Resource

Mortality Risk Profiling of *Staphylococcus aureus* Bacteremia by Multi-omic Serum Analysis Reveals Early Predictive and Pathogenic Signatures

Jacob M. Wozniak,[1,2,3,4] Robert H. Mills,[1,2,3,4,5,6,7] Joshua Olson,[2,3,5] J.R. Caldera,[3,5,9] Gregory D. Sepich-Poore,[6] Marvic Carrillo-Terrazas,[1,2,3,4] Chih-Ming Tsai,[3,5] Fernando Vargas,[2,8] Rob Knight,[3,4,5,6,7] Pieter C. Dorrestein,[2,4] George Y. Liu,[3,5] Victor Nizet,[2,3,5] George Sakoulas,[3,5] Warren Rose,[10,11] and David J. Gonzalez[1,2,3,4,12,*]

[1]Department of Pharmacology, University of California, San Diego, La Jolla, CA 92093, USA
[2]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 92093, USA
[3]Collaborative to Halt Antibiotic-Resistant Microbes, University of California, San Diego, La Jolla, CA 92093, USA
[4]Center for Microbiome Innovation, University of California, San Diego, La Jolla, CA 92093, USA
[5]Department of Pediatrics, University of California, San Diego, La Jolla, CA 92093, USA
[6]Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093, USA
[7]Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA 92093, USA
[8]Division of Biological Sciences, University of California, San Diego, La Jolla, CA 92093, USA
[9]Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA
[10]School of Pharmacy, School of Medicine and Public Health University of Wisconsin-Madison, Madison, WI 53705, USA
[11]Department of Medicine, School of Medicine and Public Health University of Wisconsin-Madison, Madison, WI 53705, USA
[12]Lead Contact
*Correspondence: djgonzalez@ucsd.edu
https://doi.org/10.1016/j.cell.2020.07.040

SUMMARY

Staphylococcus aureus bacteremia (SaB) causes significant disease in humans, carrying mortality rates of ~25%. The ability to rapidly predict SaB patient responses and guide personalized treatment regimens could reduce mortality. Here, we present a resource of SaB prognostic biomarkers. Integrating proteomic and metabolomic techniques enabled the identification of >10,000 features from >200 serum samples collected upon clinical presentation. We interrogated the complexity of serum using multiple computational strategies, which provided a comprehensive view of the early host response to infection. Our biomarkers exceed the predictive capabilities of those previously reported, particularly when used in combination. Last, we validated the biological contribution of mortality-associated pathways using a murine model of SaB. Our findings represent a starting point for the development of a prognostic test for identifying high-risk patients at a time early enough to trigger intensive monitoring and interventions.

INTRODUCTION

Overall mortality rates for *Staphylococcus aureus* bacteremia (SaB) range from 20% to 30% (Kern, 2010; van Hal et al., 2012; Wang et al., 2008) and underlying risk factors for serious *S. aureus* infections are expanding (Tong et al., 2015). SaB patients display a heterogeneous array of disease severity and patient outcomes (Holland et al., 2014; Rasmussen et al., 2011); some patients clear the pathogen on first-line therapy, while others fail to resolve the infection. Extended bacteremia leads to dysregulation of the host immune response, which is correlated with patient mortality (Minejima et al., 2016; Rose et al., 2012, 2017). This heterogeneity in SaB complicates the determination of optimal treatments. The current standard of care is to administer broad-spectrum antibiotics while awaiting pathogen susceptibilities to guide treatment decisions. However, blood

cultures are not always attainable, and it may take several days to deduce antibiotic susceptibilities. Any delay in intervention exacerbates patient mortality, especially in sepsis (Dellinger et al., 2013; Ferrer et al., 2014). In the case of vancomycin, while resistance is rare, clinical failure is common, revealing shortcomings in predictive power of standard antimicrobial susceptibility testing (Ersoy et al., 2017). Herein, we explore whether host responses measured within hours of clinical presentation can accurately predict mortality risk, which could ultimately inform appropriate and personalized therapy.

We previously identified immunological biomarkers for SaB mortality (Rose et al., 2012, 2017) and prolonged bacteremia (Rose et al., 2017), which were subsequently corroborated in independent studies (Guimaraes et al., 2019; Minejima et al., 2016). While these findings represent starting points for "precision medicine" in SaB, the host response to infection is highly
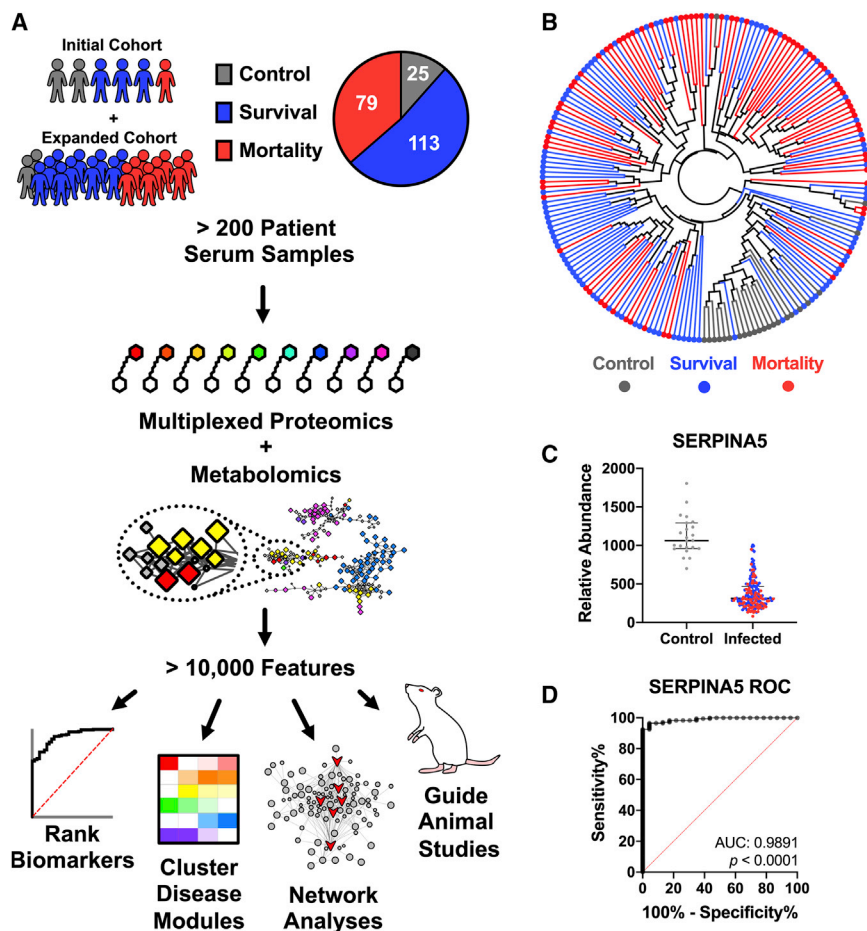
A

Initial Cohort

+

Expanded Cohort

Control 25
Survival 113
Mortality 79

> 200 Patient
Serum Samples

↓

Multiplexed Proteomics
+
Metabolomics

↓

> 10,000 Features

Rank Biomarkers

Cluster Disease Modules

Network Analyses

Guide Animal Studies

B

Control ● Survival ● Mortality ●

C

SERPINA5

Relative Abundance — Control, Infected

D

SERPINA5 ROC

Sensitivity% vs 100% - Specificity%

AUC: 0.9891
$p < 0.0001$

Figure 1. Multi-omic Analysis of SaB Patient Serum

(A) Workflow for SaB serum analysis.

(B) Hierarchical clustering (Pearson) for proteins detected across all samples.

(C) Abundance of SERPINA5 in control (gray, NN and HN) and infected samples (blue, HS; red, HM).

(D) ROC curve of SERPINA5 (control versus infected).

Error bars represent interquartile range (IQR). For all tests, significance values are denoted as follows: ****$p < 0.0001$; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; ns, not significant. See also Tables S1, S2, and S3.

who need intensified monitoring and therapy, rather than responding post hoc to failures in standard of care.

RESULTS

Overview of Multi-omic SaB Patient Serum Analysis

We employed a multi-omic approach to gain a comprehensive view of the SaB host-pathogen interaction (Figure 1A). First, a discovery cohort was analyzed by standard multiplexed proteomics to assess the ability of serum proteins to predict SaB patient phenotypes. This initial analysis identified 1,405 proteins with a false discovery rate (FDR) <1% and yielded biomarker candidates associated with various disease features including mortality (Table S1). Based on this analysis, a power calculation was performed and, together with reported recommendations (Frantzi et al., 2014), we designed a validation cohort of 200 samples (25 control, 99 SaB survival, 76 SaB mortality). To deepen our comprehension of the molecular features related to SaB, the expanded cohort was analyzed through standard proteomic workflows as well as PTM-tolerant and metabolomic approaches (Figure 1A). Overall, 1,088 proteins (294 across all samples; Table S2), 5,280 metabolomic features (720 across all samples; Table S3), and 6,700 modified peptides (332 across all samples; Table S4) were quantified in this experiment. The resultant >10,000 features were analyzed with binary comparisons to identify biomarkers, and through clustering and network based approaches to define disease associations within our primary sample groups (control groups: NN, non-hospital, non-infected; HN, hospital, non-infected; infection groups: HS, hospital, survival; HM, hospital, mortality).

Hierarchical clustering of the proteomics data showed clear segregation of the control and infected samples (Figure 1B). In contrast, SaB survival and mortality groups were intermixed, indicating that the differences between SaB mortality and survival are subtle. Nevertheless, stratification of mortality and survival groups was observed in the clustered dendrogram, indicating the potential to define mortality biomarkers. As a proof of principle for
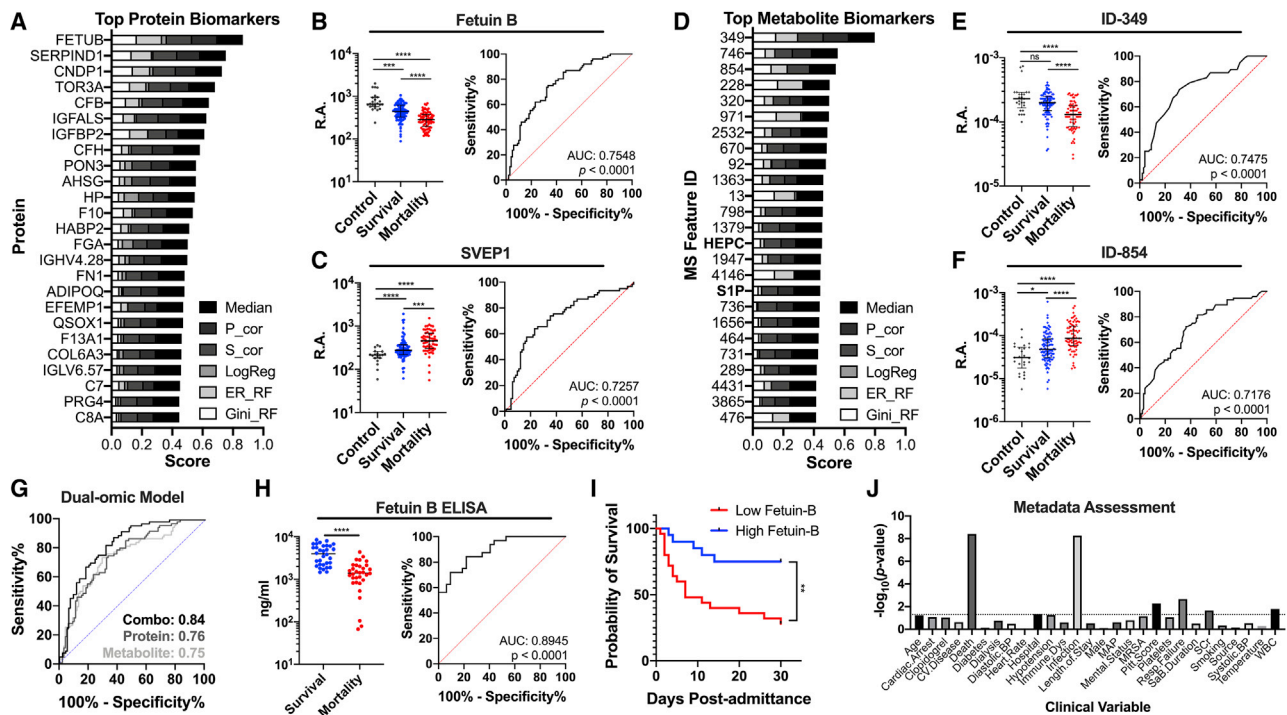
complex and extends beyond immunological parameters. For example, lactic acidosis and acute kidney injury are associated with SaB mortality and are potential markers of mitochondrial dysfunction, reflecting effects on host metabolism (Mikkelsen et al., 2009; Ralto and Parikh, 2016; Singer, 2014). Thus, a comprehensive, unbiased assessment of host factors altered during SaB may elucidate additional features that can predict patient outcomes and guide therapeutic development.

Here, we construct a molecular snapshot of SaB patient serum collected upon clinical presentation, providing the earliest view possible of the *in vivo* human response to infection. Using metabolomics and multiplexed quantitative proteomics, we analyzed samples from two cohorts (>200 individuals), including uninfected controls. Through multiple rounds of biomarker analyses, we defined features with strong individual predictive value, which increases when used in combination. The depth of analysis was enhanced through computational methods, which identified prevalent post-translational modifications (PTMs) and inferred underlying cytokine signaling networks. These techniques revealed glycopeptides as more precise biomarkers and uncovered carbamylation on serum proteins in patients who succumbed to infection. Ultimately, we provide a starting point for a rapid clinical test that identifies patients with high mortality risk. Further refinement of SaB biomarkers will enable clinicians to identify patients

Figure 2. Definition of High-Confidence Biomarkers for the Prediction of SaB Patient Mortality

(A) Top 25 EFS proteins (survival versus mortality; ER_RF, error-rate based; Gini_RF, Gini index random forests).

(B) Abundance and ROC curve of fetuin B (survival versus mortality).

(C) Abundance and ROC curve of SVEP1 (survival versus mortality).

(D) Top 25 EFS metabolites (survival versus mortality).

(E) Abundance and ROC curve of metabolite ID-349 (survival versus mortality).

(F) Abundance and ROC curve of metabolite ID-854 (survival versus mortality).

(G) Dual-omic ROC curve (survival versus mortality; Protein: FETUB; Metabolite: ID-349; Combo: FETUB + IGFBP3 + ID-349 + ID-854).

(H) ELISA abundance and ROC curve of fetuin B (survival versus mortality).

(I) Survival curves of fetuin B high (>2.2 μg/mL) and low (<2.2 μg/mL) patients.

(J) Metadata assessment of fetuin B.

Error bars represent interquartile range (IQR). For all tests, significance values are denoted as follows: ****$p < 0.0001$; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; ns, not significant. For (B), (C), (E), and (F), Kruskal-Wallis tests with Dunn's multiple comparison test significance is displayed. For (H), MWU test significance is displayed. See also Figures S1 and S2 and Tables S2, S3, S5, and S6.

distinguishing disease states, we first selected a highly discriminating protein for predicting infection, rather than mortality. SERPINA5 emerged as a top hit (Figure 1C), with a receiver operator characteristic (ROC) area under the curve (AUC) of 0.9891 (Figure 1D). This protein had a higher AUC than the standard infection marker, C-reactive protein (CRP) (WHO, 2014), in the current dataset (AUC = 0.9691) as well as reported values (maximum AUC = 0.92) (Liu et al., 2010; Park et al., 2014; Póvoa et al., 2005). This example demonstrates the power of unbiased proteomics for biomarker discovery.

Definition of High-Confidence Biomarkers to Predict SaB Patient Mortality

Our first goal was to define high-confidence biomarkers to predict SaB patient mortality. To rank biomarkers, we took an ensemble feature selection (EFS) approach (Neumann et al., 2017), which applied multiple feature selection algorithms, then aggregated and ranked the results. This method can avoid biases associated with individual feature selection algorithms

(He and Yu, 2010) and was applied to the two primary datasets to rank top biomarkers (Proteomics Figure 2A; Metabolomics Figure 2D). Due to the incompatibility of EFS with missing values, features with missing values were ranked using Mann-Whitney U (MWU) tests and the average from both strategies was taken for final biomarker rankings (Tables S2 and S3). Importantly, biomarker ranks by EFS and MWU tests were concordant (Figures S1A and S1B). The highest ranked protein biomarkers were fetuin B (Figure 2B), heparin cofactor II (SERPIND1, Figure S1C), and carnosine dipeptidase 1 (CNDP1, Figure S1D), all with decreased serum levels. The decrease in serum fetuin B was also captured in the initial cohort, despite a low number of mortality samples analyzed (Figure S1E). Our top biomarkers with increased serum levels were SVEP1 (Figure 2C), cystatin B (CSTB, Figure S1F) and pulmonary surfactant-associated protein B (SFTPB, Figure S1G). Applying a similar approach to the metabolomics data, we found that the highest ranked biomarkers were unidentified MS features (Figure 2D). However, these molecules showed considerable predictive utility (Figures

2E and 2F), similar to our top-ranked protein biomarkers (ROC AUC = ~0.75; p < 0.0001). The top ranked, identified metabolites include 2-Hexadecanoylthio-1-Ethylphosphorylcholine (HEPC, Figure S1H) and sphingosine-1-phosphate (S1P, Figure S1I) by EFS and thyroxine (T4, Figure S1J) and decanoyl-carnitine (Figure S1K) by MWU test.

The EFS approach ensures that the top-ranked biomarkers are not correlated to one another and therefore could be used in combination for the enhanced prediction of SaB patient mortality (Williams, 2009). Using the top two markers from both workflows enhanced predictive power relative to the individual markers alone (Figure 2G). To support the proteomics workflow, we validated fetuin B (Figure 2H) and other top biomarkers (Figures S1L and S1M) in a subset of samples using enzyme-linked immunosorbent assays (ELISAs) (fetuin B AUC = 0.8945). Our results indicate that patients with low fetuin B (<2.2 μg/mL) had significantly reduced survival compared to patients with high fetuin B (Figure 2I).

An important factor to consider when performing biomarker analyses is the influence of confounding factors (Ensor, 2014). To investigate this matter, we performed a metadata-wide assessment for every multi-omic feature detected (Proteomics Table S5; Metabolomics Table S6). We found that all top biomarkers (up and downregulated) are predominantly associated with infection and mortality, with minimal associations to other clinical metadata (Proteomics Figures 2J and S2A–S2G; Metabolomics Figures S2H–S2M). The next most common metadata associated with the top biomarkers are dialysis and serum creatinine levels, both related to kidney function (Levey et al., 1988) and mortality in bacteremia (Nielsen et al., 2015; Vandecasteele et al., 2009). Indeed, there was negligible influence of typical confounding variables (e.g., age, gender) on the top SaB mortality biomarkers. Overall, these results provide an extensive list of biomarkers associated with SaB mortality.

PTM-Tolerant Analysis of Serum Samples Enables the Identification of Disease-Associated PTMs

Serum is a notoriously difficult sample to analyze via proteomics (Chandramouli and Qian, 2009; Geyer et al., 2017), attributed to the large dynamic range of proteins and high numbers of PTMs. Thus, standard serum proteomic searches fail to identify greater than 90% of the spectra acquired from mass spectrometry (MS)-based proteomics (Dey et al., 2019). We hypothesized that predicating abundant PTMs could facilitate a PTM-inclusive analysis and identify more spectra. PTM identification and localization is optimally derived from high-resolution mass spectra (Chick et al., 2015; Devabhaktuni et al., 2019). Thus, all subsequent PTM analyses was performed on high-resolution MS^2 data acquired in the Orbitrap mass analyzer. The high-resolution data displayed similar results to low-resolution data when matching unmodified peptides in terms of peptide spectrum matches (PSMs) (Figure S3A), identification of peptides (Figure S3B) and proteins (Figure S3C), and quantification by both spectral counting (Figure S3D) and TMT-based quantification (Figure S3E). To identify global modifications in the serum proteome, we used molecular networking to group similar spectra that differ by regular mass shifts (Wang et al., 2016). Overall, we networked >80% of the MS^2 spectra (Figure 3A), suggesting that many of the peptides identified in the standard database search have variant forms. We observed highly abundant PTMs present in our data (Figure 3B), including expected artifacts such as oxidation of methionine (+15.99), alkylation of cysteine (+57.02) as well as unanticipated modifications such as carbamylation (+43.005), dioxidation (+31.99), and formylation (+27.99). Notably, the glycan moieties fucose (+146.06), hexose (+162.05), and sialic acid (+291.1) were also highly abundant. This suggests that the peptides captured in our MS analysis are rich in PTMs, complicating identification through traditional strategies.

To capitalize on these discovered modifications, we employed a PTM-tolerant search strategy, which achieved a doubling of the serum PSM rate (Figure 3C). Non-glycan modifications called in the PTM-tolerant search were correlated with the number of edges from the GNPS analysis (Figure 3D). Further, >85% of glycosylation sites detected have been previously reported in Uniprot (Figure S4A), supporting their identification as true glycopeptides. The distribution of mass errors from the modified peptides was nearly identical to the standard search (Figure S4B), indicating spectral identification quality is maintained. We also found that the total PSMs and unique peptides per protein were highly correlated between the standard and the PTM-tolerant search (Figures S4C and S4D). We noted that the PTM-tolerant search increased the number of unique peptides detected for low-abundant proteins; proteins with the fewest unique peptides in the standard search gained more unique peptides in the PTM-tolerant search than proteins that had many unique peptides originally detected (Figure S4E). Gene ontology (GO) analysis (Huang et al., 2007) revealed that the majority of proteins with boosted unique peptides were immunoglobulins (Figure S4F), but a number of intracellular proteins showed similar gains (Figure S4G) and had PTMs that demonstrated significant associations to hospitalization (Figure S4H) and infection (Figure S4I). Our results demonstrate that predicting abundant modifications can inform search strategies, yielding higher PSM rates and increased confidence in low-abundant proteins while also identifying disease-associated PTMs.

We reasoned that modified peptides might act as biomarkers for predicting infection or SaB mortality. Again, due to the incompatibility of EFS with sparse data, we ranked PTM biomarkers solely on MWU test p values. Intriguingly, the top biomarkers for infection and mortality were both glycans on alpha-2-HS-glycoprotein (AHSG), also known as fetuin A (Figures 3E and 3F). The support for these two glycans is strong as evidenced by >50 unique PSMs detected for each peptide, some of which have Byonic scores >400 (corresponding to an FDR <0.1%) (Bern et al., 2012). Individually, these biomarkers demonstrated higher ROC AUC values than our top unmodified protein biomarkers (0.9981 versus 0.9891 for infection and 0.8066 versus 0.7548 for mortality, Figures 3E and 3F). Unmodified fetuin A was also a top biomarker for both infection and mortality (Figure 2A), although the observed fold change was higher for the glycans than the total protein (Figure 3G). The metadata confirmed these biomarkers are primarily associated with infection and mortality (Figure S4J; Table S7). This suggests that these glycans may yield better predictive value than the unmodified protein alone. When used in concert with our top protein and metabolite biomarkers (nine total molecular features), these
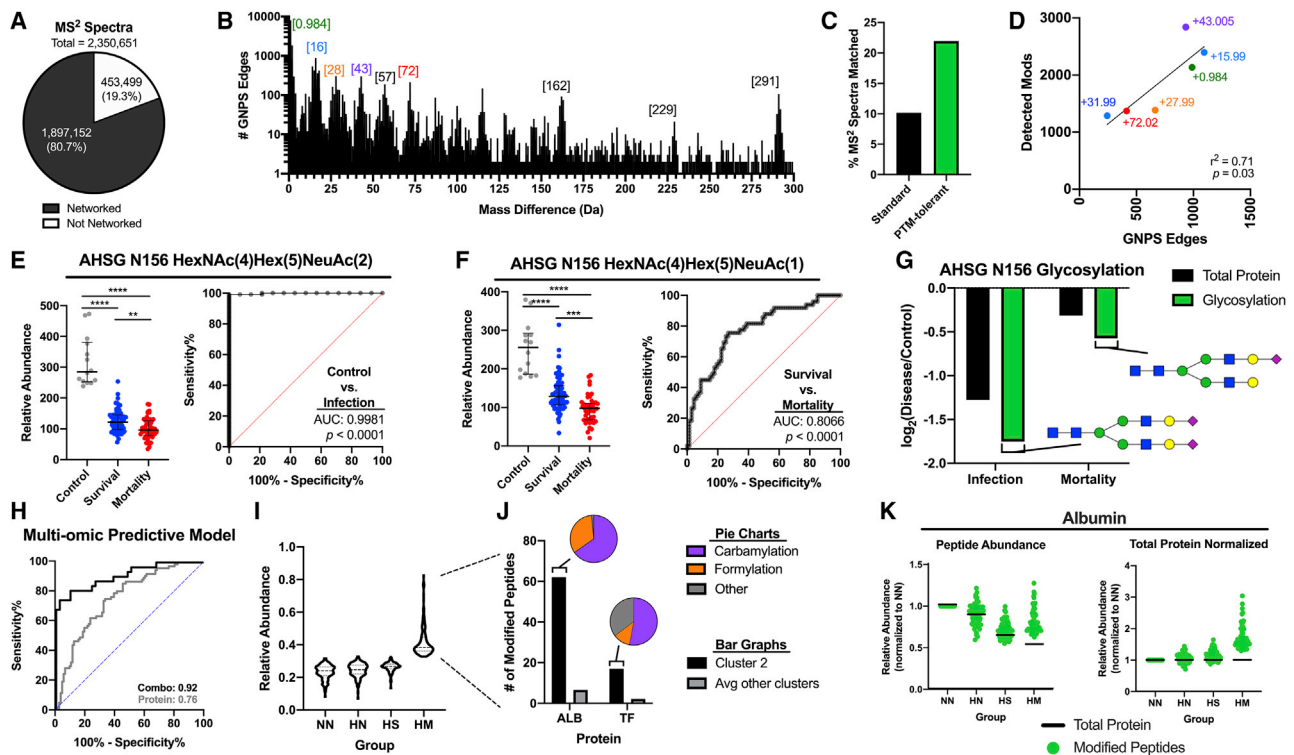
Figure 3. PTM-Tolerant Analysis of SaB Patient Serum

(A) Pie chart of networked MS^2 spectra.

(B) Network edge histogram with top mass shifts highlighted.

(C) Percentage of MS^2 spectra matched in both workflows.

(D) Correlation of network edges and PTMs detected in PTM-tolerant workflow.

(E) Abundance and ROC curve of AHSG N156 HexNAc(4)Hex(5)NeuAc(2) (control versus infected).

(F) Abundance and ROC curve of AHSG N156 HexNAc(4)Hex(5)NeuAc(1) (survival versus mortality).

(G) Fold changes of total AHSG protein and glycosylations of N156 for infection and mortality samples.

(H) Multi-omic ROC curve (survival versus mortality).

(I) Abundance of modified peptides assigned to the mortality-specific cluster.

(J) Distribution of peptide counts and modifications types of albumin (ALB) and serotransferrin (TF).

(K) Albumin mortality associated PTM plot depicting modified peptide abundance (left) and modified peptide abundance normalized to protein levels (right).

Error bars represent interquartile range (IQR). For all tests, significance values are denoted as follows: ****p < 0.0001; ***p < 0.001; **p < 0.01; *p < 0.05; ns, not significant. For (E) and (F), Kruskal-Wallis tests with Dunn's multiple comparison test significance is displayed. See also Figures S3 and S4 and Tables S2, S4, and S7.

glycans enhanced predictive power (AUC = 0.92, Figure 3H). To our knowledge, this approach generated the top model, based on AUC and *n*, for predicting mortality from any infection using patient-derived biomarkers.

The PTM analysis above was performed on the normalized raw abundances for each modified peptide without consideration for protein level changes. However, the modified peptide abundances can also be normalized to the protein level to investigate divergent regulation of the protein and associated PTMs. Overall, the modified peptides had a positive correlation to their respective protein levels (Figure S4K). Similarly, when we compare the fold changes of the modified peptide abundance to the protein normalized values for infection and mortality changes, we observe corredponding results (Figures S4L–S4M) with some exceptions. To better understand which modifications deviated from their protein level, we first filtered the protein-normalized PTMs with signif-

icant alterations in any of the primary sample groupings (i.e., NN, HN, HS, HM) using ANOVA (p < 0.05), then clustered these features based on expression, which revealed interesting trends in modified peptides (Figure S4N). Most striking was a cluster of modifications (Cluster 2) that showed a stark increase in abundance specific to the mortality samples (Figure 3I). Nearly half (46%) of the modified peptides in this cluster were assigned to only two proteins: albumin and serotransferrin (Figure 3J) with the modifications being primarily carbamylation and formylation (Figure 3J). Comparing the relative changes in the modified peptide to the total protein abundance for albumin (Figure 3K), we noted that, while total albumin levels dropped upon infection and were reduced further in mortality, albumin was modified at a higher level in the mortality group (Figure 3K). Modifications on serotransferrin demonstrated a similar trend (Figure S4O). Together, this analysis enabled a deeper interrogation of
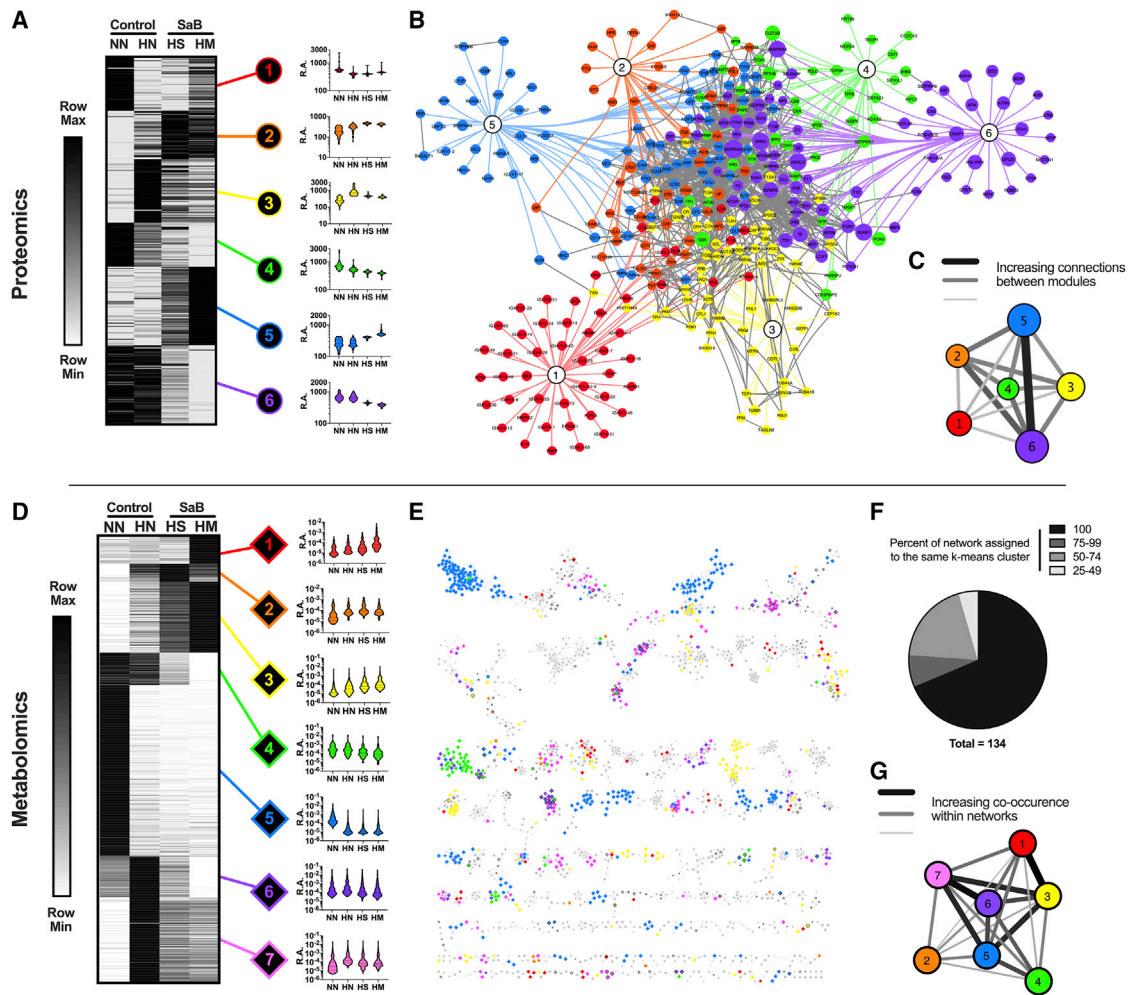
Figure 4. Clustering of Proteomics/Metabolomics Data into Disease-Relevant Modules
(A–C) K means clustered heatmap (A), protein association network (B), and module cross-talk network (C) of all significantly altered proteins (ANOVA p < 0.05) across the four primary groups. In (B) and (C), nodes are colored as in (A).
(D–G) K means clustered heatmap (D), molecular networking overview (E), within network co-regulation pie chart (F), and module cross-talk network (G) of all significantly altered metabolites (ANOVA p < 0.05) across the four primary groups. In (E) and (G), nodes are colored as in (D).
See also Figures S5 and S6 and Tables S2 and S3.

serum-derived proteomics data and linked multiple, distinct PTMs to increased SaB mortality.

Unbiased Clustering of SaB Disease Modules

Beyond defining biomarkers to predict SaB mortality, we sought to further understand the effects of *S. aureus* on the human serum landscape using our multi-omic dataset. Performing a similar clustering approach used for the PTM analysis, the proteomics data were grouped into 6 clusters (Figure 4A) and the metabolomics data into 7 clusters (Figure 4D), revealing expression profiles of interest. Proteomics cluster 2 (C2) captured the host response to infection regardless of mortality status, including CRP, serum-amyloid proteins 1 and 2, and other acute-phase components (Figures S5A–S5E). Clusters 4, 5, and 6 showed increases (C5) and decreases (C4 and C6) in the mortality group making them prime clusters for investigation. For clarity, the

mortality-associated proteomics clusters were renamed according to their expression direction and magnitude compared to control samples (C4, pMortality⁻; C5, pMortality⁺; C6, pMortality⁻; p, proteomics). Similarly, we renamed the most interesting metabolomics clusters according to their mortality expression directions and magnitude (C1, mMortality⁺⁺; C3, mMortality⁺; C4, mMortality⁻; m, metabolomics).

To examine the crosstalk of proteins between clusters, we performed a functional association analysis on the clustered proteins (Szklarczyk et al., 2019) (Figure 4B). Interestingly, we found the largest number of connections between proteins within the pMortality⁺ and pMortality⁻ clusters (Figure 4C), even though they change in opposite directions relative to the control patients. This suggests that proteins that increase in expression may impact the decrease in expression of another protein, and vice versa.

Due to their association with SaB mortality, we used GO analysis on mortality-associated protein clusters to define their functional roles (Figures S4F–S4H). While the pMortality⁻ cluster had few, low-significant enrichments (Figure S5F), the pMortality⁺ and pMortality⁻ clusters had multiple, highly significant functional groups enriched. Cluster pMortality⁺ is dominated by extracellular matrix (ECM) and insulin-like growth factor binding proteins (IGFBPs) and has a moderate enrichment for tumor necrosis factor (TNF)/ interleukin-1 (IL-1) response (Figure S5G). The ECM adhesion proteins ICAM1 and VCAM1 have previously been shown to be elevated in SaB patients (Söderquist et al., 1999), and TNF can be used as mortality biomarkers in humans (Rose et al., 2012). However, the enrichment for IGFBPs in this cluster represents an unexpected finding. In contrast, pMortality⁻ was enriched for protease inhibitors, complement/coagulation cascade members, and lipoproteins (Figure S5H). A decrease of lipoproteins is well described in sepsis (van Leeuwen et al., 2003), and the reduction in complement/coagulation is consistent with the proteolytic activation of these proteins. We also noted that a subset of IGFBPs were present in this cluster and possessed some of the most significant p values, which is particularly interesting given the presence of other family members in pMortality⁺ (discussed further below).

While functional association tools are absent for metabolic data interpretation, metabolites can be grouped based on MS² spectra using molecular networking (Wang et al., 2016). This analysis results in the formation of metabolite networks with structural similarity. For data visualization, we overlaid the K-means cluster color onto the individual metabolites in these networks (Figure 4E). A bird's eye view of the data revealed that nodes within a specific network were commonly assigned to the same expression cluster. In fact, >95% of molecular networks had at least half of their nodes co-regulated (Figure 4F). In addition, some clusters of similar expression profiles were often contained within the same networks, such as mMortality⁺ and mMortality⁺⁺ (increased in infection/mortality) and clusters 6 and 7 (increased in hospitalization) (Figure 4G). Together these findings suggest that structurally related metabolites are often co-regulated, offering more support for their importance in the host response to infection.

Using a combination of molecular networking, spectral library matching, and ClassyFire, we were able to provide identity information for nearly half (2,412/5,280, 46%) of the observed metabolomic features (Figure S5I). Spectral library matches were generally assigned the proper class of molecule by ClassyFire (83/86, 97%), indicating good agreement between the two tools. Additionally, the ontology provided by ClassyFire provided further coverage of networked metabolites (e.g., subnetwork 13, acyl-carnitines [Figure S5K]; subnetwork 15, bile acids and fatty acyls; subnetwork 45, glycerophosphocholines [Table S3]) supporting their similarities. While only 20% (481/2,412) of features with molecular information were identified using spectral libraries or ClassyFire, a much larger number of features are interpretable using molecular networks (Figure S5I).

To identify metabolites related to SaB, we first looked for compound classes that were enriched in mMortality clusters. It was found that all the significantly altered acyl-carnitines (ACs, 7/7 features) and steroid/steroid derivatives (6/6 features) were as-

signed to mMortality⁺ or mMortality⁺⁺. In contrast, significantly altered indoles were almost exclusively assigned to the mMortality⁻ cluster (4/5 features). Previous literature supports increased ACs (Puskarich et al., 2018), linked to liver dysfunction, and decreased indoles (Zeden et al., 2010) (e.g., tryptophan) in sepsis. The increased levels of steroids, including hydrocortisone, could be related to the inability of the patients to metabolize these treatments due to liver dysfunction (Schiffer et al., 2019).

Next, we screened for networks that contain multiple nodes from mMortality clusters. These include, among others, the AC network (Figure S5K), and networks containing bilirubin (Figure S5L) and biliverdin (Figure S5M). Bilirubin and biliverdin themselves were not associated with mortality; however, there were related molecules in these networks that went unidentified but had associations with mortality. Interestingly, there were many 14 Da edges in these networks, indicating methylation. Bilirubin/verdin both possess carboxylic acid functional groups, which are amenable to Fischer esterification in the presence of an alcohol. While these analogs may be the result of methanol extraction during sample preparation, their associations with mortality make them candidates for further study. We also discovered networks of co-regulated, unidentified metabolites that were associated with mortality and infection (Figure S5N). This unknown molecular network has a number of high-molecular-weight mass shifts, which may assist in determining the identity of these features. One of the benefits of using GNPS is the "living data" concept wherein the data are continuously reanalyzed (Wang et al., 2020). As GNPS matures, these molecules may be identified, which would allow for investigation into their relation to SaB.

Data Integration and Multi-group Classification

We reasoned that integration of the clustered data could be used to identify relationships between proteins and metabolites. Further, an integrated dataset could be used to support our binary biomarker analysis by determining the smallest number of features needed to classify the samples into our four primary groupings. Therefore, we took the most confidently identified features (i.e., identified in at least 50% of samples, 3,500 features), inputed missing values and then scaled the datasets before merging them. A least absolute shrinkage and selection operator (LASSO) logistic classification algorithm (Friedman et al., 2010), which penalizes large models, was employed to identify the minimum features needed to accurately classify the sample groups. A panel of 98 features was ultimately selected, which demonstrated stratification of the samples (Figure S6A) with a mean one-versus-all ROC AUC of 0.9014 for all pairwise comparisons. As expected, features important for predicting each group were primarily derived from their respective group-associated clusters in Figure 4 (proteomics Figure S6B; metabolomics Figure S6C). We noted that some of the mortality-related features were highly ranked in both the initial binary comparisons and the multi-class regression analysis including SERPIND1 and albumin from the proteomics data and T4 and hydrocortisone from the metabolomics data. This multi-omic panel reinforces the binary comparison analysis while highlighting the minimal features needed to classify all sample groups.

Next, we examined the integrated dataset for protein-metabolite relationships using the mixOmics R package (Rohart et al., 2017). Many correlations between proteins and metabolites were found (Figure S6D), which can be displayed as a network overlaid with K-means cluster values (Figure S6E). From this analysis, we identified clusters of co-regulated features associated with mortality (Figure S6E, #1) or hospitalization (Figure S6E, #2). The mortality-related features were primarily derived from pMortality⁻ and the mMortality⁻ clusters and contained many of our top biomarkers (e.g., fetuin A/B, albumin, SERPIND1, and the IGF system) as well as many nodes from the unknown molecular network in Figure S5G. Most of these proteins are involved in lipid transport or hemostasis (Figure S6E). The co-regulation observed between these proteins and metabolites suggests that the unknown metabolites function in a similar pathway and highlights potential crosstalk between the multi-omic data.

Global Characterization of Metabolic Dysfunction in SaB Mortality Patients

Many of our findings indicate that the most predictive signature for SaB mortality centers on a broad reprogramming of host metabolism. Our systems-level analysis enables a comprehensive assessment of major host metabolic pathways. For example, we quantified every member of the IGF signaling pathway, which demonstrated divergent regulation in our mortality-associated protein modules. Specifically, IGFBP1, 2, 4, and 7 were assigned to pMortality⁺ (Figure 5A), while IGFBP3, 5 and IGFALS were in pMortality⁻ (Figure 5B). IGFBPs function by binding to and stabilizing IGF-I and II in serum (Baxter, 2014). The binding of IGFs to IGFBP1, 2, 4, or 7 results in the formation of binary complexes that extend the half-life of IGFs from 2 to 30 min. However, if both IGFBP3 (or 5) and IGFALS bind IGFs, forming a ternary complex, this stabilization is increased up to 24 h. Given the expression patterns, we expect the amount of IGFs to decrease as the constituents of the ternary complex decrease. Indeed, we noted a significant decrease in IGF-II with increasing disease severity and a similar trend with IGF-I, although the latter results did not attain statistical significance (Figure 5C). Further, a comparison of the correlations between IGFI and II with all the IGFBPs detected in our dataset revealed positive correlations of IGFs with IGFBP3, 5 and ALS but negative correlations with the rest of the IGFBPs (Figure 5D). Together, these data suggest that SaB mortality is associated with a decrease in the IGFBP ternary complex and an increase in binary complexes, resulting in lower levels of circulating IGFs.

Given the striking association of the IGF system with SaB mortality and the role of IGFs in metabolism, we sought to further understand this metabolic dysfunction. Thus, we mined the data for other features related to general host metabolism. A signature derived from the GO analysis of our pMortality clusters, is the general decrease in apolipoproteins upon infection, which further decreases in SaB mortality patients (Figure 5E). A depletion of lipoproteins in response to infection is known and proposed to be a prognostic marker for severe sepsis (Christoffersen and Nielsen, 2012; Sharma et al., 2019). Our results support these findings and establish a link between this phenomenon and SaB in addition to non-specific sepsis.

We also uncovered evidence for metabolic dysfunction within our metabolomics dataset. The most prominent single feature was thyroxine (T4), a master regulator of host metabolism. T4 was a top-ranked biomarker from the binary comparison analysis and also selected for inclusion in the integrated multi-class regression model. We noted a similar trends in thyroxine-binding globulin (SERPINA7; Figure 5F) and transthyretin (TTR; Figure 5F) both of which bind to and stabilize T4 in circulation (Schussler, 2000), further supporting the reduction in T4 levels. Thyroid dysfunction during non-specific sepsis is well characterized (Bello et al., 2009; Plikat et al., 2007); however, it has not been previously associated with SaB infections or mortality.

Another connection of metabolism with SaB mortality is the increase in AC abundance in mortality patients. ACs are involved in fatty acid metabolism (McCoin et al., 2015), and the increased serum levels here are likely a signature of beta-oxidation dysfunction in the liver. Unlike T4, which was not within a molecular network, the ACs were found to be part of larger network of metabolites (Figures 5G and S5K), which was predominantly assigned to mMortality⁺ (yellow) and mMortality⁺⁺ (red) (Figures 5G and 5H). Most of these nodes were not identified by spectral library matches; however, we noted many mass shifts of 28 Da, corresponding to two links in a fatty acid chain (i.e., CH2-CH2). By following the mass shifts through the molecular network, we can assign identities to additional nodes such as hexanoyl/octanoyl-carnitine, which have stronger associations to mortality than the initially identified decanoyl-carnitine. Many of these nodes were also assigned as ACs by ClassyFire. Interestingly, there is a subset of this network that is moderately related to ACs, which is also associated with mortality and possesses 28 Da mass shifts suggestive of a fatty acid chain (Figure 5G circle). Determining the molecular structures of these compounds and how they impact fatty acid metabolism would give us a deeper understanding of metabolic dysfunction associated with SaB patient outcomes.

Knowledge-Based Analysis of Proteome Alterations Captures Underlying Cytokine Mortality Signatures

The above results describe a comprehensive assessment of the molecular features associated with SaB mortality that are amenable to MS-based analyses. However, major cytokine families, the focus of most infectious disease biomarker studies, were underrepresented in our dataset. These signaling molecules fall below the standard limit of detection in typical serum proteomic experiments (Geyer et al., 2017), even in recent attempts at ultra-deep serum proteome coverage (Dey et al., 2019; Keshishian et al., 2015) but have been shown to play major roles in disease. Therefore, we designed a computational approach to infer the relative importance of major cytokine families from our proteomics data using functional protein association networks (Figure 6A workflow description in STAR Methods). To benchmark our approach, we compared the results to ingenuity pathway analysis (IPA) applied to the same clusters. We found that the cytokine prediction scores were highly correlated between both strategies, particularly for IL-6, TGF-β1, TNF, IL-1β, and IL-10 (Core-5 cytokines, Figure 6B). Notably, four of these Core-5 cytokines have demonstrated associations with SaB mortality and/or duration in previous studies
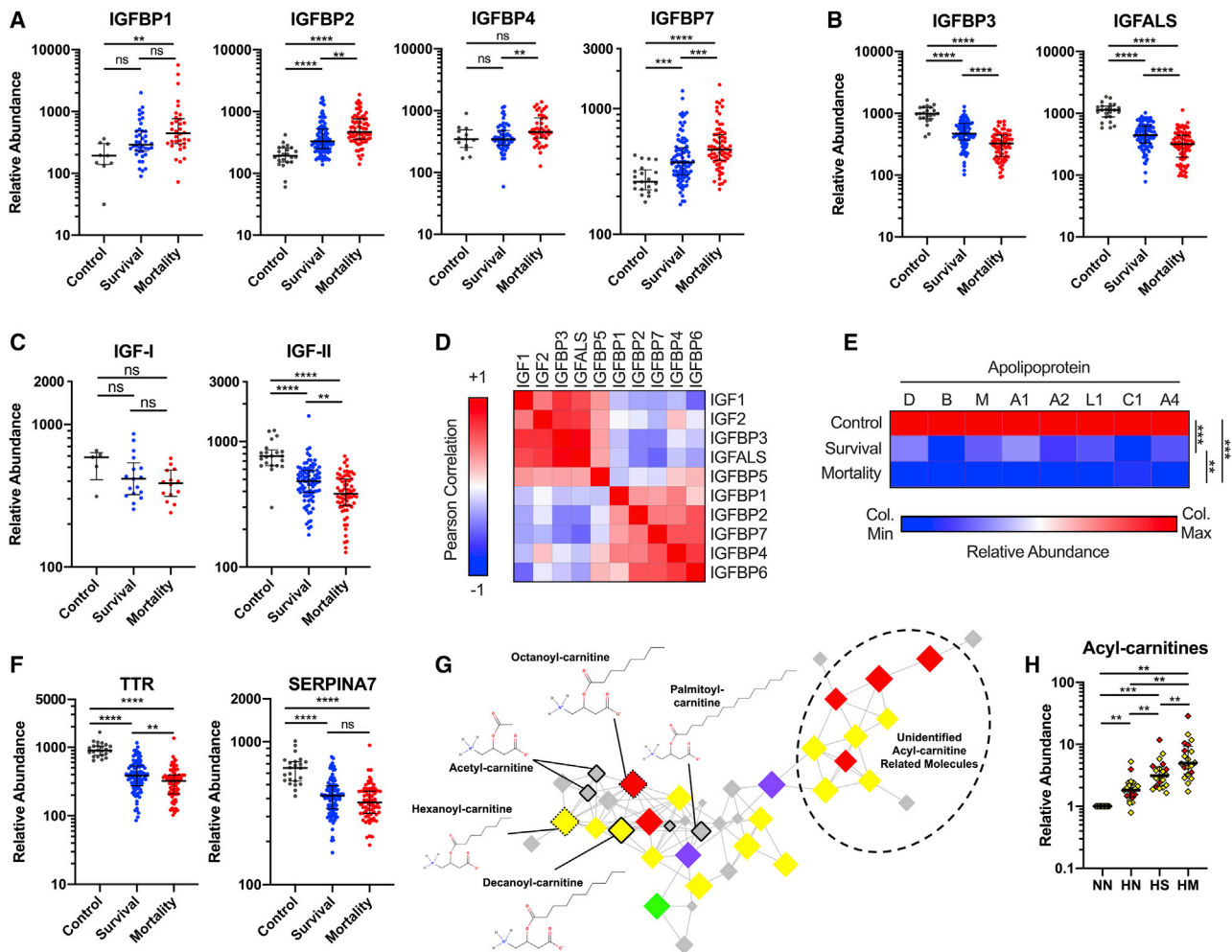
Figure 5. Detection of Metabolic Dysfunction in SaB Mortality Patients

(A and B) Abundance of IGFBP (A) binary and (B) ternary complex members.

(C) Abundance of IGFI and II.

(D) Correlation matrix of IGF-related proteins.

(E) Heatmap of apolipoprotein abundance.

(F) Abundance of thyroxine-binding serum proteins.

(G and H) Molecular network (G) and abundance of acyl-carnitines (H). In (G) and (H), nodes and points are colored according to Figure 4D. In (G), nodes are sized according to ANOVA -$\log_{10}$(p value).

Error bars represent interquartile range (IQR). For all tests, significance values are denoted as follows: ****p < 0.0001; ***p < 0.001; **p < 0.01; *p < 0.05; ns, not significant. For (A), (B), (C), and (F), ANOVA with Tukey's multiple comparison test significance is displayed. For (E) and (H), repeated-measures one-way ANOVA with Holms-Sidak's multiple comparison test significance is displayed. See also Table S2 and S3.

(TNF, IL-10, and IL-1b [Rose et al., 2012, 2017]; IL-6 [Guimaraes et al., 2019]), providing validation for this approach.

In addition to cytokines, IPA reports enrichments for a variety of categories, such as endogenous molecules, drugs, and enzymes. We found that these extended IPA analyses validated results observed in the MS data such as thyroid hormone (T3), IGF system (IGFI and II), lipid metabolism (LPL, CETP, LIPC), and infection/inflammatory responses (H_2O_2, MAPK, and MAPKK inhibitors) (Figure S7A). Although the IPA analysis largely agreed with our approach, there were differences (IPA: IL-17, IL-1α, and IL-11; cytokine inference: CCL2, CXCL8, and IL-18 [Figure S7B]). These differences can be explained by annotation

biases of each tool as even the co-predicted cytokines have marginal overlap of target proteins (~33%; Figures S7C and S7D). Nevertheless, cytokines uniquely predicted from each tool have also been linked to SaB mortality or duration by previous studies (IL-17, CCL2, CXCL8, and IL-18 [Guimaraes et al., 2019]). Thus, while the Core-5 cytokines captured the most probable contributors, a combination of both strategies provides a more complete view of the underlying cytokine signature and downstream effectors.

Focusing in on the Core-5 cytokines, we found that most of the connections were to proteins in pMortality+ (Figure 6C). Regenerating a protein association network using only proteins from
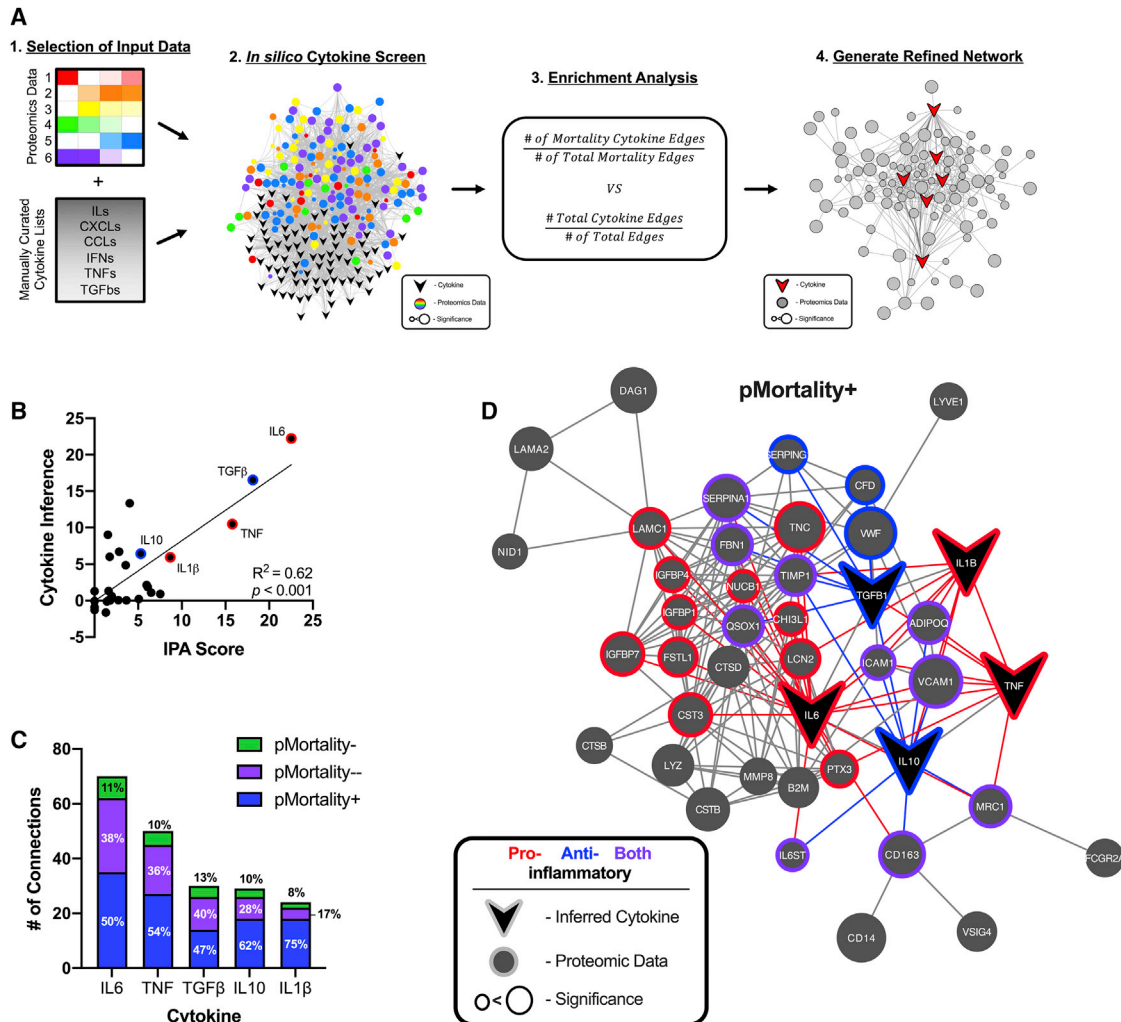
Figure 6. Knowledge-Based Analysis of Cytokines Predicts Major Contributors to Proteomic Alterations and Identifies Core of Modulated Proteins

(A) Schematic for cytokine inference analysis.

(B) Correlation of the cytokine inference score and IPA upstream regulator analysis score. The Core-5 cytokines are highlighted according to their inflammatory actions (red, pro-inflammatory; blue, anti-inflammatory).

(C) Edges between the Core-5 cytokines and each mortality-associated K-means cluster as determined by STRING-db.

(D) Refined network of Core-5 cytokines and pMortality$^+$ proteins. Protein nodes are sized according to -log$_{10}$(p value) determined via ANOVA and highlighted based their connections to pro-inflammatory (red), anti-inflammatory cytokines (blue), or both (purple). Cytokine node outlines and neighboring edges are colored based on pro-inflammatory (red) or anti-inflammatory (blue) activity.

In (A) and (C), heatmap, nodes and bars are colored as in Figure 4A. See also Figure S7.

pMortality$^+$ that are directly linked to a Core-5 cytokine yielded a refined network (Figure 6D) that is easier to interpret than the networks initially generated (Figure S7E). Similar networks can be made with the proteins from pMortality$^{-/-}$ (Figures S7F and S7G). Delving into the pMortality$^+$ network, we found a subset of proteins were connected to both pro and anti-inflammatory cytokines (Figure 6D purple circles). Due to the known imbalance of pro- and anti-inflammatory cytokines in SaB, these proteins may be the most interesting for further study. This includes several proteins that contribute to inflammation resolution (e.g., ADIPOQ, MRC1, CD163) and may represent actionable targets for new therapeutic interventions. Together, this analysis pre-

dicted the top cytokines that influence the observed proteome landscape and enables researchers to define disease-associated pathways to test in functional studies.

T4 and Adiponectin Signaling Influences SaB Outcomes *In Vivo*

While the analyses above detail the biomarkers and pathways altered during SaB, it is unclear whether they are simply bystanders or functionally contribute to disease outcomes. To address this gap, we utilized a mouse model of SaB to assess the influence thyroid hormones and adiponectin signaling on bacterial burden and overall survival.
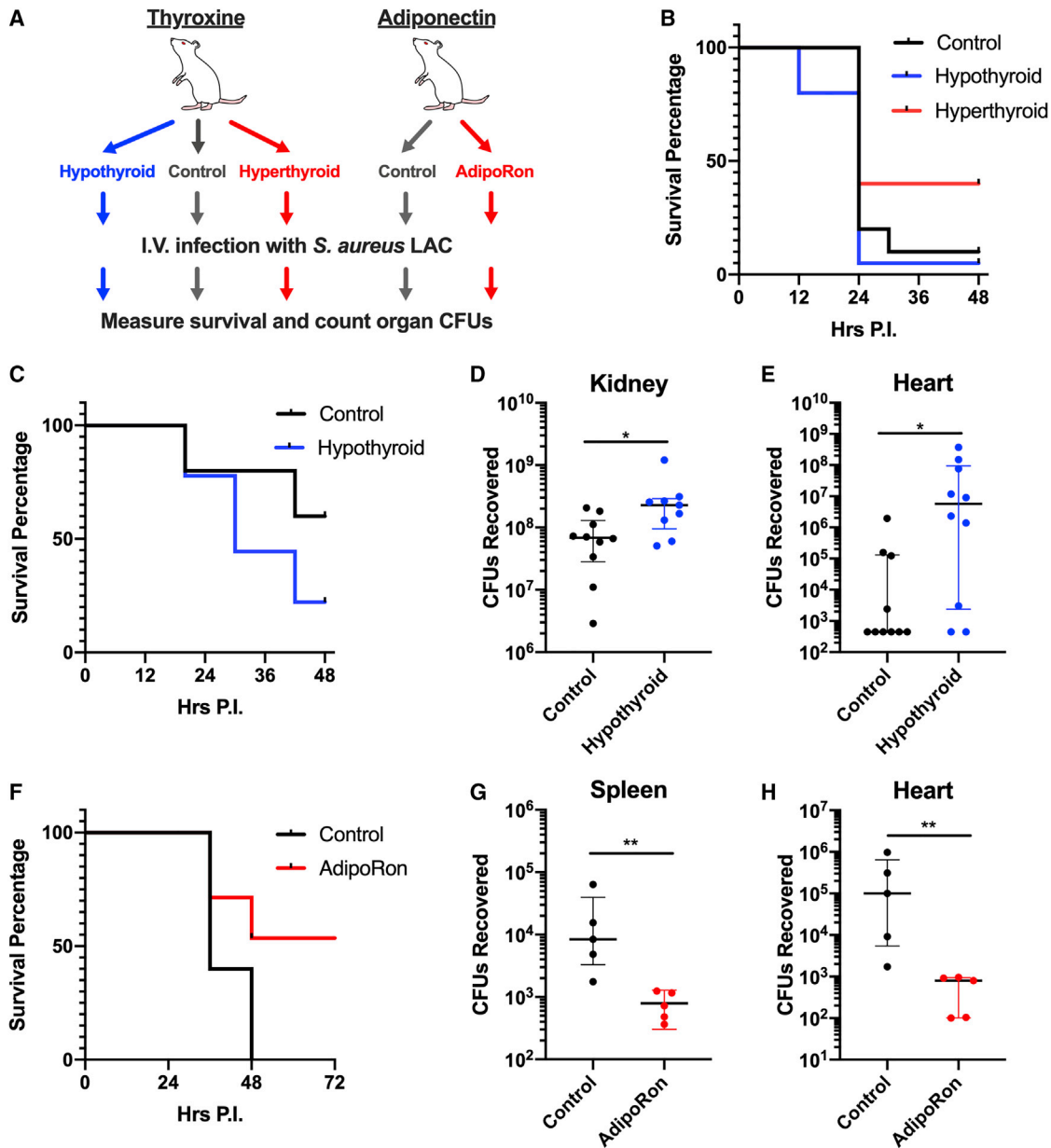
Figure 7. Thyroid and Adiponectin Signaling Contributes to SaB Mortality *In Vivo*

(A) Schematic for treatment plan and mouse model of SaB.

(B) Survival curve of mice given hyperthyroid, hypothyroid, or control treatments and then infected.

(C) Survival curve of mice given hypothyroid or control treatments and then infected.

(D and E) CFUs recovered from the kidney (D) and heart (E) in hypothyroid or control mice 48 h after infection.

(F) Survival curve of mice given AdipoRon or control treatments and then infected.

(G and H) CFUs recovered from the spleen (G) and heart (H) in AdipoRon or control mice 48 h after infection.

Error bars represent interquartile range (IQR). For all tests, significance values are denoted as follows: ****$p < 0.0001$; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; ns, not significant. All infections were with 5×10^7 CFU *S. aureus* except for (B) (1×10^8 CFU). For (D), (E), (G), and (H), MWU test significance is displayed.

Our multi-omic analysis captured a dysregulation of host metabolism (Figure 5), including reduced levels of T4 in mortality patients. Exogenous T4 has been previously shown to be protective in mouse and rat models of polymicrobial sepsis (Al-Abed et al., 2011); however, its contributions to SaB remain unclear.

Further, the impact of a hypothyroid state on SaB has not been tested. To address these questions, we designed an animal experiment to test whether altering T4 levels could affect survival in a mouse SaB model (Figure 7A). We treated mice with either a hypo- or hyperthyroid treatment (Al-Abed et al., 2011; Tsourdi

et al., 2015) and then intravenously infected the mice and assessed survival. We found that hypothyroid mice had higher mortality rate than control mice while the hyperthyroid group had a four times greater survival at 48 h post-infection (p.i.) (Figure 7B). While the hypothyroid mice died more rapidly, control animals also succumbed to infection, leaving a small window to observe differences between the groups. To clarify these results, we repeated the hypothyroid infection with a lower dose of *S. aureus* (50% original inoculum) and harvested organs for bacterial enumeration. Again, the hypothyroid mice had increased mortality than the control group (Figure 7C), and, consistently, the mice surviving at 48 h p.i. had increased bacterial load in their hearts (Figure 7D) and kidneys (Figure 7E), indicating a defect in bacterial clearance.

While the association of T4 with SaB has not been previously described, dysregulated cytokine production has been previously associated with SaB mortality and is supported by the cytokine-inference approach described above (Figure 6). Anti-inflammatory proteins associated with prominent cytokine signatures may play a role in suppressing the overwhelming immune response observed in bacteremia patients. One of these anti-inflammatory proteins, ADIPOQ (adiponectin), had not been linked to SaB, but it is known to induce IL-10 in leukocytes (Wolf et al., 2004). Given that IL-10 is protective in a mouse SaB model (Leech et al., 2017), we hypothesized that targeting adiponectin could also improve survival outcomes. To test this hypothesis, we treated mice with a small-molecule activator of the adiponectin receptor, AdipoRon, or vehicle control (Figure 7A), and utilized the same experimental scheme described above. Treatment with AdipoRon markedly enhanced mouse survival (Figure 7F) and significantly reduced organ colony-forming units (CFUs) (Figures 7G and 7H). Altogether, these *in vivo* studies demonstrate that stimulation of both the thyroid hormone system and adiponectin receptor is protective in a mouse SaB model.

DISCUSSION

The traditional strategy for defining biomarkers for infectious diseases has been based on a subset of immunological parameters. Here, we establish a new standard of infection-related biomarker assessment by examining a much broader host response profile, eliminating the assumption that all clinically relevant details are immunological. Through a multi-omic approach, we define numerous features and multi-variate models that can accurately predict SaB patient mortality. These features can be paired with previously described cytokine markers, quantified with more sensitive immunoassays, to enhance prognostic value. Further investigation is needed to determine whether these markers are specific to SaB or conserved in response to other infections. To better understand the biology underlying SaB mortality, we expanded this study through the application of additional computational analyses, an in-depth interrogation of mortality-relevant alterations, and *in vivo* validation of therapeutic relevance.

Crude predictors of SaB mortality can be based upon clinical assessments of the patient (Hawkins et al., 2007; Pastagia et al., 2012); however, they lack the sensitivity and specificity required

to serve as reliable stratification methods upon which to individualize or de-escalate therapy. Due to this ambiguity, treatment decisions in SaB are based upon "one-size-fits-all" protocols that originate from empirical clinical experience developed throughout the antibiotic era (Liu et al., 2011). As a result, improvements in mortality in MRSA bacteremia have not kept pace with other fields of medicine over the past three decades, despite better drugs and faster diagnostics (Fowler et al., 2006; Rehm et al., 2008). While combination therapies (e.g., daptomycin plus ceftaroline) offer an appealing approach to improve survival in SaB, these drugs cost >50× more than vancomycin, posing considerable economic constraints. Attempts at developing cheaper combinations have been wrought with toxicity (Burgess and Drew, 2014; Gomes et al., 2014; Tong et al., 2020). Utilizing the biomarkers uncovered herein to identify the 20%–30% of patients with high mortality risk on standard therapy would provide a compelling advance in the management of SaB.

In addition to defining standard protein and metabolite biomarkers for SaB mortality, we utilized two computational strategies for a deeper analysis. First, using a workflow for the prediction and identification of PTMs, we revealed that the paucity of identifications in the serum proteome likely derives from modified peptides, including both serum glycoproteins and small PTMs. Using refined database-searching techniques resulted in the identification of our top predictive biomarkers, glycosylated peptides derived from fetuin A. Glycosylation has been used as biomarkers for various diseases, including cancer (Silsirivanit, 2019), Alzheimer disease (Regan et al., 2019), and inflammatory conditions (Gornik and Lauc, 2008). However, this is the first time that host glycosylation patterns have been linked to human SaB mortality, which could provide a useful clinical tool in the future.

Intriguingly, our top unmodified biomarker was fetuin B, and our top modified biomarker was glycosylation of fetuin A. Fetuins belong to the cystatin superfamily of proteins (Dabrowska et al., 2015; Olivier et al., 2000) and can transport fatty acids in the bloodstream (Cayatte et al., 1990). Both fetuin A and B are studied in metabolic disorders such as obesity and diabetes; however, their expression is increased in these diseases rather than decreased as observed in SaB. Notably, fetuin A has also been shown to exert anti-inflammatory effects, and supplementation is protective in mouse models of systemic inflammation (Cayatte et al., 1990). One proposed mechanism of the protective effects of fetuin A appears to be through enhancement of spermine-mediate macrophage deactivation, potentially limiting immunopathology. Future studies are needed to determine whether similar mechanisms are at play in SaB as well as what role fetuin B plays in this process. Regardless, both proteins present an intriguing link between metabolism and bacteremia and can now be classified as biomarkers of SaB mortality.

Another PTM finding was an increase in carbamylation of albumin and serum transferrin in mortality patients. Protein carbamylation is a non-enzymatic PTM (Jaisson et al., 2018) that is related to a number of pathologies, including chronic kidney disease (Berg et al., 2013) and rheumatoid arthritis (Pruijn, 2015). In fact, multiple studies have proposed carbamylation of albumin as a prognostic factor for mortality in patients with kidney failure

(Berg et al., 2013; Kalim et al., 2013). Kidney disease and SaB are intimately related (Alobaidi et al., 2015; Nielsen et al., 2015), and the common signature of carbamylation suggests an underlying pathological process. Further, patients with rheumatoid arthritis, which can also be linked to *S. aureus* infections (Joost et al., 2017) and colonization (Goodman et al., 2019), are reported to have more anti-carbamyl antibodies (Shi et al., 2014). Whether this modification is pathological or simply a marker of disease severity requires additional experiments; nevertheless, it appears linked with a variety of disease states.

The second computational advancement in serum bio-analytics utilized in this study is the inferring of cytokine signatures from serum proteomics data. This analysis predicted major alterations in IL-6, TGF-β1, TNF, IL-1β, and IL-10 in mortality samples, all of which were validated by an orthogonal approach (i.e., IPA) and have exhibited associations to SaB human mortality in previous studies (Guimaraes et al., 2019; Minejima et al., 2016; Rose et al., 2012, 2017). This approach also enables researchers to link these major cytokine players to the observed proteomic data, facilitating the construction of testable hypotheses (such as the impact of adiponectin signaling in SaB). Together, this method refined host response pathway analysis and identified unreported potential players in SaB.

In addition to defining SaB mortality biomarkers, we sought to gain a deeper understanding of the host response to SaB through clustering and network-based analyses. Unexpectedly, the most striking findings from this were not related to the immune system but rather a dysfunction of metabolism. While some of our findings have been previously described, such as the suppression of serum lipoproteins and T4 during severe infections, we also captured surprising signatures of metabolic dysfunction, specifically linked to mortality. The most salient of these was the apparent shift from ternary to binary IGF-IGFBP complexes, resulting in lower circulating IGF levels, and the increase of ACs and related molecular species in SaB mortality patients. The ultimate functional outcome of these perturbations is unclear; however, they may help uncover alternative therapeutic avenues by which to stabilize patients while providing antimicrobial treatments.

Finally, we demonstrated that stimulation of both thyroid and adiponectin signaling pathways can enhance mouse survival in experimental SaB. Previous studies into thyroid signaling suggest inhibition of macrophage migration inhibitory factor (Al-Abed et al., 2011) or enhancement of intracellular bacterial killing (Chen et al., 2012) are responsible for its protective effects. In contrast, adiponectin is mainly studied for its role in insulin resistance and diabetes (Achari and Jain, 2017). However, it also has anti-inflammatory, cardio-protective, and vaso-protective effects (Achari and Jain, 2017), and adiponectin KO mice are more susceptible to polymicrobial sepsis (Teoh et al., 2008). In agreement, our data indicate a protective role for adiponectin signaling in SaB infection. Importantly, both T4 (Tanguay et al., 2019) and AdipoRon (Okada-Iwabu et al., 2013) are orally bioavailable, and T4 is FDA approved. If T4 or AdipoRon could offer protection in humans, they may be explored as adjunctive approaches to antibiotics for treating SaB.

Overall, we aimed to set a high standard in the infectious disease biomarker field by providing an accurate, multi-omic model (including PTMs) for predicting SaB mortality. Conducting future studies to the same depth and rigor will likely uncover additional clinically useful findings and lead to a deeper understanding of mortality in infection. Ultimately, this study sets the groundwork for a multi-marker-based tool for the rapid prediction of SaB patient mortality at the time of clinical presentation: the Rapid Index of SaB Mortality Kinetics (RISK) test.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human Studies
  - Mouse Models
- METHOD DETAILS
  - Patient and Isolate Identification and Collection
  - Clinical Measurements and Outcomes
  - Serum Metabolite Extraction
  - Metabolomic LC-MS^2 Analysis
  - Metabolite Molecular Networking and Identification by GNPS
  - Serum Protein Digestion and Labeling
  - Basic pH Reverse-phase Liquid Chromatography (bRPLC) Fractionation
  - Proteomic LC-MS^2/MS^3 Analysis
  - Peptide Identification by Proteome Discoverer
  - Statistical Analyses of MS Data
  - Gene Ontology (GO) and Network-based Analyses
  - Mouse Model of SaB
- QUANTIFICATION AND STATISTICAL ANALYSIS

AUTHOR CONTRIBUTIONS

Conceptualization, J.M.W., G.S., W.R., and D.J.G.; Methodology, J.M.W., J.O., J.R.C., G.D.S.-P., G.Y.L., V.N., and D.J.G.; Software, J.M.W.; Validation, J.M.W. and M.C.-T.; Formal Analysis, J.M.W., R.H.M., J.R.C., G.D.S.-P., and F.V.; Investigation, J.M.W., R.H.M., J.O., J.R.C., G.D.S.-P., M.C.-T., C.-M.T. and F.V.; Resources, R.K., P.C.D., G.Y.L., V.N., G.S., W.R., and D.J.G.; Data Curation, J.M.W., W.R., and D.J.G.; Writing – Original Draft, J.M.W., G.S.,

REFERENCES

Achari, A.E., and Jain, S.K. (2017). Adiponectin, a Therapeutic Target for Obesity, Diabetes, and Endothelial Dysfunction. Int. J. Mol. Sci. *18*. Published online June 21, 2017. https://doi.org/10.3390/ijms18061321.

Al-Abed, Y., Metz, C.N., Cheng, K.F., Aljabari, B., VanPatten, S., Blau, S., Lee, H., Ochani, M., Pavlov, V.A., Coleman, T., et al. (2011). Thyroxine is a potential endogenous antagonist of macrophage migration inhibitory factor (MIF) activity. Proc. Natl. Acad. Sci. USA *108*, 8224–8227.

Alobaidi, R., Basu, R.K., Goldstein, S.L., and Bagshaw, S.M. (2015). Sepsis-associated acute kidney injury. Semin. Nephrol. *35*, 2–11.

Baxter, R.C. (2014). IGF binding proteins in cancer: mechanistic and clinical insights. Nat. Rev. Cancer *14*, 329–341.

Bello, G., Pennisi, M.A., Montini, L., Silva, S., Maviglia, R., Cavallaro, F., Bianchi, A., De Marinis, L., and Antonelli, M. (2009). Nonthyroidal illness syndrome and prolonged mechanical ventilation in patients admitted to the ICU. Chest *135*, 1448–1454.

Berg, A.H., Drechsler, C., Wenger, J., Buccafusca, R., Hod, T., Kalim, S., Ramma, W., Parikh, S.M., Steen, H., Friedman, D.J., et al. (2013). Carbamylation of serum albumin as a risk factor for mortality in patients with kidney failure. Sci. Transl. Med. *5*, 175ra29.

Bern, M., Kil, Y.J., and Becker, C. (2012). Byonic: advanced peptide and protein identification software. Curr. Protoc. Bioinformatics. Published online December 2012. https://doi.org/10.1002/0471250953.bi1320s40.

Burgess, L.D., and Drew, R.H. (2014). Comparison of the incidence of vancomycin-induced nephrotoxicity in hospitalized patients with and without concomitant piperacillin-tazobactam. Pharmacotherapy *34*, 670–676.

Cao, L., Tolić, N., Qu, Y., Meng, D., Zhao, R., Zhang, Q., Moore, R.J., Zink, E.M., Lipton, M.S., Paša-Tolić, L., and Wu, S. (2014). Characterization of intact N- and O-linked glycopeptides using higher energy collisional dissociation. Anal. Biochem. *452*, 96–102.

Cayatte, A.J., Kumbla, L., and Subbiah, M.T. (1990). Marked acceleration of exogenous fatty acid incorporation into cellular triglycerides by fetuin. J. Biol. Chem. *265*, 5883–5888.

Chandramouli, K., and Qian, P.Y. (2009). Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. Hum. Genomics Proteomics *2009*, 239204.

Chen, Y., Sjölinder, M., Wang, X., Altenbacher, G., Hagner, M., Berglund, P., Gao, Y., Lu, T., Jonsson, A.B., and Sjölinder, H. (2012). Thyroid hormone enhances nitric oxide-mediated bacterial clearance and promotes survival after meningococcal infection. PLoS ONE *7*, e41445.

Chick, J.M., Kolippakkam, D., Nusinow, D.P., Zhai, B., Rad, R., Huttlin, E.L., and Gygi, S.P. (2015). A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides. Nat. Biotechnol. *33*, 743–749.

Christoffersen, C., and Nielsen, L.B. (2012). Apolipoprotein M–a new biomarker in sepsis. Crit. Care *16*, 126.

Clerc, F., Reiding, K.R., Jansen, B.C., Kammeijer, G.S., Bondt, A., and Wuhrer, M. (2016). Human plasma protein N-glycosylation. Glycoconj. J. *33*, 309–343.

Dabrowska, A.M., Tarach, J.S., Wojtysiak-Duma, B., and Duma, D. (2015). Fetuin-A (AHSG) and its usefulness in clinical practice. Review of the literature. Biomed. Pap. Med. Fac. Univ. Palacky Olomouc Czech Repub. *159*, 352–359.

Dakna, M., Harris, K., Kalousis, A., Carpentier, S., Kolch, W., Schanstra, J.P., Haubitz, M., Vlahou, A., Mischak, H., and Girolami, M. (2010). Addressing the challenge of defining valid proteomic biomarkers and classifiers. BMC Bioinformatics *11*, 594.

Dellinger, R.P., Levy, M.M., Rhodes, A., Annane, D., Gerlach, H., Opal, S.M., Sevransky, J.E., Sprung, C.L., Douglas, I.S., Jaeschke, R., et al.; Surviving Sepsis Campaign Guidelines Committee including the Pediatric Subgroup (2013). Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. Crit. Care Med. *41*, 580–637.

Devabhaktuni, A., Lin, S., Zhang, L., Swaminathan, K., Gonzalez, C.G., Olsson, N., Pearlman, S.M., Rawson, K., and Elias, J.E. (2019). TagGraph reveals vast protein modification landscapes from large tandem mass spectrometry datasets. Nat. Biotechnol. *37*, 469–479.

Dey, K.K., Wang, H., Niu, M., Bai, B., Wang, X., Li, Y., Cho, J.H., Tan, H., Mishra, A., High, A.A., et al. (2019). Deep undepleted human serum proteome profiling toward biomarker discovery for Alzheimer's disease. Clin. Proteomics *16*, 16.

Djoumbou Feunang, Y., Eisner, R., Knox, C., Chepelev, L., Hastings, J., Owen, G., Fahy, E., Steinbeck, C., Subramanian, S., Bolton, E., et al. (2016). ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. J. Cheminform. *8*, 61.

Dührkop, K., Shen, H., Meusel, M., Rousu, J., and Böcker, S. (2015). Searching molecular structure databases with tandem mass spectra using CSI:FingerID. Proc. Natl. Acad. Sci. USA *112*, 12580–12585.

Elias, J.E., and Gygi, S.P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods *4*, 207–214.

Elias, J.E., Haas, W., Faherty, B.K., and Gygi, S.P. (2005). Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. Nat. Methods *2*, 667–675.

Eng, J.K., McCormack, A.L., and Yates, J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J. Am. Soc. Mass Spectrom. *5*, 976–989.

Ensor, J.E. (2014). Biomarker validation: common data analysis concerns. Oncologist *19*, 886–891.

Ersoy, S.C., Heithoff, D.M., Barnes, L., 5th, Tripp, G.K., House, J.K., Marth, J.D., Smith, J.W., and Mahan, M.J. (2017). Correcting a Fundamental Flaw in the Paradigm for Antimicrobial Susceptibility Testing. EBioMedicine *20*, 173–181.

Ferrer, R., Martin-Loeches, I., Phillips, G., Osborn, T.M., Townsend, S., Dellinger, R.P., Artigas, A., Schorr, C., and Levy, M.M. (2014). Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program. Crit. Care Med. *42*, 1749–1755.

Fowler, V.G., Jr., Boucher, H.W., Corey, G.R., Abrutyn, E., Karchmer, A.W., Rupp, M.E., Levine, D.P., Chambers, H.F., Tally, F.P., Vigliani, G.A., et al.; S. aureus Endocarditis and Bacteremia Study Group (2006). Daptomycin versus standard therapy for bacteremia and endocarditis caused by Staphylococcus aureus. N. Engl. J. Med. *355*, 653–665.

Frantzi, M., Bhat, A., and Latosinska, A. (2014). Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development. Clin. Transl. Med. *3*, 7.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. J. Stat. Softw. *33*, 1–22.

Geyer, P.E., Holdt, L.M., Teupser, D., and Mann, M. (2017). Revisiting biomarker discovery by plasma proteomics. Mol. Syst. Biol. *13*, 942.

Gomes, D.M., Smotherman, C., Birch, A., Dupree, L., Della Vecchia, B.J., Kraemer, D.F., and Jankowski, C.A. (2014). Comparison of acute kidney injury

during treatment with vancomycin in combination with piperacillin-tazobactam or cefepime. Pharmacotherapy *34*, 662–669.

Goodman, S.M., Nocon, A.A., Selemon, N.A., Shopsin, B., Fulmer, Y., Decker, M.E., Grond, S.E., Donlin, L.T., Figgie, M.P., Sculco, T.P., et al. (2019). Increased Staphylococcus aureus Nasal Carriage Rates in Rheumatoid Arthritis Patients on Biologic Therapy. J. Arthroplasty *34*, 954–958.

Gornik, O., and Lauc, G. (2008). Glycosylation of serum proteins in inflammatory diseases. Dis. Markers *25*, 267–278.

Guimaraes, A.O., Cao, Y., Hong, K., Mayba, O., Peck, M.C., Gutierrez, J., Ruffin, F., Carrasco-Triguero, M., Dinoso, J.B., Clemenzi-Allen, A., et al. (2019). A Prognostic Model of Persistent Bacteremia and Mortality in Complicated Staphylococcus aureus Bloodstream Infection. Clin. Infect. Dis. *68*, 1502–1511.

Gupta, N., and Pevzner, P.A. (2009). False discovery rates of protein identifications: a strike against the two-peptide rule. J. Proteome Res. *8*, 4173–4181.

Haas, W., Faherty, B.K., Gerber, S.A., Elias, J.E., Beausoleil, S.A., Bakalarski, C.E., Li, X., Villén, J., and Gygi, S.P. (2006). Optimization and use of peptide mass measurement accuracy in shotgun proteomics. Mol. Cell. Proteomics *5*, 1326–1337.

Hawkins, C., Huang, J., Jin, N., Noskin, G.A., Zembower, T.R., and Bolon, M. (2007). Persistent Staphylococcus aureus bacteremia: an analysis of risk factors and outcomes. Arch. Intern. Med. *167*, 1861–1867.

He, Z., and Yu, W. (2010). Stable feature selection for biomarker discovery. Comput. Biol. Chem. *34*, 215–225.

Holland, T.L., Arnold, C., and Fowler, V.G., Jr. (2014). Clinical management of Staphylococcus aureus bacteremia: a review. JAMA *312*, 1330–1341.

Huang, D.W., Sherman, B.T., Tan, Q., Collins, J.R., Alvord, W.G., Roayaei, J., Stephens, R., Baseler, M.W., Lane, H.C., and Lempicki, R.A. (2007). The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. *8*, R183.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Jaisson, S., Pietrement, C., and Gillery, P. (2018). Protein Carbamylation: Chemistry, Pathophysiological Involvement, and Biomarkers. Adv. Clin. Chem. *84*, 1–38.

Joost, I., Kaasch, A., Pausch, C., Peyerl-Hoffmann, G., Schneider, C., Voll, R.E., Seifert, H., Kern, W.V., and Rieg, S. (2017). Staphylococcus aureus bacteremia in patients with rheumatoid arthritis - Data from the prospective INSTINCT cohort. J. Infect. *74*, 575–584.

Kalim, S., Tamez, H., Wenger, J., Ankers, E., Trottier, C.A., Deferio, J.J., Berg, A.H., Karumanchi, S.A., and Thadhani, R.I. (2013). Carbamylation of serum albumin and erythropoietin resistance in end stage kidney disease. Clin. J. Am. Soc. Nephrol. *8*, 1927–1934.

Kern, W.V. (2010). Management of Staphylococcus aureus bacteremia and endocarditis: progresses and challenges. Curr. Opin. Infect. Dis. *23*, 346–358.

Keshishian, H., Burgess, M.W., Gillette, M.A., Mertins, P., Clauser, K.R., Mani, D.R., Kuhn, E.W., Farrell, L.A., Gersten, R.E., and Carr, S.A. (2015). Multiplexed, Quantitative Workflow for Sensitive Biomarker Discovery in Plasma Yields Novel Candidates for Early Myocardial Injury. Mol. Cell. Proteomics *14*, 2375–2393.

Lapek, J.D., Jr., Greninger, P., Morris, R., Amzallag, A., Pruteanu-Malinici, I., Benes, C.H., and Haas, W. (2017a). Detection of dysregulated protein-association networks by high-throughput proteomics predicts cancer vulnerabilities. Nat. Biotechnol. *35*, 983–989.

Lapek, J.D., Jr., Lewinski, M.K., Wozniak, J.M., Guatelli, J., and Gonzalez, D.J. (2017b). Quantitative Temporal Viromics of an Inducible HIV-1 Model Yields Insight to Global Host Targets and Phospho-Dynamics Associated with Vpr. Mol. Cell. Proteomics *16*, 1447–1461.

Leech, J.M., Lacey, K.A., Mulcahy, M.E., Medina, E., and McLoughlin, R.M. (2017). IL-10 Plays Opposing Roles during *Staphylococcus aureus* Systemic and Localized Infections. J. Immunol. *198*, 2352–2365.

Levey, A.S., Perrone, R.D., and Madias, N.E. (1988). Serum creatinine and renal function. Annu. Rev. Med. *39*, 465–490.

Liu, A., Bui, T., Van Nguyen, H., Ong, B., Shen, Q., and Kamalasena, D. (2010). Serum C-reactive protein as a biomarker for early detection of bacterial infection in the older patient. Age Ageing *39*, 559–565.

Liu, C., Bayer, A., Cosgrove, S.E., Daum, R.S., Fridkin, S.K., Gorwitz, R.J., Kaplan, S.L., Karchmer, A.W., Levine, D.P., Murray, B.E., et al.; Infectious Diseases Society of America (2011). Clinical practice guidelines by the infectious diseases society of america for the treatment of methicillin-resistant Staphylococcus aureus infections in adults and children. Clin. Infect. Dis. *52*, e18–e55.

Mayampurath, A.M., Wu, Y., Segu, Z.M., Mechref, Y., and Tang, H. (2011). Improving confidence in detection and characterization of protein N-glycosylation sites and microheterogeneity. Rapid Commun. Mass Spectrom. *25*, 2007–2019.

McAlister, G.C., Huttlin, E.L., Haas, W., Ting, L., Jedrychowski, M.P., Rogers, J.C., Kuhn, K., Pike, I., Grothe, R.A., Blethrow, J.D., and Gygi, S.P. (2012). Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. Anal. Chem. *84*, 7469–7478.

McAlister, G.C., Nusinow, D.P., Jedrychowski, M.P., Wühr, M., Huttlin, E.L., Erickson, B.K., Rad, R., Haas, W., and Gygi, S.P. (2014). MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. Anal. Chem. *86*, 7150–7158.

McCoin, C.S., Knotts, T.A., and Adams, S.H. (2015). Acylcarnitines–old actors auditioning for new roles in metabolic physiology. Nat. Rev. Endocrinol. *11*, 617–625.

Mi, H., Muruganujan, A., Casagrande, J.T., and Thomas, P.D. (2013). Large-scale gene function analysis with the PANTHER classification system. Nat. Protoc. *8*, 1551–1566.

Mikkelsen, M.E., Miltiades, A.N., Gaieski, D.F., Goyal, M., Fuchs, B.D., Shah, C.V., Bellamy, S.L., and Christie, J.D. (2009). Serum lactate is associated with mortality in severe sepsis independent of organ failure and shock. Crit. Care Med. *37*, 1670–1677.

Minejima, E., Bensman, J., She, R.C., Mack, W.J., Tuan Tran, M., Ny, P., Lou, M., Yamaki, J., Nieberg, P., Ho, J., and Wong-Beringer, A. (2016). A Dysregulated Balance of Proinflammatory and Anti-Inflammatory Host Cytokine Response Early During Therapy Predicts Persistence and Mortality in Staphylococcus aureus Bacteremia. Crit. Care Med. *44*, 671–679.

Neumann, U., Genze, N., and Heider, D. (2017). EFS: an ensemble feature selection tool implemented as R-package and web-application. BioData Min. *10*, 21.

Nielsen, L.H., Jensen-Fangel, S., Benfield, T., Skov, R., Jespersen, B., Larsen, A.R., Østergaard, L., Støvring, H., Schønheyder, H.C., and Søgaard, O.S. (2015). Risk and prognosis of Staphylococcus aureus bacteremia among individuals with and without end-stage renal disease: a Danish, population-based cohort study. BMC Infect. Dis. *15*, 6.

Okada-Iwabu, M., Yamauchi, T., Iwabu, M., Honma, T., Hamagami, K., Matsuda, K., Yamaguchi, M., Tanabe, H., Kimura-Someya, T., Shirouzu, M., et al. (2013). A small-molecule AdipoR agonist for type 2 diabetes and short life in obesity. Nature *503*, 493–499.

Olivier, E., Soury, E., Ruminy, P., Husson, A., Parmentier, F., Daveau, M., and Salier, J.P. (2000). Fetuin-B, a second member of the fetuin family in mammals. Biochem. J. *350*, 589–597.

Park, J.H., Kim, D.H., Jang, H.R., Kim, M.J., Jung, S.H., Lee, J.E., Huh, W., Kim, Y.G., Kim, D.J., and Oh, H.Y. (2014). Clinical relevance of procalcitonin and C-reactive protein as infection markers in renal impairment: a cross-sectional study. Crit. Care *18*, 640.

Pastagia, M., Kleinman, L.C., Lacerda de la Cruz, E.G., and Jenkins, S.G. (2012). Predicting risk for death from MRSA bacteremia. Emerg. Infect. Dis. *18*, 1072–1080.

Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J., and Gygi, S.P. (2003). Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J. Proteome Res. *2*, 43–50.

Plikat, K., Langgartner, J., Buettner, R., Bollheimer, L.C., Woenckhaus, U., Schölmerich, J., and Wrede, C.E. (2007). Frequency and outcome of patients with nonthyroidal illness syndrome in a medical intensive care unit. Metabolism 56, 239–244.

Pluskal, T., Castillo, S., Villar-Briones, A., and Oresic, M. (2010). MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics 11, 395.

Póvoa, P., Coelho, L., Almeida, E., Fernandes, A., Mealha, R., Moreira, P., and Sabino, H. (2005). C-reactive protein as a marker of infection in critically ill patients. Clin. Microbiol. Infect. 11, 101–108.

Pruijn, G.J. (2015). Citrullination and carbamylation in the pathophysiology of rheumatoid arthritis. Front. Immunol. 6, 192.

Puskarich, M.A., Evans, C.R., Karnovsky, A., Das, A.K., Jones, A.E., and Stringer, K.A. (2018). Septic Shock Nonsurvivors Have Persistently Elevated Acylcarnitines Following Carnitine Supplementation. Shock 49, 412–419.

Ralto, K.M., and Parikh, S.M. (2016). Mitochondria in Acute Kidney Injury. Semin. Nephrol. 36, 8–16.

Rasmussen, R.V., Fowler, V.G., Jr., Skov, R., and Bruun, N.E. (2011). Future challenges and treatment of Staphylococcus aureus bacteremia with emphasis on MRSA. Future Microbiol. 6, 43–56.

Regan, P., McClean, P.L., Smyth, T., and Doherty, M. (2019). Early Stage Glycosylation Biomarkers in Alzheimer's Disease. Medicines (Basel) 6, 92.

Rehm, S.J., Boucher, H., Levine, D., Campion, M., Eisenstein, B.I., Vigliani, G.A., Corey, G.R., and Abrutyn, E. (2008). Daptomycin versus vancomycin plus gentamicin for treatment of bacteraemia and endocarditis due to Staphylococcus aureus: subset analysis of patients infected with methicillin-resistant isolates. J. Antimicrob. Chemother. 62, 1413–1421.

Rohart, F., Gautier, B., Singh, A., and Lê Cao, K.A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS Comput. Biol. 13, e1005752.

Rose, W.E., Eickhoff, J.C., Shukla, S.K., Pantrangi, M., Rooijakkers, S., Cosgrove, S.E., Nizet, V., and Sakoulas, G. (2012). Elevated serum interleukin-10 at time of hospital admission is predictive of mortality in patients with Staphylococcus aureus bacteremia. J. Infect. Dis. 206, 1604–1611.

Rose, W.E., Shukla, S.K., Berti, A.D., Hayney, M.S., Henriquez, K.M., Ranzoni, A., Cooper, M.A., Proctor, R.A., Nizet, V., and Sakoulas, G. (2017). Increased Endovascular Staphylococcus aureus Inoculum Is the Link Between Elevated Serum Interleukin 10 Concentrations and Mortality in Patients With Bacteremia. Clin. Infect. Dis. 64, 1406–1412.

Scheubert, K., Hufsky, F., Petras, D., Wang, M., Nothias, L.F., Dührkop, K., Bandeira, N., Dorrestein, P.C., and Böcker, S. (2017). Significance estimation for large scale metabolomics annotations by spectral matching. Nat. Commun. 8, 1494.

Schiffer, L., Barnard, L., Baranowski, E.S., Gilligan, L.C., Taylor, A.E., Arlt, W., Shackleton, C.H.L., and Storbeck, K.H. (2019). Human steroid biosynthesis, metabolism and excretion are differentially reflected by serum and urine steroid metabolomes: A comprehensive review. J. Steroid Biochem. Mol. Biol. 194, 105439.

Schussler, G.C. (2000). The thyroxine-binding proteins. Thyroid 10, 141–149.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504.

Sharma, N.K., Ferreira, B.L., Tashima, A.K., Brunialti, M.K.C., Torquato, R.J.S., Bafi, A., Assuncao, M., Azevedo, L.C.P., and Salomao, R. (2019). Lipid metabolism impairment in patients with sepsis secondary to hospital acquired pneumonia, a proteomic analysis. Clin. Proteomics 16, 29.

Shi, J., van de Stadt, L.A., Levarht, E.W., Huizinga, T.W., Hamann, D., van Schaardenburg, D., Toes, R.E., and Trouw, L.A. (2014). Anti-carbamylated protein (anti-CarP) antibodies precede the onset of rheumatoid arthritis. Ann. Rheum. Dis. 73, 780–783.

Silsirivanit, A. (2019). Glycosylation markers in cancer. Adv. Clin. Chem. 89, 189–213.

Singer, M. (2014). The role of mitochondrial dysfunction in sepsis-induced multi-organ failure. Virulence 5, 66–72.

Skates, S.J., Gillette, M.A., LaBaer, J., Carr, S.A., Anderson, L., Liebler, D.C., Ransohoff, D., Rifai, N., Kondratovich, M., Težak, Ž., et al. (2013). Statistical design for biospecimen cohort size in proteomics-based biomarker discovery and verification studies. J. Proteome Res. 12, 5383–5394.

Söderquist, B., Sundqvist, K.G., and Vikerfors, T. (1999). Adhesion molecules (E-selectin, intercellular adhesion molecule-1 (ICAM-1) and vascular cell adhesion molecule-1 (VCAM-1)) in sera from patients with Staphylococcus aureus bacteraemia with or without endocarditis. Clin. Exp. Immunol. 118, 408–411.

Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W., Fiehn, O., Goodacre, R., Griffin, J.L., et al. (2007). Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics 3, 211–221.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 47 (D1), D607–D613.

Tanguay, M., Girard, J., Scarsi, C., Mautone, G., and Larouche, R. (2019). Pharmacokinetics and Comparative Bioavailability of a Levothyroxine Sodium Oral Solution and Soft Capsule. Clin. Pharmacol. Drug Dev. 8, 521–528.

Tenenhaus, M., Tenenhaus, A., and Groenen, P.J.F. (2017). Regularized Generalized Canonical Correlation Analysis: A Framework for Sequential Multiblock Component Methods. Psychometrika. Published online May 23, 2017. https://doi.org/10.1007/s11336-017-9573-x.

Teoh, H., Quan, A., Bang, K.W., Wang, G., Lovren, F., Vu, V., Haitsma, J.J., Szmitko, P.E., Al-Omran, M., Wang, C.H., et al. (2008). Adiponectin deficiency promotes endothelial activation and profoundly exacerbates sepsis-related mortality. Am. J. Physiol. Endocrinol. Metab. 295, E658–E664.

Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A.K., and Hamon, C. (2003). Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal. Chem. 75, 1895–1904.

Tolonen, A.C., and Haas, W. (2014). Quantitative proteomics using reductive dimethylation for stable isotope labeling. J. Vis. Exp. Published online July 1, 2014. https://doi.org/10.3791/51416.

Tong, S.Y., Davis, J.S., Eichenberger, E., Holland, T.L., and Fowler, V.G., Jr. (2015). Staphylococcus aureus infections: epidemiology, pathophysiology, clinical manifestations, and management. Clin. Microbiol. Rev. 28, 603–661.

Tong, S.Y.C., Lye, D.C., Yahav, D., Sud, A., Robinson, J.O., Nelson, J., Archuleta, S., Roberts, M.A., Cass, A., Paterson, D.L., et al.; Australasian Society for Infectious Diseases Clinical Research Network (2020). Effect of Vancomycin or Daptomycin With vs Without an Antistaphylococcal β-Lactam on Mortality, Bacteremia, Relapse, or Treatment Failure in Patients With MRSA Bacteremia: A Randomized Clinical Trial. JAMA 323, 527–537.

Tsourdi, E., Rijntjes, E., Köhrle, J., Hofbauer, L.C., and Rauner, M. (2015). Hyperthyroidism and Hypothyroidism in Male Mice and Their Effects on Bone Mass, Bone Turnover, and the Wnt Inhibitors Sclerostin and Dickkopf-1. Endocrinology 156, 3517–3527.

van Hal, S.J., Jensen, S.O., Vaska, V.L., Espedido, B.A., Paterson, D.L., and Gosbell, I.B. (2012). Predictors of mortality in Staphylococcus aureus Bacteremia. Clin. Microbiol. Rev. 25, 362–386.

van Leeuwen, H.J., Heezius, E.C., Dallinga, G.M., van Strijp, J.A., Verhoef, J., and van Kessel, K.P. (2003). Lipoprotein metabolism in patients with severe sepsis. Crit. Care Med. 31, 1359–1366.

Vandecasteele, S.J., Boelaert, J.R., and De Vriese, A.S. (2009). Staphylococcus aureus infections in hemodialysis: what a nephrologist should know. Clin. J. Am. Soc. Nephrol. 4, 1388–1400.

Wang, F.D., Chen, Y.Y., Chen, T.L., and Liu, C.Y. (2008). Risk factors and mortality in patients with nosocomial Staphylococcus aureus bacteremia. Am. J. Infect. Control 36, 118–122.

Wang, Y., Yang, F., Gritsenko, M.A., Wang, Y., Clauss, T., Liu, T., Shen, Y., Monroe, M.E., Lopez-Ferrer, D., Reno, T., et al. (2011). Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. Proteomics *11*, 2019–2026.

Wang, M., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y., Nguyen, D.D., Watrous, J., Kapono, C.A., Luzzatto-Knaan, T., et al. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat. Biotechnol. *34*, 828–837.

Wang, M., Jarmusch, A.K., Vargas, F., Aksenov, A.A., Gauglitz, J.M., Weldon, K., Petras, D., da Silva, R., Quinn, R., Melnik, A.V., et al. (2020). Mass spectrometry searches using MASST. Nat. Biotechnol. *38*, 23–26.

WHO (2014). C-reactive protein concentrations as a marker of inflammation or infection for interpreting biomarkers of micronutrient status. https://www.who.int/nutrition/publications/micronutrients/indicators_c-reactive_protein/en/.

Williams, F.M. (2009). Biomarkers: in combination they may do better. Arthritis Res. Ther. *11*, 130.

Wolf, A.M., Wolf, D., Rumpold, H., Enrich, B., and Tilg, H. (2004). Adiponectin induces the anti-inflammatory cytokines IL-10 and IL-1RA in human leukocytes. Biochem. Biophys. Res. Commun. *323*, 630–635.

Wozniak, J.M., and Gonzalez, D.J. (2019). *PTMphinder*: an R package for PTM site localization and motif extraction from proteomic datasets. PeerJ *7*, e7046.

Zeden, J.P., Fusch, G., Holtfreter, B., Schefold, J.C., Reinke, P., Domanska, G., Haas, J.P., Gruendling, M., Westerholt, A., and Schuett, C. (2010). Excessive tryptophan catabolism along the kynurenine pathway precedes ongoing sepsis in critically ill patients. Anaesth. Intensive Care *38*, 307–316.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Lysyl Endopeptidase (LysC)	Wako Labs	129-02541
Sequencing-grade Trypsin	Promega Corporation	V5113
Anhydrous Acetonitrile	Sigma-Aldrich	271004
Tandem Mass Tags (TMT)	Thermo Fisher	90110
50% Hydroxylamine	Aldrich Chemistry	467804
C_4 5 μm Stationary Phase	Sepax	109045-0000
C_{18} 3 μm Stationary Phase	Sepax	101183-0000
C_{18} 1.8 μm Stationary Phase	Sepax	101181-0000
Critical Commercial Assays		
Pierce Quantitative Colorimetric Peptide Assay	Thermo Fisher	23275
Fetuin-B ELISA Kit	RayBiotech	ELH-FetuinB-1
IGFBP3 ELISA Kit	RayBiotech	ELH-IGFBP3-1
SERPIND1 ELISA Kit	RayBiotech	ELH-SERPIND1-1
Deposited Data		
Standard Proteomics	ProteomeXchange	PXD018030
PTM Proteomics	ProteomeXchange	PXD018031
Metabolomics	MassIVE	MSV000083593
Experimental Models: Organisms/Strains		
Mouse: 8-week old female CD1	Charles River	CD1
Software and Algorithms		
Proteome Discoverer (2.1)	Thermo Fisher	N/A
R Studio (1.1.463)	R Studio	N/A
Prism (7.0b)	GraphPad	N/A
Cytoscape (3.7.2)	(Shannon et al., 2003)	N/A
Ingenuity Pathway Analysis (01.16)	QIAGEN	N/A
Other		
Sep-Pak Cartridge 1 cc	Waters	WAT054960
HPLC Column	Thermo Fisher	720105-254630
Fused Silica Capillary Tubing	Polymicro Technologies	106815-0023
Orbitrap Fusion Tribrid Mass Spectrometer	Thermo Fisher	IQLAAEGAAPFADBMBCX

RESOURCE AVAILABILITY

Lead Contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, David J. Gonzalez (djgonzalez@ucsd.edu).

Materials Availability
This study did not generate new unique reagents.

Data and Code Availability

The proteomics data generated for this manuscript, including annotated spectra, have been deposited onto the ProteomeXchange archive through MassIVE under the following identifiers: *Standard Proteomics* (ProteomeXchange: PXD018030), *PTM-tolerant Proteomics* (ProteomeXchange: PXD018031). Metabolomics data and molecular network are available on MassIVE:MSV000083593. All other data is available upon request.

The R scripts used for analysis in this manuscript are available upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human Studies

Patient serum or plasma samples were obtained from patients at the UW Health as part of the *Staphylococcus aureus bacteremia immune response (SaBIR) study* under Health Sciences Institutional Review Board (IRB) approved protocol # 2018-0098. UW Health is a 505-bed tertiary care, academic medical center located in Madison, Wisconsin. Patients, both male and female sex, between the ages of 18-89 years of age were eligible to be included in the biobank study. Patient samples were collected as part of routine medical care, and therefore a waiver of informed consent was permitted for study inclusion. The collection of subject data is provided in the Method Details section below.

Mouse Models

All animal experiments were performed in accordance with national institutes of health (NIH) guidelines and approved by the Institutional Animal Care and Use Committee (IACUC) of the University of California San Diego. Eight-week-old female CD1 mice were used for all animal experiments. Mice were housed 5 to a cage and randomly assigned into experimental groups.

METHOD DETAILS

Patient and Isolate Identification and Collection

Patients were identified with SaB for study inclusion by electronic notification of blood cultures growing *S. aureus*, identified by Matrix Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF, Bruker Scientific LLC, Billerica, MA, USA). Methicillin-resistance was identified using GeneXpert® test (Cepheid, Sunnyvale, CA, USA). Patients were included if at least two positive blood cultures were identified, or one positive culture was congruent with a clinical diagnosis of SaB from an Infectious Diseases Physician Specialist. This study did not analyze consecutive samples from SaB patients, but rather outcomes of death (ie. hospitalized, mortality [HM]; n = 76) and survival (ie. hospitalized survival [HS]; n = 99) were selected from the SaBIR biobank for multi-omic serum analysis. The other subject groups included non-hospitalized, non-infected healthy volunteers (NN; n = 15), and hospitalized, non-infected patients (HN; n = 10) at UW Health identified through the electronic medical record.

Patient serum samples were obtained on the same day of initial presentation of SaB, before antibiotic therapy initiation and often within 1 hour of blood culture. The samples were stored at $-80°C$ until analysis

Clinical Measurements and Outcomes

Patient electronic medical records were reviewed to collect basic demographics. Metadata variables are self-explanatory or defined as follows: CV - cardiovascular, BP - blood pressure, Dys - dysfunction, MAP - mean arterial pressure, SCr - serum creatinine, WBC - white blood cell count. The mean age was $58.7 ± 15.5$ years and 49.1% of patients were male. In the SaB patient group, 33.2% were infected with MRSA and 66.8% had MSSA bacteremia, identified as above and confirmed by routine antimicrobial susceptibility testing in the clinical microbiology laboratory. Total duration of bacteremia included cases of persistent bacteremia (consecutive days of positive blood cultures) and in-hospital microbiologic relapse defined as recurrence of a positive blood culture after the first negative culture while receiving appropriate antibiotic. The median duration of bacteremia duration was 2 days with an interquartile range of 1-4 days. All included patients received appropriate antimicrobial therapy for the treatment of MSSA (anti-staphylococcal β-lactam or vancomycin/daptomycin where needed for β-lactam allergic patients) and MRSA bacteremia (vancomycin or daptomycin).

Serum Metabolite Extraction

All steps for this protocol are to be done on ice. Serum samples (100 μl) were thawed for 30 mins, then 400 μL of prechilled extraction solvent (100% MeOH with 1 μM sulfamethazine as an internal standard) was added to each sample. Samples were mix using a vortexer for 2 minutes (min) then incubated at $-20°C$ for 20 min to aid in protein precipitation. Samples were centrifuged the samples at 16,000 x g for 15 min to pellet the protein precipitate. The supernatant was transferred into 96-Well DeepWell, dried using centrifugal low-pressure system and stored at $-80°C$ once dry.

Metabolomic LC-MS^2 Analysis

Metabolomic LC-MS^2 was performed on a Bruker Daltonics® Maxis qTOF mass spectrometer (Bruker, Billerica, MA USA) with a Thermo Scientific UltiMate 3000 Dionex UPLC (Fisher Scientific, Waltham, MA USA). Plates were organized so that each row started with a blank and contained 1 - 2 controls, 4 - 7 HS samples and 3 - 5 HM samples in a random order. Metabolites were separated

using a Phenomenex C_{18} core shell (50 × 2 mm, 1.7 μm particle size) UHPLC column fitted with a C_{18} guard cartridge. The mobile phase solvents (solvent A, water-0.1% formic acid; solvent B, acetonitrile-0.1% formic acid) were run at a flow rate of 0.5 ml/min and chromatographic separation was achieved using the following elution gradient: 0 to 1 min 5% B, 1 to 10 min a linear increase from 5 to 100% B, 10 to 12 min held at 100% B, 12 to 12.5 min a linear decrease from 100 to 5% B, and 12.5 to 13 min maintained at 5% B. The mass spectrometer was calibrated twice daily using Tuning Mix ES-TOF (Agilent Technologies). For accurate mass measurements, lock mass internal calibration used a wick saturated with hexakis (2,2-difluoroethoxy)phosphazine (Synquest Laboratories, m/z 622.0289) located within the source. Ions were generated using the following parameters: nebulizer gas pressure, 2 Bar; Capillary voltage, 3,500 V; ion source temperature, 200°C; dry gas flow, 9 l/min; spectra rate acquisition, 3 spectra/s. Full scan MS spectra (m/z 50 – 1500) were acquired in the qTOF and the top five most intense ions in a particular scan were fragmented using a ramped collision induced dissociation (CID) energy from 10 - 50 eV. Data dependent automatic exclusion protocol was used so that an ion was fragmented when it was first detected, then twice more, but not again unless its intensity was 2.5x the first fragmentation. This exclusion method was cyclical, being restarted after every 30 s.

Metabolite Molecular Networking and Identification by GNPS

Metabolomics data files were converted to the .mzXML format using the Bruker Data Analysis software and uploaded to GNPS (Wang et al., 2016) through the MassIVE server (MSV000083593). Molecular networking was optimized as previously described (Scheubert et al., 2017) to an estimated false discovery rate of 1%. The data was filtered by removing all MS^2 fragment ions within ± 17 Daltons (Da) of the precursor m/z. MS^2 spectra were window filtered by choosing only the top 6 fragment ions in the ± 50 Da window throughout the spectrum. The precursor ion mass tolerance was set to 0.05 Da and an MS^2 fragment ion tolerance of 0.05 Da. A network was then created where edges were filtered to have a cosine score above 0.59 and more than 6 matched peaks. Spectra were searched against the spectral libraries contained within GNPS. The library spectra were filtered in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.7 and at least 6 matched peaks. In addition to the level 2 or 3 annotations based on the 2007 metabolomics standards initiative (Sumner et al., 2007) generated through molecular networking, CSI:FingerID (Dührkop et al., 2015) was used for molecular fingerprint identification through a fragmentation tree approach and subsequently spectra were annotated for chemical ontology through ClassyFire (Djoumbou Feunang et al., 2016). It is important to note here that our list of 5000+ metabolites is somewhat inflated due to related features that may not correspond to bone fide metabolite species; however, these occurrences are readily visible in the GNPS networks (https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp - MSV000083593 - edges with high correlation and mass shifts of 0). A more stringent estimation of total number molecular species (3310) can be derived from the total number of molecular networks (310) and non-networked features (3000).

MS1 area under the curve feature abundances were used for quantification and to produce a metabolome bucket-table with the mzMine software (Pluskal et al., 2010). MzMine modules were used with the following settings. Peak mass detection: 1E3 MS^1 noise level, 1E2 MS^2 noise level. Chromatogram deconvolution: Local minimum search algorithm: 0.2 min minimum retention time (RT) range, 3 min ratio of peak top/edge, 0.05 – 0.5 min peak duration range, 0.05 Da m/z range for MS^2, and 0.2 min RT range for MS^2. Isotopic peak grouper: 0.05 m/z tolerance, 0.1 min RT tolerance, maximum charge 4. Join aligner: 0.01 m/z tolerance, 0.3 min RT tolerance, 75% weight for m/z, 25% weight for RT, 2 minimum peaks per row. Gap filling: 20% Intensity tolerance, 0.01 m/z tolerance, 0.2 min RT tolerance. Peak filter: Area 1E3 – 1E12. The abundances of each feature in the final bucket-table were normalized first by abundance of the internal standard (1 uM sulfamethazine) within each sample and next by the total ion intensity of each sample.

Serum Protein Digestion and Labeling

Protein Digestion

100 μL of serum proteins were denatured by addition of 100 μL 8 M urea, 50 mM HEPES. Proteins were reduced and alkylated with dithiothreitol (DTT) and iodoacetamide (IAA), respectively (Haas et al., 2006), then methanol/chloroform precipitated. Proteins were re-solubilized in 1 M urea in 50 mM HEPES and 25 mM ammonium bicarbonate (pH 8.5) and digested in a two-step process (LysC and Trypsin). Digested peptides were then desalted with C_{18} Sep-Paks (Tolonen and Haas, 2014).

TMT Labeling

Samples were labeled with TMT 10-plex reagents (McAlister et al., 2012; Thompson et al., 2003) for multiplexed quantitative proteomics. TMT reagent channel 126 was reserved for bridge channels, and the remaining reagents were used to label pure sample digests. The efficient combining of MS data from 22 separate 10-plex experiments largely depends on the proper partitioning of samples into each 10-plex. Ideally, each 10-plex should be as similar as possible. Thus, we partitioned the samples in a way that ensured each 10-plex contained 1-2 control samples, 3-4 hospitalized mortality (HM) samples and 4-5 hospitalized survival (HS) samples along with the pooled bridge channel. In this way, every peptide detected in a single 10-plex will provide quantification for at least 3 samples from each of our infected groups (ie. HM and HS), By increasing the number of samples a protein is detected in per condition, this experimental design enables more robust statistical comparisons to be performed. Bridge channels consisted of an equal portion of each digest pooled together and then re-aliquoted into 50 μg portions for labeling. The bridge served as a means to control for experimental variation between mass spectrometry experiments. Labeling was conducted for 1 hr at RT and was quenched by addition of 9 μL of 5% hydroxylamine. Samples were then acidified by addition of 50 μL of 1% TFA, pooled and desalted with C_{18} Sep-Paks as described above.

Basic pH Reverse-phase Liquid Chromatography (bRPLC) Fractionation

Fractionation was carried out by bRPLC (Wang et al., 2011) with fraction combining as previously described (Lapek et al., 2017b; Tolonen and Haas, 2014). Samples were solubilized in 110 μL of 5% formic acid in 5% acetonitrile and 100 μL was separated on a 4.6 mm x 250 mm C_{18} column on an UltiMate 3000 HPLC. The resultant 96 fractions were combined into 24 distinct fractions and dried prior to multiplexed LC-MS^2/MS^3 analysis. 10 of the concatenated fractions were analyzed for each 10plex based on preliminary data indicating diminishing returns after analyzing 10 fractions.

Proteomic LC-MS^2/MS^3 Analysis

Peptides were resuspended in 5% acetonitrile/5% formic acid and analyzed on an Orbitrap Fusion Tribrid mass spectrometer with an in-line Easy-nLC 1000 System. Samples were loaded onto a 30 cm in-house pulled and packed glass capillary column (I.D. 100 μm, O.D. 350 μm). The column was packed with 0.5 cm of 5 μm C_4 resin followed by 0.5 cm of 3 μm C_{18} resin, then 29 cm of 1.8 μm of C_{18} resin. Following sample loading, peptides were eluted using a gradient ranging from 11 – 30% acetonitrile in 0.125% formic acid over 85 min at a flow rate of 300 nl/min and heating the column to 60°C. Electrospray ionization was assisted by the application of 2,000 V of electricity through a T-junction connecting the column to the nLC. All data acquired were centrioded.

MS^1 spectra were acquired in data dependent mode with a scan range of 500-1200 m/z and a resolution of 60,000. Automatic gain control (AGC) was set to 2×10^5 with a maximum ion inject time was 100 miliseconds (ms) and a lower threshold for ion intensity of 5×10^4. Ions selected for MS^2 analysis were isolated in the quadrupole at 0.5 Th. Ions were fragmented using CID with a normalized collision energy of 30% and were detected in the linear ion trap with a rapid scan rate for low resolution spectra (eight 10-plexes; standard proteomic analysis). For high-resolution spectra (fourteen 10-plexes; standard proteomic and PTM analysis), ions were fragmented using higher-energy collision-induced dissociation (HCD) with a normalized collision energy of 30% and were detected in the Orbitrap with a resolution of 3×10^4. Multiple studies have supported high-resolution scans to be beneficial for PTM analyses (Chick et al., 2015; Devabhaktuni et al., 2019). Further, for glyco-peptide matching, the detection of low mass glycan reporter ions, which are efficiently captured using the HCD fragmentation scheme(Cao et al., 2014; Mayampurath et al., 2011), is useful for selecting glyco-peptide containing spectra prior to database searching. AGC was set to 1×10^4 and the inject time was set to 35 ms.

MS^3 analysis was conducted using the synchronous precursor selection (SPS) option to maximize TMT quantitation sensitivity (McAlister et al., 2014). Up to 10 MS^2 ions were simultaneously isolated and fragmented with HCD using a normalized energy of 50%. MS^3 fragment ions were analyzed in the Orbitrap at a resolution of 6×10^4. The AGC was set to 5×10^4 using a maximum ion injection time of 150 ms. MS^2 ions 40 m/z below and 15 m/z above the MS^1 precursor ion were excluded from MS^3 selection.

Peptide Identification by Proteome Discoverer
Standard Proteomics Workflow

Resultant data files were processed using Proteome Discoverer 2.1. MS^2 data were queried against the Uniprot human database (downloaded: 03/2019; 43,518 entries; contains isoforms and unreviewed entries, but fragments were removed) using the Sequest algorithm (Eng et al., 1994). A decoy search was also conducted with sequences in reversed order (Elias and Gygi, 2007; Elias et al., 2005; Peng et al., 2003). For MS^1 spectra, a mass tolerance of 50 ppm was used and for MS^2 spectra a 0.6 Da (for low-resolution spectra) or 0.05 Da (for high-resolution spectra) tolerance was used. Static modifications included TMT 10-plex reagents (+229.162932 Da) on lysine and peptide n-termini and carbamidomethylation of cysteines (+57.02146 Da). Variable oxidation of methionine (+15.99492 Da) and deamidation of asparagine and glutamine (+0.984016 Da) were also included in the search parameters. Data were filtered to a 1% peptide and protein level false discovery rate using the target-decoy strategy (Elias and Gygi, 2007; Elias et al., 2005; Peng et al., 2003). The minimum number of peptides required to quantify a protein was one based on the logic that the estimation of protein error rates are better evidence for protein presence than two peptides (Gupta and Pevzner, 2009). Following the initial analysis, the search was repeated for a focused database (1,088 entries). For TMT experiments, reporter ion intensities were extracted from MS^3 spectra for quantitative analysis. Protein-level quantitation values were calculated by summing signal to noise values for all peptides per protein meeting the specified filters (high confidence, non-rejected spectra with an average signal:noise > 10 and isolation interference < 25%). Data were normalized in a two-step process as previously described (Lapek et al., 2017a). First, the values for each protein were normalized to the pooled bridge channel value. Then, the values were normalized to the median of each reporter ion channel.

PTM-tolerant Workflow

While using standard collisional induced dissociation (CID) and taking low-resolution scans in the ion trap (IT) is a widespread MS strategy, PTM identification and localization is optimally derived from high-resolution mass spectra taken in the Orbitrap (OT) mass analyzer (Chick et al., 2015; Devabhaktuni et al., 2019). Further, the use of glycan reporter ions to identify glycopeptide-containing spectra requires higher-energy collisional dissociation (HCD) fragmentation (Cao et al., 2014; Mayampurath et al., 2011). Therefore, we employed the use of an HCD-OT MS workflow for this PTM analysis. High resolution MS^2 data from proteomic experiments were submitted to molecular networking via GNPS as described above. Overrepresented mass shifts, as determined by the total number of network edges corresponding to each mass shift, were selected as modifications to include in a PTM-tolerant search. Mass shifts ultimately included in the search were selected based on the number of observed edges (> 100) and if the PTM identity (inferred from unimod.org) had been previously detected in proteomic experiments. MS^2 data were queried against a focused human serum proteome database (proteins detected in standard search, 1,088 entries) using Byonic (Bern et al., 2012). A decoy search was also conducted with sequences in reversed order (Elias and Gygi, 2007; Elias et al., 2005; Peng et al., 2003). For MS^1 spectra, a mass tolerance of 50

ppm was used and for MS^2 spectra a 0.05 Da tolerance was used. Static modifications included TMT 10-plex reagents (+229.162932 Da) on peptide n-termini and carbamidomethylation of cysteines (+57.02146 Da). Variable modifications were specified using Modification Fine Control. Variable modifications included: deamidation (+0.984016 Da) of asparagine and glutamine, oxidation (+15.99492 Da) of methionine, tryptophan and histidine, formylation (+27.994915 Da) of lysine, dioxidation (+31.989829 Da) of tryptophan, carbamylation (+43.005814) of lysine and arginine and dihydroxyimidazolidine (+72.021129 Da) of arginine. Spectra that contained low-mass glycan reporter ions as determined by the IMP-glycan reporter node were submitted to a glyco-peptide search with the following modification parameters. Static modifications included: TMT 10-plex reagents (+229.162932 Da) on peptide n-termini and lysines and carbamidomethylation of cysteines (+57.02146 Da). Variable modifications included: oxidation (+15.99492 Da) of methionine and glycosylation (57 common human N-glycans (Bern et al., 2012; Clerc et al., 2016) - various Da) of asparagine. Glycan structures were inferred from the monosaccharide compositions in accordance with common serum glycans. Reporter ion intensities for modified peptides were summed to the unique peptide level then normalized as above. PTMs were localized in the context of the total protein length and flanking sequences were extracted using the *PTMphinder* R package (Wozniak and Gonzalez, 2019).

Statistical Analyses of MS Data
Binary Comparisons
First, binary comparisons were used to identify biomarkers for the prediction of SaB mortality. Two types of binary analyses were used, Mann-Whitney U (MWU) tests, which has been shown to be effective for biomarker selection (Dakna et al., 2010), and an ensemble feature selection (EFS) approach, which can reduce biases of any individual feature selection method (He and Yu, 2010). MWU tests were implemented in Excel using the RealStats package and EFS was implemented in R using the *EFS* package (Neumann et al., 2017). The EFS approach combines MWU tests, logistic regression, Pearson and Spearman correlations and two random forest algorithm implementations, *cforest* and *randomForest*, into a single, rank-able score. Biomarkers were ultimately ranked by the average score from both binary comparison analyses (Tables S2 and S3).
Metadata Assessment
An R script was written to determine the association of quantified features with patient metadata. First, the metadata was split into two groups, categorical metadata and continuous metadata. Categorical metadata associations were determined using MWU test (2 categories) or Kruskall-Wallis test (> 2 categories). Continuous metadata associations were determined using Pearson correlation. All tests were performed using base R functions. Metadata associations displayed in figures represent the -log$_{10}$(p value) reported from each test.
K-means Clustering
In addition to ranking biomarkers using binary comparisons, we performed multi-class analyses to consider the control groups (NN and HN) in addition to the infected samples (HS and HM). First, we employed K-means clustering to group proteins with similar expression profiles across our four major patient groups (NN, HN, HS, HM). Prior to clustering, proteins were filtered for significant differences using uncorrected, ANOVA p < 0.05. The optimal number of clusters was determined using the elbow method. Benjamini-Hochberg corrected p values are also provided (Tables S2 and S3), but all features with uncorrected ANOVA p < 0.05 were included in figures and downstream analyses (ie. gene ontology/network analyses) because these tools work best with protein lists larger than 50 (Huang et al., 2009; Mi et al., 2013).
LASSO Regression
Second, we employed a LASSO regression algorithm to select the minimum set features required to classify the four sample groups. The initial feature space displayed significant missingness (average missingness was 48.55% for proteomics and 40.70% for all metabolomics) and thus missing values were imputed prior to regression analyses. Before imputation, features having > 50% missingness were conservatively dropped for not having enough information to confidently infer new values, leaving 504 proteomic and 3082 metabolomic features in total. Fast missing value imputation was implemented by chained random forests within each individual dataset using the missRanger R package (https://github.com/mayer79/missRanger). Briefly, each missing value was imputed by a random forest built on all other features as co-features; this process is iterative, such that it continues multiple times across all features until the average out-of-bag (OOB) prediction error plateaus. It leads to imputations with realistic variability, similar to stochastic regression imputation, but is much faster and computationally efficient across thousands of features.

After fast imputation by chained random forests on each dataset, they were separately scaled to Z-scores and concatenated for regularized (L1 norm) machine learning approaches. Subsequent regularized machine learning was done using the caret (http://topepo.github.io/caret/index.html) and glmnet (Friedman et al., 2010) R packages to train, test, and evaluate LASSO logistic classification models. The mathematical details of LASSO are comprehensively described elsewhere (Friedman et al., 2010). The basic idea is that models are penalized for both overall predictive performance and the size of the resultant feature set used to make those predictions. Due to small cohort sizes in NN and HN groups (n = 13 and n = 10, respectively), 10-fold cross-validation was used to estimate group discriminative performance rather than traditional train-test splits. Using the LASSO parameterization (alpha = 1), models were tuned on a search grid of lambda values: 0.001 to 0.3 by steps of 0.01. Models were optimized for the highest area under the ROC curve (AUROC), which for multiclass comparisons was calculated as the average AUROC across all applicable one-class-versus-all-others comparisons. Feature importance and the overall LASSO proteomic-metabolomic signature (length = 98 features: 32 proteomic, 68 metabolomic) was obtained using the varImp() in caret and keeping any features with non-zero importance values. Overall model performance and its confusion matrix was determined by examining the predictions on samples left out during their kth fold and their ground truth using the caret and MLmetrics (https://github.com/yanyachen/MLmetrics) packages.

Multi-omic data integration

Finally, for multi-omic data integration, we implemented the mixOmic data analytics pipeline (Rohart et al., 2017). This method extends the Generalized Canonical Correlation Analysis (Tenenhaus et al., 2017) framework with a generalized, supervised partial least-squares (PLS) approach to integrate multiple data types across the same group of subjects with known phenotypes while identifying key omics variables in the process. The mathematical details are described in the associated paper (Rohart et al., 2017). Imputed proteomic and metabolomic data, as described above, were utilized for this analysis with a labeled class vector composed of their sample group (NN: n = 13, HN: n = 10, HS: n = 99, HM: n = 76; n = 198 total). The design matrix was built with link weights of 0.1 between the proteomic and metabolomic data (as seen: http://mixomics.org/mixdiablo/case-study-tcga/). The model was then first fit using 10-fold cross-validation, repeated 10 times in order to determine the optimal number of components for the final model (that is, the number of components that leads to the lowest classification error rate). The optimal number of components was determined to be 4. The final model was subsequently used for plot generation.

Gene Ontology (GO) and Network-based Analyses

Proteins subsets identified through various computational approaches were subjected to GO and network-based analyses using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang et al., 2007, 2009) and STRING-db (Szklarczyk et al., 2019) tools, respectively. For GO analysis, lists of interesting proteins were submitted for enrichment analysis using all the proteins detected in the experiment as a background dataset. All GO terms displayed in the figures had a $p < 0.05$. For network-based analyses, protein lists were submitted to the STRING-db tool (all active interaction sources, interaction score > 0.8). The network was exported into a simple tabular output, reformatted to display desired parameters (e.g., K-means clusters), then imported into Cytoscape (Shannon et al., 2003) for visualization.

Network-based Cytokine Inference

Knowledge-based networks were used to infer the relative contributions of major cytokines on the observed proteomic alterations. First, a list of cytokines was manually curated (TGFb, TNF, IFN, IL1-40, CXCL1-16, CCL1-27) and submitted with the significantly altered proteins (uncorrected ANOVA $p < 0.05$) to the STRING-db tool (all active interaction sources, interaction score > 0.4). The network was exported into a simple tabular output and cytokines were profiled for their enrichment in mortality clusters (pMortality$^+$, pMortality$^-$ and pMortality$^-$) using the following approach. First, cytokines were filtered to have at least five connections to mortality networks. Then, the proportion of each cytokine within the mortality clusters (ie. number of cytokine-mortality cluster connections relative to the total connections within mortality clusters) was compared to the proportion of each cytokine in the entire dataset (ie. number of cytokine-protein connections relative to total connections within the entire dataset). Significance was determined using Chi-square tests and a fold-change was calculated by dividing the proportion of cytokine-mortality connections by the proportion of cytokine connections within the total dataset. A final enrichment score was calculated by multiplying the $-\log_{10}$(p value) from the Chi-square test by the $\log_2$(fold-change of enrichment). By including the total number of connections that each cytokine had in the entire dataset as a background, any enrichment bias due to higher annotation rates for popular cytokines in STRING-db is controlled for.

To benchmark the cytokine inference approach mentioned above, we compared the results to Ingenuity Pathway Analysis (IPA – QIAGEN). IPA receives a list of proteins and calculates upstream regulator enrichment based on the overlap of target proteins in the submitted list and reports a p value for significance. For all cytokines reported by both tools, the -log10(p value) from IPA was compared to the enrichment score from the cytokine inference approach using Pearson correlation. Additional, highly significant upstream regulators ($p < 0.001$) of non-cytokine origin were manually curated from the IPA results for inclusion in the manuscript.

Mouse Model of SaB

8-week old female CD1 mice were used for all animal experiments. Mice were treated before infections as follows or with vehicle controls. Hyperthyroid mice were given I.P. injections of 100 μg thyroxine (T4) once daily for the three days prior to infection. Hypothyroid mice were given drinking water containing hypothyroid treatment (1% (wt/vol) sodium perchlorate and 0.1% (wt/vol) methimazole) for three weeks prior to infections. Adiponectin mice were given I.P. injections of 1 mg/kg AdipoRon one day prior infection, then injected daily with AdipoRon for the duration of the experiment. Mice were then I.V. infected with *S. aureus* LAC (high dose (Figure 7B): 1×10^8 CFUs, low dose (Figures 7C–7H): 5×10^7 CFUs) and survival was monitored every 12 hours. For CFU burden experiments, mice were treated and infected as above, then euthanized 48 hours post-infection and organs were harvested for quantitation of bacterial burden.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were completed as described in the corresponding figure legends. Either GraphPad Prism, Microsoft Excel or R were used to conduct tests. Sample sizes for the validation cohort were determined based on a power calculation using example data from the preliminary cohort ($\alpha < 0.001$; $1 - \beta = 0.95$; $n \geq 45$) and a review of the literature ($n \geq 50$)(Skates et al., 2013). Significance was assessed using one or more of the following: Mann-Whitney U test, analysis of variance (ANOVA), Pearson correlation test, ROC curves, or logistic regression. For ANOVA, Tukey's multiple comparisons test was used. For all tests, significance values are denoted as follows: ****$p < 0.0001$; ***$p < 0.001$, **$p < 0.01$; *$p < 0.05$, ns - not significant.
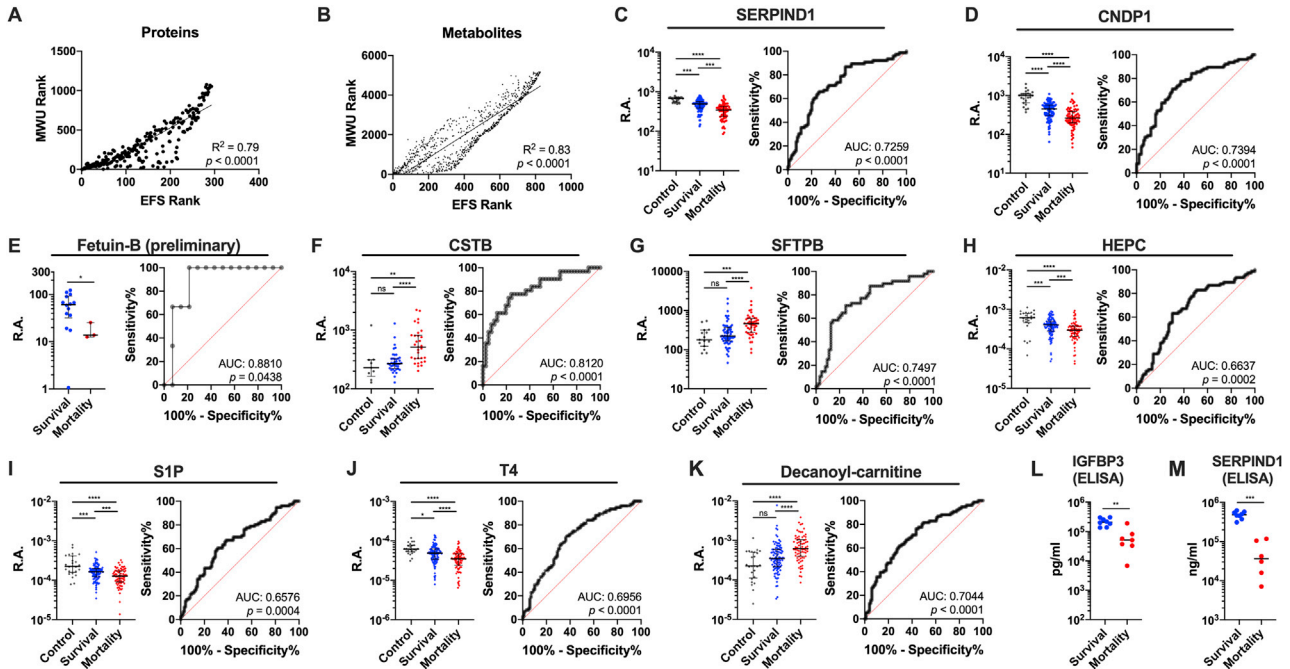
Supplemental Figures

Figure S1. Additional Biomarkers Analysis and Validation of Top Markers, Related to Figure 2

Correlation of Mann-Whitney U (MWU) test and ensemble feature selection (EFS) rankings for (**A**) proteins and (**B**) metabolites. (**C**) Abundance and ROC curve of SERPIND1 (survival versus mortality). (**D**) Abundance and ROC curve of CNDP1 (survival versus mortality). (**E**) Abundance and ROC curve of Fetuin B (survival versus mortality - preliminary cohort). (**F**) Abundance and ROC curve of CSTB (survival versus mortality). (**G**) Abundance and ROC curve of SFTPB (survival versus mortality). (**H**) Abundance and ROC curve of HEPC (survival versus mortality). (**I**) Abundance and ROC curve of S1P (survival versus mortality). (**J**) Abundance and ROC curve of T4 (survival versus mortality). (**K**) Abundance and ROC curve of decanoyl-carnitine (survival versus mortality). (**L**) Abundance of IGFBP3 measured via ELISA. (**M**) Abundance of SERPIND1 measured via ELISA. Error bars represent interquartile range (IQR). For all tests, significance values are denoted as follows: ****$p < 0.0001$; ***$p < 0.001$, **$p < 0.01$; *$p < 0.05$, ns - not significant. For C, D, F, G, H, and I, significance is displayed based on Kruskal-Wallis tests with Dunn's multiple comparison test. For E, J and K, significance is displayed based on Mann Whitney U tests.
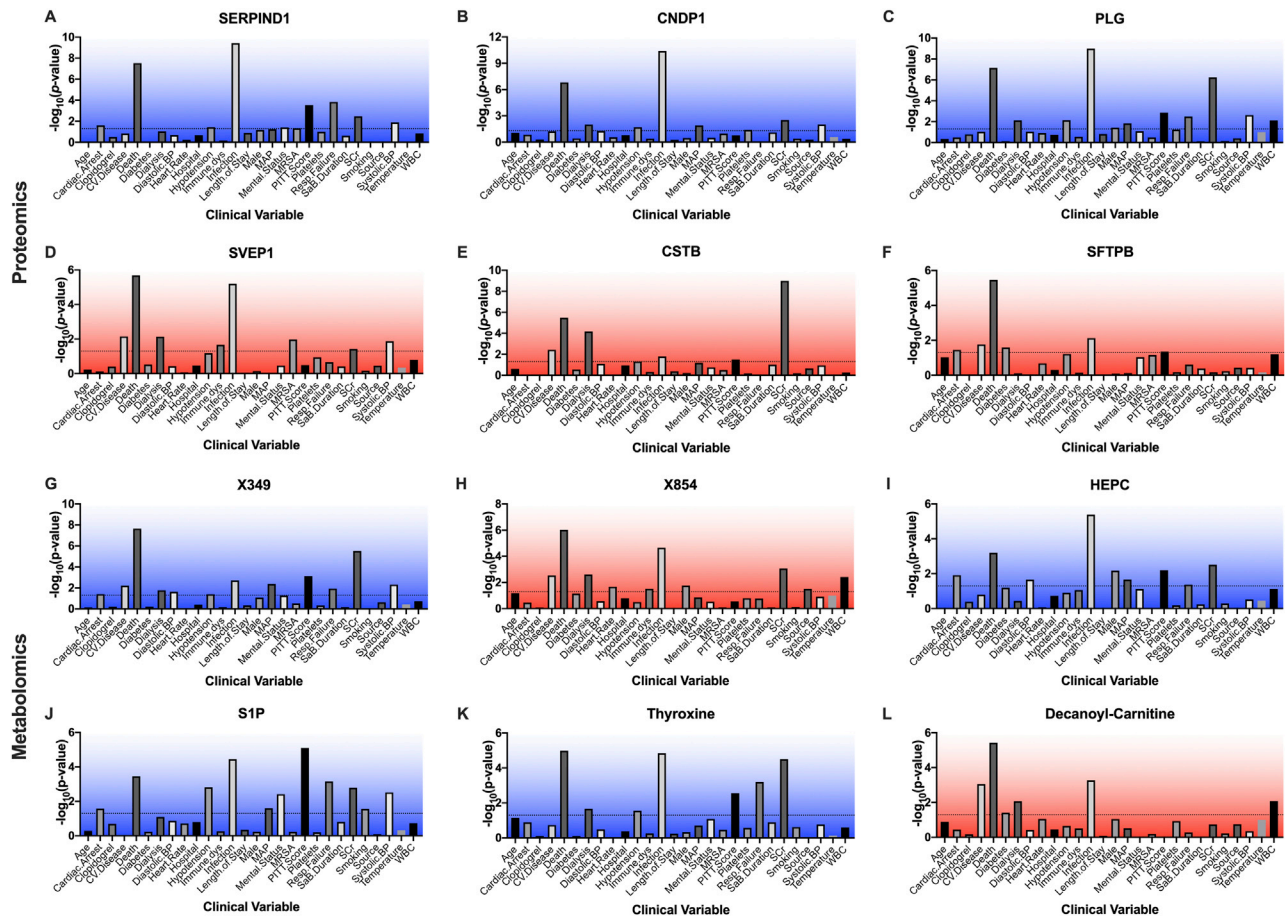
Figure S2. Metadata Associations of Top Biomarkers, Related to Figure 2

Metadata assessments of top biomarkers including: decreased proteins (**A**) - SERPIND1, (**B**) - CNDP1, (**C**) - PLG), increased proteins (**D** - IGFBP2, (**E**) - ADIPOQ, (**F**) - EFEMP1), decreased metabolites (**G**) - X349, (**H**) - X228, (**I**) - X320) and increased metabolites (**J**) - X746, (**K**) - X854, (**L**) - X2532). Plots are highlighted red for increased expression in mortality or blue for decreased expression in mortality.
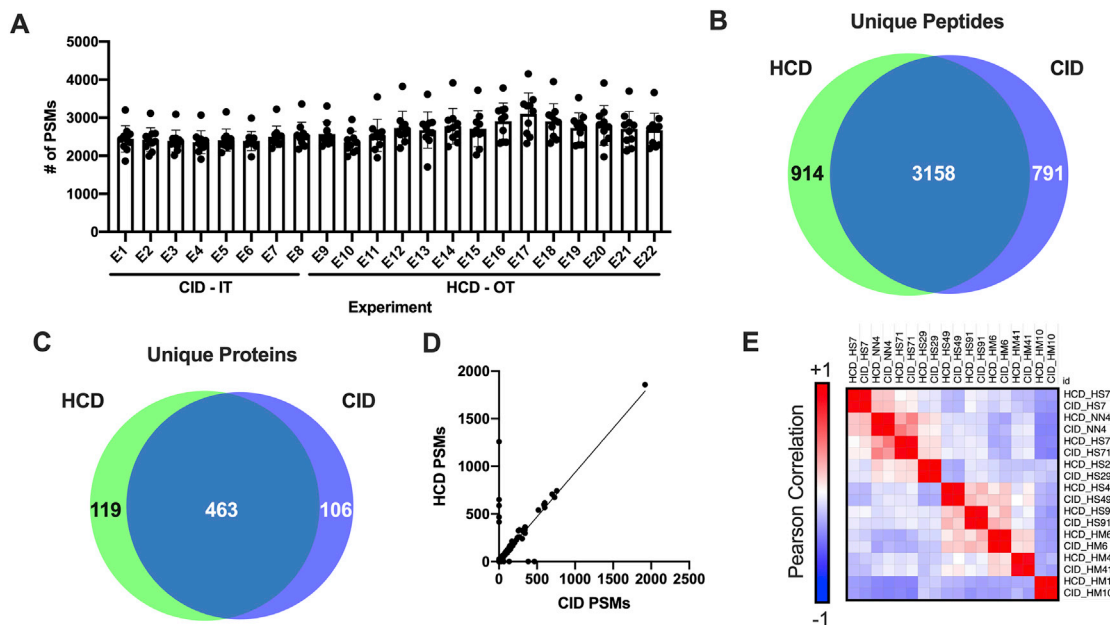
Figure S3. Comparison of Low- and High-Resolution Mass Spectrometer Methods, Related to Figure 3
(A) Number of PSMs detected across each 10plex experiment. (B) Venn diagram of peptides identified by each method for experiment 8 (E8). (C) Venn diagram of proteins identified by each method for E8. (D) Correlations of PSMs assigned to each protein by each method for E8. (E) Correlations of TMT-based quantitation for every protein in each sample by each method in E8.
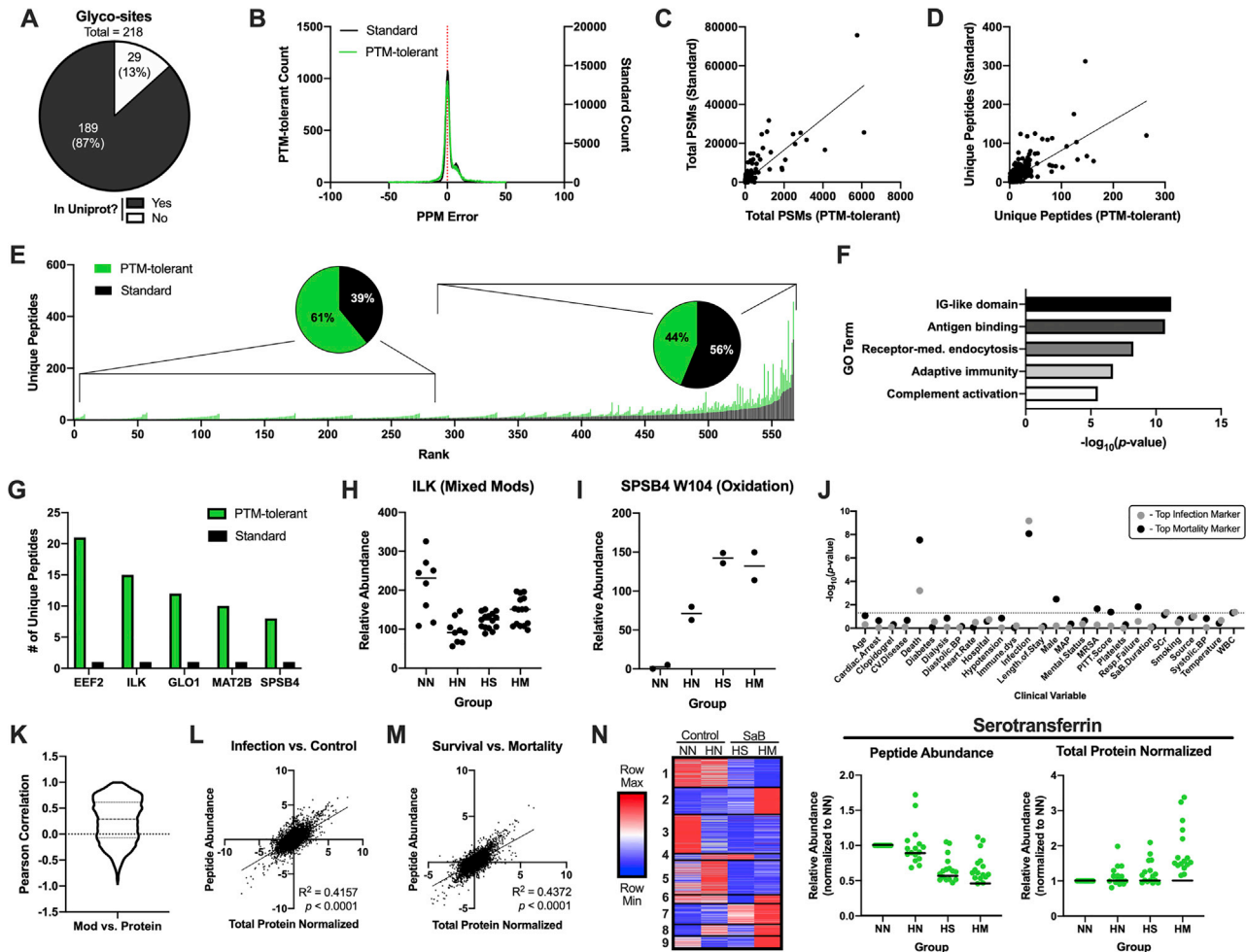
Figure S4. Extended PTM-Tolerant Search Analysis, Related to Figure 3

(**A**) Proportion of detected glyco-sites present in Uniprot. (**B**) MS[1] mass errors for standard and PTM-tolerant database searches. Correlations of total PSMs (**C**) and unique peptides (**D**) per protein detected in the standard and PTM-tolerant database searches. (**E**) Unique peptides detected in the standard and PTM-tolerant database search ranked by number of unique unmodified peptides then number of unique modified peptides. Pie charts depict unique peptide proportions of top and bottom 50% of proteins detected in the standard and PTM-tolerant workflows. (**F**) GO analysis of proteins with bottom 50% of unique peptides in the standard search. (**G**) Proteins with the largest gain in unique peptides detected in the PTM-tolerant search. (**H**) Abundance of modified ILK peptides detected in PTM-tolerant search. (**I**) Abundance of dioxidation of SPSB4 104W detected in PTM-tolerant search. (**J**) Metadata assessment of top modified biomarkers for infection and mortality. (**K**) Correlation of modified peptide (Mod) and total protein relative abundances. Scatterplot of fold-changes comparing (**L**) control versus infected and (**M**) survival versus mortality. (**N**) K means clustered heatmap of all significantly altered, protein-normalized, modified peptides (-ANOVA p < 0.05) across the four primary groups (Control groups: NN – Non-hospital, Non-infected, HN – Hospital, Non-infected; Infection groups: HS –Hospital, Survival, HM – Hospital, Mortality). (**O**) Serotransferrin mortality-associated PTM plot depicting modified peptide abundance (left) and modified peptide abundance normalized to total protein levels (right).
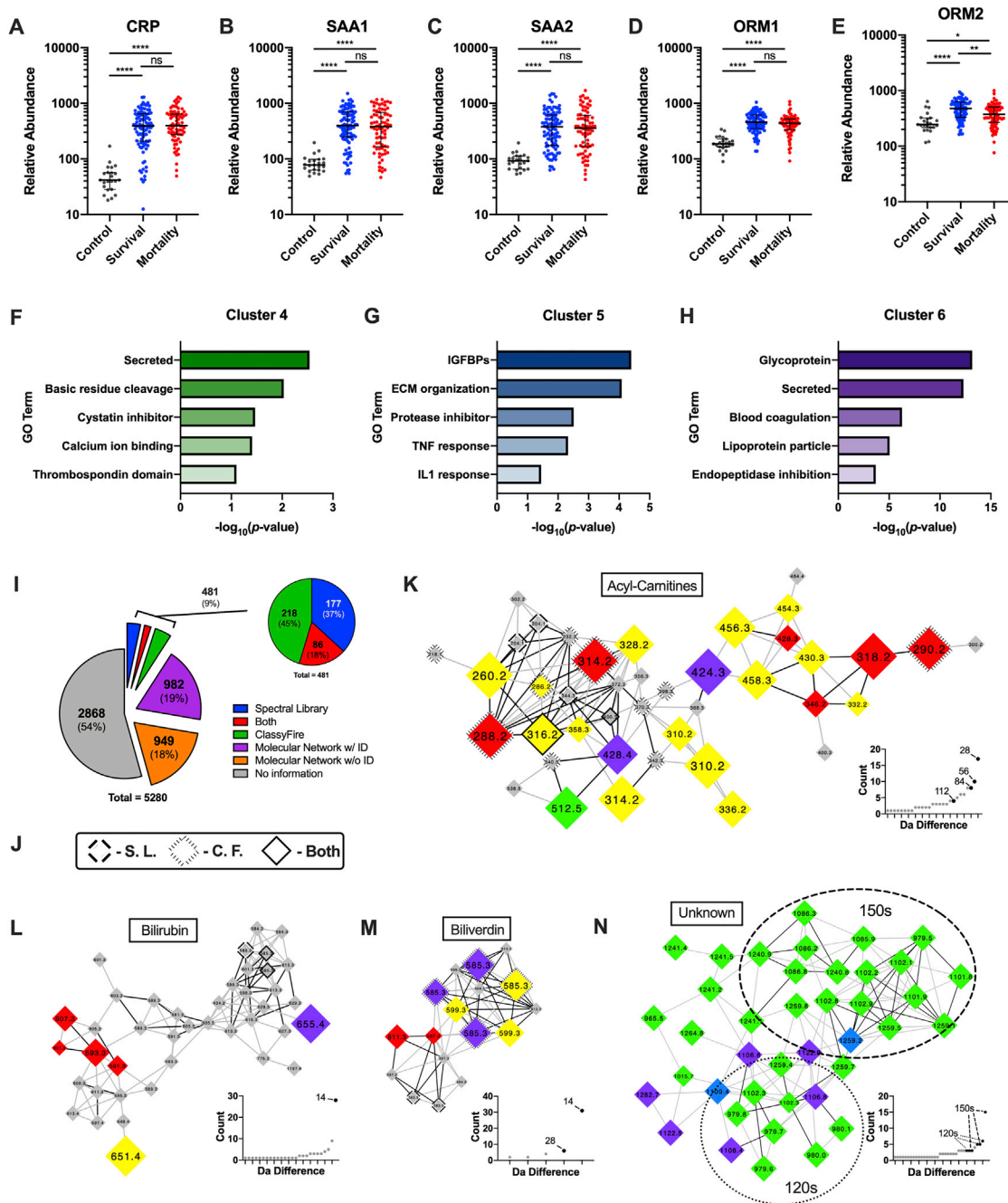
Figure S5. Extended Analysis of SaB Disease Modules, Related to Figure 4

Individual plots for major acute-phase reactant proteins contained within proteomics infection-associated cluster 2 from Figure 4A including: (**A**) CRP, (**B**) SAA1, (**C**) SAA2, (**D**) ORM1, and (**E**) ORM2. (**F-H**) GO analysis of proteomics mortality-associated clusters: pMortality⁻ (**F**), pMortality⁺ (**G**) and pMortality⁻ (**H**). (**I**) Pie chart for sources of molecular information for all metabolomic features detected in this experiment. (**J**) Key for source of molecular identity used in all molecular networks in figure (S. L. - Spectral library, C.F. - ClassyFire). Molecular networks that are associated with mortality and contain identified nodes including: (**K**) acyl-carnitines, (**L**) bilirubin, and (**M**) biliverdin. (**N**) Mortality-associated molecular network that did not containing any identified nodes. Nodes are colored according to cluster designations in Figure 4D and sized according to -log$_{10}$(*p-value*) determined via ANOVA. Mass shifts in networks are displayed in plots to the lower right of each network (Da - Daltons). High-occurring mass shifts are highlighted in the networks as black edges and annotated in plots. Error bars represent interquartile range (IQR). For all tests, significance values are denoted as follows: ****$p < 0.0001$; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; ns, not significant. For A - E, significance is displayed based on ANOVA with Tukey's multiple comparison test.
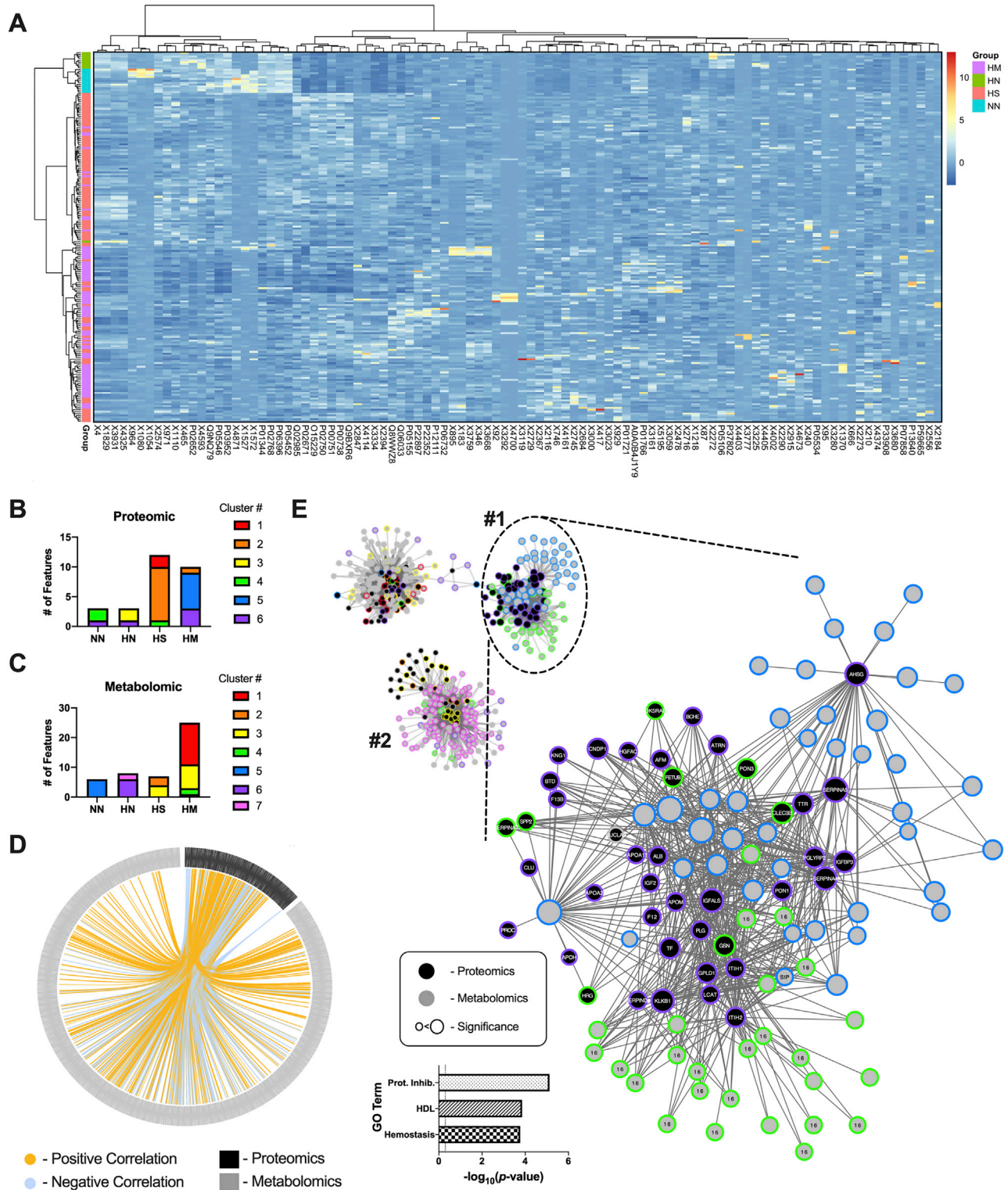
Figure S6. Data Integration and Multi-Group Classification, Related to Figure 4

(**A**) Heatmap of multi-omic, multi-group classification model final features (blue - low expression, red - high expression). Bar charts displaying number of features important for each group colored according their K-means cluster membership for (**B**) proteomics data and (**C**) metabolomics data. (**D**) Circos plot of correlations

(legend continued on next page)

across proteomic and metabolomic datasets. (**E**) Correlation network of proteins and metabolites overlaid with K-means cluster information (Focus #1: mortality-associated; Focus #2: hospital-associated). Node borders are colored according to K-means clusters defined for proteins (Figure 4A) and metabolites (Figure 4D). GO analysis of proteins in Focus #1 is displayed as a bar chart in the lower left region of the network. Metabolites belonging to the unknown, mortality-associated network (Figure S5N - Subnetwork 16) are noted.
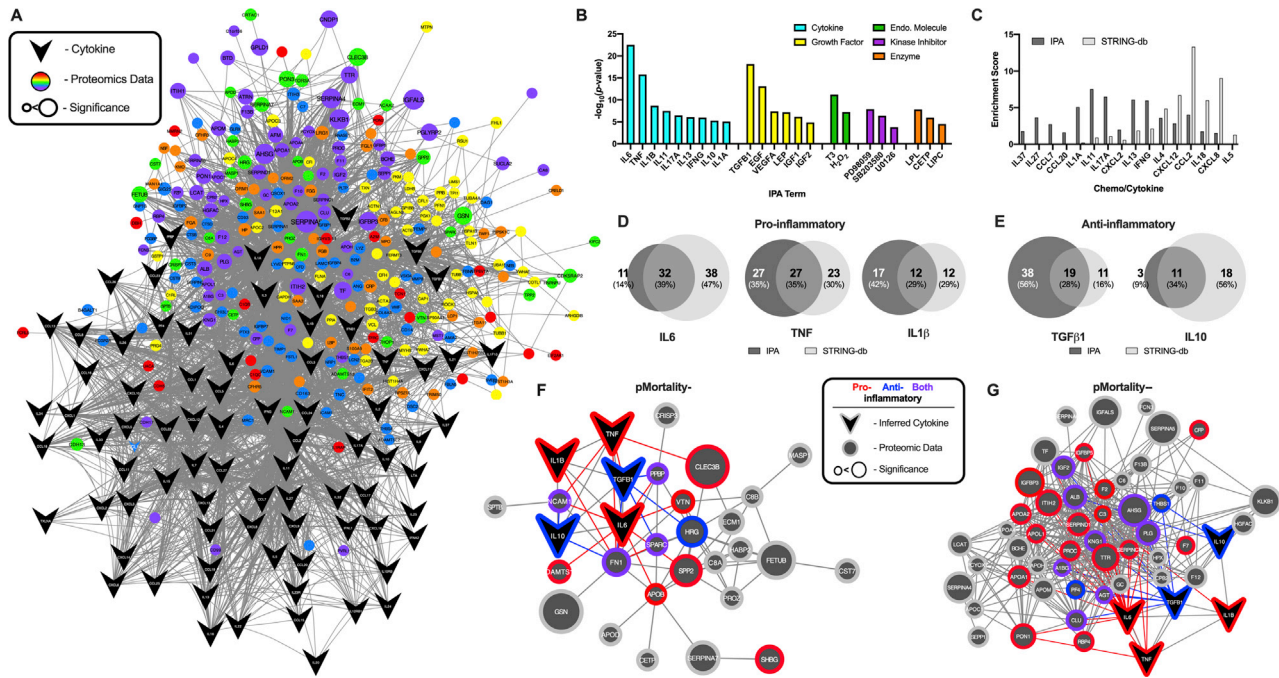
Figure S7. Extended Knowledge-Based Analysis of Cytokines, Related to Figure 6

(**A**) Ingenuity Pathway Analysis (IPA) terms enriched in mortality proteomics clusters ordered by category then by -log$_{10}$(*p-value* of overlap). (**B**) Comparison of cytokines preferentially enriched in IPA or Cytokine Inference method. Venn diagrams of target proteins of the commonly predicted pro-inflammatory (**C**) and anti-inflammatory (**D**) cytokines determined by IPA and STRING-db. (**E**) First-pass analysis with all input cytokines and all proteins significantly altered (ANOVA p < 0.05) in the standard proteomics data. Proteomics nodes are colored according to designations in Figure 4A. Refined networks of top 5 commonly predicted cytokines and pMortality⁻ (**F**) and pMortality⁻ (**G**) data. Proteomics node outlines are colored based their connections to pro-inflammatory cytokines (red), anti-inflammatory cytokines (blue) or both (purple). Cytokine node outlines and neighboring edges are colored based on pro-inflammatory (red) or anti-inflammatory (blue) activity. In all networks, proteomics nodes are sized according to -log$_{10}$(*p-value*) determined via ANOVA.