

Multi-omic signatures of host response associated with presence, type, and outcome of enterococcal bacteremia

Charlie Bayne,^{1,2,3} Dominic McGrosso,^{1,2,3} Concepcion Sanchez,^{1,2,3} Leigh-Ana Rossitto,^{1,2,3} Maxwell Patterson,² Carlos Gonzalez,^{2,3} Courtney Baus,⁴ Cecilia Volk,⁵ Haoqi Nina Zhao,³ Pieter Dorrestein,^{2,3,6,7} Victor Nizet,^{2,3,6} George Sakoulas,^{6,8} David J. Gonzalez,^{2,3,7} Warren Rose⁵

AUTHOR AFFILIATIONS See affiliation list on p. 23.

ABSTRACT Despite the prevalence and severity of enterococcal bacteremia (EcB), the mechanisms underlying systemic host responses to the disease remain unclear. Here, we present an extensive study that profiles molecular differences in plasma from EcB patients using an unbiased multi-omics approach. We performed shotgun proteomics and metabolomics on 105 plasma samples, including those from EcB patients and healthy volunteers. Comparison between healthy volunteer and EcB-infected patient samples revealed significant disparities in proteins and metabolites involved in the acute phase response, inflammatory processes, and cholestasis. Several features distinguish these two groups with remarkable accuracy. Cross-referencing EcB signatures with those of *Staphylococcus aureus* bacteremia revealed shared reductions in cholesterol metabolism proteins and differing responses in platelet alpha granule and neutrophil-associated proteins. Characterization of *Enterococcus* isolates derived from patients facilitated a nuanced comparison between EcB caused by *Enterococcus faecalis* and *Enterococcus faecium*, uncovering reduced immunoglobulin abundances in *E. faecium* cases and features capable of distinguishing the underlying microbe. Leveraging extensive patient metadata, we now have identified features associated with mortality or survival, revealing significant multi-omic differences and pinpointing histidine-rich glycoprotein and fetuin-B as features capable of distinguishing survival status with excellent accuracy. Altogether, this study aims to culminate in the creation of objective risk stratification algorithms—a pivotal step toward enhancing patient management and care. To facilitate the exploration of this rich data source, we provide a user-friendly interface at https://gonzalezlab.shinyapps.io/EcB_multiomics/.

IMPORTANCE *Enterococcus* infections have emerged as the second most common nosocomial infection, with enterococcal bacteremia (EcB) contributing to thousands of patient deaths annually. To address a lack of detailed understanding regarding the specific systemic response to EcB, we conducted a comprehensive multi-omic evaluation of the systemic host response observed in patient plasma. Our findings reveal significant features in the metabolome and proteome associated with the presence of infection, species differences, and survival outcome. We identified features capable of discriminating EcB infection from healthy states and survival from mortality with excellent accuracy, suggesting potential practical clinical utility. However, our study also established that systemic features to distinguish *Enterococcus faecalis* from *Enterococcus faecium* EcB show only a moderate degree of discriminatory accuracy, unlikely to significantly improve upon current diagnostic methods. Comparisons of differences in the plasma proteome relative to healthy samples between bacteremia caused by *Enterococcus* and *Staphylococcus aureus* suggest the presence of bacteria-specific responses alongside conserved inflammatory reactions.

Editor Aleksandra Nita-Lazar, NIAID, NIH, Bethesda, Maryland, USA

Peer Reviewer Małgorzata Barbara Łobocka, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland

Address correspondence to Warren Rose, warren.rose@wisc.edu, or David J. Gonzalez, djgonzalez@health.ucsd.edu.

P.D. is a scientific advisor and holds equity in Cybele and bileOmix, and he is a scientific co-founder and advisor and holds equity in Ometa, Arome, and Enveda with prior approval by UC-San Diego.

See the funding table on p. 23.

Received 2 November 2024

Accepted 6 December 2024

Published 21 January 2025

Copyright © 2025 Bayne et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

KEYWORDS proteomics, metabolomics, *Enterococcus*, multi-omics, bacteremia, microbiology

Enterococci are widely distributed in the environment and have co-evolved as common microorganisms within the gastrointestinal microbiota since terrestrial animals first transitioned from water to land (1). From this diverse lineage, the distantly related species *Enterococcus faecalis* and *Enterococcus faecium* have independently evolved to become members of the human gut microbiome. These microbes establish their niche as commensals within the first 10 days following birth (2, 3) and typically compose <0.1% of the gut microbiome (4). Under certain circumstances, *E. faecalis* and *E. faecium* can become pathogenic. These conditions are commonly encountered in healthcare settings, where several features acquired during their evolution have allowed *E. faecalis* and *E. faecium* to become important healthcare-associated pathogens (5).

These microbes possess the inherent ability to survive commonly used disinfection routines and persist on surfaces in healthcare settings, thereby facilitating patient-to-patient transmission (6). Additionally, they are increasingly resistant to antibiotics, owing both to an intrinsic resistance to commonly used broad-spectrum antibiotics, such as cephalosporins and carbapenems (7), and to an impressive capacity to acquire mobile genetic elements through horizontal gene transfer, which enhances their fitness (8). Even more concerning is that newly developed antibiotics targeting Gram-positive pathogens are either ineffective against enterococci or rapidly lead to the emergence of resistance (9–12). Furthermore, antibiotics that are active against enterococci, such as beta-lactams in the case of *E. faecalis*, are bacteriostatic, necessitating prolonged courses of combination antibiotic therapy to prevent relapse (13). These traits, combined with inadequate antibiotic stewardship and the increase in aggressive medical treatments in an aging patient population, have led to a rise in severe, invasive infections with a mortality rate of 25%–50% in enterococcal bacteremia (EcB) (6, 7, 9–11, 14).

Host factors are widely recognized as critical determinants of the outcome of host-microbe interactions (15) and have been used as prognostic biomarkers to predict patient outcomes in a variety of diseases, ranging from coronavirus disease 2019 (COVID-19) to cancer (16, 17). While molecular features associated with mortality have been described in *Staphylococcus aureus* bacteremia (18), in the context of EcB, the prediction of mortality has so far been limited to crude clinical metrics such as severity of illness and age (19). A comprehensive profile of the molecular features of the systemic response in a well-documented EcB patient cohort would enable the discovery of associations with successful and suboptimal outcomes.

To describe the systemic host response in EcB patients, we employed high-resolution tandem mass tag (TMT) LCMS3 mass spectrometry (MS)-based proteomics and metabolomics to profile plasma samples collected from clinical EcB cases and healthy controls. Furthermore, we utilized previously published results from our group to compare deviations from homeostasis observed in EcB and *S. aureus* bacteremia. This data set provides an initial assessment of the ability to use unbiased molecular features of the host response to predict the presence of EcB, whether the bacteremia is caused by *E. faecalis* or *E. faecium*, and the outcome of the infection. The enhanced understanding provided by this multi-omic resource can serve as a starting point for developing novel therapeutic strategies aimed at improving patient outcomes in EcB.

MATERIALS AND METHODS

Experimental design and statistical rationale

The study conducted here utilized human plasma collected from patients between 2018 and 2021 at UW Health, a 450-bed tertiary academic medical center in Madison, WI. The sample size was determined based on technical considerations for multiplexed proteomics and metabolomics approaches, as well as logistical constraints (20, 21).

Further details regarding statistical analysis, demographics, and clinical data are provided in the sections below.

Human plasma samples

Upon admission, plasma was obtained from 32 patients with *E. faecium* bacteremia and 44 patients with *E. faecalis* bacteremia. Enterococcal bacteremia was diagnosed through positive blood cultures and treated with antibiotics. *E. faecalis* and *E. faecium* bacteremia were differentiated using matrix assisted laser desorption ionization-time of flight (MALDI-TOF)-based rapid identification and confirmed using standard culture and biochemical methods. Additionally, plasma from 29 healthy controls was collected from blood bank volunteers.

Clinical data collection

Patient electronic medical records were reviewed to collect basic demographics, including age, gender, and comorbidities. Data on the infection and treatment (antibiotic use and source control methods) and clinical course included the organism type (*E. faecalis*/*E. faecium*), source of the bloodstream infection (endovascular, urine, abdominal fluid, etc.), antibiotic susceptibility, and laboratory values and markers of infection (e.g., serum creatinine, white blood cell count with differential, temperature, vital signs). The mean age of the patients was 59.6 ± 16.3 years, and 59% were male. Among the 76 patients with enterococcal bacteremia, 30.1% were infected with vancomycin-resistant *Enterococcus*, as confirmed by routine antimicrobial susceptibility testing in the clinical microbiology laboratory. The total duration of bacteremia included cases of persistent bacteremia (consecutive days of positive blood cultures) and in-hospital microbiologic relapse, defined as the recurrence of a positive blood culture after the first negative culture while receiving appropriate antibiotic therapy. The mean duration of bacteremia was 2.7 ± 1.9 days, with a median of 2 days and a range of 1–14 days. The mortality rate during hospitalization and within 1 year of infection onset was 21.7% and 38.6%, respectively.

Proteomics arm

Protein preparation

A 25 μ L aliquot of plasma from each patient was added to 200 μ L of lysis buffer containing 6 M urea, 7% SDS, 50 mM tetraethylammonium bromide (TEAB), one protease inhibitor tablet (Roche cat # 06538282001), and one PhosStop tablet (Roche cat # 04906845001), with the final pH adjusted to 8.1 using phosphoric acid. The samples were then sonicated in a water bath and vortexed for 5 min each. Disulfide bonds were reduced using 100 μ M dithiothreitol (DTT) for 30 min at 47°C, followed by brief cooling on ice. Alkylation was performed with 300 μ M iodoacetamide for 45 min at room temperature in the dark. The reactions were quenched with 100 μ M DTT at room temperature for 5 min. Samples were loaded onto S-Trap 96-well plates (Protifi) per the manufacturer's instructions. Samples were then digested with 5 μ g of trypsin in 115 μ L 50 mM TEAB for 3 hours at 47°C. Peptides were eluted sequentially with 125 μ L of 50 mM TEAB, then 5% formic acid (FA), and finally 50% acetonitrile with 5% FA and were then dried under vacuum centrifugation. The peptides were desalted using 50 mg Sep-Pak tC-18 cartridges (Waters cat # WAT054960) per manufacturer instructions and were quantified using a Pierce Quantitative Colorimetric Peptide Assay (ThermoFisher cat # 23275). A 50 μ g aliquot of each sample was designated for TMT labeling, with an additional 10 μ g from each sample combined into a final 50 μ g aliquot for use as the bridge channel.

TMT labeling

Aliquots were labeled with TMT-Pro 16-Plex reagents (cat # A44520, ThermoFisher Scientific, lot # WJ327115) as described previously with channel 134N reserved as the bridge channel (22).

Chromatography and mass spectrometry

Basic pH reverse-phase LC, followed by data acquisition through LC–MS2/MS3, was performed as previously described (22). Briefly, 75-min linear gradients of 22% to 35% acetonitrile and 10 mM ammonium bicarbonate were used to elute peptides on HPLC C18 columns (Biobasic). The resulting 96 fractions were combined as previously described (23). These fractions were then analyzed using tandem mass spectrometry (MS2/MS3) on an Orbitrap Fusion mass spectrometer (ThermoFisher) equipped with an in-line EASY-nLC 1000 (ThermoFisher). Separation and acquisition settings were performed using previously defined methods (24).

Statistics and data analysis

Thermo .raw files were converted to mzML files using MSConvert, with a peak picking (centroiding) filter applied to the conversion. The resulting mzML files were then input into Fragpipe version 20.0 and analysis was performed using MSFragger version 3.8, IonQuant version 1.9.8, and Philosopher version 5.0.0. Fragpipe was run using the TMT-16 MS3 quantification (TMT16-MS3) workflow. Briefly, this workflow specifies the use of DDA mode with low mass accuracy MS2 (Ion trap) for identification and quantification utilizing a high mass accuracy (Orbitrap) MS3. MSFragger settings included a precursor mass tolerance of 20 ppm for both upper and lower limits, strict trypsin digestion rules allowing for two missed cleavages, and the search was set to include static modifications of carbamidomethylation of cysteines, TMTpro on lysines and N termini, and variable modifications of oxidized methionine. A database containing the human reference proteome (UP000005640) concatenated with the reference proteomes from *Enterococcus faecalis* (UP000001415) and *Enterococcus faecium* (UP000321556) was used to match spectral IDs. All reference proteomes were downloaded from UniProt on 11 September 2023.

Proteins identified as differentially abundant were subjected to gene ontology (GO) analysis using all proteins identified in the experiment as the universe, as implemented in ClusterProfiler4.0 (25). All GO terms described have adjusted *P*-values ≤ 0.05 unless otherwise described.

Metabolomics arm

Plasma metabolite extraction

All steps were done on ice unless otherwise indicated. Plasma samples (50 μ L) were thawed for 30 min, and then 200 μ L of prechilled extraction solvent (100% methanol with 1 mM sulfamethazine as an internal standard) was added to each sample. Samples were mixed by vortexing for 2 min and then incubated at 20°C for 20 min to aid in protein precipitation. Samples were centrifuged at 16,000 $\times g$ for 15 min to pellet the protein precipitate. The supernatant was then transferred into a 96-well deep well plate, dried using a centrifugal low-pressure system, and stored at -80°C once dry.

Metabolomics LC-MS2 analysis

Metabolomics LC-MS2 was performed on a Q-Exactive mass spectrometer coupled to a Thermo Vanquish HPLC system. The chromatographic analysis was carried out on a Polar C18 100A LC Column 100 $\times$ 21 mm (catalog no. 00D-4759-AN). 5 μ L of each sample was injected and run on a 10-min gradient. 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 0.500 mL/min, and chromatographic separation was achieved using the following

gradient: 0 to 1 min 5% B; 1 to 7 min a linear increase from 5 to 100% B; 7 to 7.5 min held at 100% B; 7.5 min to 8 min a linear decrease from 100% to 5% B; and then 5% B from 8 min to 10 min.

Metabolite identification

Full scan MS spectra (m/z 100–1,500) were acquired and the top 5 most intense ions from a unique scan were fragmented. Dynamic exclusion was set to 10.0 s. The isolation window was set to 3.0 m/z with an isolation offset of 0.5 m/z , and the intensity threshold was set to $5e4$.

Raw file processing

Thermo .raw files were first converted to mzML using MSConvert (3.0.22155-0ff594) (26). mzML files were processed together using MzMine (3.5.0) (27) to identify metabolite features. Parameters for individual sample metabolite feature identification were as follows: mass detection (MS1 noise level: 1.0E3, MS2 noise level: 5.0E2), feature detection through ADAP chromatogram builder (min group size in # of scans = 4, group intensity threshold = 3,000, min highest intensity = 1,000, m/z tolerance = 0.005 Da or 10 ppm), feature detection chromatogram resolving (MS/MS scan pairing with retention time (RT) tolerance = 0.10 min and MS1–MS2 precursor tolerance = 0.0100 m/z ; local min search used with chromatographic threshold = 90%, min RT range 0.50 min, min relative height 0.01%, min absolute height 1,000, min ratio of peak top/edge = 1.7, peak duration 0.05–1 min, and min # data points = 4), ^{13}C isotope filter (m/z tolerance = 0.01 m/z , RT tolerance = 0.30 min, and maximum charge = 5). Parameters for the metabolome feature bucket table were as follows: join aligner (m/z tolerance = 0.01 m/z , m/z weight = 80, RT tolerance = 0.30 min, RT weight = 20), feature list filtering (at least two peaks per row), and gap filling (intensity tolerance = 10%, m/z tolerance = 0 m/z , RT tolerance = 0.4). Metabolite feature tables were then exported using the GNPS FBMN option to generate the required files for online GNPS FBMN analysis.

Feature annotations

Feature annotation done through the Global Natural Products Social Molecular Networking (GNPS) feature-based molecular networking (FBMN) workflow (version 28.2) (28). MS2 MGF files were exported from MzMine3 along with a feature quantification table and imported to the FBMN workflow. Spectral library files were set to search all GNPS spectral libraries. Library search min matched peaks were set to 6 and the score threshold to 0.7. Precursor mass tolerance was set to 0.02 Da and fragment ion mass tolerance to 0.02 Da. Features with annotation were then appended to the feature quantitation table exported from MzMine.

Molecular networking

A molecular network was created using the online workflow (<https://ccms-ucsd.github.io/GNPSDocumentation/>) on the GNPS website (<http://gnps.ucsd.edu>). The data were filtered by removing all MS/MS fragment ions within ± 17 Da of the precursor m/z . MS/MS 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 2.0 Da and an MS/MS fragment ion tolerance of 0.5 Da. A network was then created where edges were filtered to have a cosine score above 0.7 and more than six matched peaks. Furthermore, edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. Finally, the maximum size of a molecular family was set to 100, and the lowest scoring edges were removed from molecular families until the molecular family size was below this threshold. The spectra in the network were then searched against GNPS' spectral libraries. 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 six matched peaks.

Data processing of metabolite features

Metabolomics data were normalized using sulfamethazine as a single internal standard as described previously (29). Briefly, any feature containing a 0 value in any sample was discarded, and values for each remaining feature were divided by the observed value for sulfamethazine within each sample. This resulting value was then log₁₀ transformed and multiplied by 1E6.

Biomarker identification

An ensemble feature selection approach, which combines the Mann-Whitney U-test, Pearson and Spearman correlations, logistic regression, and four variable importance measures derived from two different implementations of the random forest algorithms *cforest* and *randomforest*, was employed as an R package to select features with minimal bias (30).

Metadata assessment

Metadata correlations were assessed in the following manner. First, categorical metadata associations were determined using Mann-Whitney *U* or Kruskal-Wallis tests when appropriate. Continuous metadata associations were determined using Pearson correlation. Associations in the figures represent the $-\log_{10}$ (*P*-value) of each test. All tests were performed in R.

Statistical analysis

All statistical analyses were completed as reported in the corresponding figure legends or methods details. R was used to conduct all tests. For all tests, significance is denoted as follows: *****P* < 0.0001; ****P* < 0.001; ***P* < 0.01; **P* < 0.05, ns—not significant.

Machine learning analysis

Machine learning analysis was performed using the tidymodels framework in R (31). Models built with proteomics and metabolomics data were restricted to putatively annotated metabolites and features that had no missing values across the 105 samples included in the study. Models built using the clinical metadata excluded race, death_at_one_year, peripheral_vascular_disease, dementia hemiplegia, highly_active_antiretroviral_therapy, and source, as these columns had limited variability across our set of patients.

Lasso regression models were trained on 80% of the data, using fivefold cross-validation to optimize and evaluate model performance. In cases where multiple models achieved identical performance on the training set, the first reported model was selected. These models were then evaluated on the remaining 20% of the data, which served as an unseen test set. Performance on the test set was visualized using receiver operator characteristic (ROC) curves, and estimates of the model coefficients were provided.

Nanopore sequencing

Strain isolation

E. faecium and *E. faecalis* clinical isolates were recovered from blood culture vials by plating patient blood onto solid media. Single isolated colonies were then recovered and inoculated into liquid culture and once cultures reached turbidity, a final concentration of 15% glycerol was added, and stocks were frozen at -80°C . Strain identification was differentiated via MALDI-TOF rapid identification and was confirmed using traditional culture-based and biochemical methods.

DNA extraction and QC

Enterococcus faecium and *Enterococcus faecalis* from the frozen glycerol stocks were inoculated into 1 mL of brain heart infusion (BHI) broth. These cultures were subsequently incubated at 37°C shaking at 220 RPM overnight. Bacteria were pelleted by centrifugation at $5,000 \times g$ for 5 min, after which supernatant was discarded and the pellets were resuspended in a modified lysis buffer containing 1 mL of QIAGEN B1 buffer supplemented with 2.29 mg/mL lysozyme, 0.29 mg/mL Labiase (cat no: OZ-30EX OZEKI Ci., Ltd.), and 0.2 mg/mL RNAaseA (cat # 1007885 QIAGEN). Labiase was added due to our observation that lysozyme alone was insufficient to lyse many of these clinical strains, due to well-documented lysozyme resistance among enterococcal clinical isolates (32). These cells were then incubated overnight at 37°C to facilitate lysis. The next day, 45 μ L of proteinase K solution (cat # RP107B-10 QIAGEN) was added to each sample and incubated for 1 hour at 37°C. Next, 0.35 mL of QIAGEN Buffer B2 was added to each sample, tubes were mixed several times by inversion and then incubated at 50°C for 30 min. The resulting high molecular weight DNA was then purified using QIAGEN Genomic-tip 20/G (cat # 10223 Qiagen). A genomic tip was equilibrated with 1 mL of buffer QBT. Samples were vortexed for 10 s at maximum speed and then applied to the equilibrated genomic tips. After all liquid had passed through, each genomic tip was washed three times with 1 mL of QIAGEN Buffer QC. Genomic DNA was then eluted by applying 1 mL of Buffer QF twice. Also, 1.4 mL of room temperature isopropanol was then added to the eluate, and it was inverted several times to precipitate the DNA. To collect the DNA, samples were centrifuged at $12,000 \times g$ for 15 min at 4°C. The supernatant was then carefully removed, and then the DNA pellet was washed with 1 mL of 70% ethanol. The samples were then vortexed briefly and then centrifuged at $12,000 \times g$ for 10 min at 4°C. This wash and centrifugation step was then repeated. The supernatant was then removed, taking great care not to disturb the pellet, and was air-dried for 10 min before being resuspended in 50 μ L of 10 mM Tris-Cl, pH 8.5. The DNA was then dissolved by shaking at room temperature overnight, followed by gentle pipetting with a wide bore pipette tip.

The resulting DNA was then checked for purity by Nanodrop and DNA concentration was assessed using Qubit dsDNA Quantification Assay Kit, broad range (ThermoFisher cat # Q32853). A subset of purified DNA samples was run on a Genomic DNA ScreenTape (Agilent, cat # 5365) to assess DNA length and integrity.

Barcoding and pooling

Four hundred nanograms of purified DNA from each enterococcal strain was barcoded using Native Barcoding Kit 96 V14 (Oxford Nanopore cat # SQK-NBD114.96). The manufacturer's ligation sequencing genomic DNA (gDNA) – Native Barcoding Kit 96 V14 version NBE_9171_v114_rev1_15Sep2022 protocol was followed as described to perform DNA repair and end prep, native barcode ligation, adapter ligation and clean up, priming and loading the SpotOn flow cell.

Data acquisition

A 10.4.1 flow cell was loaded with ~20 fmol pooled barcoded DNA (estimated from the library concentration combined with the TapeStation-derived distribution of molecular weight). Minknow 23.07.5 was used to acquire the data in high accuracy basecalling mode, facilitated by a custom-built desktop computer equipped with a Nvidia GeForce RTX 3070 GPU.

Sequencing analysis

Unwanted *Streptomyces fulvissimus* reads present in our Labiase preparation were bioinformatically decontaminated by mapping all reads to a reference database using centrifuge (33) and reads mapping to streptomyces were then removed using seqtk (34). Decontaminated reads were then used as input into the Epi2me labs graphic interface

which was used to run the wf-bacterial-genomes nextflow workflow (35). Briefly, this workflow concatenates input files and generates per read stats via bamstats (36), performs *de novo* assembly via flye (37), polishes with medaka (38), performs multi locus sequence typing (39) to infer the identity of the isolate, and lastly performs antimicrobial resistance calling via ResFinder (40) to identify genes associated with antimicrobial resistance. Basic assembly statistics such as N50 and the number of contigs were also generated using assembly_stats (41). The resulting FASTA assemblies were then used as input to Roary (42) to generate a pangenome.

RESULTS

Overview of multi-omic EcB patient plasma analysis

Using a comprehensive multi-omic strategy, we aimed to profile the host response to EcB (Fig. 1A). We analyzed plasma samples from cohorts consisting of 29 healthy volunteers, 32 patients diagnosed with *E. faecium* bacteremia, and 44 patients with *E. faecalis* bacteremia. Extensive clinical metadata were collected corresponding to each EcB patient, including gender, the Charlson comorbidity index, duration of bacteremia, day of blood draw, and mortality during admission (Fig. 1B). Additionally, antimicrobial susceptibility testing was performed on the enterococcal isolates obtained from each patient. As expected, based on the reported susceptibilities of clinical isolates, all *E. faecalis* isolates were resistant to vancomycin, while only a minority of *E. faecium* isolates displayed vancomycin resistance (43). This highlights that initial identification of the species can inform the selection of optimal antibiotic therapy for EcB.

In the TMT-guided proteomics data set, we quantified a total of 589 proteins mapping to the human proteome, with 278 proteins quantified across all samples. No proteins were found to map to the *Enterococcus* proteomes. In our metabolomics data set, we quantified a total of 10,092 features, with 693 of these features being putatively identified based on their MS2 profiles using GNPS molecular networking. Of these putatively identified metabolites, 212 were quantified across all samples. This comprehensive profiling enabled us to examine the differences between healthy individuals and those infected with *E. faecalis* or *E. faecium*, while also providing a unique opportunity to explore associations with mortality vs. survival.

Unsupervised hierarchical clustering of both the global proteomics and metabolomics data revealed a clear delineation between the healthy and infected groups (Fig. 1C and D). Significant differences between *E. faecalis* and *E. faecium* bacteremia were observed at the level of individual features (see Fig. 4A and 5A), yet the global dissimilarities between the two types of bacteremia were more subtle. This indicates large-scale similarity in the host response elicited by these two related pathogens.

Differences in plasma profiles of bacteremia types compared to healthy

We initially focused on identifying proteins most effective at differentiating EcB from healthy populations. Compared to healthy volunteers, 204 proteins were identified to be more abundant in individuals with EcB, while 85 proteins were significantly less abundant (Fig. 2A, with false discovery rate (FDR) adjusted *P*-value ≤ 0.05). Some proteins showed highly significant Benjamini-Hochberg adjusted *P*-values, reaching as low as 1×10^{-25} . Similarly, evaluation of the metabolomics data revealed that 427 features were significantly increased in the infected group compared to the healthy group, while 968 features were significantly decreased. FDR adjusted *P*-values for these features were as low as 1×10^{-50} (Fig. 3A). Of the metabolite features identified in our study, 29% of them were putatively identified using GNPS molecular networking (Fig. 3B). This partial annotation highlights a well-documented limitation inherent to current untargeted metabolomics analysis approaches (44).

To further understand the differences between these bacteremia types, we investigated how many of the specific proteins identified as significantly altered relative to healthy were shared among the different types of bacteremia. Additionally, we compared these

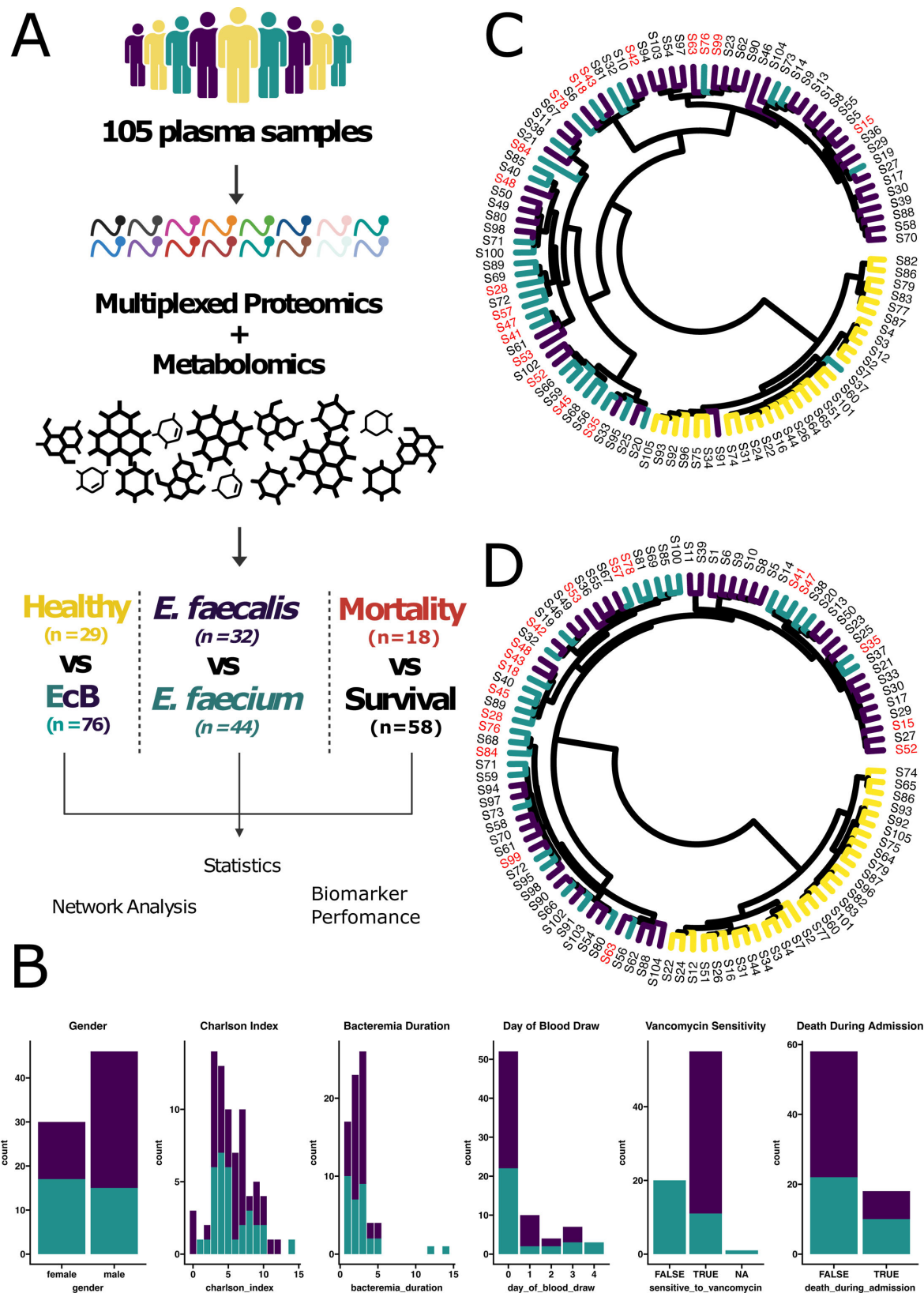


FIG 1 Multi-omic analysis of enterococcal bacteremia patient plasma. (A) Workflow for enterococcal bacteremia plasma analysis. (B) Distribution of values collected from selected clinical metadata fields across enterococcal bacteremia patients. (C) Unsupervised hierarchical clustering of proteomics data as visualized after calculating Euclidean distance and utilizing the Ward.D2 agglomeration method. Colors of branches indicate the species of infection (yellow = healthy (Continued on next page)

Fig 1 (Continued)

volunteer, navy = *E. faecalis*, teal = *E. faecium*) and color of sample ID labels represent mortality or survival (black = survival, red = mortality). (D) Unsupervised hierarchical clustering of metabolomics data as visualized after calculating Euclidean distance and utilizing the Ward.D2 agglomeration method. Colors of branches indicate the species of infection (yellow = healthy volunteer, navy = *E. faecalis*, teal = *E. faecium*) and color of sample ID labels represent mortality or survival (black = survival, red = mortality).

deviations from homeostasis observed upon infection to another clinically relevant pathogen by analyzing previously published proteomic differences between *S. aureus* bacteremia patients and healthy volunteers (18). We found that 24% (53) of the significant proteins that increased upon infection were common across all types of bacteremia, while 13% (30) were specific to *E. faecium* bacteremia, none were specific to *E. faecalis* bacteremia, and 42% (93) were specific to *S. aureus* bacteremia (Fig. 2D). In terms of proteins that were found to be significantly decreased in infections, 28% (94) were shared across all types of bacteremia, 2% (6) were specific to EcB, 3% (10) specific to *E. faecium* bacteremia, 2% (7) to *E. faecalis* bacteremia, and 32% (106) were found only in *S. aureus* bacteremia (Fig. 2D). This analysis underscores the complex and often overlapping protein expression profiles in bacterial infections and highlights specific protein markers that could potentially differentiate between these infections.

We explored the biological processes involving the proteins identified as statistically significant in comparisons between infected and healthy individuals, focusing on how these processes varied among bacteremia caused by *Enterococcus*, *E. faecalis*, *E. faecium*, and *S. aureus*. To facilitate this analysis, we conducted GO enrichment analysis on the proteins identified as significantly different in binary comparisons (Fig. S5). For both types of EcB, we observed an enrichment in biological processes such as neutrophil chemotaxis, tertiary granule lumen, focal adhesion, extracellular exosome, and inflammatory response (Fig. 2B). These findings indicate that despite the microbial differences, there are common host responses involving critical immune and structural cellular responses.

Enterococcal and *S. aureus* bacteremia share several biological processes that are significantly depleted upon infection. Notably, proteins involved in cholesterol metabolism were reduced in infected patients, as evidenced by the significant reduction in GO terms such as reverse cholesterol transport, cholesterol homeostasis, cholesterol metabolic process, very low-density lipoprotein particle, high-density lipoprotein particle, and blood microparticle. Similarly, we observed commonalities in processes related to blood clotting between EcB and *S. aureus* bacteremia, with significant depletion noted in the GO terms for blood coagulation, heparin binding, and zymogen activation. Of note, the platelet alpha granule lumen was the only GO term observed to have the opposing effects in different types of bacteremia. It was found to be enriched in the infected samples from EcB but depleted in those from *S. aureus* bacteremia (Fig. 2B).

The metabolites that were significantly differentially abundant (adjusted *P*-value ≤ 0.05) underwent enrichment analysis to determine if any class of molecule was significantly enriched. Interestingly, steroids were found to be significantly enriched in infected patients compared to all features detected in the experiment (Fig. 3C). A closer examination revealed that 10 out of the 12 molecules annotated as steroids could be more specifically described as bile acids. These included the primary bile acid cholic acid, the conjugated primary bile acids glycocholic acid (GCA), taurocholic acid (TCA), glycochenodeoxycholate (GCDCA), glycochenodeoxycholic acid, glycohyocholic acid, taurochenodeoxycholic acid as well as the conjugated secondary bile acids tauroursodeoxycholic acid (TUDCA), taurohyodeoxycholic acid, and glyoursodeoxycholic acid. The same functional enrichment was observed in *S. aureus*-infected samples (Fig. S2), where the abundances of TCA, GCA, TUDCA, 3beta-hydroxy-5-cholenoic acid, and 12-ketodeoxycholic acid were enriched. In addition, several phosphatidylcholine molecules were identified as being significantly depleted in *S. aureus* bacteremia (Fig. S2). While this class of molecules did not rise to the level of statistical significance in GO enrichment analysis from healthy to EcB patients, we did note that several phosphatidylcholines were

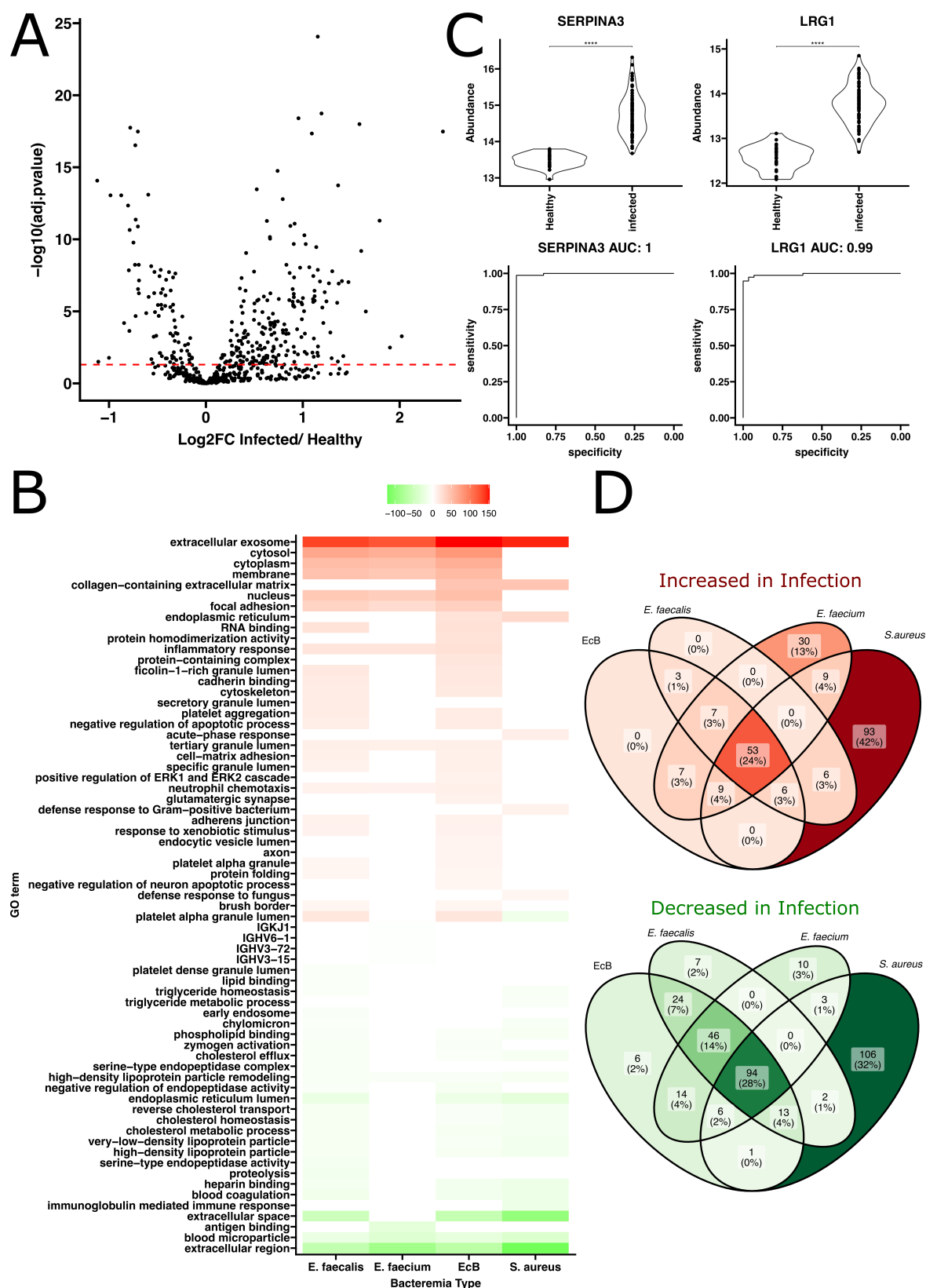


FIG 2 Untargeted proteomics reveal dramatic differences between bacteremia types and healthy plasma. (A) Volcano plot comparing $\log_2$ fold change and FDR adjusted P -values of protein abundances observed when comparing infected to healthy. (B) Significantly enriched GO terms from plasma proteomics of patients suffering from enterococcal, *E. faecalis*, *E. faecium*, and *S. aureus* bacteremia. (C) Top 2 performing protein biomarkers as ranked using ensemble feature selection. (Continued on next page)

Fig 2 (Continued)

Violin plot statistics indicate results of *t*-tests. (D) Venn diagram displaying the numbers of significantly different proteins (FDR adjusted *P*-value ≤ 0.05) shared between enterococcal, *E. faecalis*, *E. faecium*, and *S. aureus* bacteremia when compared to healthy.

also significantly depleted in the case of EcB bacteremia (Table S2). This highlights a potentially common lipid metabolism disruption in bacteremia, regardless of the bacterial species involved.

We next set out to evaluate the potential utility of features within our data set as biomarkers to distinguish healthy from infected samples. To rank these biomarkers, we used ensemble feature selection (EFS), an unbiased approach that integrates outcomes from eight distinct feature selection algorithms (30, 45), subsequently aggregating and assigning ranks to the scores. This approach helps mitigate many of the inherent biases often associated with individual algorithms (45, 46). We set the correlation threshold at 0 in the EFS, ensuring that well-performing biomarkers that were highly correlated with others would still be ranked highly.

Our analysis identified the top-ranked protein and metabolite biomarkers—serpin A3-1 (SERPINA3), leucine-rich alpha-2-glycoprotein (LRG1), threonylcarbamoyladenine, and 13-OXO-ODE—as highly effective at distinguishing infected from non-infected samples. When evaluated using logistic regression, these biomarkers produced area under the curve (AUCs) ranging from 0.99 to 1.0, indicating nearly perfect discrimination (Fig. 2C and 3D). Other top biomarkers identified in our study also were able to distinguish healthy and EcB plasma extremely well, as shown in Fig. S4, highlighting the extreme differences between the metabolic and proteomic profiles of healthy and infected individuals. A complementary machine learning model using Lasso regression showed similar performance when evaluated on an unseen test set, achieving a receiver operating characteristic (ROC) AUC of 1 (Fig. S11). There was notable overlap in the features important to the model and those identified with the alternate EFS-based approach (Fig. S5 and S11).

To benchmark these findings, we evaluated two well-known clinical biomarkers of inflammation, C-reactive protein (CRP) and serum amyloid A (SAA1), which are commonly used to monitor infection progression. Both proteins were significantly elevated in infected samples compared to healthy controls (Fig. S3). As expected, there were no significant differences in the levels of these proteins between samples infected with *E. faecalis* and *E. faecium*. When analyzed using the EFS approach, these clinically validated biomarkers demonstrated good performance, though they were not among the top performers in our data set, ranking 12th and 47th, respectively. Subsequent ROC analysis confirmed their ability to differentiate infected from uninfected samples with a high degree of sensitivity and specificity as expected, displaying AUC values of 0.97 for CRP and 0.92 for SAA1 (Fig. S3). These results reaffirm the reliability of CRP and SAA1 as indicators of infection while highlighting the potential of using EFS to discover additional biomarkers that could further enhance diagnostic accuracy.

Due to the critical role of cytokines as modulators of the immune system, we examined potential differences in cytokine profiles by bacteremia type within our data set. Given that cytokines are not readily detected in plasma using untargeted MS-based proteomics due to their low levels of absolute abundance (47), we utilized a previously reported method to infer cytokine profiles. This method leverages known interactions between proteins and cytokines to infer cytokine abundance (18). Applying this technique, we found that the inferred level of Tumor necrosis factor alpha (TNF- α) was significantly increased in EcB-infected samples compared to healthy samples, with an adjusted *P*-value ≤ 0.05 (Fig. S6). Additionally, we observed a trend indicating an increase in many other cytokines in infected samples compared to healthy samples, with interleukin 6 (IL-6) being notably higher, but they did not reach the statistical significance threshold set at an alpha level of 0.05.

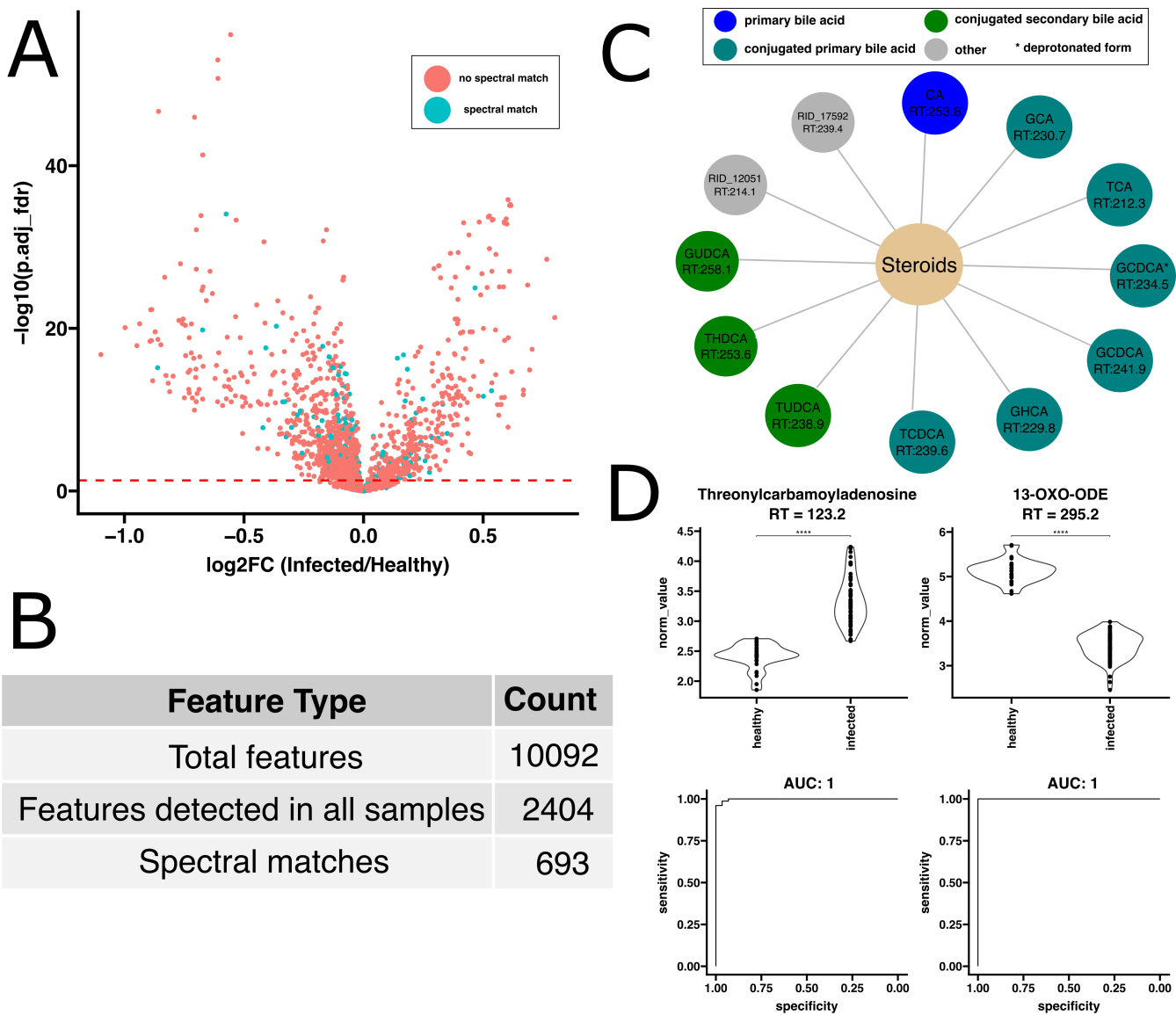


FIG 3 Untargeted metabolomics reveal dramatic differences between enterococcal bacteremia and healthy plasma. (A) Volcano plot comparing $\log_2$ fold change and FDR adjusted P -values resulting from t -tests of normalized metabolite abundances comparing infected to healthy. No spectral match (red points) indicates features that did not produce a spectral match to any of the metabolites in the GNPS database, while spectral match (blue points) indicates features that produced a spectral match to metabolites in the GNPS database. (B) Count of the number of metabolite features identified in this study. (C) Enrichment analysis of metabolite spectral matches found to be significantly different ($P_{\text{adj}} \leq 0.05$) in infected patients relative to healthy. Colors show bile acid class (blue = primary bile acid, teal = conjugated primary bile acid, and green = conjugated secondary bile acid). RID12051 and RID17592 are features corresponding to spectral matches to (((3 a,6b,7b)-3,6,7-trihydroxy-24-oxocholan-24-yl)amino)ethanesulfonic acid and (((3 a,12b)-3,12-dihydroxy-24-oxocholan-24-yl)amino)ethanesulfonic acid, respectively. (D) Evaluation of the top 2 performing metabolite spectral matches as ranked using ensemble feature selection for distinguishing enterococcal bacteremia patients from healthy. Violin plot statistics indicate results of t -tests.

Differences in systemic plasma profiles between EcB caused by *E. faecalis* and *E. faecium*

To further delineate the systemic plasma profiles caused by different enterococcal species, we focused on cases of EcB where $>90\%$ of the isolates were identified as either *E. faecalis* or *E. faecium*. We identified 70 proteins that were significantly enriched (adjusted P -values < 0.05) in plasma from patients infected with *E. faecalis* compared to those with *E. faecium*. Conversely, 30 proteins were significantly enriched in *E. faecium* relative to *E. faecalis* (Fig. 4A). Similarly, in the metabolomics data, we found 11

10.1128/msystems.01471-24 **14**

Fig 4 (Continued)

relative to *Enterococcus faecium*. (C) Immunoglobulin abundances compared across patients with *E. faecalis* or *E. faecium* bacteremia and healthy volunteers. Statistics indicate result of *t*-tests adjusted for multiple comparisons. (D) GO term enrichment analysis of the proteins found to be significantly enriched in proteins significantly more abundant in *Enterococcus faecium* patients relative to *Enterococcus faecalis*. (E) Top 2 performing protein biomarkers as ranked using ensemble feature selection for distinguishing patients infected with *Enterococcus faecalis* from those infected with *Enterococcus faecium*. Violin plot statistics indicate results of *t*-tests.

metabolites significantly increased in *E. faecalis* relative to *E. faecium*, and 33 significantly increased in *E. faecium* relative to *E. faecalis* (Fig. 5A).

The GO term enrichment analysis of proteins significantly increased in *E. faecalis*-infected samples revealed dramatic differences in immunoglobulin abundances (Fig. 4B). Upon further investigation, it was found that these increased abundances in *E. faecalis* relative to *E. faecium* were attributable to reduced levels of antibodies in *E. faecium*-infected samples, as the levels of immunoglobulins in *E. faecalis* infections and healthy samples were comparable (Fig. 4C). These differential abundances of immunoglobulins were also apparent in the GO term analysis comparing *E. faecium*-infected samples to healthy ones, where there was a notable reduction in immunoglobulin-related GO terms (Fig. 2B; Fig. S2).

When analyzing the GO term enrichment for proteins found to be more abundant in samples from patients infected with *E. faecalis*, we observed a significant enrichment in proteins associated with cholesterol metabolism. These included processes such as reverse cholesterol transport, cholesterol efflux, chylomicron, and very low-density lipoprotein particle dynamics (Fig. 4D).

In our analysis to distinguish between *E. faecalis* and *E. faecium* infections, the top proteomic biomarkers identified were immunoglobulin kappa variable 2–30 and retinol binding protein 4 (RBP4), which produced ROC AUCs of 0.76 and 0.79, respectively (Fig. 4E). Additionally, the top metabolomic features identified were retinol and PC(16:1/0:0), which had ROC AUCs of 0.77 and 0.82, respectively. The machine learning model showed

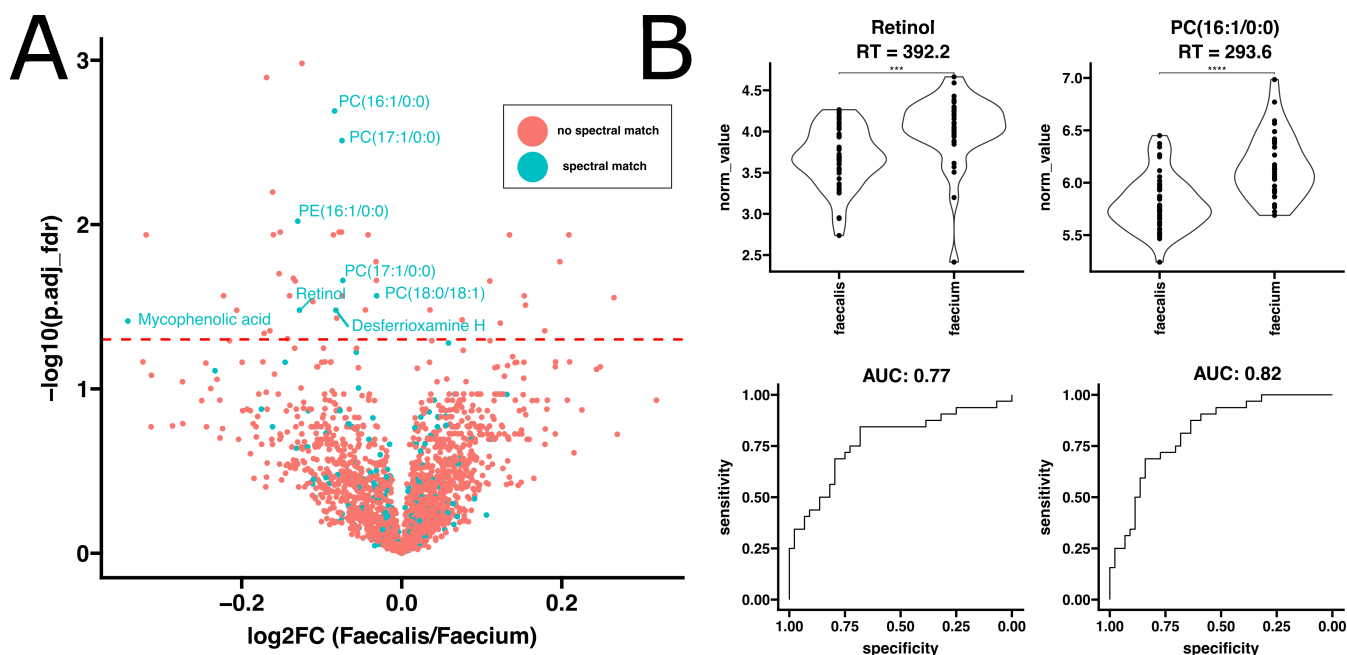


FIG 5 Untargeted metabolomics reveals differences between *Enterococcus faecalis* and *Enterococcus faecium* bacteremia. (A) Volcano plot comparing log₂ fold change and FDR adjusted *P*-values resulting from *t*-tests of normalized metabolite abundances comparing *E. faecalis* to *E. faecium*. (B) Evaluation of top 2 performing putatively identified metabolite spectra matches as ranked using ensemble feature selection for distinguishing enterococcal bacteremia patients from healthy. Violin plot statistics indicate results of *t*-tests.

similar performance when evaluated on an unseen test set, achieving an ROC AUC of 0.71 (Fig. S11). There was notable overlap in the features important to the model and those identified with the alternate EFS-based approach (Fig. S5 and S11). Additionally, we noted that no significant differences were observed in the inferred cytokine profiles between *E. faecalis* and *E. faecium* (Fig. S6).

In the realm of biomarker discovery, it is important to ensure that the differences observed in biomarkers are truly attributable to the variable of interest and not confounded by additional factors (48, 49). Our evaluation of clinical metadata suggested the potential presence of confounding variables within our data set. Specifically, transplant type was significantly associated with the levels of 9 out of the top 10 ranked protein biomarkers (Fig. S7). Further investigation revealed that this association could be challenging to disentangle from the type of pathogen causing the infection, as patients infected with *E. faecium* were more likely to have organ transplants in our cohort. To address this, we refined our analysis to include only patients who had not undergone a transplant. Under these conditions, we found that antibody levels remained significantly reduced in *E. faecium*-infected samples compared to *E. faecalis* (Fig. S10A). However, we observed a loss of significance for the remaining protein biomarkers APOC1, AZGP1, PCOX1, RBP4, and SERPINC1, indicating that their significance may be confounded by transplant status (Fig. S10C). Nevertheless, the direction of enrichment for APOC3 and RBP4 was preserved, narrowly missing our threshold for statistical significance with *P*-values of 0.098, and 0.064, respectively. When examining only non-transplant patients, metabolites like PE(16:1/0:0), mycophenolic acid, and PC(16:1/0:0) no longer displayed statistically significant differences between *E. faecalis* and *E. faecium*, suggesting the observed significance may also be driven by transplant type and treatment. However, PC(16:1/0:0), PC(17:1/0:0), and retinol still showed significant differences, indicating these conclusions were not confounded by transplant status (Fig. S10D).

Additionally, smoking status was significantly associated with the abundances of 5 out of the top 10 protein biomarkers (Fig. S7). When filtering for non-smoking patients, significant differences in the levels of these biomarkers between *E. faecalis* and *E. faecium* persisted, suggesting that these biomarkers were not confounded by smoking status (Fig. S10B). Interestingly, AZGP1, a top biomarker for distinguishing *E. faecalis* from *E. faecium* and a gene reported to be overexpressed in the airway upon smoking (50), was not found to be influenced by smoking status in our data set.

A machine learning model trained solely on clinical metadata effectively utilized antibiotic sensitivity profiles to predict the type of EcB with high accuracy, achieving an ROC AUC of 0.94 on the test set (Fig. S12A). However, when features not available at the time of admission were excluded, model performance decreased significantly, resulting in an ROC AUC of 0.75 (Fig. S12B).

Prediction of clinical outcome

Next, we set out to combine the multi-omic data with the extensive medical metadata associated from our patient cohort to identify biomarkers predictive of mortality at the time of hospital presentation and admission. Our unsupervised hierarchical clustering analyses of proteomics (Fig. 1C) or metabolomics data (Fig. 1D) revealed no high-level associations of mortality with overall proteomic or metabolomic profiles. However, when conducting feature-level analysis, we identified specific proteins that were significantly associated with mortality outcomes. We found that 84 proteins were significantly enriched in patients who died, while 25 proteins were significantly enriched in patients who survived (Fig. 6A).

The enrichment analysis of proteins significantly associated with survival revealed an association with biological processes and components linked to cellular structures and signaling functions. Specifically, the significant proteins were enriched for terms associated with the Golgi apparatus, external side of the plasma membrane, calcium ion binding, and extracellular matrix (Fig. 6B). These terms suggest a role for cellular trafficking, membrane interactions, and structural integrity in influencing patient

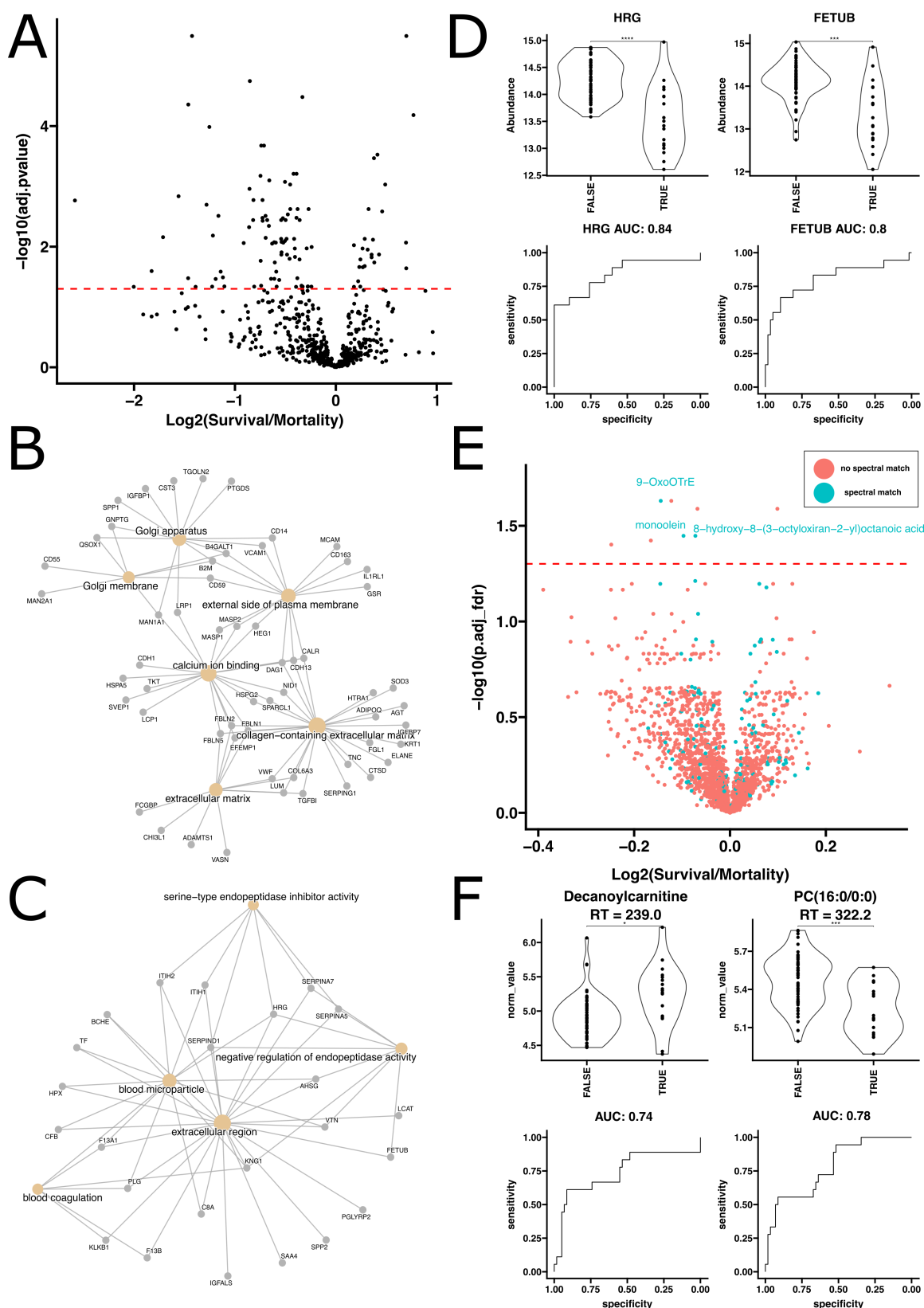


FIG 6 Multi-omic discrimination of mortality from survival. (A) Volcano plot comparing $\log_2$ fold change and FDR adjusted P -values of protein abundances observed when comparing patients who suffered mortality during admission to those who survived. (B) GO term enrichment analysis of the proteins found to be significantly enriched (FDR adjusted P -value ≤ 0.05) in mortality relative to survival. (C) GO term enrichment analysis of the proteins found to be significantly enriched (FDR adjusted P -value ≤ 0.05) in survival relative to mortality. (Continued on next page)

Fig 6 (Continued)

enriched (FDR adjusted P -value ≤ 0.05) in survival relative to mortality. (D) Evaluation of top 2 performing protein biomarkers as ranked using ensemble feature selection for distinguishing mortality from survival. Violin plot statistics indicate results of t -tests. (E) Volcano plot comparing log₂ fold change and FDR adjusted P -values resulting from t -tests of normalized metabolite abundances observed when comparing patients who suffered mortality during admission to those who survived. (F) Evaluation of top 2 performing metabolite spectral matches as ranked using ensemble feature selection for distinguishing mortality from survival. Violin plot statistics indicate results of t -tests.

survival, potentially through mechanisms that enhance cellular resilience and communication.

Conversely, proteins significantly associated with mortality were enriched in terms related to acute response and homeostasis disruptions. Notably, these proteins were associated with blood coagulation, blood microparticle, and negative regulation of endopeptidase activity serine-type endopeptidase inhibitor activity, (Fig. 6C). This enrichment highlights the importance of coagulation processes and immune response regulation in the context of mortality, indicating that disruptions in these pathways might contribute to poorer outcomes in patients.

Histidine-rich glycoprotein (HRG) and fetuin-B (FETUB) were identified as the two best performing biomarkers for distinguishing mortality from survival in our study, showing excellent predictive accuracy when evaluated by logistic regression where they produced ROC AUCs of 0.84 and 0.80, respectively (Fig. 6D). Our complementary machine learning model showed similar performance when evaluated on an unseen test set, achieving an ROC AUC of 0.83 (Fig. S11). Once again, we noted considerable overlap in the features important to the model and those identified with the alternate EFS-based approach (Fig. S5 and S11).

In our investigation into the impact of metabolites on mortality outcomes, we found that very few metabolite features showed significant differences. Specifically, seven metabolites were significantly increased in patients who succumbed to mortality, while only two showed significant increases in patients that survived (Fig. 6E). Among these, the top two identified metabolite biomarkers, decanoylcarnitine and PC(16:0/0:0), showed moderate discriminatory accuracy, producing ROC AUCs of 0.74 and 0.78, respectively (Fig. 6F). Further assessment of our top 10 protein and metabolite biomarkers of mortality for potential confounding variables revealed no significant associations, with the exception of ICU admission status, an expected association with mortality (Fig. S7). Highlighting the degree to which ICU admission status is confounded with mortality, a machine learning model trained with clinical metadata was able to leverage ICU admission status to predict mortality with a ROC AUC of 0.91 (Fig. S13A). When ICU admission status was withheld from the model, the model lost all predictive accuracy as expected (Fig. S13B).

Lastly, we asked if there were any specific genomic features of the strains of *E. faecalis* or *E. faecium* that were associated with mortality. Using nanopore sequencing and *de novo* assembly, we analyzed the gDNA from *E. faecium* and *E. faecalis* isolates collected from patients in our study. However, hierarchical clustering based on the presence or absence of gene content across the clinical strains revealed no significant correlations with mortality outcomes (Fig. S9).

DISCUSSION

The progress achieved in medical and surgical interventions has brought about a demographic shift toward an aging population, which often presents with increased comorbidities and compromised immune systems. Simultaneously, there is a rising prevalence of drug-resistant pathogens, necessitating the frequent use of broad-spectrum antibiotics. This combination has created a “perfect storm,” paving the way for the emergence of less virulent but intrinsically antibiotic-resistant commensals, such as *Enterococcus* spp., to become significant pathogens in invasive diseases such as bacteremia. Despite the increasing relevance of enterococci as pathogens, there has

been a lack of comprehensive, unbiased research describing the host systemic response to enterococcal bacteremia using a multi-omics approach.

Beyond advancing our understanding, a comprehensive evaluation of the systemic response to EcB also serves a practical purpose. Patients with EcB often do not present with the classical signs and symptoms of infection such as fever, elevated white blood cell count, or localized inflammation and pain. Instead, their symptoms can be vague and non-specific, including generalized weakness, malaise, and weight loss, which complicates diagnosis and assessment of their condition. Currently, molecular tests that analyze the host systemic response are not commonly used in clinical settings to diagnose bacteremia or inform therapeutic decisions. However, the data from our multi-omic analysis of the systemic response to EcB provide valuable insights that could potentially be leveraged to help predict the presence of enterococcal bacteremia, distinguish whether the infection is caused by *E. faecalis* or *E. faecium*, and assess if a patient is responding well to the current treatment strategy.

In principle, molecular diagnostics that leverage differences in systemic host response have several desirable qualities. They utilize plasma, a clinically accessible and easy-to-obtain biospecimen, and can detect disease-relevant host proteins without any need for signal amplification, which is necessary in blood culture-based testing. As a result, these tests offer the potential to provide informative results within hours rather than days. If true differences in host systemic response exist, they represent attractive targets to exploit for developing novel diagnostics.

Considering these concepts, our work employs high-resolution plasma profiling using TMT proteomics and metabolomics to begin dissecting the host responses of patients afflicted with EcB, focusing on the two most common species: *E. faecalis* and *E. faecium*. This approach is driven by three overarching goals: (i) to characterize the host response EcB relative to homeostasis (i.e., healthy); (ii) to determine the systemic differences between *E. faecalis* and *E. faecium* bacteremia; (iii) to define the systemic response associated with increased mortality, offering a starting point for the future development of molecular methods that could be used to stratify patients based on predicted outcomes and subsequently intervene to improve.

We observed that the systemic response to EcB is dramatically different from homeostasis, both globally and on the scale of individual protein and metabolite features. We identified several individual protein and metabolites based on spectral matches (annotation at level 2 or 3 according to 2007 metabolomics standards initiative guidelines [51]) that were able to predict the presence of EcB relative to healthy volunteers with near-perfect discriminatory power. This indicates the potential for molecular diagnostics to predict EcB, but for these diagnostics to be more clinically useful than general molecular markers of inflammation such as CRP or SAA1, they must have some specificity to EcB rather than solely being indicators of inflammation. When we leveraged a complementary data set comparing *S. aureus* bacteremia to healthy volunteers to examine this possibility, we found several significant proteins and biological processes that differed across these two types of bacteremia, in addition to many expected conserved responses. Notably, proteins associated with the platelet alpha granule lumen were decreased in *S. aureus* bacteremia. Platelet alpha granules contain proteins with direct microbicidal properties as well as chemokine functions. Two of the proteins associated with platelet alpha granules, PPBP (CXCL7) and PF4 (CXCL4), are potent chemokines that attract neutrophils. This, combined with the observation that neutrophil-associated proteins are significantly enriched in EcB but not in *S. aureus* bacteremia, suggests differences in the role of platelet/neutrophils between EcB and *S. aureus* bacteremia and/or the differential effect of these organisms on platelet and neutrophil function (52, 53)

Although both the *S. aureus* and EcB studies were conducted in our lab using similar high-resolution TMT proteomic workflows, we cannot directly compare the two data sets due to limitations inherent in TMT-based proteomics. TMT normalization requires a “bridge” channel that consists of a pooled aliquot from all samples in the study. Since our

studies were conducted independently at different times, there was no shared pooled sample. Nevertheless, the observed differences in relation to healthy patients suggest distinct features of the host response that could help differentiate between these types of bacteremia. A larger study specifically designed to compare these and other types of bacteremia is warranted to confirm this through direct comparisons.

When we applied previously established methods of proteomics-based cytokine inference, we inferred there to be a significant increase in TNF-alpha in EcB. Notably, IL-6 showed a similar increase in EcB, but narrowly missed our threshold for statistical significance. TNF-alpha and IL-6 are two major inflammatory cytokines that are elevated in patients with bacteremia and sepsis (54), and their production by the innate immune system in response to bacterial infection is likely a major driver of the general inflammatory responses observed. As such, these cytokines are unlikely to be useful for specifically predicting EcB, instead functioning as markers of general inflammation.

As expected, several proteins involved in the acute phase response or inflammatory processes were observed to be enriched during systemic infections with both analyzed enterococcal species and *S. aureus*. Most of the features we identified as being most effective at distinguishing healthy from infected individuals have been previously reported to be biomarkers of other inflammatory processes; these include gelsolin (55, 56), leucine-rich alpha-2-glycoprotein 1 (LRG1) (57), and lipopolysaccharide binding protein (LBP) (58). Furthermore, proteins involved in cholesterol metabolism were noted to be reduced upon infection in all bacteremia types. This observation aligns with cholesterol's involvement in a myriad of biological processes, including immunity, cellular membrane functions, signaling, pathway regulation, and as a precursor for the synthesis of steroid hormones, bile acids, vitamin D, and oxysterols (59). Both low-density lipoprotein (LDL) and high-density lipoprotein (HDL) cholesterol levels are reported to be reduced in cases of sepsis, regardless of the causative organism (60), and here, we show that hypocholesterolemia is also a major feature of EcB.

Our observations that primary and secondary bile acids were significantly enriched upon infection across all bacteremia types clearly indicate cholestasis, where inflammation caused by proinflammatory cytokines results in impaired bile acid flow and increased bile acid concentrations in serum/plasma (61). This process is commonly observed during systemic inflammation, which also has the capacity to activate and amplify coagulation. Consequently, changes in proteins associated with coagulation have been reported as a general response to sepsis, regardless of the causative organism (62). We observed that several of our top identified features for distinguishing healthy from infected individuals are known to be involved in the degradation of fibrin clots or platelet aggregation. These include alpha 1-antichymotrypsin (SERPINA3) (63, 64) and enolase 1 (ENO1) (65), which were increased in infections, as well as protein C inhibitor (SERPINA5) (66, 67), CLEC3B (tetranection) (68), and 13-OXO-ODE (69), which were decreased.

Most of the top metabolite features associated with the presence of EcB, as reported in this study, represent novel associations with bacteremia, suggesting the potential to be specific markers for EcB. Two of the top features associated with EcB, 9-oxooctre and cyclo(l-phe-d-pro), have noted antimicrobial activities. 9-Oxooctre, an oxylipin produced by the LOX, COX, and CYP450 pathways (70), has been shown to possess antimicrobial activity against various bacterial and fungal species (71), along with anti-inflammatory properties. Cyclo(l-phe-d-pro) is a diketopiperazine, the smallest cyclic peptides known, isolated from Gram-positive bacteria, fungi, and higher organisms (72). It has demonstrated strong antibiotic activity against *Vibrio anguillarum* (73) and likely possesses antimicrobial activity in other contexts as well. The observed reduction in these metabolites during infection could indicate the consumption of these antimicrobial metabolites produced by the host in combating the infection.

Another top biomarker, threonylcarbamoyladenosine, a nucleoside modification found in all kingdoms of life, has been noted to restrict translation initiation to the AUG start codon and suppress frameshifting at tandem ANN codons (74). This modification

has also been proposed as a strong candidate biomarker for COVID-19 infection and severity (75). To determine the specificity of these metabolite biomarkers for EcB, it would be essential to conduct a comprehensive study directly comparing plasma from multiple types of bacteremia patient populations.

In our metabolomics analysis, we note that three of the most effective metabolites at differentiating healthy from infected individuals have roles as industrial plasticizers: phthalic anhydride, 7-bis(2-ethylhexyl) phthalate, and 5-tris(2-butoxyethyl) phosphate (76, 77). These differences were likely due to logistical constraints that resulted in subtle differences in the plastics used for plasma collection between sources. This highlights the importance of seemingly innocuous components of experimental design in highly sensitive, untargeted multi-omic approaches, as has been reported elsewhere (78). These findings suggest that some metabolite features distinguishing healthy from infected samples may be confounded by differences in sample collection. Therefore, caution is advised when interpreting metabolites that show differences between healthy and infected samples in this data set. Metabolites identified as industrial chemicals, which are not typically expected to be found in blood, are likely unwanted artifacts.

Differences between *E. faecalis*-infected samples from *E. faecium*-infected samples proved to be less stark than those observed when comparing healthy to infected individuals. However, we identified several features that could distinguish the two infection types with moderate performance, correctly predicting the outcome in around ~80% of cases in our data set. Distinguishing *E. faecalis* from *E. faecium* based on host response has the potential to be a clinically important diagnostic tool, as it could inform the optimal choice of antimicrobial therapy more quickly than the current state-of-the-art methods, potentially improving patient outcomes as a result. However, our data suggest that the ability to distinguish these highly related infections based on systemic host response, as observed in plasma, is limited to a best-case scenario of around 80% accuracy. This degree of predictive power may not be sufficient for clinical use, where clinicians would certainly opt for slower, but more accurate, current methods.

Even so, our approach uncovered interesting differences between *E. faecium* and *E. faecalis* bacteremia. Most strikingly, consistent reductions in immunoglobulins were observed in *E. faecium* relative to both *E. faecalis* and healthy samples. Statistical assessment of the extensive clinical metadata collected indicated that this association was unlikely to be driven by any confounding differences, for example, organ transplant status, offering strong evidence that immunoglobulins are truly reduced in *E. faecium* bacteremia compared to *E. faecalis*. This raises the question of whether this observation is due to the manipulation of host processes by the pathogen after infection, or whether they predate the infection and instead influence susceptibility. The potential for this reduction in immunoglobulins to be a direct consequence of *E. faecium* infection exists, as there are many reports of important human pathogens reducing immunoglobulins through various mechanisms, including the direct degradation of antibodies, as demonstrated to be important virulence strategies in several bacterial species (79, 80). Alternatively, the reduction we observe in *E. faecium* bacteremia may be an indication that patients with lower titers of immunoglobulins are particularly susceptible to *E. faecium*, perhaps indicating that antibodies are more important for preventing *E. faecium* bacteremia than they are for *E. faecalis*. *E. faecium* is highly prevalent in patients with immunosuppressive diseases, notably solid organ transplant patients, supporting this notion (81). More research is necessary to conclusively determine which of these two possibilities explains the difference in immunoglobulins reported between the two cohorts.

Both the metabolomics and proteomics data pointed to differences in retinol (vitamin A) abundance and transport between the two types of EcB, with both retinol and RBP4 being increased in *E. faecium* bacteremia relative to *E. faecalis*. These were among the top proteomic and metabolomic features capable of discriminating the infections. While RBP4 levels were potentially confounded by transplant status in our data set, narrowly missing our threshold for significance when considering only patients

without a transplant, retinol levels remained significantly different. This suggests that the differences in these retinol-associated features were most likely due to *E. faecalis* and *E. faecium*. Retinol levels decline during the acute-phase response to infection as a consequence of reduced RBP transcription in the liver (82) and increased urinary loss, suggesting that the acute-phase response to *E. faecium* bacteremia may be muted relative to what is observed in *E. faecalis*-driven bacteremia. Retinol is important for the function of various aspects of the innate and adaptive immune system, and the differences observed in this study may impact immune system function (83). Perhaps as a consequence of a reduced acute-phase response, we also noted significant decreases in proteins associated with cholesterol metabolism in *E. faecalis* compared to *E. faecium* bacteremia and healthy volunteers. While the suppression of serum lipoproteins in response to infection has been reported previously, differences have never been reported between two types of closely related bacteremia. It is possible that these differences function as an indicator of the underlying disease severity, as lipoprotein levels in serum have been shown to correlate with infection severity (84), and *E. faecium* tends to be less virulent than *E. faecalis* (85). If this were indeed the case, we would expect to see lipoproteins also associated with mortality, a finding not observed in our study. Thus, the underlying reasons for these interesting differences in proteins involved in cholesterol metabolism remain unclear.

The evaluation of the systemic responses comparing mortality to survival uncovered several important differences among these patients. We found that features derived from proteomics data were more effective at predicting mortality status than those derived from metabolomics data. The biological processes observed to decrease in patients who succumbed to mortality, compared to those who survived, primarily consisted of the same ones found to be significantly different when comparing healthy individuals to those infected. This suggests that the degree of alteration relative to baseline for proteins involved in blood coagulation, blood microparticles (lipoproteins), and negative regulation of endopeptidase (SERPINS) activity correlates with the severity of EcB. Conversely, the biological processes that were found to be significantly increased in mortality were vaguer, encompassing calcium ion binding, the extracellular matrix, the external side of the plasma membrane, and the Golgi apparatus. The potential biological relevance of these increases is challenging to discern, indicating that while they are associated with mortality, their specific roles in the pathophysiology of the disease remain unclear and require further investigation.

A number of features from the proteomics data were able to predict mortality, with classification accuracies greater than 80%. Notably, decreases in two cysteine protease inhibitors belonging to the type-3 cystatins class of structurally and functionally related proteins (86), HRG and FETUB, were the best predictors of mortality. These proteins have roles in blood coagulation, and HRG has been shown to possess antibacterial properties (87). They have also been identified as biomarkers of mortality in various studies, with evidence indicating that lower levels are associated with mortality in conditions such as COVID-19 (88), *S. aureus* bacteremia (18), as even in mouse models of sepsis, where administration of exogenous HRG was able to improve outcomes (89). This evidence suggests that a similar therapeutic approach involving HRG may be beneficial in the context of EcB.

The degree of accuracy displayed by these potential biomarkers could be clinically relevant, especially since there is currently no existing quantitative method to stratify patients specifically in the setting of EcB. Diagnostic tests that achieve ROCs ranging from 0.8 to 0.9 are considered to have excellent discriminatory power (90), indicating that the accuracy levels we observed could be highly useful in clinic settings. Further research is necessary to determine whether the predictive accuracy observed in this study can be replicated in an independent cohort. Additionally, it would be important to investigate whether predicting patient outcomes can be effectively paired with targeted interventions.

In conclusion, our study has significantly advanced a powerful multi-omics framework, as outlined in previous research (18) to produce the most granular profile of the systemic response to EcB to date. We have reported significant differences between the multi-omic plasma profiles of EcB relative to healthy volunteers, differences driven by infections with *E. faecalis* versus *E. faecium*, and distinctions between outcomes related to mortality and survival. These data are easily explorable at https://gonzalezlab.shinyapps.io/EcB_multiomics/. The culmination of these efforts represents a significant stride toward characterizing the systemic response to EcB and its relationship to other bacteremia types, while enhancing future abilities to risk-stratify patients for improved antimicrobial and/or immunotherapy approaches.

ACKNOWLEDGMENTS

C.B. and L.A.-R. were supported in part by the UCSD Graduate Training Program in Cellular and Molecular Pharmacology through an institutional training grant from the National Institute of General Medical Sciences, T32 GM007752. C.S. was supported in part by the NIH/NIAMS T32 AR064194 Rheumatic Diseases Training Grant. D.J.G. is supported by R01AI148417. This study was also supported by the UCSD Collaborative Center of Multiplexed Proteomics.

AUTHOR AFFILIATIONS

- ¹Biomedical Sciences Graduate Program, UC San Diego, La Jolla, San Diego, California, USA
- ²Department of Pharmacology, University of California San Diego, La Jolla, San Diego, California, USA
- ³Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, San Diego, California, USA
- ⁴Department of Pharmacy, UW Health, Madison, Wisconsin, USA
- ⁵School of Pharmacy, University of Wisconsin-Madison, Madison, Wisconsin, USA
- ⁶Department of Pediatrics, UC San Diego, La Jolla, San Diego, California, USA
- ⁷Center for Microbiome Innovation, University of California at San Diego, La Jolla, San Diego, California, USA
- ⁸Sharp Rees Stealy Medical Group, San Diego, California, USA

AUTHOR ORCIDs

- Charlie Bayne  <http://orcid.org/0000-0003-1870-5298>
- Leigh-Ana Rossitto  <http://orcid.org/0000-0002-7744-3630>
- Carlos Gonzalez  <http://orcid.org/0000-0002-4673-4048>
- Haoqi Nina Zhao  <http://orcid.org/0000-0003-3908-630X>
- Victor Nizet  <http://orcid.org/0000-0003-3847-0422>
- George Sakoulas  <https://orcid.org/0000-0002-6063-6217>
- David J. Gonzalez  <http://orcid.org/0000-0003-1423-5970>
- Warren Rose  <http://orcid.org/0000-0001-5012-5993>

FUNDING

Funder	Grant(s)	Author(s)
HHS NIH National Institute of General Medical Sciences (NIGMS)	GM007752	Charlie Bayne Leigh-Ana Rossitto
HHS NIH National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)	AR064194	Concepcion Sanchez
HHS NIH National Institute of Allergy and Infectious Diseases (NIAID)	R01AI148417	David J. Gonzalez

AUTHOR CONTRIBUTIONS

Charlie Bayne, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review and editing | Dominic McGrosso, Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review and editing | Concepcion Sanchez, Formal analysis, Methodology, Visualization | Leigh-Ana Rossitto, Formal analysis, Methodology | Maxwell Patterson, Formal analysis, Visualization, Writing – review and editing | Carlos Gonzalez, Funding acquisition, Methodology, Supervision, Writing – review and editing | Courtney Baus, Data curation, Methodology | Cecilia Volk, Investigation, Methodology | Haoqi Nina Zhao, Investigation, Methodology | Pieter Dorrestein, Data curation, Funding acquisition, Methodology, Resources | Victor Nizet, Conceptualization, Funding acquisition, Methodology, Writing – review and editing | George Sakoulas, Conceptualization, Methodology, Project administration, Resources, Writing – review and editing | David J. Gonzalez, Conceptualization, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review and editing | Warren Rose, Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review and editing

DATA AVAILABILITY

The data reported in this manuscript are available on MassIVE under the following identifiers: proteomics, [MSV000096728](https://massive.ucsf.edu/MSV000096728); metabolomics, [MSV000096729](https://massive.ucsf.edu/MSV000096729). All analysis of data downstream of Fragpipe (proteomics) or MzMine3 (metabolomics) processing can be found in the GitHub repo at https://github.com/baynec2/EcB_multiomics. An interactive data analysis app to explore the data contained in this paper can be found hosted at https://gonzalezlab.shinyapps.io/EcB_multiomics/. The source code to build the shiny web application can be found in the 04_shiny_app directory within the GitHub repo. Any other data are available upon request.

ETHICS APPROVAL

All human samples were approved by UW Madison/UW Health under protocol number 2018-0098.

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Figure S2 (mSystems01471-24-S0001.pdf). Functional networks of significantly different features.

Supplemental Material (mSystems01471-24-S0002.docx). Supplemental figures and captions for supplemental tables.

Table S1 (mSystems01471-24-S0003.xlsx). Protein level comparisons between groups.

Table S2 (mSystems01471-24-S0004.xlsx). Metabolite comparisons between groups.

Open Peer Review

PEER REVIEW HISTORY (review-history.pdf). An accounting of the reviewer comments and feedback.

REFERENCES

1. Lebreton F, Willems RJL, Gilmore MS. 2014. *Enterococcus* diversity, origins in nature, and gut colonization. In *Enterococci: from commensals to leading causes of drug resistant infection*. Massachusetts Eye and Ear Infirmary, Boston.
2. Krawczyk B, Wityk P, Gałęcka M, Michalik M. 1900. The many faces of *Enterococcus* spp.—commensal, probiotic and opportunistic pathogen. *Microorganisms* 9:1900. <https://doi.org/10.3390/microorganisms9091900>

3. Hufnagel M, Liese C, Loescher C, Kunze M, Proempeler H, Berner R, Krueger M. 2007. Enterococcal colonization of infants in a neonatal intensive care unit: associated predictors, risk factors and seasonal patterns. *BMC Infect Dis* 7:107. <https://doi.org/10.1186/1471-2334-7-107>
4. Schloissnig S, Arumugam M, Sunagawa S, Mitreva M, Tap J, Zhu A, Waller A, Mende DR, Kultima JR, Martin J, Kota K, Sunyaev SR, Weinstock GM, Bork P. 2013. Genomic variation landscape of the human gut microbiome. *Nature New Biol* 493:45–50. <https://doi.org/10.1038/nature11711>
5. Lebreton F, Manson AL, Saavedra JT, Straub TJ, Earl AM, Gilmore MS. 2017. Tracing the Enterococci from paleozoic origins to the hospital. *Cell* 169:849–861. <https://doi.org/10.1016/j.cell.2017.04.027>
6. Fiore E, Van Tyne D, Gilmore MS. 2019. Pathogenicity of Enterococci. *Microbiol Spectr* 7:7. <https://doi.org/10.1128/microbiolspec.gpp3-0053-2018>
7. Higueta NIA, Huycke MM. 2014. Enterococcal disease, epidemiology, and implications for treatment. In *Enterococci: from commensals to leading causes of drug resistant infection*. Massachusetts Eye and Ear Infirmary, Boston.
8. García-Solache M, Rice LB. 2019. The *Enterococcus*: a model of adaptability to its environment. *Clin Microbiol Rev* 32:e00058-18. <https://doi.org/10.1128/CMR.00058-18>
9. Herrero IA, Issa NC, Patel R. 2002. Nosocomial spread of linezolid-resistant, vancomycin-resistant *Enterococcus faecium*. *N Engl J Med* 346:867–869. <https://doi.org/10.1056/NEJM200203143461121>
10. Chow JW, Davidson A, Sanford E, Zervos MJ. 1997. Superinfection with *Enterococcus faecalis* during quinupristin/dalfopristin therapy. *Clin Infect Dis* 24:91–92. <https://doi.org/10.1093/clinids/24.1.91>
11. Sabol K, Patterson JE, Jorgensen JH, Lewis JS, Owens A, Cadena J. 2005. Emergence of daptomycin resistance in *Enterococcus faecium* during daptomycin therapy. *Antimicrob Agents Chemother* 49:1664–1665. <https://doi.org/10.1128/AAC.49.4.1664-1665.2005>
12. Arias CA, Murray BE. 2012. The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* 10:266–278. <https://doi.org/10.1038/nrmicro2761>
13. Rogers R, Rice LB. 2024. State-of-the-art review: persistent enterococcal bacteremia. *Clin Infect Dis* 78:e1–e11. <https://doi.org/10.1093/cid/ciad612>
14. Elizaga ML, Weinstein RA, Hayden MK. 2002. Patients in long-term care facilities: a reservoir for vancomycin-resistant enterococci. *Clin Infect Dis* 34:441–446. <https://doi.org/10.1086/338461>
15. Casadevall A, Pirofski L. 2001. Host-pathogen interactions: the attributes of virulence. *J Infect Dis* 184:337–344. <https://doi.org/10.1086/322044>
16. Bivona G, Agnello L, Ciccio M. 2021. Biomarkers for prognosis and treatment response in COVID-19 patients. *Ann Lab Med* 41:540–548. <https://doi.org/10.3343/alm.2021.41.6.540>
17. Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, Pienta KJ, Rubin MA, Chinnaiyan AM. 2001. Delineation of prognostic biomarkers in prostate cancer. *Nature New Biol* 412:822–826. <https://doi.org/10.1038/35090585>
18. Wozniak JM, Mills RH, Olson J, Caldera JR, Sepich-Poore GD, Carrillo-Terrazas M, Tsai C-M, Vargas F, Knight R, Dorrestein PC, Liu GY, Nizet V, Sakoulas G, Rose W, Gonzalez DJ. 2020. Mortality risk profiling of *Staphylococcus aureus* bacteremia by multi-omic serum analysis reveals early predictive and pathogenic signatures. *Cell* 182:1311–1327. <https://doi.org/10.1016/j.cell.2020.07.040>
19. Lautenbach E, Bilker WB, Brennan PJ. 1999. Enterococcal bacteremia: risk factors for vancomycin resistance and predictors of mortality. *Infect Control Hosp Epidemiol* 20:318–323. <https://doi.org/10.1086/501624>
20. Frantzi M, Bhat A, Latosinska A. 2014. Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development. *Clin Transl Med* 3:7. <https://doi.org/10.1186/2001-1326-3-7>
21. Skates SJ, Gillette MA, LaBaer J, Carr SA, Anderson L, Liebler DC, Ransohoff D, Rifai N, Kondratovich M, Težak Ž, Mansfield E, Oberg AL, Wright I, Barnes G, Gail M, Mesri M, Kinsinger CR, Rodriguez H, Boja ES. 2013. Statistical design for biospecimen cohort size in proteomics-based biomarker discovery and verification studies. *J Proteome Res* 12:5383–5394. <https://doi.org/10.1021/pr400132j>
22. Lapek JD, Mills RH, Wozniak JM, Campeau A, Fang RH, Wei X, van de Groep K, Perez-Lopez A, van Sorge NM, Raffatellu M, Knight R, Zhang L, Gonzalez DJ. 2018. Defining host responses during systemic bacterial infection through construction of a murine organ proteome atlas. *Cell Syst* 6:579–592. <https://doi.org/10.1016/j.cels.2018.04.010>
23. Wang Y, Yang F, Gritsenko MA, Wang Y, Clauss T, Liu T, Shen Y, Monroe ME, Lopez-Ferrer D, Reno T, Moore RJ, Klemke RL, Camp DG, Smith RD. 2011. Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. *Proteomics* 11:2019–2026. <https://doi.org/10.1002/pmic.201000722>
24. Wierzbicki IH, Campeau A, Dehaini D, Holay M, Wei X, Greene T, Ying M, Sands JS, Lamsa A, Zuniga E, Pogliano K, Fang RH, LaRock CN, Zhang L, Gonzalez DJ. 2019. Group A streptococcal S protein utilizes red blood cells as immune camouflage and is a critical determinant for immune evasion. *Cell Rep* 29:2979–2989. <https://doi.org/10.1016/j.celrep.2019.11.001>
25. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X, Yu G. 2021. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb)* 2:100141. <https://doi.org/10.1016/j.xinn.2021.100141>
26. Holman JD, Tabb DL, Mallick P. 2014. Employing ProteoWizard to convert raw mass spectrometry data. *CP in Bioinformatics* 46. <https://doi.org/10.1002/0471250953.bi1324s46>
27. Schmid R, Heuckeroth S, Korf A, Smirnov A, Myers O, Dyrland TS, Bushuiev R, Murray KJ, Luzzatto-Knaan T, et al. 2023. Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat Biotechnol* 41:447–449. <https://doi.org/10.1038/s41587-023-01690-2>
28. Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, Nguyen DD, Watrous J, Kapono CA, 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. <https://doi.org/10.1038/nbt.3597>
29. Campeau A, Mills RH, Stevens T, Rossitto L-A, Meehan M, Dorrestein P, Daly R, Nguyen TT, Gonzalez DJ, Jeste DV, Hook V. 2022. Multi-omics of human plasma reveals molecular features of dysregulated inflammation and accelerated aging in schizophrenia. *Mol Psychiatry* 27:1217–1225. <https://doi.org/10.1038/s41380-021-01339-z>
30. Neumann U, Genze N, Heider D. 2017. EFS: an ensemble feature selection tool implemented as R-package and web-application. *BioData Min* 10:21. <https://doi.org/10.1186/s13040-017-0142-8>
31. Max K, Wickham H. 2020. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles
32. Hébert L, Courtin P, Torelli R, Sanguinetti M, Chapot-Chartier M-P, Auffray Y, Benachour A. 2007. *Enterococcus faecalis* constitutes an unusual bacterial model in lysozyme resistance. *Infect Immun* 75:5390–5398. <https://doi.org/10.1128/IAI.00571-07>
33. Kim D, Song L, Breitwieser FP, Salzberg SL. 2016. Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res* 26:1721–1729. <https://doi.org/10.1101/gr.210641.116>
34. Li H. 2023. seqtk: Toolkit for processing sequences in FASTA/Q formats. Available from: <https://github.com/lh3/seqtk>
35. Wright C, Griffiths S, Parker M. 2024. wf-bacterial-genomes. Small variant calling for haploid samples. Available from: <https://github.com/epi2me-labs/wf-bacterial-genomes>
36. Palumbo E. 2019. A command line tool to compute mapping statistics from a BAM file. *Bamstats*. Available from: <https://github.com/guigolab/bamstats>
37. Freire B, Ladra S, Parama JR. 2021. Memory-efficient assembly using Flye. *IEEE/ACM Trans Comput Biol Bioinf* 1–1. <https://doi.org/10.1109/TCBB.2021.3108843>
38. Wright C, Wykes Mm. 2023. Sequence correction provided by ONT research. Available from: <https://github.com/nanoporetech/medaka>
39. Seeman T. 2022. Scan contig files against PubMLST typing schemes. mlst. Available from: <https://github.com/tseemann/mlst>
40. Florensa AF, Kaas RS, Clausen PTLC, Aytan-Aktug D, Aarestrup FM. 2022. ResFinder - an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microb Genom* 8:40. <https://doi.org/10.1099/mgen.0.000748>
41. Trizna M. 2020. assembly_stats. Calculates both scaffold and contig statistics (N50, L50, Etc.) from a scaffold FASTA file. Available from: https://github.com/MikeTrizna/assembly_stats
42. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>
43. Rosselli Del Turco E, Bartoletti M, Dahl A, Cervera C, Pericàs JM. 2021. How do I manage a patient with enterococcal bacteraemia? *Clin Microbiol Infect* 27:364–371. <https://doi.org/10.1016/j.cmi.2020.10.029>

44. da Silva RR, Dorrestein PC, Quinn RA. 2015. Illuminating the dark matter in metabolomics. *Proc Natl Acad Sci USA* 112:12549–12550. <https://doi.org/10.1073/pnas.1516878112>
45. Neumann U, Riemenschneider M, Sowa J-P, Baars T, Kältsch J, Canbay A, Heider D. 2016. Compensation of feature selection biases accompanied with improved predictive performance for binary classification by using a novel ensemble feature selection approach. *BioData Min* 9:36. <https://doi.org/10.1186/s13040-016-0114-4>
46. He Z, Yu W. 2010. Stable feature selection for biomarker discovery. *Comput Biol Chem* 34:215–225. <https://doi.org/10.1016/j.compbiolchem.2010.07.002>
47. Kupcova Skalnikova H, Cizkova J, Cervenka J, Vodicka P. 2017. Advances in proteomic techniques for cytokine analysis: focus on melanoma research. *Int J Mol Sci* 18:2697. <https://doi.org/10.3390/ijms18122697>
48. Pichet Binette A, Janelidze S, Cullen N, Dage JL, Bateman RJ, Zetterberg H, Blennow K, Stomrud E, Mattsson-Carlsson N, Hansson O. 2023. Confounding factors of Alzheimer's disease plasma biomarkers and their impact on clinical performance. *Alzheimers Dement* 19:1403–1414. <https://doi.org/10.1002/alz.12787>
49. Ain QU, Sarfraz M, Prasesti GK, Dewi TI, Kurniati NF. 2021. Confounders in identification and analysis of inflammatory biomarkers in cardiovascular diseases. *Biomolecules* 11:1464. <https://doi.org/10.3390/biom11101464>
50. Vanni H, Kazeros A, Wang R, Harvey B-G, Ferris B, De BP, Carolan BJ, Hübner R-H, O'Connor TP, Crystal RG. 2009. Cigarette smoking induces overexpression of a fat-depleting gene AZGP1 in the human. *Chest* 135:1197–1208. <https://doi.org/10.1378/chest.08-1024>
51. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW-M, Fiehn O, Goodacre R, Griffin JL, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane AN, Lindon JC, Marriott P, Nicholls AW, Reilly MD, Thaden JJ, Viant MR. 2007. Proposed minimum reporting standards for chemical analysis. *Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI)*. *Metabolomics* (Los Angel) 3:211–221. <https://doi.org/10.1007/s11306-007-0082-2>
52. Sun J, Uchiyama S, Olson J, Morodomi Y, Cornax I, Ando N, Kohno Y, Kyaw MMT, Aguilar B, Haste NM, Kanaji S, Kanaji T, Rose WE, Sakoulas G, Marth JD, Nizet V. 2021. Repurposed drugs block toxin-driven platelet clearance by the hepatic Ashwell-Morell receptor to clear *Staphylococcus aureus* bacteremia. *Sci Transl Med* 13:e01181–21. <https://doi.org/10.1126/scitranslmed.abd6737>
53. Golden GJ, Toledo AG, Marki A, Sorrentino JT, Morris C, Riley RJ, Spliid C, Chen Q, Cornax I, Lewis NE, Varki N, Le D, Malmström J, Karlsson C, Ley K, Nizet V, Esko JD. 2021. Endothelial heparan sulfate mediates hepatic neutrophil trafficking and injury during *Staphylococcus aureus* sepsis. *MBio* 12:e0118121. <https://doi.org/10.1128/mBio.01181-21>
54. Chaudhry H, Zhou Y, Zhong Y, Ali MM, McGuire J, Nagarkatti PS, Nagarkatti M. 2013. Role of cytokines as a double-edged sword in sepsis. *In Vivo* 27:669–684.
55. Piktet L, Levental I, Durnas B, Janmey PA, Bucki R. 2018. Plasma gelsolin: indicator of inflammation and its potential as a diagnostic tool and therapeutic target. *Int J Mol Sci* 19:2516. <https://doi.org/10.3390/ijms19092516>
56. Yang Z, Bedugnis A, Levinson S, Dinubile M, Stossel T, Lu Q, Kobzik L. 2019. Delayed administration of recombinant plasma gelsolin improves survival in a murine model of penicillin-susceptible and penicillin-resistant pneumococcal pneumonia. *J Infect Dis* 220:1498–1502. <https://doi.org/10.1093/infdis/jiz353>
57. Hashida T, Nakada T-A, Satoh M, Tomita K, Kawaguchi R, Nomura F, Oda S. 2017. Proteome analysis of hemofilter adsorbates to identify novel substances of sepsis: a pilot study. *J Artif Organs* 20:132–137. <https://doi.org/10.1007/s10047-016-0936-3>
58. Blairon L, Wittebole X, Laterre P-F. 2003. Lipopolysaccharide-binding protein serum levels in patients with severe sepsis due to gram-positive and fungal infections. *J Infect Dis* 187:287–291. <https://doi.org/10.1086/346046>
59. Hofmaenner DA, Kleyman A, Press A, Bauer M, Singer M. 2022. The many roles of cholesterol in sepsis: a review. *Am J Respir Crit Care Med* 205:388–396. <https://doi.org/10.1164/rccm.202105-1197TR>
60. Fraunberger P, Schaefer S, Werdan K, Walli AK, Seidel D. 1999. Reduction of circulating cholesterol and apolipoprotein levels during sepsis. *Clin Chem Lab Med* 37:357–362. <https://doi.org/10.1151/CCLM.1999.059>
61. Bhogal HK. 2013. The molecular pathogenesis of cholestasis in sepsis. *Front Biosci* E5:87–96. <https://doi.org/10.2741/E598>
62. Davis RP, Miller-Dorey S, Jenne CN. 2016. Platelets and coagulation in infection. *Clin Transl Immunol* 5:e89. <https://doi.org/10.1038/cti.2016.39>
63. Sánchez-Navarro A, González-Soria I, Caldiño-Bohn R, Bobadilla NA. 2021. An integrative view of serpins in health and disease: the contribution of serpinA3. *Am J Physiol Cell Physiol* 320:C106–C118. <https://doi.org/10.1152/ajpcell.00366.2020>
64. Bianchini EP, Auditeau C, Razanakolona M, Vasse M, Borgel D. 2020. Serpins in hemostasis as therapeutic targets for bleeding or thrombotic disorders. *Front Cardiovasc Med* 7:622778. <https://doi.org/10.3389/fcvm.2020.622778>
65. Liu K-J, Shih N-Y. 2007. The role of enolase in tissue invasion and metastasis of pathogens and tumor cells
66. Yaron JR, Zhang L, Guo Q, Haydel SE, Lucas AR. 2021. Fibrinolytic serine proteases, therapeutic serpins and inflammation: fire dancers and firestorms. *Front Cardiovasc Med* 8:648947. <https://doi.org/10.3389/fcvm.2021.648947>
67. Maas C, de Maat S. 2021. Therapeutic SERPINS: improving on nature. *Front Cardiovasc Med* 8:648349. <https://doi.org/10.3389/fcvm.2021.648349>
68. Iram S, Rahman S, Choi I, Kim J. 2024. Insight into the function of tetranectin in human diseases: a review and prospects for tetranectin-targeted disease treatment. *Heliyon* 10:e23512. <https://doi.org/10.1016/j.heliyon.2023.e23512>
69. Fang X, Kaduce TL, Spector AA. 1999. 13-(S)-hydroxyoctadecadienoic acid (13-HODE) incorporation and conversion to novel products by endothelial cells. *J Lipid Res* 40:699–707. [https://doi.org/10.1016/S0022-2275\(20\)32149-0](https://doi.org/10.1016/S0022-2275(20)32149-0)
70. Cambiaggi L, Chakravarty A, Nouredine N, Hersberger M. 2023. The role of α -linolenic acid and its oxylipins in human cardiovascular diseases. *Int J Mol Sci* 24:6110. <https://doi.org/10.3390/ijms24076110>
71. Prost I, Dhondt S, Rothe G, Vicente J, Rodríguez MJ, Kift N, Carbonne F, Griffiths G, Esquerré-Tugayé M-T, Rosahl S, Castresana C, Hamberg M, Fournier J. 2005. Evaluation of the antimicrobial activities of plant oxylipins supports their involvement in defense against pathogens. *Plant Physiol* 139:1902–1913. <https://doi.org/10.1104/pp.105.066274>
72. P. de Carvalho M, Abraham W-R. 2012. Antimicrobial and biofilm inhibiting diketopiperazines. *CMC* 19:3564–3577. <https://doi.org/10.2174/092986712801323243>
73. Fdhila F, Vázquez V, Sánchez JL, Riguera R. 2003. DD-diketopiperazines: antibiotics active against *Vibrio anguillarum* isolated from marine bacteria associated with cultures of *Pecten maximus*. *J Nat Prod* 66:1299–1301. <https://doi.org/10.1021/np030233e>
74. Swinehart W, et al. 2020. Specificity in the biosynthesis of the universal tRNA nucleoside N⁶-threonylcarbamoyl adenosine (t⁶A)—TsaD is the gatekeeper. *RNA* 26:1094–1103. <https://doi.org/10.1261/rna.075747.120>
75. Nagayoshi Y, Nishiguchi K, Yamamura R, Chujo T, Oshiumi H, Nagata H, Kaneko H, Yamamoto K, Nakata H, Sakakida K, Kunisawa A, Adachi M, Kakizoe Y, Mizobe T, Kuratsu J, Shimada S, Nakamori Y, Matsuoka M, Mukoyama M, Wei F-Y, Tomizawa K. 2022. t6A and ms2t6A modified nucleosides in serum and urine as strong candidate biomarkers of COVID-19 infection and severity. *Biomolecules* 12:1233. <https://doi.org/10.3390/biom12091233>
76. Wang Y, Qian H. 2021. Phthalates and their impacts on human health. *Healthcare (Basel)* 9:603. <https://doi.org/10.3390/healthcare9050603>
77. Win-Shwe T-T, Yanagisawa R, Lwin T-T, Kawakami F, Koike E, Takano H. 2022. Dietary exposure to flame retardant tris (2-butoxyethyl) phosphate altered neurobehavior and neuroinflammatory responses in a mouse model of allergic asthma. *IJMS* 23:655. <https://doi.org/10.3390/ijms23020655>
78. Cruickshank-Quinn C, Zheng LK, Quinn K, Bowler R, Reisdorph R, Reisdorph N. 2018. Impact of blood collection tubes and sample handling time on serum and plasma metabolome and lipidome. *Metabolites* 8:88. <https://doi.org/10.3390/metabo8040088>
79. Naegeli A, Bratanis E, Karlsson C, Shannon O, Kalluru R, Linder A, Malmström J, Collin M. 2019. *Streptococcus pyogenes* evades adaptive immunity through specific IgG glycan hydrolysis. *J Exp Med* 216:1615–1629. <https://doi.org/10.1084/jem.20190293>
80. Kagawa TF, Cooney JC, Baker HM, McSweeney S, Liu M, Gubba S, Musser JM, Baker EN. 2000. Crystal structure of the zymogen form of the group A *Streptococcus* virulence factor SpeB: an integrin-binding cysteine protease. *Proc Natl Acad Sci U S A* 97:2235–2240. <https://doi.org/10.1073/pnas.040549997>
81. Moradigaravand D, Gouliouris T, Blane B, Naydenova P, Ludden C, Crawley C, Brown NM, Török ME, Parkhill J, Peacock SJ. 2017. Within-host evolution of *Enterococcus faecium* during longitudinal carriage and

- transition to bloodstream infection in immunocompromised patients. *Genome Med* 9:119. <https://doi.org/10.1186/s13073-017-0507-0>
82. Rosales FJ, Ritter SJ, Zolfaghari R, Smith JE, Ross AC. 1996. Effects of acute inflammation on plasma retinol, retinol-binding protein, and its mRNA in the liver and kidneys of vitamin A-sufficient rats. *J Lipid Res* 37:962–971.
 83. Stephensen CB. 2001. Vitamin A, infection, and immune function. *Annu Rev Nutr* 21:167–192. <https://doi.org/10.1146/annurev.nutr.21.1.167>
 84. Sammalkorpi K, Valtonen V, Kerttula Y, Nikkilä E, Taskinen M-R. 1988. Changes in serum lipoprotein pattern induced by acute infections. *Metab Clin Exp* 37:859–865. [https://doi.org/10.1016/0026-0495\(88\)90120-5](https://doi.org/10.1016/0026-0495(88)90120-5)
 85. Miller W. 2024. Treatment of enterococcal infections. UpToDate. Available from: <https://www.uptodate.com/contents/treatment-of-enterococcal-infections>
 86. Lee C. 2009. Type 3 cystatins; fetuins, kininogen and histidine-rich glycoprotein. *Front Biosci* 14:2911. <https://doi.org/10.2741/3422>
 87. Rydengård V, Olsson A-K, Mörgelin M, Schmidtchen A. 2007. Histidine-rich glycoprotein exerts antibacterial activity. *FEBS J* 274:377–389. <https://doi.org/10.1111/j.1742-4658.2006.05586.x>
 88. Völlmy F, van den Toorn H, Zenezini Chiozzi R, Zucchetti O, Papi A, Volta CA, Marracino L, Viecei Dalla Sega F, Fortini F, Demichev V, Tober-Lau P, Campo G, Contoli M, Ralser M, Kurth F, Spadaro S, Rizzo P, Heck AJ. 2021. A serum proteome signature to predict mortality in severe COVID-19 patients. *Life Sci Alliance* 4:e202101099. <https://doi.org/10.26508/lsa.202101099>
 89. Kuroda K, Ishii K, Mihara Y, Kawanoue N, Wake H, Mori S, Yoshida M, Nishibori M, Morimatsu H. 2021. Histidine-rich glycoprotein as a prognostic biomarker for sepsis. *Sci Rep* 11:10223. <https://doi.org/10.1038/s41598-021-89555-z>
 90. Mandrekar JN. 2010. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* 5:1315–1316. <https://doi.org/10.1097/JTO.0b013e3181ec173d>