**Multi-omic Signatures of Host Response Associated with Presence, Type, and Outcome of Enterococcal Bacteremia**

Charlie Bayne1,2,3, Dominic McGrosso1,2,3, Concepcion Sanchez1,2,3, Leigh-Ana Risotto1,2,3, Carlos Gonzalez2,3, Peiter Dorrestein2,3,4, Victor Nizet2,3,4, George Sakoulas4,5, David J. Gonzalez2,3,6, Warren Rose7

Institutions:

1Biomedical Sciences Graduate Program, UC San Diego, La Jolla, California 92093, USA.

2Department of Pharmacology, University of California San Diego, La Jolla, California 92093, USA

3Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, California, USA

4Department of Pediatrics, UC San Diego, La Jolla, California 92093, USA

5Sharp Rees Stealy Medical, San Diego, CA

6Center for Microbiome Innovation, University of California at San Diego, La Jolla, California 92093, USA

7School of Pharmacy, University of Wisconsin, Madison, Wisconsin 53705, USA

**Importance:**

We utilize untargeted quantitative proteomics and metabolomics to provide the first described deep multi-omic plasma profiles capturing the systemic host response observed in Enterococcal bacteremia patients. We report significant differences between Enterococcal bacteremia and healthy volunteers while leveraging extensive clinical metadata associated with our patient cohort to characterize the differences in host response observed between infections caused by *E. faecalis* or *E. faecium*, as well as those associated with mortality and survival. Altogether, data reported in this study provides a first report of the systemic responses observed in clinically relevant manifestations of Enterococcal bacteremia and associations with survival or mortality. To facilitate the exploration of this rich data source, we also provide an easy-to-use companion user interface hosted at https://gonzalezlab.shinyapps.io/EcB\_multiomics/.

**Abstract:**

Despite the prevalence and severity of Enterococcal bacteremia (EcB), systemic host responses to EcB have never been comprehensively examined using an unbiased multi-omics approach. Here we present an extensive resource profiling molecular differences as observed in EcB patient plasma. To accomplish this, we performed untargeted shotgun proteomics and metabolomics on a total of 105 plasma samples encapsulating patients with EcB and healthy volunteers. Comprehensive profiling of the Enterococcus isolates derived from patients in our cohort allowed the unique opportunity to examine enterococcal bacteremia with an added layer of granularity comparing EcB caused by *E. faecalis* and *E. faecium*. Additionally, we cross-referenced our findings with previously published molecular signatures of bacteremia caused by another clinically relevant human pathogen, *Staphylococcus aureus.* The collection of extensive patient metadata in our EcB cohort allowed us to examine features linked with mortality or survival and test them for associations with potentially confounding variables. We report significant differences between the plasma profiles of healthy volunteers when compared to EcB patients and how these differences to healthy compare to those observed in *S. aureus* bacteremia. We also report significant differences between EcB caused by *E. faecalis* and *E. faecium*, as well as between mortality and survival. Altogether, this endeavor aspires to culminate in the creation of objective risk stratification algorithms —a pivotal step toward enhancing patient management and care.

**Introduction:**

Enterococci are widely distributed in the environment and have co-evolved as common microorganisms within the gastrointestinal microbiota since the transition of terrestrial animals from water to land 1. From this diverse lineage, the distantly related species *Enterococcus faecalis* and *Enterococcus faecium,* evolved independently 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* have the potential to become pathogenic. These conditions are commonly met in the healthcare setting, where several features acquired over the course of their evolution have allowed *E. faecalis* and *E. faecium* to become important healthcare-associated pathogens 5.

These microbes have the inherent ability to survive commonly used disinfection routines and persist on surfaces in healthcare settings, from which they can go on to spread and infect the large supply of immunocompromised patients that are in close proximity 6. They are also increasingly antibiotic resistant, both due to an intrinsic resistance to commonly used broad-spectrum antibiotics such as cephalosporins and carbapenems 7 as well as an impressive capacity to acquire mobile genetic elements through horizontal gene transfer to increase fitness 8. These traits combined with inadequate antibiotic stewardship and increases in aggressive medical treatment undertaken by increasingly older patient populations have resulted in a rise in severe, invasive infections caused by enterococci 7,9,10,11,12. Enterococcal infections are both prevalent and deadly, causing around 7% of hospital- acquired infections and a mortality rate ranging from 25-50% in EcB 6.

Historically, two distinct waves of enterococcal infections have occurred in the United States. During the first wave brought on by the introduction of third generation cephalosporins in the late 1970s, *E. faecalis* accounted for 90-95% of clinical enterococcal isolates. We are currently in the second wave, which started in the early 1990s and has been characterized by an increase in *E. faecium* infections to the point that now *E. faecium* is almost as common a cause of nosocomial infections as *E. faecalis*13*.* Of the two microbes, *E. faecium* is more challenging to treat clinically as ~90% of these isolates are vancomycin resistant compared to ~2% of *E. faecalis* isolates 14 .

As a result of their evolutionary history, *E. faecalis* and *E. faecium* have many conserved features involved in the host-pathogen relationship, as well as distinct differences in addition to their antibiotic sensitivity profiles. They are closely related gram-positive bacteria, and as such have largely conserved structural features. Shared features between *E. faecalis*, *E. faecium*, and other bacteria interact with components of the innate immune system once these bacteria escape their commensal niche, resulting in a conserved host response. For example, shared components such as peptidoglycan and flagellin activate toll-like receptors leading to inflammatory cytokine signaling that produces the acute phase response readily observable in plasma15. Microbes also have various differences in the strategies that they use to circumvent the host response16. These are most obvious when comparing distantly related bacteria, yet important differences also exist at taxonomic resolutions as granular as the strain level. Some of these differences are documented between Enterococcus species; *E. faecium* isolates have a reduced susceptibility to neutrophil-mediated phagocytosis13, while *E. faecalis* isolates have presence of virulence factors such as cytolysin, gelatinase, extracellular superoxide, and extracellular surface protein17. It is conceivable that these similarities and differences (known and unknown) may be reflected in the host response observable in plasma, but to our knowledge this has never been comprehensively profiled. As a result, the current picture of the deviations from homeostasis in response to EcB are restricted to what is reported in bacteremia in general and it is unclear any differences in the host responses are invoked between *E. faecalis* and *E. faecium*.

Diagnosis of enterococcal bacteremia is currently performed through the use of blood cultures, where optimal treatment can additionally be informed by the determination of the identity and antibiotic sensitivity of the isolate responsible for the infection 14. This approach is powerful but imperfect because it takes several days to perform, and increased time to effective therapy has been shown to lead to worse clinical outcomes 18. On the other hand, general markers of inflammation such as c-reactive protein (CRP) and serum amyloid A (SAA) can be very quickly measured, but suffer from offering results that are non-specific to the underlying cause of infection and are thus unable to inform on optimal antibiotic treatment strategy 19. Altogether, the inability to quickly pair the infection with the optimal treatment results in clinicians treating under the assumption that the strain causing the infection is resistant to first line antibiotic treatments when enterococcal bacteremia is suspected. This is suboptimal, because it results in increased usage of “last resort” antibiotics relative to what would be required if the antibiotic resistance profiles of the isolates were known *a priori*. As time progresses, enterococci will develop increased resistance to these antibiotics, making future treatment progressively more challenging. This has been well documented previously, where vancomycin resistance in *E. faecium* rose from 0% in the mid 1980s to over 80% by 2007 13. In addition to promoting increases in antibiotic resistance, suboptimal antibiotic use in EcB is associated with worse clinical outcomes 20,21. If substantial differences in host response are capable of being detected in plasma as a result of the underlying pathogen, in principle they could be leveraged to develop diagnostics that could quickly predict the specific type of bacteremia and inform on the optimal treatment method. In the context of EcB, this would allow epidemiological data to inform the clinician on whether treatment with vancomycin would likely be effective (as would be expected for a majority of *E. faecalis* isolates but very few *E. faecium* isolates). To evaluate the feasibility of such a strategy, the systemic response to EcB must first be profiled in an unbiased manner.

Host factors are well appreciated to be critical determinants of the outcome of host-microbe interactions16, and they have been utilized as prognostic biomarkers to inform on the trajectory of patient outcome in a variety of diseases ranging from COVID-19 to cancer 22,23. If these tests suggest that a patient’s trajectory is likely to be poor, the therapeutic strategy can be altered to improve the chance of a successful outcome. Molecular features associated with mortality have been described in *S. aureus* bacteremia 24, but in the context of EcB prediction of mortality has thus far been limited to crude clinical metrics such as severity of illness, and age 25. A broad profile of the molecular features of systemic response in a well-documented EcB patient cohort would facilitate the discovery of any associations with mortality or survival.

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 as well as healthy controls. Furthermore, we utilized previously published results from our group to compare the deviations from homeostasis observed in EcB and *S. aureus* bacteremia. This approach allows us to identify individual features as well as biological processes that are altered in aspects of enterococcal bacteremia through three primary comparisons (1) EcB and *S. aureus* compared to healthy, (2) bacteremia caused by *E. faecalis* compared to *E. faecium*, and (3) EcB patients that succumbed to mortality compared to those that survived. The information gleaned from such data offers deep insights into the systemic molecular and biological features that underly the host-pathogen relationship in EcB. Importantly, this data set also 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 presented here was conducted on human plasma collected from patients under UW Madison / UW Health IRB# 2018-0098. Sample size was selected based on technical considerations for multiplexed proteomics and metabolomics approaches as well as logistical constraints 26,27. Further details regarding statistical analysis, demographics, and clinical data are provided below.

*Human Plasma Samples* – All human samples were approved by UW Madison / UW Health IRB# 2018-0098. On admission, plasma was obtained from, 32 patients with *E. faecium* bacteremia and 44 patients with *E. faecalis* bacteremia. Enterococcal bacteremia was diagnosed by positive blood cultures and treated with antibiotics. *E. faecalis* and *E. faecium* bacteremia were differentiated using MALDI-TOF based rapid identification and confirmed using standard culture and biochemical methods. Plasma from 29 healthy controls was collected from blood bank volunteers.

**Proteomics Arm:**

*Protein Preparation* – A 25 µL aliquot of plasma from each patient was added to 200 µL lysis buffer containing 6M urea, 7% SDS, 50mM tetraethylammonium bromide (TEAB), and one protease inhibitor tablet (Roche cat # 06538282001) and PhosStop tablet (Roche cat # 04906845001), final pH 8.1 adjusted with phosphoric acid. Then samples were sonicated in a water bath and vortexed for 5 minutes each. Disulfide bonds were reduced with 100 µM dithiothreitol (DTT) for 30 minutes at 47℃, cooled briefly on ice, followed by alkylation with 300 µM iodoacetamide (IAA) for 45 minutes at room temperature in the dark. The reactions were quenched with 100 µM DTT at room temperature for 5 minutes. Samples were loaded on S-Trap 96-well plates (Protifi) per the manufacturer’s instructions. Samples were digested with 5 µg of trypsin in 115 µL 50 mM TEAB over 3 hours at 47℃. Peptides were eluted serially with 125 µL of 50 mM TEAB, then 5% formic acid (FA), finally 50% acetonitrile (ACN), 5% FA, and then dried under vacuum centrifugation. Peptides were desalted using 50 mg Sep-Pak tC-18 cartridges (Waters cat # WAT054960) using manufacturer instructions and were then quantified with a Pierce Quantitative Colorimetric Peptide Assay (ThermoFisher cat # 23275). 50 µg of each sample were aliquoted for TMT labeling with a further 10 µg of each sample being combined and aliquoted into a final 50 µg aliquot for usage 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 28 .

*Proteomics LC-LC-MSn* - Basic pH reverse-phase LC, followed by data acquisition through LC–MS2/MS3, was performed as previously described 28. Briefly, 75 minute linear gradients of 22% to 35% acetonitrile and 10 mM ammonium bicarbonate were passed on HPLC C18 columns (Biobasic) with the resulting ninety-six fractions combined as previously described 29. Fractions were next analyzed using tandem mass spectrometry (MS2/MS3) on an Orbitrap Fusion mass spectrometer (ThermoFisher) with an in-line EASY-nLC 1000 (ThermoFisher). Separation and acquisition settings were performed using previously defined methods 30.

*Statistics and Data Analysis* - Thermo .raw files were converted to mzML files using MSConvert with a peak picking (centroiding) filter applied to the conversion. 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 was run with precursor mass upper and lower tolerance set to 20 ppm, strict trypsin digestion rules allowing for 2 missed cleavages, and the search was set to include static modifications of carbamidomethylation of cystines, 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 September 11th, 2023.

Proteins identified as differentially abundant were subjected to Gene Ontology Analysis using all proteins identified in the experiment as the universe as implemented in ClusterProfiler4.0 31. 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 µl) were thawed for 30 minutes, and then 200 µL 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 minutes and then incubated at 20C for 20 minutes to aid in protein precipitation. Samples were centrifuged at 16,000 x g for 15 minutes 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℃ once dry.

*Metabolomics LC-MS2 Analysis* – Metabolomic 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 100x21mm (Catalog no. 00D-4759-AN). 5 µL of each sample was injected and run on a 10-minute 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 0.500mL/min and chromatographic separation was achieved using the following gradient: 0 to 1 minute 5% B, 1 to 7 minute a linear increase from 5 to 100% B, 7 to 7.5 minute held at 100% B, 7.5 minute to 8 minute a linear decrease from 100% to 5% B, and then 5% B from 8 minute to 10 minute.

*Metabolite Identification* – Full scan MS spectra (m/z 100 – 1500) were acquired and the top five most intense ions from a unique scan were fragmented. Dynamic exclusion was set to 10.0s. 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) 32. mzML files were processed together using MzMine (3.5.0)33 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=3000, min highest intensity=1000, m/z tolerance = 0.005 Da or 10 ppm), Feature Detection Chromatogram Resolving (MS/MS scan pairing with 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 0.50min, min relative height 0.01%, min absolute height 1000, min ratio of peak top/edge=1.7, peak duration 0.05-1 min, and min # data points=4), 13C Isotope filter (m/z tolerance= 0.01 m/z, RT tolerance=0.30min, 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= 0.30min, RT weight= 20), Feature list filtering (at least 2 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 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) 34. 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 (speclibs). 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 was 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 +/- 50Da window throughout the spectrum. The precursor ion mass tolerance was set to 2.0 Da and a 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 6 matched peaks. Further, 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 6 matched peaks.

*Data Processing of Metabolite Features* - Metabolomics data was normalized using sulfamethazine as a single internal standard as described previously.35 Briefly, any feature containing a 0 value across 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 36.

*Metadata Assessment* – Metadata correlations were assessed in the following manner. First, categorical metadata associations were determined using MWU or Kruskall-Wallis tests when appropriate. Continuous metadata associations were determined using Pearson correlation. Associations in the figures represent the -log10 (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.

**Nanopore Sequencing:**

*Strain isolation* - Enterococcus faecium and Enterococcus faecalis clinical isolates were recovered from blood culture vials by streaking out 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℃ . 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 BHI. These cultures were subsequently incubated at 37℃ shaking at 220 RPM overnight. Bacteria were pelleted by centrifugation at 5000xg for 5 minutes, 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, likely due to well-documented lysozyme resistance among enterococcal clinical isolates 10. These cells were then incubated overnight at 37℃ to facilitate lysis. The next day, 45 μL of proteinase K solution (Cat # RP107B-10 QIAGEN) was added to each sample and incubated for 1 hour at 37C. Next, 0.35 mL of QIAGEN Buffer B2 was added to each sample, tubes were mixed several times by inversion and then incubated at 50C for 30 minutes. The resulting high molecular weight DNA was then purified using QIAGEN Genomic-tip 20/G. A genomic tip was equilibrated with 1mL of buffer QBT. Samples were vortexed for 10 seconds 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. 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 12,000xg for 15 minutes at 4℃. 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 x g for 10 minutes at 4C. 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 minutes before being resuspended in 50 μL 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-* 400ng 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 gDNA – Native Barcoding Kit 96 V14 version NBE\_9171\_v114\_revl\_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 concertation combined with the TapeStation derived distribution of molecular weight.) Minknow 23.07.5 was used to acquire the data in SAC 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 Centrifuge37 and reads mapping to streptomyces were then removed using seqtk38. Decontaminated reads were then used as input into the Epi2me labs graphic interface which was used to run the wf-bacterial-genomes nextflow workflow39. Briefly, this workflow concatenates input files and generate per read stats via bamstats40, performs *denovo* assembly via flye41, polishes with medaka42, performs multi locus sequence typing (MLST)43 to infer the identity of the isolate, and lastly performs antimicrobial resistance (AMR) calling via ResFinder 44 to identify genes associated with antimicrobial resistance. Basic assembly statistics such as N50 and the number of contigs were also generated using assembly\_stats45. The resulting FASTA assemblies were then used as input to Roary 46 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 enterococcal bacteremia (Figure 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 corresponding to each EB patient was collected, including gender, the Charleston index, duration of bacteremia, day of blood draw, and mortality during admission (Figure 2B). Additionally, anti-microbial susceptibility testing was performed on the enterococcal isolates that were isolated from each patient. As expected based on reported susceptibilities of clinical isolates, we observed that all *E. faecalis* isolates were resistant to vancomycin, while only a minority of *E. faecium* isolates displayed resistance 14, highlighting that the identification of species can inform the selection of optimal antibiotic therapy.

In the TMT-guided proteomics dataset, we quantified a total of 589 proteins mapping to the human proteome, of which 278 were quantified across all samples. No proteins were found to map to the Enterococcous proteomes. In our metabolomics dataset, we quantified a total of 10092 features, of which 693 were able to be putatively identified based on their MS2 profiles using GNPS molecular networking. Of these putatively identified metabolites, 212 were quantified across all samples. This allowed us to examine the differences between healthy and infected*, E. faecalis* and *E. faecium*, while providing a unique opportunity to explore associations with mortality or survival.

Unsupervised hierarchical clustering of both the global proteomics data and metabolomics data revealed a clear delineation between the healthy and infected groups (Figure 1C) (Figure 1D). Significant differences between *E. faecalis* and *E. faecium* bacteremia were observed at the level of individual features (Figure 3A) (Figure 4A) but global dissimilarities between the two types of bacteremia were more subtle, indicating large-scale similarity in the host response elicited by the two closely related pathologies.

**Differences In Plasma Profiles of Bacteremia Types Compared to Healthy**

We first queried our proteomics data to determine which proteins are most effective at differentiating enterococcal bacteremia in general from healthy populations. Relative to healthy volunteers, 204 proteins were identified to be more abundant in Enterococcal bacteremia while 85 were found to be significantly less abundant (Figure 2A, FDR adjusted p-value <= 0.05). The Benjamini-Hochberg adjusted p values for several proteins were highly significant, reaching values as extreme as 1×10-25. Evaluation of the metabolomic data showed similar results, where we observed that 427 features were significantly increased in infected relative to healthy, 968 were significantly decreased, and FDR-adjusted p values were as low as 1×10-50 (Figure 3A). Of the metabolite features identified in our study, 29% of them were putatively identified using GNPS molecular networking (Figure 3B). The ability to only annotate a subset of features is a well-documented limitation inherent to current untargeted metabolomics analysis approaches 34.

To further understand the differences between these bacteremia types, we investigated the degree to which the specific proteins identified as significant relative to healthy were shared among the types of bacteremia. To reference the deviations from homeostasis observed upon infection to another clinically relevant pathogen, we also analyzed previously published 24 proteomic differences between *S. aureus* bacteremiapatientsand healthy volunteers. We found that 24% (53) of the significant proteins identified to be increased upon infection were shared across all types of bacteremia, while 30% of the significant proteins were specific to *E. faecium* bacteremia, 0% were specific to *E. faecalis* bacteremia and 42% were specific to *S. aureus* bacteremia (Figure 2D). When considering proteins that were found to be significantly decreased in infection, we found that 28% (94) of the proteins were shared across all types of bacteremia, while 6% were specific to EcB, 10% were specific to *E. faecium* bacteremia, 7% were specific to E. faecalis bacteremia, and 32% were only significant in *S. aureus* bacteremia (Figure 2D).

We then asked what biological processes the proteins identified as statistically significant when comparing infected to healthy were involved in, and how these processes differed when comparing bacteremia caused by *Enterococcus*, *E. faecalis*, *E. faecium*, and *S. aureus*. To facilitate this, we conducted GO enrichment analysis on the proteins identified as significantly different by binary comparisons (Supplementary Figure 5). In both EcB types, we observed an enrichment in neutrophil chemotaxis, tertiary granule lumen, focal adhesion, extracellular exosome, and inflammatory response (Figure 2B).

Enterococcal and *S. aureus* bacteremia had several conserved biological processes that were significantly depleted upon infection. We saw evidence that proteins involved in cholesterol metabolism were reduced in infected patients, noting the significant reduction in GO terms reverse cholesterol transport, cholesterol homeostasis, cholesterol metabolic process, very-low-density lipoprotein particle, high-density lipoprotein particle, and blood microparticle. We also observed similarities in processes related to blood clotting between EcB and *S. aureus* bacteremia noting significant depletion in the GO terms upon infection: blood coagulation, heparin binding, and zymogen activation. Platelet alpha granule lumen was the only GO Term that was observed to have the opposite effect in different types of bacteremia, where it was found to be enriched in the infected samples from EcB and depleted in *S. aureus* infected samples (Figure 2B).

The metabolites that were significantly differentially abundant (p.adj value <= 0.05) were also subjected to enrichment analysis to see if any class of molecule was significantly enriched. Interestingly, steroids were found to be significantly enriched in infected patients relative to all features detected in the experiment (Figure 3C). A closer examination of these molecules showed that 10/12 of these molecules annotated as steroids could be more specifically described as bile acids. These bile acids included the primary bile acid cholic acid as well as the secondary bile acids taurodeoxycholic acid, glycochenodeoxycholate, glycochenodeoxycholic acid, glycohyocholic acid, tauroursodeoxycholic acid, glycocholic acid, taurocholic acid, and taurohyodeoxycholic acid. The same functional enrichment was observed for *S. aureus* infected samples (Supplementary Figure 2) where the abundances of Taurocholic acid, Glycocholic acid, Tauroursodeoxycholic acid, 3beta−Hydroxy−5−cholenoic acid, and 12−Ketodeoxycholic acid were enriched. Several phosphatidylcholine molecules were identified as being significantly depleted in *S. aureus* bacteremia (Supplementary Figure 2). While this class of molecules did not rise to the level of statistical significance when utilizing GO enrichment in healthy to EcB patients, we did note that several phosphatidylcholines were also significantly depleted in the case of EcB bacteremia (Supplementary Table 2).

We next set out to evaluate the potential utility of features collected within our dataset as biomarkers to distinguish healthy from infected samples. To rank these biomarkers within our dataset, we used ensemble feature selection (EFS) as an unbiased approach to rank protein biomarkers for EcB17. This method integrates the outcomes of 8 distinct feature selection algorithms, subsequently aggregating and assigning ranks to the scores. This approach mitigates many of the inherent biases often associated with individual algorithms 47, 48. Since EFS can be utilized as a feature selection tool, it is capable of reducing redundant measurements by removing features that are highly correlated with each other. Here, we opted to set the correlation threshold to 0, ensuring that a well-performing biomarker would still score highly, even if it was highly correlated with another potential biomarker. We observed that the top-ranked protein and metabolite biomarkers, **Serpin A3-1 (**SERPINA3), Leucine-rich alpha-2-glycoprotein (LRG1), n6−threonylcarbamoyladenosine, and 13−keto−9z,11e−octadecadienoic acid were highly effective at distinguishing infected from non-infected samples when evaluated using logistic regression, producing AUCs ranging from 0.99- 0.1 (Figure 2C) (Figure 3D). Other top biomarkers identified in our study were also able to distinguish infected from healthy almost perfectly, highlighting the extreme differences between healthy and EcB plasma (Supplementary Figure 4).

To benchmark these findings, we queried our dataset for two biomarkers of infection commonly used in the clinic, c-reactive protein (CRP), and serum amyloid A (SAA1). Both of these proteins were found to be significantly increased in infected samples compared to healthy controls (Supplementary Figure 3). As expected, they also displayed no significant differences between samples infected with *E. faecalis* and *E. faecium*. When evaluated via the EFS approach, these clinically validated biomarkers showed good, but not the best performance in our dataset, ranking 12rd and 47th respectively. Subsequent ROC analysis showed these proteins were able to differentiate infected from uninfected samples with a high degree of sensitivity and specificity as expected, displaying AUC values of 0.97 and 0.92 respectively (Supplementary Figure 3).

Due to the importance of cytokines as modulators of the immune system, we were interested in querying our data to see if we could uncover any indication of differences in cytokine profiles by bacteremia type*.* Since cytokines are not readily detected in plasma using untargeted mass spectrometry-based proteomics due to their low levels of absolute abundance 49, we inferred cytokine profiles using previously reported methods in which known interactions between proteins and cytokines are leveraged to infer cytokine abundance 24. Using this technique, we saw that the inferred amount of TNF-alpha was significantly increased in EcB infected samples compared to healthy samples (p.adj <= 0.05) (Supplementary Figure 6). We also observed a trend where many other cytokines were increased in infected relative to healthy, most notably in IL-6, but they did not rise to our alpha level of 0.05 denoting statistical significance.

**Differences in Systemic Plasma Profiles Between EcB caused by E. *faecalis* and *E. faecium***

We next set out to define the difference in systemic plasma profile the causative organism in cases of enterococcal bacteremia, defined as cases where >90% of the isolates were identified as either *E. faecalis* or *E. faecium*. We observed 70 proteins that were significantly enriched (adjusted p values <= 0.05) in *E. faecalis* infected plasma relative to *E. faecium*, while 30 were found to be significantly enriched in *E. faecium* relative to *E. faecalis* (Figure 4A). Comparable results were seen when investigating the metabolomic data, where we identified 11 significantly increased metabolites in *E. faecalis* relative to *E. faecium*, and 33 significantly increased in *E. faecium* relative to *E. faecalis* (Figure 5A).

GO term enrichment analysis of the proteins that were found to be significantly increased in *E. faecalis* showed dramatic differences in immunoglobulin abundances (Figure 4B). Further investigation revealed that these increased abundances *in E. faecalis* relative to *E. faecium* were due to reduced levels of antibodies in *E. faecium* infected samples as *E. faecalis* and healthy had the same levels of immunoglobulins (Figure 4C). These differences 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 (Figure 2B) (Supplementary Figure 2).

When investigating the GO term enrichment of the 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, including reverse cholesterol transport, cholesterol efflux, chylomicron, and very low-density lipoprotein particle (Figure 4D).

The top proteomic features for distinguishing *E. faecalis* infections from *E. faecium* were identified as immunoglobulin kappa variable 2-30 (IGKV2-30) and retinol binding protein 4 (RBP4), which produced ROC AUCs of 0.76 and 0.79 respectively (Figure 4E). The top metabolomic features were retinol and C24H49N1O7P1, which had ROC AUCs of 0.77 and 0.82 respectively. When we queried the cytokine inference data for differences between *E. faecalis* and *E. faecium* inferred cytokine profiles, no significant differences were observed (Supplementary Figure 6).

An important factor in biomarker discovery is the consideration of additional factors to best ensure that the difference in biomarker observed is truly due to the variable of interest and not confounded by additional factors ,50,51 . Evaluation of our clinical metadata for associated correlations suggested the potential presence of confounding variables in our dataset. Transplant type was found to be significantly associated with the levels of 9 out of the top 10 protein ranked biomarkers (Supplementary Figure 7). Further investigation suggested that this would be challenging to separate from the type of pathogen causing the infection as *E. faecium* infected patients were more likely to have organ transplants in our cohort. Thus, we considered the possibility that the biomarkers we observed as being associated with *E. faecalis* or *E. faecium* in our study were a function of unbalanced groupings. To assess this, we filtered our data to only consider the patients who did not have a transplant. We found that antibody levels were still significantly reduced in *E. faecium* infected samples relative to *E. faecalis* (Supplementary Figure 10A). We observed a loss of significance for the remaining protein biomarkers APOC1, AZGP1, PCOX1, RBP4 and SERPINC1, suggesting that the significance we observed for these biomarkers may be confounded by transplant status (Supplementary Figure 10C). Importantly, the direction of enrichment for APOC3 and RBP4 was preserved, and they narrowly missed our threshold for statistical significance with p values of 0.098, and 0.064 respectively. When restricted to only include patients that had not had a transplant, the metabolites C21H42N1O7P1, mycophenolic acid, and lyso PC also no longer showed any statistically significant differences between *E. faecalis* and *E. faecium,* suggesting that the significance observed may have been driven by transplant type for these features as well. C24H49N1O7P1, C25H51N1O7P1, and retinol still showed significant differences, suggesting these conclusions were not confounded by other variables (Supplementary Figure 10D). Additionally, we found smoking status to be significantly associated with the abundances of 5/10 of the top protein biomarkers (Supplementary Figure 7). Upon filtering our data to only include patients who did not smoke, we observed that significant differences in the levels of these biomarkers between *E. faecalis* and *E. faecium* remained, suggesting that these biomarkers were not confounded by smoking status (Supplementary Figure 10B). Interestingly AZGP1, one of our top biomarkers for distinguishing *E. faecalis* from *E. faecium* and a gene reported to be overexpressed in the airway upon smoking 52, was not found to be influenced by smoking status in our dataset.

**Prediction of Clinical Outcome**

Next, we set out to combine the multi-omic data with the extensive medical metadata associated with our patient cohort to determine whether we could identify biomarkers predictive of mortality upon admission to the ER. No association of mortality with overall proteomic or metabolomic profiles was observed in the unsupervised hierarchical clustering of proteomics data (Figure 1C) or metabolomics (Figure 1D).

87 proteins were found to be significantly enriched in patients who succumbed to mortality, while 29 proteins were found to be significantly enriched in patients who survived (Figure 6A).

Enrichment analysis of the significant proteins found to be associated with survival showed enrichment for terms associated with the Golgi apparatus, external side of the plasma membrane, calcium ion binding, and extracellular matrix (Figure 6B). Proteins significantly associated with mortality on the other hand were found to have more proteins enriched in blood coagulation, blood microparticle, complement activation, alternative pathway, negative regulation of endopeptidase activity serine-type endopeptidase inhibitor activity, and endopeptidase inhibitor activity (Figure 6C).

HRG and 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 0.86 and 0.84 respectively (Figure 6D).

Very few metabolite features showed significant differences in mortality outcomes. 6 were found to be significantly increased in mortality, while only 3 were found to be significantly increased in patients that survived (Figure 6E). The top 2 identified metabolite biomarkers, decanoylcarninine and 1−hexadecyl−sn−glycero−3−phosphocholine showed moderate predictive accuracy, producing ROC AUCs of 0.75, and 0.79 respectively (Figure 6F). Assessing our top 10 protein and metabolite biomarkers of mortality for significant associations with potentially confounding variables showed no significant associations other than ICU admission status, an expected association with mortality (Supplementary Figure 7).

Lastly, we asked if there were any genomic features of the strains of *E. faecalis* or *E. faecium* that were associated with mortality. To assess this, we performed nanopore sequencing and *de novo* assembly on gDNA from the *E. faecium* and *E. faecalis* isolates isolated from patients in our study. We observed different gene content across the clinical strains assessed, but we observed no correlations with mortality through hierarchical clustering of gene content by presence or absence (Supplemental Figure 9).

**Discussion:**

The progress achieved in medical and surgical interventions has brought about a demographic shift characterized by an aging population with more comorbidities and compromised immune systems. This, coupled with an increasing rate of drug-resistant pathogens necessitating the use of broad-spectrum antibiotics, has created a ‘perfect storm’ paving the way for the emergence of generally less virulent but intrinsically antibiotic-resistant commensals like *Enterococcus* spp. to become significant pathogens in invasive diseases such as bacteremia. Despite the increasing relevance of enterococci as pathogens, the host systemic response to enterococcal bacteremia has never been unbiasedly described using a muti-omics approach.

Outside of advancing our understanding, a comprehensive evaluation of the systemic response to EcB also serves a practical purpose. Patients with EcB may not present the classical signs and symptoms of infection, such as fever, elevated white blood cell count, and pain. Instead, they often manifest with vague symptoms such as generalized weakness, malaise, and weight loss, making diagnosis and assessment challenging. Molecular tests of host systemic response are infrequently used in the clinic to diagnose bacteremia or inform on therapy, and the data presented here can be used to assess the feasibility of leveraging the systemic host response to predict the presence of Enterococcal bacteremia, whether an infection is driven by *E. faecalis* or *E. faecium*, or whether a patient is responding well to the current treatment strategy.

In principle, molecular diagnostics leveraging differences in systemic host response have several desirable qualities. They utilize plasma, which is a clinically accessible, easy-to-obtain biospecimen, and disease-relevant host proteins can be detected without any amplification of signal, as is necessary when performing blood culture-based testing. In result, these tests offer the potential to provide informative results on the time scale of hours rather than days. If true differences in host systemic response exist, then they represent attractive targets to exploit for developing novel diagnostics.

In light of these concepts, our work uses high-resolution plasma profiling TMT multi-omics to begin to dissect the host responses of patients afflicted with enterococcal bacteremia, focusing on the two most common species: *E. faecalis* and *E. faecium*. This approach is driven by three overarching goals: (1) characterizing the host response to enterococcal bacteremia relative to homeostasis; (2) determining the systemic differences between *E. faecalis* and *E. faecium* bacteremia; (3) defining 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 outcome 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. As such, we identified several individual protein and metabolite biomarkers that were able to predict the presence of EcB relative to healthy volunteers with near-perfect discriminatory power. This shows there is the potential for molecular diagnostics to predict Enterococcal bacteremia, 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 Enterococcal bacteremia rather than solely being an indicator of inflammation. When we leveraged a complementary dataset comparing *S. aureus* bacteremia to healthy volunteers to examine this possibility, we found that there were indeed several significant proteins and biological processes that differed across these two types of bacteremia in addition to many expected conserved responses. Of particular interest was the observation that proteins associated with the platelet alpha granule lumen were decreased in *S. aureus* bacteremia. Platelet α-granules contain proteins with direct microbicidal properties as well as chemokine functions. Two of the proteins associated with platelet α-granules, PPBP (CXCL7) and PF4 (CXCL4) are potent chemokines that attract neutrophils. This is in combination with the observation that neutrophil associated proteins are significantly enriched in EcB but not in *S. aureus* bacteremia suggests differences in the role of neutrophils between Enterococcal and *S. aureus* bacteremia. Technical considerations as to the experimental design of our *S. aureus* and EcB study prevents the direct comparison of these two data sets, but the differences we observe relative to healthy patients suggest the existence of features of host response that could conceivably be exploited to distinguish these types of bacteremia. A larger study designed to directly compare these, and other types of bacteremia is warranted to uncover these differences with greater confidence 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 of statistical significance. TNF-alpha and IL-6 are two of the major inflammatory cytokines that are increased in patients with bacteremia and sepsis 53, and production of these cytokines by the innate immune system in response to the presence of bacteria are likely major drivers of the general inflammatory responses that we observed. As such, these features are unlikely to be useful for predicting EcB specifically, instead functioning as markers of general inflammation.

As expected, several proteins generally involved in the acute phase response, or inflammatory processes were observed to be enriched during systemic infection with both analyzed enterococcal species and *S. aureus*. Most of the features we identified as being most successful at distinguishing healthy from infected have been previously reported to be biomarkers of other inflammatory processes; including Gelsolin 54 55, LRG156, and LBP 57. Furthermore, proteins involved in cholesterol metabolism were noted to be reduced upon infection in all bacteremia types. This makes sense given cholesterol’s involvement in a myriad of biological processes, where it has roles in immunity, cellular membrane processes, signaling, pathway regulation, precursor for the synthesis of steroid hormones, bile acids, vitamin D, and oxysterols 58. Both LDL and HDL are reported to be reduced in cases of sepsis, regardless of the causative organism 59, and here we show that hypocholesteremia is also one of the major features of Enterococcal bacteremia.

Our observations that primary and secondary bile acids were significantly enriched upon infection in all bacteremia types is a clear indicator of cholestasis, where the increase in inflammation caused by proinflammatory cytokines results in impaired bile acid flow and increased bile acid concentrations in serum/plasma 60. Systemic inflammation also has the capacity to activate and amplify coagulation, and as such, changes in proteins associated with coagulation have been reported as a general response to sepsis regardless of the causative organism 61. We observed that several of our top identified features for distinguishing healthy from infected are known to be involved in degradation of fibrin clots or platelet aggregation including SERPINA3 62,63, and ENO164 which were increased in infection, as well as SERPINA5 65,66 ,CLEC3B (Tetranectin)67, and 13−keto−9z,11e−octadecadienoic acid 68 which were decreased.

Most of the top metabolite features associated with the presence of EcB reported in this study are novel associations with bacteremia to our knowledge. This suggests there is potential for these to be specific markers to Enterococcal bacteremia. Two of the top features associated with EcB, 9-oxootre and cyclo(l-phe-d-pro) have been reported to have antimicrobial activity. 9-oxootre belongs to a class of molecules called oxylipins that can be produced by the LOX, COX and CYP450 pathways 69, has been reported to possess antimicrobial activity on a variety of bacterial and fungal species 70, while also possessing anti-inflammatory activity. Cyclo(l-phe-d-pro) belongs to a class of molecules called Diketopiperazines which are the smallest cyclic peptides known and have been isolated from Gram-positive bacteria, fungus, and higher organisms 71. It has been demonstrated to possess strong antibiotic activity against *Vibrio anguillarum* 72 and likely has antimicrobial activity in other contexts. The reduction in these metabolites observed in infection could indicate consumption of these antimicrobial metabolites produced by the host throughout combating infection. One of the other top biomarkers, N6−threonylcarbamoyladenosine, is a nucleoside modification that is found in all kingdoms of life and has been reported to restrict translation initiation to AUG and suppress frameshifting at tandem ANN codons 73, having previously been proposed to be a strong candidate biomarker for COVID-19 infection and severity 74.Three of the most effective metabolites at differentiating healthy from infected have roles as plasticizers including phthalic anhydride, 7-bis(2−ethylhexyl) phthalate, and 5-tris(2−butoxyethyl) phosphate 75,76. 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 when utilizing highly sensitive, untargeted multi-omic approaches, as has been reported elsewhere 77. In order to determine the specificity of these metabolite biomarkers for EcB, a comprehensive study directly comparing plasma from multiple types of bacteremia patient populations should be conducted.

Differences between *E. faecalis* infected samples from *E. faecium* proved to be less stark than those observed when comparing healthy to infected but 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 dataset. Distinguishing *E. faecalis* from *E. faecium* by host response has the potential to be a clinically important diagnostic, as it could inform the optimal choice of antimicrobial therapy more quickly than the current state of the art and conceivably improve patient outcomes as a result. However, our data suggests that the ability to distinguish these highly related infections based on systemic host response as observed in plasma is limited to a best case of around 80% accuracy, a degree of predictive power not sufficient to be useful for clinicians who would certainly utilize the slower, but more accurate current methods.

Even so, interesting differences between *E. faecium* and *E. faecalis* bacteremia were uncovered through our approach. 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 such as 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 78,79. 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 highlighting that antibodies are more important for preventing *E. faecium* bacteremia than they are for *E. faecalis*. 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 bacteremia, as both retinol and retinol binding protein 4 (RBP4) were increased in *E. faecium* bacteremia relative to *E. faecalis* bacteremia and were among the top 2 proteomic and metabolomic features capable of discriminating the infections. While RBP4 levels were found to be potentially confounded by transplant status in our data set, narrowly missing our threshold for significance when only considering patients without a transplant, retinol levels remained significantly different. Together, this suggests that the differences in these retinol associated features were most likely truly 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 80 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 aspects of the innate and adaptive immune system, and the differences observed in this study may have an impact on immune system function 81. 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 among two types of closely related bacteremia. It is possible that these differences function as an indication of the underlying disease severity, as lipoprotein levels in serum have been shown to correlate with infection severity 82 and *E. faecium* tend to be less virulent than *E. faecalis* 83. If this was indeed the case, we would expect to see lipoproteins also be associated with mortality, a finding not observed in our study. Thus, the underlying reasons for these interesting differences in proteins involved in cholesterol metabolism are currently opaque.

Evaluation of the systemic responses comparing mortality to survival uncovered a number of important differences among these patients. We found features derived from proteomics data to be better at predicting mortality status than those derived from metabolomics data. The biological processes observed to be decreased in mortality compared to survival were primarily composed of the same ones that were found to be significantly different when comparing healthy to infected. This suggests that the degree of alteration relative to baseline for proteins involved in blood coagulation, blood microparticle (lipoproteins), and negative regulation of endopeptidase (SERPINS) activity correlate with the severity of enterococcal bacteremia. The biological processes that were found to be significantly increased in mortality were vague, including calcium ion binding, extracellular matrix, external side of the plasma membrane, and the Golgi apparatus making any potential biological relevance hard to discern.

A number of features from the proteomics data were able to predict mortality, with classification accuracies approaching 90%. Decreases in two cysteine protease inhibitors belonging to the type-3 cystatins class of structurally and functionally related proteins,84 HRG and FETUB were the best predictors of mortality. These proteins have roles in blood coagulation, and HRG has been shown to have antibacterial properties 85. They have been identified as biomarkers of mortality in a variety of other studies, with evidence showing that lower levels were associated with mortality in COVID-19 86, *S. aureus* bacteremia 24, as well as mouse models of sepsis- where delivery of external HRG was even able to improve outcome 87. This opens the possibility that a similar approach may be useful in the context of EcB.

The degree of accuracy displayed by these potential biomarkers may be of clinical relevance, as there is no existing quantitative method to stratify patients in the setting of EcB. Furthermore, diagnostic tests with ROCs ranging from 0.8 – 0.9 are considered to have excellent discriminatory power 88, suggesting the degree of accuracy we observed could be useful in the clinic. Further work would need to explore whether the predictive accuracy observed in this study would hold up in an independent cohort, and whether predicting patient outcomes could be paired with an intervention to have a meaningful impact on patient outcomes.

In conclusion, our study has advanced a powerful multi-omics framework outlined previously24 to produce the most granular profile of the systemic response to EcB to date, reporting significant differences between multi-omic plasma profiles of EcB relative to healthy volunteers, bacteremia driven by *E. faecalis* and *E. faecium*, as well as mortality and survival. The culmination of these efforts represents a significant stride toward characterizing the systemic response to enterococcal bacteremia and its relationship to other bacteremia types.

Acknowledgements:

CB and LR 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. CS was supported in part by the NIH/NIAMS T32 AR064194 Rhematic Diseases Training Grant. D.J.G. is supported by R01AI148417. This study was also supported by the UCSD Collaborative Center of Multiplexed Proteomics.

Data Availability: The data reported in this manuscript are available on MassIVE under the following identifier: Proteomics (MSV000094414), Metabolomics (MSV000094413). 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 manuscript 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 is available upon request.

Tables: See Attached Files

Figure Legends: See Attached Files

Figure 1. See Attached Files

Figure 2. See Attached Files

Figure 3. See Attached Files

Figure 4. See Attached Files

Figure 5. See Attached Files

Figure 6. See Attached Files

Supplemental Figure 1. See Attached Files

Supplemental Figure 2. See Attached Files

Supplemental Figure 3. See Attached Files

Supplemental Figure 4. See Attached Files

Supplemental Figure 5. See Attached Files

Supplemental Figure 6. See Attached Files

Supplemental Figure 7. See Attached Files

Supplemental Figure 8. See Attached Files

Supplemental Figure 9. See Attached Files

Supplemental Figure 10. See Attached Files

Supplemental Table 1: See Attached Files

Supplemental Table 2: See Attached Files

**References**

1. Lebreton, F., Willems, R. J. L. & Gilmore, M. S. Enterococcus Diversity, Origins in Nature, and Gut Colonization.

2. Krawczyk, B., Wityk, P., Gałęcka, M. & Michalik, M. The Many Faces of Enterococcus spp.—Commensal, Probiotic and Opportunistic Pathogen. *Microorganisms* **9**, 1900 (2021).

3. Hufnagel, M. *et al.* Enterococcal colonization of infants in a neonatal intensive care unit: associated predictors, risk factors and seasonal patterns. *BMC Infect Dis* **7**, 107 (2007).

4. Schloissnig, S. *et al.* Genomic variation landscape of the human gut microbiome. *Nature* **493**, 45–50 (2013).

5. Lebreton, F. *et al.* Tracing the Enterococci from Paleozoic Origins to the Hospital. *Cell* **169**, 849-861.e13 (2017).

6. Fiore, E., Van Tyne, D. & Gilmore, M. S. Pathogenicity of Enterococci. *Microbiol Spectr* **7**, 7.4.9 (2019).

7. Higuita, N. I. A. & Huycke, M. M. Enterococcal Disease, Epidemiology, and Implications for Treatment.

8. García-Solache, M. & Rice, L. B. The Enterococcus: a Model of Adaptability to Its Environment. *Clin Microbiol Rev* **32**, e00058-18 (2019).

9. Elizaga, M. L., Weinstein, R. A. & Hayden, M. K. Patients in Long-Term Care Facilities: A Reservoir for Vancomycin-Resistant Enterococci. *Clinical Infectious Diseases* **34**, 441–446 (2002).

10. Chow, J. W., Davidson, A., Sanford, E. & Zervos, M. J. Superinfection with Enterococcus faecalis During Quinupristin/Dalfopristin Therapy. *Clinical Infectious Diseases* **24**, 91–92 (1997).

11. Herrero, I. A., Issa, N. C. & Patel, R. Nosocomial Spread of Linezolid-Resistant, Vancomycin-Resistant Enterococcus faecium. *N Engl J Med* **346**, 867–869 (2002).

12. Sabol Kathryn *et al.* Emergence of Daptomycin Resistance in Enterococcus faecium during Daptomycin Therapy. *Antimicrobial Agents and Chemotherapy* **49**, 1664–1665 (2005).

13. Arias, C. A. & Murray, B. E. The rise of the Enterococcus: beyond vancomycin resistance. *Nat Rev Microbiol* **10**, 266–278 (2012).

14. Rosselli Del Turco, E., Bartoletti, M., Dahl, A., Cervera, C. & Pericàs, J. M. How do I manage a patient with enterococcal bacteraemia? *Clinical Microbiology and Infection* **27**, 364–371 (2021).

15. Jain, S., Gautam, V. & Naseem, S. Acute-phase proteins: As diagnostic tool. *J Pharm Bioall Sci* **3**, 118 (2011).

16. Casadevall, A. & Pirofski, L. Host‐Pathogen Interactions: The Attributes of Virulence. *J INFECT DIS* **184**, 337–344 (2001).

17. Mundy, L. M., Sahm, D. F. & Gilmore, M. Relationships between Enterococcal Virulence and Antimicrobial Resistance. *Clin Microbiol Rev* **13**, 513–522 (2000).

18. Russo, A., Picciarella, A., Russo, R., d’Ettorre, G. & Ceccarelli, G. Time to Effective Therapy Is an Important Determinant of Survival in Bloodstream Infections Caused by Vancomycin-Resistant Enterococcus spp. *IJMS* **23**, 11925 (2022).

19. McCabe, R. E. & Remington, J. S. C-reactive protein in patients with bacteremia. *J Clin Microbiol* **20**, 317–319 (1984).

20. Bhatti, J. M. *et al.* Antibiotic choices among healthcare professionals for enterococcal bacteremia with patterns of resistance and risk factors of mortality, in settings of poor antibiotic stewardship program — a five-year retrospective cohort study. *BMC Infect Dis* **23**, 514 (2023).

21. Suppli, M. *et al.* Mortality in enterococcal bloodstream infections increases with inappropriate antimicrobial therapy. *Clinical Microbiology and Infection* **17**, 1078–1083 (2011).

22. Bivona, G., Agnello, L. & Ciaccio, M. Biomarkers for Prognosis and Treatment Response in COVID-19 Patients. *Ann Lab Med* **41**, 540–548 (2021).

23. Dhanasekaran, S. M. *et al.* Delineation of prognostic biomarkers in prostate cancer. *Nature* **412**, 822–826 (2001).

24. Wozniak, J. M. *et al.* Mortality Risk Profiling of Staphylococcus aureus Bacteremia by Multi-omic Serum Analysis Reveals Early Predictive and Pathogenic Signatures. *Cell* **182**, 1311-1327.e14 (2020).

25. Lautenbach, E., Bilker, W. B. & Brennan, P. J. Enterococcal Bacteremia: Risk Factors for Vancomycin Resistance and Predictors of Mortality. *Infect Control Hosp Epidemiol* **20**, 318–323 (1999).

26. Frantzi, M., Bhat, A. & Latosinska, A. Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development. *Clinical & Translational Med* **3**, e7 (2014).

27. Skates, S. J. *et al.* Statistical Design for Biospecimen Cohort Size in Proteomics-based Biomarker Discovery and Verification Studies. *J. Proteome Res.* **12**, 5383–5394 (2013).

28. Lapek, J. D. *et al.* Defining Host Responses during Systemic Bacterial Infection through Construction of a Murine Organ Proteome Atlas. *Cell Systems* **6**, 579-592.e4 (2018).

29. Wang, Y. *et al.* Reversed‐phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. *Proteomics* **11**, 2019–2026 (2011).

30. Wierzbicki, I. H. *et al.* Group A Streptococcal S Protein Utilizes Red Blood Cells as Immune Camouflage and Is a Critical Determinant for Immune Evasion. *Cell Reports* **29**, 2979-2989.e15 (2019).

31. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, 100141 (2021).

32. Holman, J. D., Tabb, D. L. & Mallick, P. Employing ProteoWizard to Convert Raw Mass Spectrometry Data. *CP in Bioinformatics* **46**, (2014).

33. Schmid, R. *et al.* Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat Biotechnol* **41**, 447–449 (2023).

34. Wang, M. *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* **34**, 828–837 (2016).

35. Campeau, A. *et al.* Multi-omics of human plasma reveals molecular features of dysregulated inflammation and accelerated aging in schizophrenia. *Mol Psychiatry* **27**, 1217–1225 (2022).

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

37. Kim, D., Song, L., Breitwieser, F. P. & Salzberg, S. L. Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res.* **26**, 1721–1729 (2016).

38. Li, H. seqtk. (2023).

39. Wright, C., Griffiths, Sarah & Parker, Matthew. wf-bacterial-genomes.

40. Palumbo, E. bamstats.

41. Freire, B., Ladra, S. & Parama, J. R. Memory-Efficient Assembly using Flye. *IEEE/ACM Trans. Comput. Biol. and Bioinf.* 1–1 (2021) doi:10.1109/TCBB.2021.3108843.

42. Wright, C. & Wykes, M. Medaka. (2023).

43. Seeman, T. mlst.

44. Florensa, A. F., Kaas, R. S., Clausen, P. T. L. C., Aytan-Aktug, D. & Aarestrup, F. M. ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microbial Genomics* **8**, (2022).

45. Trizna, M. assembly\_stats.

46. Page, A. J. *et al.* Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693 (2015).

47. He, Z. & Yu, W. Stable feature selection for biomarker discovery. *Computational Biology and Chemistry* **34**, 215–225 (2010).

48. Neumann, U. *et al.* Compensation of feature selection biases accompanied with improved predictive performance for binary classification by using a novel ensemble feature selection approach. *BioData Mining* **9**, 36 (2016).

49. Kupcova Skalnikova, H., Cizkova, J., Cervenka, J. & Vodicka, P. Advances in Proteomic Techniques for Cytokine Analysis: Focus on Melanoma Research. *IJMS* **18**, 2697 (2017).

50. Pichet Binette, A. *et al.* Confounding factors of Alzheimer’s disease plasma biomarkers and their impact on clinical performance. *Alzheimer’s & Dementia* **19**, 1403–1414 (2023).

51. Ain, Q. U., Sarfraz, M., Prasesti, G. K., Dewi, T. I. & Kurniati, N. F. Confounders in Identification and Analysis of Inflammatory Biomarkers in Cardiovascular Diseases. *Biomolecules* **11**, 1464 (2021).

52. Vanni, H. *et al.* Cigarette Smoking Induces Overexpression of a Fat-Depleting Gene AZGP1 in the Human. *Chest* **135**, 1197–1208 (2009).

53. Chaudhry, H. *et al.* Role of Cytokines as a Double-edged Sword in Sepsis. *in vivo* (2013).

54. Piktel, E., Levental, I., Durnaś, B., Janmey, P. & Bucki, R. Plasma Gelsolin: Indicator of Inflammation and Its Potential as a Diagnostic Tool and Therapeutic Target. *IJMS* **19**, 2516 (2018).

55. Yang, Z. *et al.* Delayed Administration of Recombinant Plasma Gelsolin Improves Survival in a Murine Model of Penicillin-Susceptible and Penicillin-Resistant Pneumococcal Pneumonia. *The Journal of Infectious Diseases* **220**, 1498–1502 (2019).

56. Hashida, T. *et al.* Proteome analysis of hemofilter adsorbates to identify novel substances of sepsis: a pilot study. *J Artif Organs* **20**, 132–137 (2017).

57. Blairon, L., Wittebole, X. & Laterre, P. Lipopolysaccharide‐Binding Protein Serum Levels in Patients with Severe Sepsis Due to Gram‐Positive and Fungal Infections. *J INFECT DIS* **187**, 287–291 (2003).

58. Hofmaenner, D. A., Kleyman, A., Press, A., Bauer, M. & Singer, M. The Many Roles of Cholesterol in Sepsis: A Review. *Am J Respir Crit Care Med* **205**, 388–396 (2022).

59. Fraunberger, P., Schaefer, S., Werdan, K., Walli, A. K. & Seidel, D. Reduction of Circulating Cholesterol and Apolipoprotein Levels during Sepsis. *cclm* **37**, 357–362 (1999).

60. Bhogal, H. K. The molecular pathogenesis of cholestasis in sepsis. *Front Biosci* **E5**, 87–96 (2013).

61. Davis, R. P., Miller‐Dorey, S. & Jenne, C. N. Platelets and coagulation in infection. *Clin &amp; Trans Imm* **5**, e89 (2016).

62. Sanchez-Navarro, A., González-Soria, I., Caldiño-Bohn, R. & Bobadilla, N. A. Integrative view of serpins in health and disease: the contribution of serpinA3. *American Journal of Physiology-Cell Physiology* ajpcell.00366.2020 (2020) doi:10.1152/ajpcell.00366.2020.

63. Bianchini, E. P., Auditeau, C., Razanakolona, M., Vasse, M. & Borgel, D. Serpins in Hemostasis as Therapeutic Targets for Bleeding or Thrombotic Disorders. *Front. Cardiovasc. Med.* **7**, 622778 (2021).

64. Liu, K.-J. & Shih, N.-Y. The Role of Enolase in Tissue Invasion and Metastasis of Pathogens and Tumor Cells. (2007).

65. Yaron, J. R., Zhang, L., Guo, Q., Haydel, S. E. & Lucas, A. R. Fibrinolytic Serine Proteases, Therapeutic Serpins and Inflammation: Fire Dancers and Firestorms. *Front. Cardiovasc. Med.* **8**, 648947 (2021).

66. Maas, C. & De Maat, S. Therapeutic SERPINs: Improving on Nature. *Front. Cardiovasc. Med.* **8**, 648349 (2021).

67. Iram, S., Rahman, S., Choi, I. & Kim, J. Insight into the function of tetranectin in human diseases: A review and prospects for tetranectin-targeted disease treatment. *Heliyon* **10**, e23512 (2024).

68. Fang, X., Kaduce, T. L. & Spector, A. A. 13-(S)-Hydroxyoctadecadienoic acid (13-HODE) incorporation and conversion to novel products by endothelial cells. *Journal of Lipid Research* **40**, 699–707 (1999).

69. Cambiaggi, L., Chakravarty, A., Noureddine, N. & Hersberger, M. The Role of α-Linolenic Acid and Its Oxylipins in Human Cardiovascular Diseases. *IJMS* **24**, 6110 (2023).

70. Prost, I. *et al.* Evaluation of the Antimicrobial Activities of Plant Oxylipins Supports Their Involvement in Defense against Pathogens. *Plant Physiology* **139**, 1902–1913 (2005).

71. P. De Carvalho, M. & Abraham, W.-R. Antimicrobial and Biofilm Inhibiting Diketopiperazines. *CMC* **19**, 3564–3577 (2012).

72. Fdhila, F., Vázquez, V., Sánchez, J. L. & Riguera, R. dd -Diketopiperazines: Antibiotics Active against *Vibrio* *a* *nguillarum* Isolated from Marine Bacteria Associated with Cultures of *Pecten* *m* *aximus*. *J. Nat. Prod.* **66**, 1299–1301 (2003).

73. Swinehart, W. *et al.* Specificity in the biosynthesis of the universal tRNA nucleoside *N* 6 -threonylcarbamoyl adenosine (t 6 A)—TsaD is the gatekeeper. *RNA* **26**, 1094–1103 (2020).

74. Nagayoshi, Y. *et al.* t6A and ms2t6A Modified Nucleosides in Serum and Urine as Strong Candidate Biomarkers of COVID-19 Infection and Severity. *Biomolecules* **12**, 1233 (2022).

75. Wang, Y. & Qian, H. Phthalates and Their Impacts on Human Health. *Healthcare* **9**, 603 (2021).

76. Win-Shwe, T.-T. *et al.* Dietary Exposure to Flame Retardant Tris (2-Butoxyethyl) Phosphate Altered Neurobehavior and Neuroinflammatory Responses in a Mouse Model of Allergic Asthma. *IJMS* **23**, 655 (2022).

77. Cruickshank-Quinn, C. *et al.* Impact of Blood Collection Tubes and Sample Handling Time on Serum and Plasma Metabolome and Lipidome. *Metabolites* **8**, 88 (2018).

78. Naegeli, A. *et al.* *Streptococcus pyogenes* evades adaptive immunity through specific IgG glycan hydrolysis. *Journal of Experimental Medicine* **216**, 1615–1629 (2019).

79. Kagawa, T. F. *et al.* 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 (2000).

80. Rosales, F. J., Ritter, S. J., Zolfaghari, R., Smith, J. E. & Ross, A. C. Effects of acute inflammation on plasma retinol, retinol-binding protein, and its mRNA in the liver and kidneys of vitamin A-sufficient rats. *Journal of Lipid Research* **37**, 962–971 (1996).

81. Stephensen, C. B. VITAMIN A, INFECTION, AND IMMUNE FUNCTION. (2001).

82. Sammalkorpi, K., Valtonen, V., Kerttula, Y., Nikkilä, E. & Taskinen, M.-R. Changes in serum lipoprotein pattern induced by acute infections. *Metabolism* **37**, 859–865 (1988).

83. Miller, W. Treatment of enterococcal infections. *UpToDate*.

84. Lee, C. Type 3 cystatins; fetuins, kininogen and histidine-rich glycoprotein. *Front Biosci* **Volume**, 2911 (2009).

85. Rydengård, V., Olsson, A., Mörgelin, M. & Schmidtchen, A. Histidine‐rich glycoprotein exerts antibacterial activity. *The FEBS Journal* **274**, 377–389 (2007).

86. Völlmy, F. *et al.* A serum proteome signature to predict mortality in severe COVID-19 patients. *Life Sci. Alliance* **4**, e202101099 (2021).

87. Kuroda, K. *et al.* Histidine-rich glycoprotein as a prognostic biomarker for sepsis. *Sci Rep* **11**, 10223 (2021).

88. Mandrekar, J. N. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *Journal of Thoracic Oncology* **5**, 1315–1316 (2010).