Title Page: 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 Dorrestein, 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

Present Address For: Carlos Gonzalez

Running Title: Differentiating Host Responses in Enterococcus faecium vs faecalis Bacteremia

Abbreviations:

Importance (150 words):

We utilize quantitative un-targeted proteomics and metabolomics to provide the first described plasma profiles capturing the systemic host response observed in Enterococcal bacteremia patients. We report significant multi-omic differences between Enterococcal bacteremia and healthy volunteers while leveraging extensive clinical metadata associated with our patient cohort to describe differences in host-response observed between infections caused by *E. faecalis* or *E. faecium*, as well as between mortality and survival. Altogether, the 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 exploration of this rich, first of its kind data source, we also provide a companion data analysis application that can be accessed by the research community at: <https://charliebayne.shinyapps.io/EcB_Multiomics/>

Abstract <300 words:

While typically members of the commensal microbiome, *Enterococcus faecalis* and *Enterococcus faecium* have the potential to become clinically relevant pathogens under certain circumstances. These circumstances are met in the healthcare setting, where these microbes commonly persist and infect immunocompromised patients due to their intrinsic capability to survive disinfection routines and ability to acquire foreign DNA through horizontal gene transfer to increase their fitness. As such, Enterococcus infections have become one of the most prevalent hospital acquired infections. A clinically relevant manifestation of these infections is Enterococcal bacteremia (EcB), where mortality rates as high as 46% have been reported.

Despite the prevalence and severity of EcB, systemic host responses to EcB have never been comprehensively examined. Here we present an extensive multi-omic resource profiling these 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 us the unique opportunity to examine enterococcal bacteremia with an added layer of granularly, comparing EcB caused by *E. faecalis* to that caused by *E. faecium* as well as to bacteremia caused by *Staphylococcus aureus*. Additionally, the collection of extensive patient metadata allowed us to examine features associated with mortality or survival and test them for associations with confounding variables. We report significant differences between the plasma profiles of healthy and infected patients, individuals with *E. faecalis* and *E faecium* bacteremia, and 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. The gastrointestinal (GI) tract is the predominant reservoir of enterococci, and in recent decades, enterococci have emerged as important healthcare associated pathogens. This has been largely attributable to the increased usage of broad-spectrum antibiotics such as cephalosporins and carbapenems, to which enterococci are intrinsically resistant 2. The human associated enterococcal species, *Enterococcus faecalis* and *Enterococcus faecium*, establish their niche as commensals within the first 10 days following birth 3,4. The combination of invasive procedures and aggressive medical care in increasingly older and complex patients, along with inadequate antibiotic stewardship, have resulted in a rise in severe, invasive infection caused by enterococci, which are increasingly multi-drug resist 2,5,6,7,8. Enterococcal infections are both prevalent and deadly, with around 7% of hospital acquired infections being due to enterococcus and a mortality rate of up to 46% being reported in enterococcal bacteremia (EB).

Diagnosis of enterococcal bacteremia is currently performed through the use of blood cultures, where optimal treatment can be aided by the determination of the antibiotic sensitivity of the isolate responsible for the infection. General markers of inflammation such as CRP and SAA can be used to assist the diagnostic process of bacteremia but suffer from offering inconclusive results and are unable to inform on optimal antibiotic treatment strategy. As such, they are infrequently used.

Despite the clinical importance of EcB there have been no dedicated efforts to profile systemic features of host response to it. To address this gap in the field, we employed high resolution mass spectrometry (MS)-based proteomics and metabolomics to profile plasma samples collected from clinical EB cases as well as healthy controls. 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) enterococcal bacteremia compared to healthy, (2) bacteremia caused by *E. faecalis* compared to *E. faecium*, and (3) EB patients that succumbed to mortality compared to those that survived. The information gleaned from such data offers the first insights of their kind into the systemic molecular and biological features that underly the host-pathogen relationship in EB. Importantly, this data set also provides an initial assessment of our ability to use features of the host response to predict the presence of enterococcal bacteremia, 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.

**Experimental Procedures:**

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 for the discovery cohort was selected based on technical considerations for multiplexed proteomics and metabolomics approaches as well as logistical constraints9 10. 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, 30 patients with *E. faecium* bacteremia and 46 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 healthy controls was collected from blood bank volunteers.

**Proteomics Arm:**

Protein Preparation – A 25uL aliquot of plasma from each patient was added to 200uL lysis buffer containing 6M urea, 7% SDS, 50mM TEAB, and one each protease inhibitor tablet and PhosStop tablet (Roche), final pH 8.1 adjusted with phosphoric acid. Then samples were bath sonicated and vortexed for 5 minutes each. Disulfide bonds were reduced with 100uM DTT for 30 minutes at 47C, cooled briefly on ice, followed by alkylation with 300uM IAA for 45 minutes at room temperature in the dark. The reactions were quenched with 100uM DTT at room temperature for 5 minutes. Samples were loaded on S-Trap 96-well plates (Protifi) per manufacturer’s instructions. Samples were digested with 5ug of trypsin in 115uL 50mM TEAB over 3 hours at 47C. Peptides were eluted serially with 125uL of 50mM TEAB, then 5%FA, finally 50% ACN 5% FA, then dried under vacuum centrifugation. Peptides were desalted using 50mg Sep-Pak tC-18 cartridges (Waters) using manufacturer instructions and were then quantified with a colorimetric peptide assay (Pierce). Fifty micrograms of each sample were aliquoted for TMT labeling with a further 10ug of each sample being combined and aliquoted in 50ug for usage as the bridge channel.

TMT Labeling – Aliquots were labeled with TMT-Pro 16-Plex reagents (Cat # A44520, ThermoFisher Scientific) as described previously with channel 134N reserved as the bridge channel 11 .

Proteomics LC-LC-MSn - Basic pH reverse-phase LC, followed by data acquisition through LC–MS2/MS3, was performed as previously described 11. Briefly, 75 min linear gradients of 22% to 35% acetonitrile and 10mM ammonium bicarbonate were passed on HPLC C18 columns (Biobasic) with the resulting ninety-six fractions combined as previously described12. 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.13

Statistics and Data Analysis - Thermo .raw files were converted to mzML files using MSConvert with a peakpicking (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, stricttrypsin digestion rules allowing for 2 missed cleavages, and the search was set to include static modifications of carbidomethylation of cystines, TMTpro on lysine 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 14. 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 ul) were thawed for 30 mins, then 200uL of prechilled extraction solvent (100% MeOH with 1 mM sulfamethazine as an internal standard) was added to each sample. Samples were mixed by vortexing for 2 minutes then incubated at 20C for 20 min to aid in protein precipitation. Samples were centrifuged at 16,000 x g for 15 min to pellet the protein precipitate. Supernatant was then transferred into 96-well deep well plate, dried using centrifugal low-pressure system and stored at -80C 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). 5uL 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 min 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) 15. mzML files were processed together using MzMine (3.5.0)16 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 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 table were then exported using GNPS FBMN option to generate required files for online GNPS FBMN analysis.

Feature Annotations – Feature annotations was done through the Global Natural Products Social molecular networking (GNPS) Feature-Based Molecular Networking workflow (Version 28.2) 17. MS2 MGF files were exported from MzMine 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 was set to 6 and 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 – Metabolite features Metabolomics data was normalized using sulfamethazine as a single internal standard as described previously.18 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 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 19.

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, freezer stocks were created by the addition of a final concentration of 15% glycerol. Strain identify 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* glycerol stocks were inoculated into 1mL of BHI. These cultures were subsequently incubated at 37C 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 (Mat No 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 left at 37C overnight to facilitate lysis. The next day, 45 μL of proteinase K solution ( Cat # 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. 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 3x using 1 mL of QIAGEN buffer QC. Genomic DNA was then eluted by applying 1mL of buffer QF twice. 1.4 mL of room temperature isopropanol was then added to the eluate, and it was inverted several times in order to precipitate the DNA. To collect the DNA, samples were centrifuged at 12,000xg for 15 minutes at 4C. 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 50μL 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.

Resulting DNA was then checked for purity by Nanodrop and DNA concentration was assessed using Qubit™ dsDNA Quantification Assay Kit, broad range (Q32853, ThermoFisher). A subset of purified DNA samples were run on a Genomic DNA ScreenTape (Agilent, 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 (SQK-NBD114.96). The manufacturers Ligation sequencing gDNA – Native Barcoding Kit 96 V14 version NBE\_9171\_v114\_revl\_15Sep2022 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- Unexpectedly, initial inspection of the sequencing data showed that a large number of reads mapped to the streptomyces genus. Further investigation uncovered that the Labiase ordered from the vendor had been “purified” from *Streptomyces fulvissimus TU-6*, and subsequent qubit quantification of the Labiase stock used showed that it contained a significant concentration of DNA. Thus, we concluded that the streptomyces reads we observed resulted from a preparation of Labiase contaminated with *Streptomyces fulvissimus* gDNA and opted to take a bioinformatic approach to decontaminating the reads. Briefly, all reads were matched to a reference database using Centrifuge20 and reads mapping to streptomyces were then removed using seqtk21. Validation of remaining read identity was performed using an independent database (BLAST) after which we saw that reads assigned to *Streptomyces* were almost completely eliminated after our decontamination steps.

These decontaminated reads were then used as input into the Epi2me labs graphic interface which was used to run the wf-bacterial-genomes nextflow workflow22. Briefly, this workflow concatenates input files and generate per read stats via bamstats23, performs denovo assembly via flye24, polishes with medaka25, performs multi locus sequence typing (MLST)26 to infer identity of the isolate, and lastly performs antimicrobial resistance (AMR) calling via ResFinder 27to identify genes associated with antimicrobial resistance. Basic assembly statistics such as N50 and the number of contigs were also generated using assembly\_stats28.

The resulting FASTA assemblies were then used as input to Roary 29 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 individuals, 30 diagnosed with *E. faecium* bacteremia, and 46 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 30, highlighting that the identification of species can inform the selection optimal antibiotic therapy.

In our proteomics dataset, we quantified a total of 589 proteins, of which 278 were quantified across all samples. 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 rich feature space combined with patient-centric clinical and microbiological metadata allows us to examine the differences between healthy and infected*, E. faecalis* and *E. faecium*, while providing an unique opportunity to explore associates 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). However, dissimilarities between the two types of bacteremia were more subtle, indicating an anticipated large-scale similarity between the two closely related pathologies. The moderate stratification observed suggested the potential for identifying proteomic and metabolomic differences between the profiles produced by *E. faecalis* bacteremia and *E. faecium* bacteremia.

**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 more abundant in Enterococcal bacteremia while 85 were found to be significantly less abundant (Figure 2A. FDR adjusted pvalue <= 0.05). The adjusted p values for many of these proteins were highly significant, reaching values as extreme as 1 x 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 10 x 10-50 (Figure 3A). Of the metabolite features identified in our study, a minority (29%) of them were able to be putatively identified using GNPS molecular networking, displaying a limitation inherent to current untargeted metabolomics analysis approaches 34 (Figure 3B).

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 all EcB types, we observed an enrichment in neutrophil chemotaxis, tertiary granule lumen, focal adhesion, extracellular exosome, and inflammatory response (Figure 2B).

EcB and S. aureus bacteremia had a number of 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 a 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 patients.

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 the background of the experiment (Figure 3C). 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 (Supplemental Figure X) where the abundances of Taurocholic acid, Glycocholic acid, Tauroursodeoxycholic acid, 3beta−Hydroxy−5−cholenoic acid, and 12−Ketodeoxycholic acid were enriched. Interestingly, several phostidylcholine 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 a number of phostidylcholines were also significantly depleted in the case of EcB bacteremia (Supplementary Table X).

We next set out to evaluate the potential utility of features collected within our dataset as biomarkers to predict healthy from infected samples. To rank these biomarkers within our dataset, we used ensemble feature selection (EFS) as an unbiased approach to rank proteins capable of distinguishing healthy from infected patients17. 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 31, 32. 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, SERPINA3, 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 C-reactive protein and serum amyloid A were found to be significantly increased in infected samples compared to healthy (Supplementary Figure 3). As expected, they also displayed no significant differences between *E. faecalis* and *E. faecium* infected samples. 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 33, we inferred cytokine profiles using previously reported methods in which known interactions between proteins and cytokines are leveraged to infer cytokine abundance 34. Using this technique, we saw that the inferred amount of TNF-a was significantly increased in infected samples compared to healthy samples (p.adj <= 0.05) (Supplementary Figure 6). We also observed a trend in many other cytokines to be increased in infected relative to healthy, notably in IL-6, but they did not rise to the level of statistical significance.

**Differences in Systemic Plasma Profiles in 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, where >90% of the isolates are 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). We observed comparable results when investigating the metabolomic data, observing that 11 metabolites that were significantly increased in *E. faecalis* relative to *E. faecium*, and 33 that were significantly increased in *E. faecium* (Figure 5A).

GO term enrichment analysis of the proteins that were found to be significantly increased in *E. faecalis* showed dramatic difference 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 rather than increased levels in *E. faecalis* as there were no differences in antibody abundances between healthy and *E. faecalis*, however there was a reduction in antibody abundances in *E. faecium* relative to healthy (Figure 4C). These differences were also apparent in the GO TERM analysis comparing *E. faecium* infected samples to healthy, 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 *E. faecalis* infected samples, 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).

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 factors9,35,36 . Evaluation of the clinical metadata for associated correlations showed a few confounding clinical variables were associated with the abundances of these biomarkers other than type of pathogen. Transplant type was found to be significantly associated with the levels of 9/10 of the biomarkers (Supplementary Figure 7). Further investigation suggested that this would be hard to separate from the type of pathogen infected as *E. faecium* infected patients were more likely to have organ transplants in our cohort. Thus, we considered the possibility that the reduction in antibodies we observed in our study were a function of unbalanced groupings. In order to assess this, we only considered the patients that did not have a transplant and found that antibody levels were still significantly reduced in *E. faecium* infected samples relative to *E. faecalis* (Supplementary Figure 5). Additionally, smoking was found to be significantly associated with the abundances of 5/10 of the top biomarkers (Supplementary Figure 4). 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 37, was not found to be influenced by smoking status in our dataset . 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 our conclusions were not confounded by additional factors (Supplementary Figure 10).

When we queried the cytokine inference data for differences between *E. faecalis* and *E. faecium* inferred cytokine profiles, we did not see any significant differences (Supplementary Figure 6).

**Prediction of Host Response**

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 that succumbed to mortality, while 29 proteins were found to be significantly enriched in patients that survived (Figure 6A). Very few metabolite features showed significant differences by mortality outcome. 6 were found to be significantly increased in mortality, while only 3 were found to be significantly increased in patients that survived (Figure 6E).

Enrichment analysis of the significant proteins found to be associated with survival showed an enrichment for terms associated with the Golgi apparatus, external side of 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 best 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). The top 2 identified metabolite biomarkers, Decanoylcarninine and 1−hexadecyl−sn−glycero−3−phosphocholine showed moderate predictive accuracy, producing ROC AUS of 0.75, and 0.79 respectively (Figure 6F).

The top biomarkers protein biomarkers showed good predictive accuracy. MAN1A1, MASP1, CTSD, and FBLN1 were increased in cases of mortality, while HRG, FETUB, SERPINA7, KNG1, and SERPINA5 were increased in survival (Supplementary Figure 8). Assessing our top 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 then 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. Hierarchical clustering of these genes by presence or absence did not show distinct clustering by mortality or survival (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 and endocarditis. Despite the increasing relevance of enterococci as pathogens, the host systemic response to enterococcal bacteremia has never been broadly described.

Outside of advancing knowledge by describing the molecular responses to Enterococcal bacteremia, comprehensive evaluation of these responses also serves a practical purpose. Patients with enterococcal bacteremia 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 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 a number of desirable qualities. They utilize plasma, which is an extremely easy to obtain biospecimen, and do not requiring any amplification, as is necessary to perform blood culture-based testing. As a results, 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 the purposes of developing novel diagnostics.

In light of these concepts, our work uses high-resolution plasma profiling 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 were able to predict the presence of EcB relative to healthy volunteers with near perfect discriminatory power. This finding 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. We noticed that proteins associated with extracellular exosomes were highly enriched in EcB, while no such association was identified in S. aureus bacteremia. Extracellular exosomes are well established to have a variety of roles in a variety of immunological processes, and the differential abundance of proteins associated with them may indicate immunological differences in the host response to EcB and *S. aureus* bacteremia. Furthermore, we observed evidence for the presence of immunological differences between these two types of bacteremia include the observation that a variety of processes associated with neutrophils were observed to be significantly enriched in EcB, while no such association was observed in *S. aureus* bacteremia, potentially indicating that neutrophils are more activated in EcB relative to *S. aureus*. Furthermore proteins associated with the platelet alpha granule lumen were enriched in EcB, but decreased in *S. aureus* bacteremia, representing the only shared GO term between these two types of bacteremia that was found to be enriched in opposite directions relative to healthy when comparing the two pathogens. This is particularly interesting as platelet α-granules contain proteins with direct microbicidal properties. Technical considerations as to the experimental design of our study prevents the direct comparison of EcB to *S. aureus* bacteremia plasma, but the differences we observe relative to healthy patients suggest the existence of features of host response that could 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.

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-a and IL-6 are two of the major inflammatory cytokines that are increased in patients with bacteremia and sepsis 38, 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 Enterococcal bacteremia specifically, instead functioning as markers of general inflammation.

The conserved responses of EcB and *S. aureus* bacteremia to heathy also revealed conserved biological responses to infection across bacteremia types. A number of proteins generally involved in the acute phase response, or inflammatory processes were observed to be enriched in infection to each type of bacteremia. As such, 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 39 40, LRG141, and LBP 42. Furthermore, proteins involved in cholesterol metabolism were noted to be reduced upon infection in all bacteremia types. Cholesterol is involved in a myriad of biological processes, with roles in immunity, cellular membrane processes, signaling, pathway regulation, precursor for the synthesis of steroid hormones, bile acids, vitamin D, and oxysterols 43. Both LDL and HDL are reported to be reduced in cases of sepsis, regardless of the causative organism 44, and here we show evidence of hypocholesterolemia also occurring in Enterococcal bacteremia. Additionally, our study observed indications of cholestasis in both EcB and *S. aureus* bacteremia, as a variety of primary and secondary bile acids were significantly enriched upon infection. The increase in inflammation caused by proinflammatory cytokines results in impaired bile acid flow, leading to increased bile acid concentrations in serum/plasma 45 . Notably we saw a significant depletion of phostidylcholines in *S. auerus* bacteremia in our enrichment analysis that was not observed in EcB, however closer inspection showed a number of phostidylcholines were reduced in EcB as well. Phostidylcholines have significant roles in inflammation, and our study . Systemic inflammation 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 46. We observed that proteins involved in coagulation and these associations with several of our top identified features for distinguishing healthy from infected including SERPINA3 47,48, SERPINA5 49,50, ENO1, CLEC3B, and 13−keto−9z,11e−octadecadienoic acid 51

Most of the top metabolite features associated with the presence of enterococcal bacteremia reported in this study represent novel associations with bacteremia, suggesting there is potential for these to be specific markers of Enterococcal bacteremia. Interestingly, 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 52, has been reported to possess antimicrobial activity on a variety of bacterial and fungal species 53, 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 54. It has been demonstrated to possess strong antibiotic activity against *Vibrio Anguillarum* 55 and likely has antimicrobial activity in other contexts. The reduction in these metabolites observed in infection could indicate a consumption of these antimicrobial metabolites produced by the host over the course of 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 56, having previously been proposed to be a strong candidate biomarker for COVID-19 infection and severity 57.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 58,59. 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 60. Altogether, while these metabolites represent novel associations with bacteremia to our knowledge, it is unclear if this is due to a lack of metabolic profiling of plasma derived from bacteremic patients in general, or if these are unique to EcB. To investigate this further, a comprehensive resource allowing for the direct comparison of plasma from multiple types of bacteremia patient populations would be needed.

Differences between *E. faecalis* infected samples from *E. faecium* proved to be less stark than those observed when comparing healthy to infected but we were able to identify a number of features that were able to 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 thus improve patient outcomes. 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 potentially confounding variables 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 61,62. 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 the two possibilities could account for the difference in immunoglobulins reported between the two cohorts.

We also noted marked differences in cholesterol metabolism when comparing *E. faecalis* to *E. faecium* bacteremia, observing that the proteins involved in these processes were generally enriched in *E. faecium* bacteremia. 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 possible that these differences function as an indication of the underlying disease severity, as *E. faecium* infections tend to be less virulent than *E. faecalis* ones, but if this was the case we would expect to see lipoproteins also be associated with mortality- a finding not observed in our study. Thus, reasons for these differences in proteins associated with cholesterol metabolism are currently unclear.

Evaluation of the systemic responses comparing mortality to survival showed that there were a number of important differences in the systemic host response to EcB among these patients. Interestingly, proteomics data was better at predicting mortality status than metabolomic data, both at the global level, and on a per feature basis.

Our observation that biological processes associated with calcium ion binding, extracellular matrix, external side of plasma membrane, and the Golgi apparatus were significantly upregulated in mortality, suggested that there may be some ….

The biological processes observed to be decreased in infection included blood microparticle, blood coagulation, complement activation, and regulation of endopeptidase activity.

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,63 HRG and FETUB were the best predictors of mortality. These proteins have roles in blood coagulation, and HRG has been displayed to have antibacterial properties. 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 64, *S. aureus* bacteremia34, as well as mouse models of sepsis- where delivery of external HRG was even able to improve outcome 65. 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 66, 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 outcome.

In conclusion, our study has advanced a potent multi-omics framework outlined previously 34 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: This work was funded by

Data Availability: The data generated for this manuscript are available on the ProteomeXchange archive or MassIVE under the following identifier: Proteomics (ProteomeXchange Number). Metabolomics data is available on MassIVE (MassIVE Number).

An interactive data analysis app to explore the data contained in this manuscript can be found at : <https://charliebayne.shinyapps.io/EcB_Multiomics/> All analysis of downstream data reported in the manuscript can be found in github repo at <https://github.com/baynec2/EcB_multiomics>.

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

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

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

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

4. 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).

5. 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).

6. 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).

7. 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).

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

9. 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).

10. 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).

11. 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).

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

13. 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).

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

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

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

17. 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).

18. 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).

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

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

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

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

23. Palumbo, E. bamstats.

24. 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.

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

26. Seeman, T. mlst.

27. 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).

28. Trizna, M. assembly\_stats.

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

30. 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).

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

32. 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).

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

34. 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).

35. 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).

36. 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).

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

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

39. 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).

40. 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).

41. 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).

42. 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).

43. 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).

44. 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).

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

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

47. 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.

48. 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).

49. 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).

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

51. 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).

52. 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).

53. 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).

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

55. 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).

56. 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).

57. 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).

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

59. 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).

60. 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).

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

62. 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).

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

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

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

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