are involved in regulation of lipid metabolism, especially cholesterol absorption, efflux, transport and excretion and play an important role in modulation of immune response and inflammation [28,29].

The plasma proteome is composed of proteins originating from surrounding cells and tissues and characterized by its high dynamic range and variability [30]. Thus, proteins identified in plasma can be assigned to various biological processes and pathways in different organs. Proteins which play an important role in the above mentioned biological processes involved in the CAP disease progression can be negatively or positively associated to SOFA score or parameters of SOFA.

The initiation of pneumonia is caused by microbial pathogens like Streptococcus pneumoniae followed by several physiological reactions in different cell types and organ systems, starting with an innate defense response. In our analysis, we found histidin-rich glycoprotein (HRG), an abundant plasma protein with antibacterial activity against Gram-positive and Gram-negative bacteria [31], to be negatively associated to severity of CAP. Furthermore, proteins like CRP, LBP and PLM showed higher abundances in patients with higher SOFA scores. PML, which is induced by interferons after viral infection, plays an important role in host antiviral defense. For Gram-negative pneumonia lipopolysaccharide-binding protein is known to be a key player in the innate immune response, sensitizing cells to bacterial lipopolysaccharides [32]. The expression of the acute phase protein C-reactive protein is stimulated by cytokines due to infection and tissue inflammation. It was discovered that the concentration of CRP can vary in CAP patients with different infections [32].

The complement system, as part of the immune system, also protects against microbial infection and stimulates several cellular and humoral interactions. For the present cohort of CAP patients, we have found indications for positive association to SOFA for proteins of the complement system like complement component C8 beta chain, mannann-binding lectin serine protease 1 and 2, complement 3 and negative association for vitronectin. In transgenic mice, it was demonstrated that complement 3 plays an important protective role in the initial phase of infection with Streptococcus pneumoniae [34]. Furthermore, deficiency of complement 3 led to induction of an inflammatory response during pneumococcal pneumonia. In sepsis patients, a high expression of vitronectin in inflamed lung tissue was reported [35]. Particularly for infection with Streptococcus pneumoniae, a specific internalization strategy of the Gram-positive bacterium by interacting with vitronectin was described [36,37].

The acute phase response, as a reaction to infection and inflammation which is regulated by proinflammatory cytokines, involves proteins like protease inhibitors, coagulation proteins, complement proteins, transporters and others [38]. Altered levels of acute phase proteins were reported in two proteome studies related to CAP and sepsis [9,10]. In our analysis, acute phase proteins (APPs) with different physiological function for the immune system like the above-mentioned LBP and CRP and protease inhibitor AACT were altered in association to different SOFA scores. Concerning the SOFA subscore parameters, most of these proteins showed also associations to bilirubin levels and oxidation index, which are criteria for liver and lung dysfunction, respectively. It has been reported that PEDF is involved in inflammatory processes due to different pulmonary disease pattern [39,40]. In condition of sepsis PEDF was found in significantly elevated levels in plasma playing an important role in induction of hyperpermeability [41].

In addition to altered protein levels, the observed metabolic changes might reflect the pathophysiology of inflammation. Previous studies have shown that metabolic patterns as well as single metabolites can be used for discrimination of patients with non-infectious systemic inflammation from patients with sepsis and revealed an association of metabolites with unfavorable outcome or an inflammatory reaction, reflected by defervescence [11,23,42].

Observed alterations of amino acids and biogenic amines, e.g. decrease of tryptophan and increase of kynurenine with higher SOFA scores, were in accordance with previous studies which demonstrated low tryptophan concentrations in patients with severe sepsis and showed a correlation between elevated kynurenine concentrations and increased mortality [43–45]. Proinflammatory cytokines induce the expression of indoleamine 2,3-dioxygenase resulting in a higher metabolization of tryptophan to kynurenine. Further investigations demonstrated that elevated kynurenine levels are related to hypotension during sepsis [46], and a dysregulated immune response and an impaired microvascular reactivity [43]. This is in line with the association between kynurenine and tryptophan and low MAP found in this study. Furthermore, a link between indoleamine 2,3-dioxygenase activity, immunosuppression and tolerance could be shown [47,48]. Cells expressing IDO can suppress T-cell responses and promote tolerance.

In addition to the increase due to changes in arginine metabolism, putrescine plays a different role in pneumonia. CAP is mainly caused by Gram-positive bacteria like Streptococcus pneumoniae. Weber et al. [49] showed that polyamines are key metabolites for growth and virulence of Streptococcus pneumoniae. Moreover, elevated putrescine concentrations have been observed in Streptococcus pneumoniae infected mice and as prognostic marker in patients with CAP [23,50]. These findings are in accordance with the observed association of putrescine with higher SOFA score values. In addition to the well-known marker creatinine, also SDMA correlates with renal dysfunction. SDMA is fully removed by renal excretion, associates with the glomerular filtration rate [51,52] and can be used as prognostic marker in critical illness and sepsis [53,54].

It is well known that acylcarnitines can lead to induction of inflammatory pathways [55]. In our study, alterations in fatty acid metabolism, such as serum concentrations of 13 short- and medium-chain acylcarnitines, correlated prominently with the SOFA score. Similar findings have been previously reported in human [44,56,57], primate [58] and rat samples [59–61]. In contrast, three long-chain acylcarnitines decreased with higher SOFA score. Acylcarnitines are byproducts of mitochondrial β-oxidation, but in contrast to acyl-CoAs, which cannot cross the mitochondrial membrane, acylcarnitines can transfer efficiently into the cytosol and subsequently into the bloodstream. Transfer of long-chain fatty acids into mitochondria is the rate-limiting step of fatty acid oxidation which is decreased in animal sepsis [62,63]. This leads to a decreased conversion of long-chain fatty acids into acylcarnitines and thus might explain why long-chain acylcarnitines are decreased with increasing pneumonia severity. In contrast, short- and medium-chain fatty acids can permeate the outer and inner mitochondrial membranes by simple diffusion. In animal sepsis β-oxidation of short- and medium-chain fatty acyl-CoAs is impaired [64,65], resulting in a subsequent conversion into short- and medium-chain acylcarnitines. The observed acylcarnitine changes with increasing SOFA were also associated with poor oxygenation, low MAP and renal dysfunction. Recently, elevated short- and medium-chain acylcarnitine concentrations in acute kidney injury [66] and diabetic kidney disease [67] were reported, which might result from reduced expression and activity of mitochondrial and peroxisomal fatty acid oxidation enzymes in kidney tissue [68]. Hypoxia [69] and cardiovascular events [70,71] can also lead to increased acylcarnitine levels. This suggests, that
acylcarnitine changes are signals of general mitochondrial dysfunction [72].

Besides acylcarnitine alterations, other members of fatty acid metabolism were associated with the SOFA score. The concentration of 13 LPCs, 47 PCs and 13 sphingolipids were decreased with increasing pneumonia severity. This is in line with former studies [23,44,56,57]. Low sphingolipid concentrations at high SOFA score values are in accordance with changes in the sphingomyelin pathway, which have been shown to be mediated by inflammatory cytokines, such as TNF-α and IL-1β, due to an increase in secretory sphingomyelinase activity within 3 h after endotoxin treatment [73]. This leads to increased hydrolysis of sphingomyelins to ceramides which are subsequently converted to sphingosine with a concomitant decrease of sphingolipids [74]. The increased sphingolipid hydrolysis due to inflammation causes a twofold increase of ceramide in the surfactant leading to impaired biophysical properties of the alveolar surfactant film [75]. In airway epithelia cells, the direct relation between cytokine-mediated catabolism of sphingomyelins with inflammatory ceramide release and the anti-inflammatory phosphatidylcholine synthesis could be demonstrated. Cytokine (TNF-α) derived ceramide and sphingosine inhibits the rate-limiting enzymatic step in de novo PC synthesis by reducing the phosphocholine cytidylyltransferase activity [76,77].

Besides low PC and sphingolipid concentrations, the enhanced conversion to lysophosphatic acid by plasmatic lysophospholipase D or the exertion of immune-suppressive function by binding to immune-regulatory receptor G2A contributes to low LPC levels [78,79]. Low LPC concentrations are associated with a poor outcome in sepsis patients [23,44,80,81]. Our findings of low levels of LPCs are in accordance with previous findings in humans and primates [56–58,82]. LPCs seemed to have potential as inflammatory and prognostic markers [11,23,42]. The observed changes in fatty acid metabolism with increasing SOFA were also associated with poor oxygenation, low MAP and liver dysfunction. Hypoxia influences several enzymes of de novo synthesis or hydrolysis of sphingolipid metabolism, leading to increased ceramide levels [83]. The LPC/PC imbalance is certainly related to hepatic dysfunction due to the origin of these lipids in the liver. Low circulating levels of LPC have also been reported in inflammatory liver disease [84,85]. Moreover, LPCs were shown to inhibit LPS-induced release of TNF-α from neutrophils (major producing cell type) in response to LPS [86]. This explains the association between LPCs and TNF-α related circulatory collapse, hepatic failure and inflammation.

In our analyses, most PCs are negatively associated with severity of CAP (Fig. 2). PCs are main components of membranes and also of lipoproteins, which transport hydrophobic substances in the hydrophilic environment of plasma [87]. Apolipoproteins are important components of lipoproteins, which act as regulatory and receptor binding proteins [87]. For lipoproteins, association of proteins like LCAT, PLTP, CETP, TRF and transthyretin was reported [87–89]. Consistent with our observations of altered metabolomics pattern, we have seen decreased levels of apolipoprotein A-I, A-II, D, F, M; TRF and LBP in the proteome analyses in combination with increased levels of phospholipid transfer protein (PLTP), apolipoprotein B-100 and E, in relation to severity of disease. Association of PCs and LPCs to TRF were also described in a previous study by [90]. Our findings in the metabolome and proteome of CAP patients confirm reports which showed a reduction of circulating lipoproteins, especially high density lipoproteins (HDL), low density lipoproteins (LDL) and phospholipids as well as changes in their composition during infection and inflammation [91] [92,93]. In case-control studies, changes of apolipoproteins in sepsis patients have been reported, which support our results. Comparing CAP patients to healthy controls [10] showed lower levels of HDL and apolipoproteins like A-II and D in septic patients. Cao et al. [9] described higher concentration of the cholesterol transporter apolipoprotein B-100 in young sepsis patients in comparison to CAP patients in the same age range.

5. Conclusion

In the present study, a comparative analysis of a comprehensively phenotyped cohort of community acquired pneumonia patients using a combination of proteomics and metabolomics analyses was performed. Association analysis to SOFA indicated tremendous changes in proteome as well as in metabolome profiles in relation to the severity of the disease. Analysis of these associations to variables of SOFA allows precise specification of metabolome and proteome signatures involved in immune response related processes, particularly in lipid metabolism, acute phase response, coagulation, complement activation and inflammation. CAP-specific signatures could be characterized not only in relation to severity of the disease but also in association to the extent of dysfunction of the respiratory, renal, coagulation and cardiovascular system as well as to liver dysfunction. The comprehensive characterization of severity-associated changes in the proteome and metabolome pattern presented, might be helpful for future improvement of diagnosis and follow up of patients with community-acquired pneumonia.

Ethics approval and consent to participate

Data from the PROGRESS study (clinicaltrials.gov: NCT02782013) were used in this manuscript. The approval of PROGRESS was authorized in accordance of the ethics committee of the University of Jena (2403–10/08) and by locally responsible ethics committees of each study center. All participants or their legal guardians gave written informed consent for participation in the study. Requirements of the Declaration of Helsinki [94].

Funding

The PROGRESS study is funded by the German Federal Ministry of Education and Research (BMBF), grant numbers 01K07110 (Gisßen), 01K07111 (Jena), 01K07113 (Leipzig), 01K07114 (Berlin), 01K11010 (Leipzig), and 01K11010D (Greifswald). Additional funding is provided by the German Center for Lung Research (DZL, grant number 82DZLJ19A2). BMBF and DZL have no influence on the design of the study, on collection, analysis, and interpretation of data, or on writing the manuscript.

Declaration of Competing Interest

The authors have declared no conflict of interest.

Acknowledgement

We want to express our appreciation to the patients and their relatives or legal guardians for participating in the PROGRESS study and the PROGRESS Study Group for recruiting and data collection. We wish to thank Jette Anklam, Cora Richert, Julia Köhler, Kay Stötzer for technical assistance and Vishnu Dhople and Elke Hammer for helpful discussion and support in the mass spectrometric analyses for proteome analyses.

The PROGRESS study group comprised: Stefan Angermair, Charité-Universitätsmedizin Berlin, Benjamin Franklin, Klinik für Anaesthesiologie und Intensivtherapie, Hindenburgdamm 30, 12203 Berlin; Christoph Arntzen, Krankenhaus Angermünde, Klinik f. Innere Medizin/Pneumologie, Rudolf-Breitscheid-Str. 37, 16278 Angermünde; Lorenz Balke, Universitätsklinikum Schleswig-Holstein-Campus Kiel, Medizin 1, Arnold-Heller-Straße 3, 24105 Kiel; Robert Balz, Universitätsklinikum des Saarlandes, Innere Medizin V, Kirbergerstr. Gebäude 91, 66421 Homburg/Saar; Michael Benze, Charité-Universitätsmedizin Berlin, Medizinische Klinik, Infektiologie und Pneumologie, Charitéplatz 1, 10117 Berlin; Ayhan Berber, Mathias-Spittal Rheine, Klinik für Pneumologie und Thoraxonkologie, Med. Klinik V, Frankenburgstr. 31, 48431 Rheine; Frank Bloos,

Appendix A. Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.jprot.2019.103627.

References


