UNIVERSITÀ DEL PIEMONTE ORIENTALE

SCUOLA DI MEDICINA

DIPARTIMENTO DI MEDICINA TRASLAZIONALE/SCIENZE DELLA SALUTE

Class of degrees in Medical, veterinary and pharmaceutical biotechnology (LM-9)

Degree Thesis

Longitudinal Gut Microbiota Profiling by 16S Metagenomics in Patients with Sepsis: A Representative Cohort from the SURVEIL Project

Supervisor: **Prof. Sandra D'ALFONSO**

Co-Supervisor: **Dr. Marta MELLAI**

Candidate:

Alaa JAMMOUL

Matricola 20063520

Academic Year 2025/2026

Table of Contents

Abstract

Sepsis is a life-threatening organ dysfunction caused by a dysregulated host response to infection, frequently accompanied by gut microbiome alteration and colonization with multidrug-resistant (MDR) bacteria. How the gut community changes over the course of sepsis remains incompletely described. As the metagenomic component of the multicenter SURVEIL (SvilUppo di una piattafoRma di Vigilanza gEnomica per IL contrasto della pandemia covid-19) project (Principal Investigator: Prof. Sandra D'Alfonso), this study used 16S rRNA gene sequencing to characterize the gut microbiome of critically ill patients by sepsis, comparing Gram-negative (GN) and Gram-positive (GP) sepsis and MDR colonization status.

A subgroup of 16 patients of the SURVEIL project was selected from the original cohort for a multi-omics study including 16S metagenomics. From this subgroup, 14 septic patients were successfully analyzed. Their rectal swabs, collected at admission, sepsis onset, and recovery (n = 34), were sequenced by Next Generation Sequencing (NGS) on an Illumina MiSeq platform across the V3, V4, and V6 hypervariable regions of the 16S rRNA gene. Primary and secondary bioinformatic analyses included alpha and beta diversity, the *Bacillota-to-Bacteroidota* ratio, and linear discriminant analysis effect size (LEfSe) were computed in MicrobiomeAnalyst across three complementary comparisons.

The gut microbial community was markedly disturbed during the acute phase and only partially reorganized by recovery. The most consistent finding was the expansion of the opportunistic GP anaerobic coccus *Finegoldia magna* at the acute phase and its pronounced reduction at recovery, most marked in GP patients (median species-level abundance from 27.85% to 1.22%). This signal appeared from the species up to the class and phylum levels and was mirrored by a shift from *Bacillota* to *Bacteroidota*. GP patients showed lower microbial richness than GN patients, partly confounded by their more acute sampling timing. Individual trajectories varied widely and in some patients the gut community remained or got more imbalanced despite clinical recovery. No change survived correction for multiple testing, so the findings are exploratory; nevertheless, the *Finegoldia* signal and a recovery-associated *Enterococcus* enrichment were concordant with the full SURVEIL cohort.

Overall, this study provides a longitudinal, 16S-based description of gut dysbiosis in sepsis, identifies *Finegoldia magna* as a candidate acute-phase marker, provides a multi-omic perspective to a parallel transcriptomic profiling study, and establishes a metagenomic basis for future host-microbe data integration through machine learning within the SURVEIL project.

Keywords: sepsis; gut microbiota; 16S rRNA gene sequencing; dysbiosis; *Finegoldia magna.*

1. Introduction

1.1 Sepsis: Pathophysiology

Sepsis is currently defined, according to the Third International Consensus Definitions (Sepsis-3), as a life-threatening organ dysfunction caused by a dysregulated host response to infection[1]. Septic shock represents its most severe subset, in which underlying circulatory, cellular, and metabolic abnormalities are profound enough to substantially increase mortality, and it is identified clinically by the need for vasopressors to maintain an adequate mean arterial pressure together with a raised serum lactate despite adequate fluid resuscitation[1]. This conceptual shift moved the focus away from the older view of sepsis as infection plus a systemic inflammatory response, and placed organ dysfunction and the disordered host response at the center of the definition[1].

The clinical burden of sepsis is considerable. The most recent estimate from the Global Burden of Disease study reported approximately 166 million sepsis cases and 21.4 million sepsis-related deaths worldwide in 2021, corresponding to roughly one third of all global deaths[2]. These figures update earlier estimates, which had described tens of millions of cases and around 11 million deaths in 2017, and they indicate that the long decline in sepsis mortality was reversed during the COVID-19 pandemic[2,3]. In high-income settings, sepsis remains one of the leading reasons for admission to intensive care units and a frequent cause of in-hospital death, with a substantial proportion of cases acquired during the hospital stay[4].

From a pathophysiological standpoint, sepsis is not a single disease but a complex syndrome that can be triggered by bacterial, viral, or fungal infection at almost any anatomical site[5]. The host response combines an early, often exaggerated pro-inflammatory phase with a concurrent and frequently prolonged immunosuppressive phase[5,6]. During the pro-inflammatory phase, recognition of microbial components by innate immune receptors drives the release of cytokines, activation of complement and coagulation, and widespread endothelial dysfunction, which together promote microvascular injury, tissue hypoperfusion, and progressive organ failure[5,6]. In parallel, many patients develop a state of immune exhaustion, characterized by lymphocyte apoptosis, reduced antigen presentation, and a shift toward anti-inflammatory mediators, which leaves them vulnerable to secondary infections[6,7]. The balance between these processes varies widely between individuals and over the course

of time, which contributes to the marked clinical heterogeneity of the syndrome and to the difficulty of developing targeted therapies[5,7].

1.2 Sepsis: Clinical Diagnosis

Because sepsis is defined by organ dysfunction rather than by the infection alone, its clinical recognition depends on identifying that dysfunction in a patient with suspected or confirmed infection[1]. In the Sepsis-3 framework, organ dysfunction is determined through the Sequential Organ Failure Assessment (SOFA) score, which grades the function of the respiratory, cardiovascular, hepatic, coagulation, renal, and neurological systems; an acute increase of two points or more in the SOFA score, in the context of infection, is taken to represent sepsis[1]. For rapid bedside screening outside the intensive care unit, the quick SOFA (qSOFA) was proposed, based on three simple criteria: altered mental status, a respiratory rate of at least 22 breaths per minute, and a systolic blood pressure of 100 mmHg or lower[1].

The transition to these criteria has been accompanied by ongoing discussion about how best to balance sensitivity, specificity, and ease of use at the bedside[8]. The earlier Systemic Inflammatory Response Syndrome (SIRS) criteria, although sensitive, lacked specificity for a dysregulated response, whereas qSOFA is more specific but can miss patients at an early stage[8,9]. Comparative analyses of SIRS, SOFA, qSOFA, and the National Early Warning Score have reported differing performance for the diagnosis of sepsis and for the prediction of adverse outcomes, and no single score performs best across all settings[9]. In practice, clinical diagnosis therefore combines these structured tools with the overall clinical picture, and it remains a time-critical judgment, since delays in recognition translate directly into worse outcomes[1,9].

1.3 Sepsis: Microbiological Laboratory Diagnosis

Laboratory investigation supports the clinical diagnosis of sepsis in two complementary ways: by documenting the infection and its causative organism, and by quantifying the host response and organ dysfunction. Microbiological confirmation depends primarily on blood cultures (hemocultures), which remain the reference standard for detecting bloodstream infection[10]. Current recommendations are to collect at least two sets of blood cultures, each comprising an aerobic and an anaerobic bottle, from separate peripheral venipunctures using strict aseptic technique[10]. Sampling before the first dose of antimicrobials is important, because prior antibiotic exposure

markedly reduces the probability of recovering the responsible organism[11]. Once a bottle flags positive, the diagnostic workflow typically proceeds through Gram staining, subculture on selective and enriched media, biochemical or molecular species identification, increasingly by mass spectrometry such as Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight (MALDI-TOF MS), and antimicrobial susceptibility testing to define the resistance profile[10].

Blood cultures nonetheless have well-recognized limitations, including a relatively low positivity rate, a turnaround time of up to several days, and the risk of contamination by skin commensals, which can complicate interpretation[10,11]. Alongside microbiological testing, routine laboratory parameters help to characterize the severity of the response. The complete blood count, and in particular the leukocyte count and its differential count, contributes both diagnostic and prognostic information[12]. Serum lactate is measured as an indicator of tissue hypoperfusion and is incorporated into the definition of septic shock, where a concentration above 2 mmol/L despite adequate resuscitation carries prognostic weight[1]. Additional markers of organ dysfunction, such as creatinine, bilirubin, and coagulation indices, feed directly into the SOFA score and into the day-to-day assessment of the septic patient[1,12].

1.4 Clinical Biomarkers

Given the limitations of culture and clinical scores, considerable effort has been directed toward circulating biomarkers that can support the early recognition, risk stratification, and monitoring of sepsis[13]. The most widely used are acute-phase reactants and inflammatory mediators. C-reactive protein (CRP) rises in response to infection and inflammation, and persistently high levels have been associated with an increased risk of death, although it is not specific to bacterial sepsis[13]. Procalcitonin is more closely linked to systemic bacterial infection and is used, in some cases, to support decisions about starting and stopping antibiotics, even if its diagnostic accuracy varies across populations[13]. Interleukin-6 (IL-6) is an early pro-inflammatory cytokine that reflects the intensity of the host response, while other indices, including the erythrocyte sedimentation rate, ferritin, and fibrinogen, provide complementary but non-specific information[13].

Several additional markers have been proposed to improve on these established analytes. Calprotectin, a protein released by activated neutrophils and monocytes, has

been reported to perform, at least as well as procalcitonin, as a sepsis marker and predictor of short-term mortality in critically ill patients[14]. Neutrophil CD64 expression increases markedly during bacterial infection and has been explored as a diagnostic indicator[13]. More recently, the monocyte distribution width has been validated and cleared for use as an early indicator of sepsis in patients admitting to the emergency department as a measure of the variability in monocyte cell volume that can be obtained from a routine blood count[15]. Despite this expanding panel, no single biomarker is sufficiently sensitive or specific to serve as a stand-alone test[13]. This limitation has motivated growing interest in combining several markers and in host-response-based strategies, which aim to capture the complexity of the septic status more completely than any individual analyte[13].

1.5 Treatment of Sepsis

The management of sepsis is built around early recognition followed by the rapid delivery of a small number of time-critical interventions, as set out in the Surviving Sepsis Campaign guidelines[16]. Prompt administration of broad-spectrum antimicrobials is essential, since each hour of delay in appropriate therapy is associated with worse outcomes; the initial regimen is chosen to cover the most likely pathogens for the suspected source and the local resistance pattern[16]. This is combined with control of the infectious source where feasible, for example through drainage or removal of an infected device, and with hemodynamic resuscitation[16].

Resuscitation involves intravenous fluids to restore perfusion, guided in part by serum lactate, and the early use of vasopressors, with norepinephrine recommended as the first-line agent, to maintain a mean arterial pressure of at least 65 mmHg[16]. Supportive care for failing organs, such as mechanical ventilation or renal replacement therapy, is provided as required[16]. An important principle is that the initial broad antimicrobial therapy should be reassessed and narrowed, or de-escalated in order to limit unnecessary exposure, once culture and susceptibility results identify the organism[16]. This last point connects treatment directly to a central problem of modern intensive care: the broad and often prolonged use of antimicrobials, while necessary, exerts strong selective pressure that favors the emergence and persistence of resistant organisms, which is the subject of the following section[16].

1.6 Multi-Drug-Resistant Organisms

Antimicrobial resistance is now recognized as one of the major threats to global health, and it is particularly relevant in the setting of sepsis and intensive care, where antibiotic use is intense and patients are highly vulnerable[17]. The World Health Organization has formalized the most pressing targets in its Bacterial Priority Pathogens List, which was updated in 2024 and ranks organisms by the urgency of the need for new treatments, the critical-priority group includes carbapenem-resistant *Acinetobacter baumannii* and carbapenem-resistant and third-generation cephalosporin-resistant *Enterobacterales*[17]. Many of the most problematic species are captured by the ESKAPEE acronym, a group of nosocomial pathogens notable for their capacity to escape the effects of commonly used antibiotics[18].

Multidrug-resistant organisms are responsible for a large share of healthcare-associated infections, which affect a substantial proportion of hospitalized patients and add considerably to morbidity, mortality, and cost[18]. Gram-negative (GN) bacteria are especially important in this context, both because of the severity of the infections they cause and because of the limited therapeutic options that remain once resistance is established[19]. The spread of carbapenemase-producing *Enterobacterales*, including strains carrying enzymes such as *Klebsiella Pneumoniae* Carbapenemase (KPC) and New Delhi Metallo-β-lactamase (NDM), has been a particular concern, since carbapenems are often among the last reliable agents for serious GN infections[20]. The pathogenesis of GN bloodstream infection, from initial colonization through tissue invasion to systemic disease, illustrates why resistant members of this group are so closely associated with severe sepsis[21]. Crucially, the human body, and the gastrointestinal tract in particular, can act as a silent reservoir in which such resistant organisms are carried before they cause invasive infection, a link that is developed further in the final section of this chapter[19,21].

1.7 The Gut Microbiome

The human microbiome is the collective community of microorganisms, together with their genes, that inhabit the body, and the gastrointestinal tract harbors by far its largest and most diverse component[22]. Because the great majority of gut bacteria are difficult or impossible to grow in the laboratory, knowledge of this community has advanced mainly through culture-independent, sequencing-based methods, which have

made it possible to describe the composition of the healthy gut in detail[22,23]. At the level of bacterial phyla, the healthy adult gut is dominated by *Bacillota* (formerly *Firmicutes*) and *Bacteroidota* (formerly *Bacteroidetes*), with smaller and more variable contributions from *Actinomycetota*, *Pseudomonadota* (formerly *Proteobacteria*), and *Verrucomicrobiota*[22,23].

Far from being passive passengers, these microorganisms perform functions that are essential to host physiology[24]. They contribute to the development and regulation of both innate and adaptive immunity, reinforce the integrity of the intestinal barrier, and participate in metabolism, including the fermentation of dietary fiber[24,25]. A key product of this fermentation is the group of short-chain fatty acids, principally acetate, propionate, and butyrate, which serve as an energy source for colonocytes and help to shape immune and inflammatory responses[26]. Butyrate, for example, has been shown to imprint an antimicrobial program in macrophages, linking microbial metabolism directly to host defense[27]. A further fundamental function is colonization resistance, the capacity of an intact community to prevent incoming pathogens from establishing themselves, through competition for nutrients and space and through interactions with the host immune system[28].

The structure of the gut microbiome is commonly described using ecological diversity measures[29]. Alpha diversity refers to the diversity within a single sample and captures two related properties: richness, the number of distinct taxa present, and evenness, how uniformly they are distributed. Several indices summarize these properties in different ways. The number of observed taxa and the Chao1 estimator describe richness, with Chao1 giving particular weight to rare taxa in order to estimate the total number of species likely present, including those missed by sampling. The Shannon index combines richness and evenness in a single value, so that a community with many taxa in balanced proportions scores higher than one dominated by a few. The Simpson index also accounts for evenness but is more strongly influenced by the most abundant taxa. Beta diversity, on the other hand, describes the difference in composition between samples and is typically explored using ordination methods, such as principal coordinates analysis, which place compositionally similar samples close together. A high and balanced alpha diversity is generally associated with stability, resilience, and the ability of the community to resist and recover from disturbance [29]. The relative balance between the two dominant phyla is often expressed as the

Bacillota-to-*Bacteroidota* (B:B) ratio, a simple summary indicator of community state, and shifts in this ratio have been associated with a range of physiological and pathological conditions[23,29]. Beyond these summary measures, composition varies considerably across individuals and is shaped by factors including diet, geography, and age, with characteristic changes reported in older adults, a point of relevance to the typically elderly intensive care population[30,31,32].

1.8 16S rRNA Sequencing

The profiling of bacterial communities such as the gut, relies heavily on sequencing of the 16S ribosomal RNA gene, an approach whose foundations were laid when this gene was first proposed as a molecular marker for reconstructing microbial phylogeny[33]. The 16S rRNA gene is present in all bacteria and archaea and is approximately 1,500 base pairs in length[33]. Its usefulness as a taxonomic marker arises from its structure: highly conserved gene regions, which are shared across distantly related organisms, alternate with nine hypervariable regions (designated V1 to V9) whose sequences differ between taxa[34] (Figure 1).

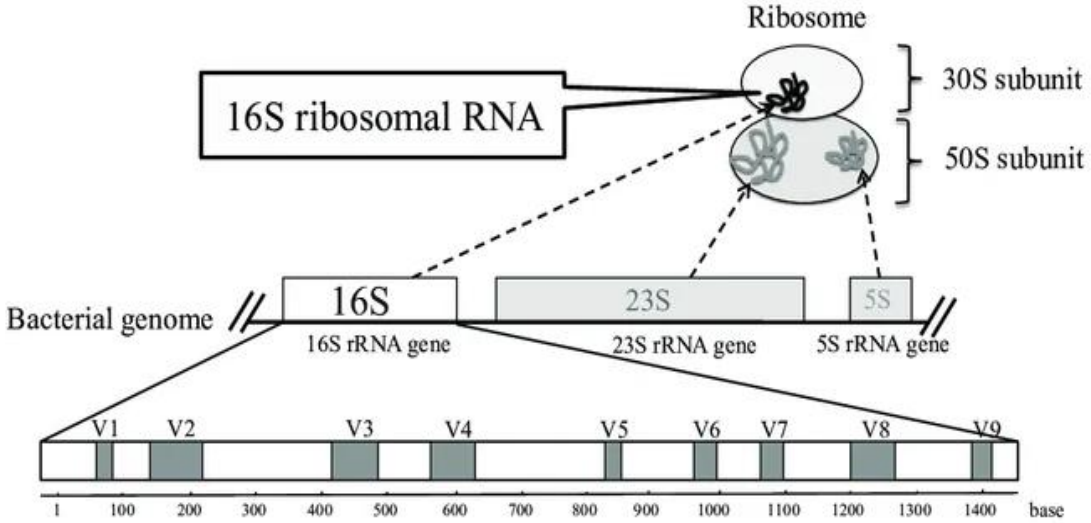


Figure 1: Schematic figure representing the 16S rRNA gene. White boxes represent the constant regions. Dark gray boxes represent the hypervariable regions. Abbreviations: V, Variable. Reproduced from [35].

The conserved regions provide binding sites for broad-range, so-called universal primers, while the variable regions carry the sequence information needed to distinguish one organism from another[34].

By Next Generation Sequencing (NGS), the analysis of the 16S gene is an amplicon-based workflow, where one or more hypervariable regions are amplified by polymerase chain reaction (PCR) and then sequenced on a high-throughput sequencing platform[34]. The choice of region matters, because different regions and primer pairs

vary in their coverage of taxa and in the taxonomic resolution they provide, and combining adjacent regions can improve the ability to discriminate between bacteria[34]. The present experimental workflow targets three main regions (V3, V4, and V6), as described in the Materials and Methods section. The resulting sequence reads are first subjected to quality control, in which low-quality bases and reads are removed, commonly with reference to Phred quality scores that estimate the probability of a base-calling error[36]. Sequences are then grouped into units for analysis: in the traditional approach they are clustered into operational taxonomic units (OTUs) at a fixed similarity threshold (97%), whereas newer denoising methods resolve exact amplicon sequence variants (ASVs) that can differ by as little as a single nucleotide[37]. Taxonomy is assigned by comparing these sequences against curated reference databases, such as the Ribosomal Database Project (RDP)[38]. Several integrated software pipelines, including QIIME 2, are widely used to carry out these steps in research settings and represent alternatives to the dedicated pipeline adopted in this study[39].

It is important to note both strengths and limitations of the method. Sequencing of the 16S rRNA gene is comparatively rapid and inexpensive, does not require the organisms to be cultured, and is well-suited to characterizing whole communities; however, it typically resolves bacteria only to the genus or species level, can be affected by primer and region bias, and yields relative rather than absolute abundance data, meaning that the given insight is compositional and not quantitative[34,37]. These properties distinguish it from shotgun metagenomic sequencing, which sequences total community DNA and offers higher taxonomic and functional resolution but at greater cost and analytical complexity[37].

1.9 Gut Microbiome in Sepsis and Critical Illness

The functions of the gut microbiome described above, the colonization resistance and the support of immune and barrier integrity, make it directly relevant to sepsis and critical illness[40]. Critically ill patients are exposed to a combination of factors, including broad-spectrum antibiotics, acid suppression, interrupted enteral nutrition, altered motility, and impaired perfusion, that together disrupt the normal community[41]. The result is gut dysbiosis: a sharp loss of overall biodiversity, depletion of beneficial commensal anaerobes, and overgrowth of a small number of potentially pathogenic taxa[41]. In many critically ill patients this progresses to domination of the

community by a single genus or family, frequently belonging to the *Pseudomonadota*, such as the *Enterobacteriaceae*, or to *Enterococcus*, a state that has been conceptualized as the emergence of a harmful pathobiome from the collapse of the normal microbiome[42].

This picture is supported by a series of observational studies. Communities of extremely low diversity, sometimes reduced to one or two organisms, have been described in the gut of patients during prolonged critical illness[43]. Pilot work using sequencing has shown that intestinal dysbiosis in the intensive care unit is already present on admission and varies greatly between individuals[44], and altered gut flora have been associated with septic complications and death in critically ill patients with a systemic inflammatory response[45]. The clinical consequences of dysbiosis follow from the loss of colonization resistance and the weakening of the intestinal barrier[28]. The gastrointestinal tract serves as a reservoir from which nosocomial pathogens can be transmitted and cause infection[46], and a high relative abundance, or domination, of a given organism in the gut has been linked to a subsequent risk of bloodstream infection with that same organism[47]. In line with this, carriage of multidrug-resistant bacteria in the gut at the time of intensive care admission or during hospitalization has been identified as a risk factor for later infection and death[48], which directly connects the problem of resistant organisms introduced earlier to the state of the gut community.

More recent work has begun to integrate these observations with measures of the host response. A landmark prospective study using serial 16S rRNA sequencing in intensive care patients linked dysbiosis of a combined microbiota and immune system to the development of nosocomial infections, framing the gut and the immune system as a single interacting metasystem[49].

The relevance of the gut community extends beyond the abdomen, as illustrated by the gut-lung axis, in which the intestinal microbiota contributes to host defense against pneumonia[50]. Contemporary sequencing studies in critically ill and septic patients have repeatedly reported enrichment of *Enterobacteriaceae* together with a reduction in microbial diversity, in designs closely comparable to the present work[51], and recent reviews have synthesized the mechanisms, outcomes, and possible therapeutic implications of intensive care dysbiosis[52,53]. Taken together with the broader understanding of the microbiota in pathogen colonization and immunity[54], this study provides the rationale for examining the gut microbiome of septic patients in

detail. It is within this framework that the present work is situated as the 16S rRNA-based component of the multicenter SURVEIL project, which investigates the gut microbiota as a possible key modulator in the pathophysiology of sepsis[55] in the context of a multi-omic perspective and complementary to a transcriptomic profiling of the same dataset.

The gut microbiome captures only the microbial side of sepsis, whose course and severity are ultimately shaped by the host response that the infection induces. Therefore, a complete description benefits from information on both sides of this interaction. For this reason, within the same SURVEIL cohort and across the same admission, sepsis-onset, and two weeks after the onset of sepsis timepoints analyzed here, the host response of these patients was profiled in parallel by RNA sequencing (RNA-Seq), an NGS technique that measures genome-wide gene expression in peripheral blood and resolves the immune and inflammatory programs activated during sepsis[55]. The 16S metagenomic analysis presented in this study describes the microbiological component of this multi-omics design: it is informative on its own as a longitudinal description of the gut microbiome over the course of sepsis, and it is intended to be integrated with the parallel transcriptomic layer toward a future combined host-and-microbe characterization of the same patients.

2. Aim of Study

Sepsis is accompanied by a profound disturbance of the gut microbiome, and the gut is increasingly regarded both as a reservoir of opportunistic pathogens and as a contributor to systemic complications in critically ill patients. Within the multicenter SURVEIL project, which investigates the relationship between the gut microbiota, multidrug-resistant (MDR) colonization, and the development of sepsis, the present study constitutes the metagenomic component, complementary to the concurrent analysis of the host transcriptional response. Its general aim was to characterize the composition and the longitudinal dynamics of the gut microbiome in patients with sepsis, using 16S rRNA gene sequencing of rectal swabs collected across the course of the disease.

To this end, the study pursued four objectives: (i) to describe the taxonomic composition of the gut microbiota at the phylum, genus, and species levels in septic patients; (ii) to compare the gut microbiome between patients with Gram-Negative (GN) and Gram-positive (GP) sepsis, and between patients colonized (COL) and not colonized (Non-COL) by MDR bacteria at admission; (iii) to follow the changes in the gut community from the onset of sepsis to recovery, using alpha diversity, beta diversity, the *B:B* ratio, and differential abundance analysis; and (iv) to identify candidate microbial features associated with the acute phase of sepsis.

Beyond this descriptive characterization, the study was conceived as the metagenomic foundation for the future integration of microbial and host data within the SURVEIL cohort, in support of the longer-term goal of identifying combined host-microbe signatures of sepsis through multi-omics and machine-learning approaches.

3. Materials and Methods

3.1. Study Design and Experimental Plan

This thesis represents the metagenomic component of the SURVEIL project, a prospective, longitudinal, integrated multi-omics study of the gut microbiome in critically ill adults[55]. A total of 134 patients were consecutively enrolled at hospital admission, between June 2022 and October 2023, across two University Hospitals of the University of Piemonte Orientale: the "SS. Antonio e Biagio e Cesare Arrigo" University Hospital in Alessandria and the "Maggiore della Carità" University Hospital in Novara. Participants were recruited from the internal medicine, cardiac surgery intensive and sub-intensive care, and resuscitation care units. The study was approved by the local Ethics Committees of both hospitals and was conducted in accordance with the Declaration of Helsinki, with written informed consent obtained from each participant or their legal representative before enrollment.

Eligible participants were adults older than 18 years, of either sex or any ethnicity, admitted to the recruiting units with a minimum hospitalization of 48 hours; the only exclusion criterion was the absence of signed informed consent. Sepsis was diagnosed according to the Sepsis-3 criteria, combining clinical evaluation (quick Sequential Organ Failure Assessment, qSOFA, score ≥ 2) and the laboratory biomarker procalcitonin with microbiological confirmation of bacteremia, detected on blood by Gram staining, FilmArray, direct MALDI-TOF mass spectrometry, and the T2Dx system; antimicrobial susceptibility of the isolates was characterized with the Vitek 2 system[1,55].

Sampling for the multi-omics analyses was scheduled at two or three timepoints using eNAT™ collection tubes (COPAN Italia S.p.A., Brescia, Italy), assigned through a stratification algorithm based on the twice-weekly rectal-swab surveillance used to monitor gut colonization by MDR organisms (Figure 2). The timepoints were defined as hospital admission (T0), suspected sepsis (T1a), sepsis confirmed by microbiological analysis (T1b), and recovery (T2); in patients who were already septic at admission, T0 coincided with the onset of sepsis. For the metagenomic analysis presented in this thesis, a rectal swab was collected in an eNAT preservation tube at T0, T1, and T2 and stored at -80 °C until processing.

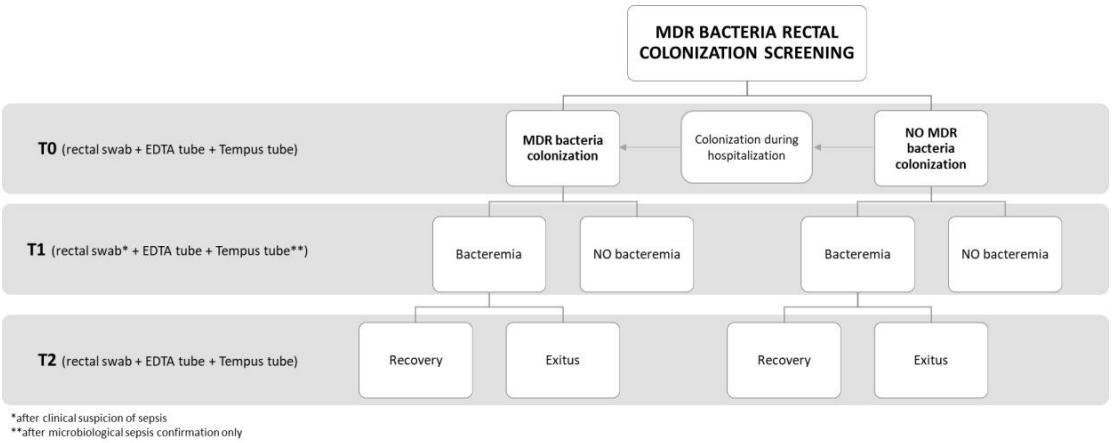


Figure 2: Sampling algorithm based for patient stratification on the twice-weekly swab screening for monitoring the presence of multi-drug resistant bacteria in patients' gut. Abbreviations: MDR, multi-drug resistant.

3.2. Sample Collection

This study was conducted within the SURVEIL project, which was reviewed and approved by the Clinical Trial Center of the Azienda Ospedaliera SS. Antonio e Biagio e Cesare Arrigo of Alessandria (Italy) under protocol number 0009474 of 29 April 2022. All procedures were carried out in accordance with the Declaration of Helsinki and the principles of Good Clinical Practice, and the study was registered on ClinicalTrials.gov (registration number NCT 07183722). Written informed consent was obtained from all participants or their legal representatives prior to sample collection[55].

Rectal swab samples were collected from patients using eNAT™ collection tubes (COPAN Italia S.p.A.). These tubes contain a stabilizing buffer that preserves nucleic acids and maintains sample integrity during transport and storage. All samples were handled as potentially infectious materials and processed in accordance with standard biosafety guidelines. The samples were stored at −80°C until further processing.

3.3. Bacterial DNA Extraction

Microbial DNA was extracted from the rectal swab samples using the QIAamp® PowerFecal® Pro DNA Kit (QIAGEN, Hilden, Germany), following the manufacturer's instructions. This kit was selected for its suitability for stool and gut-derived samples, combining bead-beating mechanical lysis with chemical lysis to an Inhibitor Removal Technology® step, which efficiently depletes the PCR inhibitors commonly co-purified from stool material. All DNA extraction procedures were performed under sterile

conditions within a laminar flow cabinet, employing sterile reagents and consumables, in compliance with good scientific practices to prevent microbial contamination.

Briefly, each sample was transferred into a PowerBead Pro Tube together with 800 µl of Solution CD1 and homogenized by horizontal vortexing at maximum speed for 10 minutes on a Vortex Adapter, according to the manufacturer's protocol. The tube was then centrifuged at 15,000 × g for 1 minute, and approximately 650 µl of the resulting supernatant was transferred directly into the appropriate position of a rotor adapter pre-loaded onto a QIAcube® Connect instrument (QIAGEN, Hilden, Germany). All subsequent steps were performed automatically by the QIAcube Connect, including inhibitor removal by addition of Solution CD2, DNA binding onto an MB Spin Column in the presence of Solution CD3, two sequential washes with Solutions EA and C5, and final elution in Solution C6 (10 mM Tris). The purified DNA was recovered in 50–100 µl of eluate and stored at −20 °C until further use.

3.4. DNA Quantification

The concentration and purity of the extracted DNA were assessed by UV–Vis spectrophotometry on a Synergy™ HT multi-mode microplate reader (BioTek Instruments, Winooski, VT, USA) equipped with a Take3™ Micro-Volume Plate (BioTek Instruments). This configuration requires only 2 µl of each sample and provides absorbance measurement at 260 nm, from which DNA concentration is calculated. The A260/A280 ratio was used to assess potential contamination with proteins or phenol, with a value of approximately 1.8 considered acceptable for pure DNA. The A260/A230 ratio was additionally used as an indicator of carryover from salts, chaotropic agents, or EDTA from the extraction kit buffers, with an acceptable range of 2.0–2.2. Samples showing a low A260/A230 ratio were considered at risk of inhibiting downstream enzymatic reactions. Based on these measurements, each extracted DNA sample was diluted to a working concentration of 1 ng/µl in DNase/RNase-free water immediately before use in the library preparation steps.

3.5. 16S rRNA Gene Library Preparation

Library preparation was performed using the AD4SEQ Microbiota Solution B kit (Arrow Diagnostics S.r.l., Genoa, Italy), which targets the hypervariable regions V3–V4–V6 of the bacterial 16S rRNA gene. This kit was selected for its specificity for stool-derived samples and its compatibility with the Illumina sequencing platform. The

workflow consists of two successive PCR amplification steps (Target PCR and Index PCR), each followed by a purification step, and concludes with normalization and pooling of the indexed libraries.

3.5.1 Target PCR

A master mix (Target Mix) was prepared by combining Enzyme Mix 1, Amp Mix V3–V6, and DNase/RNase-free water following the volumes specified in the kit protocol. The mix was prepared in the dedicated PCR room under the laminar flow hood to minimize contamination risk. For each sample, 15 µl of Target Mix was aliquoted into 0.2 ml PCR strip tubes, followed by the addition of 5 µl of the diluted DNA template at 1 ng/µl. A negative control, in which the DNA template was replaced by 5 µl of DNase/RNase-free water, was included in each amplification run. Samples were briefly centrifuged and placed on the Applied Biosystems 2720 Thermal Cycler (Applied Biosystems, Foster City, CA, USA) running the following program: initial denaturation at 95°C for 5 minutes, followed by 25 cycles of denaturation (95°C, 30 seconds), annealing (55°C, 30 seconds), and extension (72°C, 30 seconds), with a final extension at 72°C for 5 minutes, and hold at 4°C.

3.5.2 Quality Control by Agarose Gel Electrophoresis

The quality of the Target PCR amplicons was verified by agarose gel electrophoresis. A 1.5% agarose gel was prepared in 1X TAE buffer and stained with SYBR® Safe DNA gel stain (Thermo Fisher Scientific, Waltham, MA, USA). Each amplified product (2 µl) was mixed with 2X DNA Loading Dye (Thermo Fisher Scientific) and loaded onto the gel. Electrophoresis was run at 150 mA for 20 minutes. The gel was then imaged using a ChemiDoc™ MP Imaging System (Bio-Rad Laboratories, Hercules, CA, USA). The absence of any visible band in the negative control confirmed the absence of contamination.

3.5.3 Purification of Target PCR Products

Following quality control (QC), the Target PCR amplicons were purified using AMPure XP magnetic beads (Beckman Coulter, Inc., Brea, CA, USA) to remove residual primers and short non-specific fragments. Briefly, each amplified product (17 µl) was diluted with 19 µl of DNase/RNase-free water (Thermo Fisher Scientific), and 29.5 µl of pre-homogenized AMPure XP beads were added and mixed thoroughly by pipetting. After a 5-minute incubation at room temperature, the tubes were placed

on a magnetic rack for 3 minutes to allow bead separation. The supernatant was discarded, and the beads were washed twice with 180 µl of freshly prepared 80% ethanol (1-minute incubation per wash). After complete removal of the ethanol and a 5-minute air-drying step, the purified DNA was eluted by resuspending the beads in 40 µl of DNase/RNase-free water (20 µl for samples with a weak gel band), incubating for 2 minutes, and recovering approximately 35 µl (or 18 µl for samples with a weak band) of supernatant after magnetic separation.

3.5.4 Index PCR

A second amplification step was performed to incorporate a unique barcoded index into each sample. An Index Mix was prepared by combining DNase/RNase-free water and Enzyme Mix 2 at the volumes specified in the kit protocol (18 µl total per sample). Hence, 18 µl of Index Mix was added to the corresponding well of the Index Strip (provided in the kit), mixed well to resuspend the lyophilized index pellet, and transferred to a newly labelled 0.2 ml PCR strip tube (this step was also conducted away from the samples in the PCR room under the hood). Subsequently, 2 µl of the purified Target PCR product was added to each tube. The Index PCR was run on the thermal cycler using the following program: initial denaturation at 95°C for 5 minutes, followed by 10 cycles of denaturation (95°C, 30 seconds), annealing (65°C, 30 seconds), and extension (72°C, 30 seconds), with a final extension at 72°C for 5 minutes, and hold at 4°C. Each sample was assigned a unique single index to allow unambiguous identification after sequencing.

3.5.5 Purification of Index PCR Products

Index PCR products were purified following the same AMPure XP bead-based protocol described in Section 3.5.3, with the following minor adjustments: the elution volume was 27 µl of DNase/RNase-free water (Thermo Fisher Scientific), and approximately 25 µl of the recovered supernatant was transferred to a new labelled tube. A last QC check, like the procedure described in Section 3.5.2, was performed to verify the presence of the indexed amplicons before library normalization.

3.6. Quantification and Normalization of Indexed Libraries

The concentration of each purified indexed library was measured using the Qubit™ 4.0 Fluorometer with the Qubit™ 1X dsDNA HS Assay Kit (Thermo Fisher

Scientific). The molar concentration (nM) of each sample was calculated using the formula provided by Arrow Diagnostics, considering the average amplicon length of 900 bp for the V3–V4–V6 regions:

$$DNA\ conc.[nM] = DNA\ conc.[ng/\mu l] \times 106 / (656.6 \times 900\ bp)$$

Each sample was subsequently diluted to a working concentration of 10 nM in DNase/RNase-free water (Thermo Fisher Scientific). The volume of DNase/RNase-free water (Thermo Fisher Scientific) required for each sample was calculated using a pre-prepared Excel spreadsheet specifically designed for this purpose, which automatically determined the correct dilution volumes based on each sample's measured concentration.

3.7. Library Pooling and Sequencing

Pooling of the normalized libraries was performed by transferring 3 µl of each 10 nM indexed sample into a single 1.5 ml low-binding tube. The pool was diluted from 10 nM to 4 nM and then denatured by combining equal volumes of the 4 nM pool and freshly prepared 0.2 N NaOH, followed by a 5-minute incubation at room temperature. The denatured pool was then diluted to a final loading concentration of 5 pM using the Hybridization Buffer HT1 (Illumina Inc., San Diego, CA, USA).

PhiX Control v3 (Illumina Inc.) was prepared in parallel following the same denaturation and dilution procedure and spiked into the final pool at 20% by volume. The use of PhiX at this proportion serves as an internal sequencing quality control and helps to compensate for the low sequence diversity that is typical of amplicon-based 16S rRNA gene libraries.

A total volume of 600 µl of the final pool was loaded onto the Illumina MiSeq™ platform (Illumina Inc., San Diego, CA, USA) using the MiSeq® v2 Nano Reagent Kit (500 cycles, Illumina Inc.). Sequencing was performed in paired-end mode with a read length of 2 × 250 bp, using the Generate FASTQ run module. The sample sheet was prepared using the Illumina Experiment Manager software, with parameters set according to the AD4SEQ kit specifications. Raw sequencing data were subsequently demultiplexed based on the unique single index assigned to each sample during the Index PCR step. The processing of the samples was carried out at Integrative Genomics Core Facility, Center for Translational Research on Autoimmune and Allergic Disease (CAAD), University of Eastern Piedmont (Novara, Italy), and the sequencing was

carried out in the human genetics laboratory, University of Eastern Piedmont (Novara, Italy).

3.8. Bioinformatics Processing and Data Analysis

The demultiplexed FASTQ files were analyzed using MicrobAT (Microbiota Analysis Tool) software v. 1.3.0 (SmartSeq S.r.l., Novara, Italy). MicrobAT is a standalone, client/server-based application with a Java graphical interface that allows users to load FASTQ files, retrieve metadata, and generate per-sample quality reports. The software applied a quality-filtering step that removed short reads (length < 200 nt) and low-quality reads (average Phred quality score < 25)[36]. Reads passing these thresholds were aligned against the Ribosomal Database Project (RDP) reference database[38]; only alignments with a minimum coverage of ≥ 80% and a sequence similarity of ≥ 97% were assigned at the species taxonomic level. The software produced relative abundance tables, as well as OTU, taxonomy, and metadata files as output for use in downstream analyses.

These output files were subsequently imported into MicrobiomeAnalyst v3.0 (https://microbiomeanalyst.ca)[56], a web-based platform for comprehensive statistical, visual, and meta-analysis of microbiome data. MicrobiomeAnalyst is built on the phyloseq R package[57] and includes a marker gene profiling module specifically designed for community-level profiling and comparative analyses of 16S rRNA gene datasets. It was used here with a threshold of 15 as a minimum count during data filtering for all the subsequent comparisons. This platform provides graphical representations and statistical analyses, including alpha- and beta-diversity assessments, differential abundance testing, linear discriminant analysis effect size (LEfSe) analysis with the linear discriminant analysis (LDA) score, which expresses the effect-size value returned by LEfSe for each discriminant taxon, expressed on a base-10 logarithmic scale, where a higher score reflects a larger and more consistent difference in abundance between the compared groups; only taxa with an LDA score above 2.0 (default configuration) were retained as significant. A 0.05 cutoff for *p* value was used in this study while adjusting for the False Discovery Rate (FDR) to determine the significant features across all the taxonomic ranks in each comparison.

4. Results

4.1 SURVEIL Study Population

According to the enrolment criteria of the SURVEIL project, 134 participants were recruited at two University Hospitals of UPO. Forty-eight patients (35.8%) were enrolled at the "Maggiore della Carità" University Hospital of Novara, of whom 22 (45.8%) developed sepsis; among these, 21 (95.5%) were Non-Col by multidrug-resistant (MDR) bacteria and one (4.5%) was COL. The "SS. Antonio e Biagio e Cesare Arrigo" University Hospital of Alessandria enrolled 86 patients (64.2%), of whom 21 (24.4%) developed sepsis; among these, 16 (76.2%) were MDR-COL and five (23.8%) were not. Overall, 43 of 134 patients (32.1%) developed clinical and microbiologically confirmed sepsis, while 91 (67.9%) did not (Table 1).

	MDR bacteria rectal colonization at T0	**Sepsis, *n* (%)**	**No Sepsis, *n* (%)**	**Total, *n* (%)**
Novara University Hospital	Colonized	1 (0.8)	1 (0.8)	2 (1.5)
	Non-colonized	21 (15.7) (9/21, 42.9%, septic at T0)	25 (18.7)	46 (34.3)
Alessandria University Hospital	Colonized	16 (11.9) (7/16, 43.8%, septic at T0)	19 (14.2)	35 (26.1)
	Non-colonized	5 (3.7) (4/5, 80%, septic at T0)	46 (34.3)	51 (38.1)
Total		43 (32.1)	91 (67.9)	134

Table 1: Study Population, MDR: multi-drug resistance.

From this cohort, 16 patients with confirmed sepsis were selected for a multi-omics study including RNA-Seq and 16S rRNA gut microbiome analysis, based on the availability of rectal swab samples across the scheduled timepoints, the type of microorganism responsible for sepsis (GN *vs* GP). Patients have been enrolled according to inclusion and exclusion criteria of the study including the absence of cancer, hematological malignancies, immunodeficiencies, or autoimmune pathologies. Except for two cases (sepsis caused by *Klebsiella pneumoniae* KPC and *Klebsiella pneumoniae* NDM), all episodes of sepsis were caused by antibiotic-sensitive bacteria. Four non-septic patients who did not develop sepsis during hospitalization were also included as controls. The demographic and clinical characteristics of these patients, including hospital unit, MDR colonization status at admission, and the microorganisms isolated from hemocultures, are summarized in Table 2.

Sample	Sex	Age (Years)	Unit	Sepsis, GP/GN	Col/Not Col	T0	MDR screening at admission	MDR screening during hospitalization	T1	T0 = T1b	MO isolated from hemocultures	T2
S8	F	65.53	RC	Yes, GN	Col	08/01/2023	COL1	*P. aeruginosa* & *K. pneumoniae*/ *K. pneumoniae* & *E. faecium*	11/01/2023	No	*S. marcescens*	24/01/2023
S14	M	80.93	RC	Yes, GN	Col	24/03/2023	COL1	*K. pneumoniae*/COL2/COL1/ *K. pneumoniae*	03/04/2023	No	*K. aerogenes*	17/04/2023
S5	M	81.88	RC	Yes, GN	Col	07/12/2022	COL1	COL1/COL2/*P. aeruginosa*	12/12/2022	No	*K. oxytoca*	14/12/2022
S35	M	68.78	SIC	Yes, GN	Col	01/11/2023	COL2	*K. pneumoniae NDM* (x6)	14/11/2023	No	*K. pneumoniae NDM*	20/11/2023
S34	M	76.96	SIC	Yes, GN	Col	01/11/2023	COL1	COL1 x 2/K. *pneumoniae NDM (*x4)	09/11/2023	No	*E. coli*	20/11/2023
S45 NO	M	48	RC	Yes, GN	Not Col	21/09/2023	COL2	COL2	30/09/2023	No	*K. pneumoniae*	02/10/2023
S16	M	67.98	RC	Yes, GN	Not Col	28/02/2023	COL1	COL1/COL2	02/03/2023	Yes	*E. coli*	10/03/2023
S2	F	75.91	RC	No	Not Col	03/12/2022	COL2	COL2	NA	No	NA	08/02/2022
S19	M	66.89	RC	No	Col	20/04/2023	*E. coli*	*E. coli*	NA	No	NA	24/04/2023
S24	M	62.38	INT MED	No	Col	14/07/2023	*E. coli*	*E. coli* & *E. faecium*	NA	No	NA	24/07/2023
S11	M	62.51	RC	No	Not Col	05/01/2023	COL2	COL1	NA	No	NA	11/01/2023
S45 AL	M	77.3	SIC	Yes, GN	Not Col	16/10/2023	COL1	COL2/COL2	16/10/2023	Yes	*E. coli*	23/10/2023
S28	F	83.57	SIC	Yes, GN	Col	16/10/2023	*K. pneumoniae KPC & P. aeruginosa*	COL1	16/10/2023	Yes	*K. pneumoniae KPC*	23/10/2023
S4	F	70.56	RC	Yes, GP	Col	04/12/2022	*E. coli*	*E. coli*	04/12/2022	Yes	*S. aureus*	06/12/2022
S13	M	75.09	RC	Yes, GP	Col	23/03/2023	COL2	*K. pneumoniae*	23/03/2023	Yes	*S. aureus*	27/03/2023
S9	M	73	RC	Yes, GP	Col	08/02/2023	*E. faecium*	*E. faecium*	10/02/2023	Yes	*E. faecalis*	17/02/2023
S10	M	54	RC	Yes, GP	Not Col	08/02/2023	COL2	COL2	18/02/2023	Yes	*E. faecalis*	13/03/2023
S23	M	69	RC	Yes, GP	Not Col	27/03/2023	COL2	COL2	28/03/2023	Yes	*E. faecalis*	28/03/2023

Table 2: Patient demographics correlated with unit, developing of sepsis and microorganisms isolated, timing of sampling (T0 = admission, T1 = microbiological diagnosis of sepsis, T2 = recovery), rectal colonization status at Hospital admission and hospitalization. GN, Gram-negative; GP, Gram-positive; Col, colonized; Not Col, non-colonized; M, male; F, female; RC, resuscitation care unit; SIC, sub-intensive care unit; INT MED, internal medicine unit; COL1, isolation of normal flora; COL2, absence of bacterial growth; MO, microorganism; NA, not applicable.

4.2 16S Metagenomics Study Population

The cohort was previously analyzed by RNA-seq for a differential gene expression (DEG) analysis to decipher longitudinal transcriptomic changes from sepsis onset and clinical recovery. Moving forward with the same cohort to the 16S metagenomic study, two patients were excluded from the analysis: S22, due to a mixed GP and GN bloodstream infection, and S21, whose rectal swab yielded Staphylococcus epidermidis, considered a possible environmental contaminant. The final cohort, therefore, comprised 14 patients. Of these, nine suffered from GN sepsis and five from GP sepsis. With respect to multidrug-resistant (MDR) bacterial gut colonization at hospital admission, nine patients were COL group and five were Non-COL group.

We considered two subgroups of patients according to their status at admission: the first group is composed of six patients (42.9%) with three time points and eight patients (57.1%) with two time points, thus including a total of 34 samples. In particular, the eight patients were septic at the time of hospital admission. Therefore, a T1a rectal swab was not collected, and was taken at the point of clinical suspicion of sepsis only for the six patients who were not septic when they were admitted in the hospital.

For this reason, the main longitudinal comparison in this analysis is between T0 (hospital admission) and T2 (recovery), with T1a data described where available to highlight transient changes during the acute manifestation of sepsis. This timepoint corresponds to the microbiological diagnosis of the disease by hemocultures.

Four gut microbiome parameters were assessed at each available timepoint: the Biodiversity Index, two alpha diversity indices (Shannon index and Chao1 index), and the *Bacillota-to-Bacteroidota* Ratio (B:B ratio). Individual values for each patient across timepoints are shown in Table 3.

Patient	Group	Biodiv T0	Biodiv T1	Biodiv T2	Shannon T0	Shannon T1	Shannon T2	Chao1 T0	Chao1 T1	Chao1 T2	B:B T0	B:B T1	B:B T2
S5	GN/COL	424	395	307	3.55	4.03	3.25	460	346	283	0.76	1.29	2.57
S8	GN/COL	252	225	200	3.53	3.74	3.14	283	260	248	0.87	0.53	0.58
S14	GN/COL	304	387	206	3.77	4.08	3.90	409	444	226	1.74	0.86	0.57
S34	GN/COL	114	230	220	2.53	2.78	3.39	118	292	306	1.50	19.95	0.54
S35	GN/COL	129	401	326	3.08	4.13	3.64	174	487	393	0.38	0.38	0.15
S45_NO	GN/Non-COL	282	282	265	3.77	3.47	3.54	358	323	303	5.89	1.10	0.47
S28	GN/COL	430	–	158	3.97	–	2.96	522	–	217	0.96	–	9.32
S16	GN/Non-COL	284	–	322	3.81	–	4.18	342	–	376	1.70	–	0.77
S45_AL	GN/Non-COL	391	–	200	3.72	–	3.44	540	–	234	1.11	–	1.02
S4	GP/COL	224	–	252	4.18	–	3.69	267	–	291	1.38	–	0.71
S9	GP/COL	173	–	230	2.75	–	3.27	213	–	277	52.41	–	0.36
S13	GP/COL	423	–	218	4.53	–	3.04	481	–	243	1.47	–	0.22
S10	GP/Non-COL	175	–	264	2.65	–	2.73	101	–	303	15.49	–	64.73
S23	GP/Non-COL	153	–	346	3.57	–	3.86	102	–	410	0.48	–	0.67

Table 3: Gut microbiome parameters at T0, T1a (where available), and T2 for each patient. COL = MDR colonized at admission; Non-COL = not colonized. A dash (–) indicates that no T1a sample was available, either because the patient was already septic at admission (T0 = T1b) or because the sample was not collected. Biodiv = Biodiversity Index; B:B = Bacillota / Bacteroidota ratio, GP = Septic with Gram-Positive bacteria; GN = Septic with Gram-Negative bacteria.

4.3. Sequencing Quality Control

By 16S metagenomics with NGS, we successfully analyzed all the 34 biological samples from the 14 patients of this cohort. The quality of the sequencing raw data was assessed by sequencing metrics and statistics, and the rarefaction curves (shown in table 4, figure 3). The Phred score (Quality Score 30) is a metric used in DNA sequencing to measure the accuracy of a base call. It indicates a 0.1% probability of an incorrect base call, meaning the sequencing accuracy is 99.9%[36].

Sequencing was carried out across two independent runs (October 2023 and February 2024). Good-quality read counts ranged from 25,172 to 221,407 per sample, with a median of 50,793 reads. All 34 samples exceeded the minimum threshold of 25,000 good-quality reads required by the DGPRE 0018191-P-15/06/2018 national guidelines for reliable 16S rRNA metagenomic results[59]. The rarefaction curves of all samples collected at each timepoint reached a plateau, confirming that sequencing depth was adequate to fully capture the microbial diversity of each sample (Figure 3).

To investigate changes in gut microbiota composition across the course of sepsis and to assess differences between GN and GP patients, samples were stratified according to the timepoint and clinical presentation at hospital admission. We also evaluate the possible relationships with the colonization status of the patients.

Sequencing Metrics			
	Cluster density (K/mm^2): 600 - 800	Cluster passing filter	Q-score ≥ 30
Run #1	695	59%	77.1%
Run #2	1191	47.2%	72.5%

Sequencing Statistics									
	Total Reads			Good Quality Reads (%)			Unclassified (%)		
	Min	Max	Median	Min	Max	Median	Min	Max	Median
All samples	25172	221407	50793	40.35%	91.75%	81.72%	1.54%	44.77%	10.24%

Table 4: Sequencing metrics and statistics.

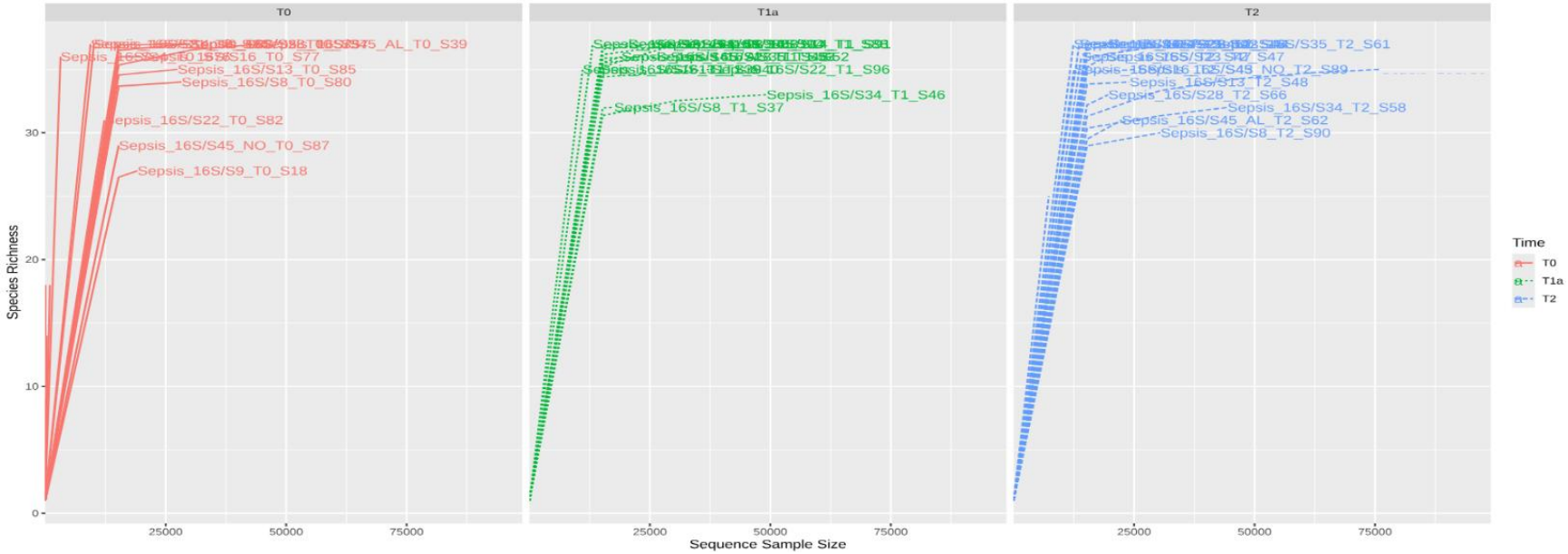


Figure 3: Rarefaction curves of the samples at each timepoint.

4.4 Gut Microbiome Community Profiling

The four non-septic patients (S2, S11, S19, and S24) did not develop sepsis during their hospital stay. They were originally selected for the concurrent transcriptomic study, and their gut microbiome profiles are reported here only for completeness. Each patient was sampled at admission (T0) and at discharge (T2), and the samples were sequenced and classified together with the others. Overall, these patients showed a mixed bacterial community without a single strongly dominant taxon. At the phylum level, the profiles were mainly dominated by *Bacteroidota* and *Bacillota*, together with a large fraction of *Unclassified_Bacteria*, and smaller proportions of *Actinomycetota* and *Pseudomonadota*. The most common genera were typical gut and mucosal bacteria, including *Prevotella* and other *Prevotellaceae*, *Porphyromonas*, and members of the *Bacteroidales*.

The following subsections describe the gut microbiome of the fourteen septic patients across the study timepoints.

4.4.1 Biodiversity Index

The Biodiversity Index indicates the total number of taxa captured in the gut community of each patient at the respective time point.

At T0, values ranged from 114 (S34, GN/COL) to 430 (S28, GN/COL). The patients with the lowest Chao1 at baseline also showed the lowest Biodiversity Index values: S34 (114), S35 (129), S23 (153), and S10 (175). In the GN group, six of nine patients (66.7%) showed a decrease from T0 to T2, with the largest drops in S28 (430 to 158), S45_AL (391 to 200), S14 (304 to 206), and S5 (424 to 307). Three GN patients showed substantial increases: S34 (114 to 220), S35 (129 to 326), and S16 (284 to 322), all starting from lower baseline values. In the GP group, four of five patients (80.0%) showed an increase by T2: S23 (153 to 346), S10 (175 to 264), S9 (173 to 230), and S4 (224 to 252). S13 (423 to 218) was the only exception in the GP subgroup.

4.4.2 Shannon Alpha Diversity Index

The Shannon index measures the alpha diversity (*i.e.* variety and balance of the microbiome within a single patient) of the gut microbial community by considering both the number of different species and how evenly they are distributed. Higher values indicate a more diverse and balanced gut community.

At T0, Shannon values ranged from 2.53 (S34, GN/COL) to 4.53 (S13, GP/COL), with most patients clustering between 3.5 and 4.2. Three patients entered the hospital with notably low diversity: S34 (2.53), S10 (2.65), and S9 (2.75). These values suggest a compromised baseline microbial ecosystem in a subset of patients, independent of the type of sepsis.

Between T0 and T2, Shannon diversity showed no consistent directional trend at the cohort level, with seven patients (50.0%) showing an increase and seven (50.0%) showing a decrease. In the GN group, five of nine patients (55.6%) showed a decrease by T2, the most pronounced being S28 (from 3.97 to 2.96). Four of nine GN patients (44.4%) showed an increase, with S34 (2.53 to 3.39) and S16 (3.81 to 4.18) showing the clearest gains. In the GP group, three of five patients (60.0%) showed an increase: S9 (2.75 to 3.27), S23 (3.57 to 3.86), and S10 (2.65 to 2.73, a marginal change). Two GP patients showed a decrease: S4 (4.18 to 3.69) and S13 (4.53 to 3.04). The drop in S13 represented the largest single decline observed across the whole cohort.

Where T1a data was available, several patients showed a transient increase in Shannon diversity during the acute phase that did not persist to T2. These included S35 (3.08 then 4.13 back to 3.64), S14 (3.77 to 4.08 back to 3.90), S5 (3.55 then 4.03 back to 3.25), and S8 (3.53 to 3.74 back to 3.14). This pattern may reflect a short-term reorganization of the gut microbiota during active infection.

4.4.3 Chao1 Species Richness Index

Chao1 estimates the total number of species present in the gut microbial community, with special sensitivity to rare taxa. Lower values indicate a community with fewer species overall. This index mirrors the Biodiversity index results closely across the cohort, as expected given the overlap in the type of information both measures capture.

At T0, Chao1 values ranged widely, from 101 (S10, GP/Non-COL) to 540 (S45_AL, GN/Non-COL). Several patients entered the hospital with low species richness: S10 (101), S23 (102), S34 (118), and S35 (174). At the other end of the range, S45_AL (540), S28 (522), and S5 (460) showed the highest baseline values.

Between T0 and T2, the GN and GP subgroups followed different patterns. In the GN group, six of nine patients (66.7%) showed a decrease in Chao1 by T2. The largest drops were observed in S45_AL (540 to 234), S28 (522 to 217), S14 (409 to 226), and S5 (460 to 283). Three GN patients showed an increase: S34 (118 to 306),

S35 (174 to 393), and S16 (342 to 376). All three had started from comparatively lower baseline values. In the GP group, four of five patients (80.0%) showed an increase in Chao1 at T2, most notably S23 (102 to 410) and S10 (101 to 303), followed by S9 (213 to 277) and S4 (267 to 291). S13 was the only GP patient with a clear decrease (481 to 243).

Across both subgroups, patients who entered the hospital with low Chao1 values at T0 tended to show the largest gains by T2, while patients with higher initial species richness more frequently showed declines. This observation may reflect a partial recovery of a previously depleted gut community in some patients, and the disruptive effect of antibiotics and hospitalization on a richer baseline microbiota in others.

Where T1a data were available, S35 showed a large transient Chao1 peak at T1a (487), well above its T0 value of 174, followed by a partial decrease to 393 at T2. S34 also showed a marked increase at T1a (292 vs. T0 = 118), and S14 a more modest one (444 vs. T0 = 409). These transient peaks during active sepsis are consistent with rapid and short-lived shifts in species composition at the time of infection.

4.4.4 Bacillota-to-Bacteroidota Ratio

The *Bacillota-to-Bacteroidota* Ratio (B:B ratio) reflects the balance between two major bacterial phyla in the gut. While a normal, healthy ratio generally fluctuates between 0.8 and 1.2, values above the healthy range indicate *Bacillota* predominance and are associated with fermentative dysbiosis, while values below the healthy range indicate *Bacteroidota* predominance and are associated with putrefactive dysbiosis[58].

At T0, the degree of dysbiosis varied considerably across the cohort, probably due to the status of the patients upon admission. In the GN group, most patients presented with mild dysbiosis at admission. S14 (1.74), S34 (1.50), S16 (1.70), and S45_AL (1.11) showed very mild fermentative dysbiosis, while S5 (0.76), S8 (0.87), and S28 (0.96) were slightly putrefactive. S45_NO (5.89) stood out with marked fermentative dysbiosis, and S35 (0.38) showed the most observed putrefactive dysbiosis within the GN subgroup. In the GP group, S9 (52.41) and S10 (15.49) presented with severe fermentative dysbiosis at admission, while S13 (1.47) and S4 (1.38) showed mild fermentative dysbiosis, and S23 (0.48) was the only sample with putrefactive dysbiosis in the GP group at the baseline.

Where T1a data was available, the acute phase of sepsis was linked to clear shifts in the degree of dysbiosis in several patients. S34 showed the most striking

change, moving from mild fermentative dysbiosis at T0 (1.50) to severe fermentative dysbiosis at T1a (19.95). This pattern suggests a sharp disruption of the phylum balance at the onset of sepsis. S5 showed a moderate worsening of fermentative dysbiosis (from 0.76 to 1.29), while S8 and S35 remained relatively stable at T1a (0.53 and 0.38, respectively). In contrast, S45_NO shifted from marked fermentative dysbiosis toward a near-balanced state (from 5.89 to 1.10) and this is the only patient which was not septic upon hospital admission and had no MDR bacteria colonization, and S14 moved slightly from a fermentative to a putrefactive state (from 1.74 to 0.86).

Between T1a and T2, or directly from T0 to T2 when T1a was not available, the trajectories diverged further across patients. In the GN group, S34 resolved from its extreme fermentative spike at T1a back to putrefactive dysbiosis at T2 (from 19.95 to 0.54). S14 moved further into putrefactive dysbiosis by T2. S45_NO shifted from a near balanced state to a slightly putrefactive state (0.47). S35 deepened its putrefactive dysbiosis between T0 and T2 (from 0.38 to 0.15). In contrast, S5 progressed into moderate fermentative dysbiosis by T2 (2.57). S28 showed the most marked deterioration within the GN group, shifting from a near-balanced state at T0 (0.96) to severe fermentative dysbiosis at T2 (9.32).

In the GP group, S9 resolved from extreme fermentative dysbiosis at T0 (52.41) to putrefactive dysbiosis at T2 (0.36), representing the largest improvement observed in the cohort. S13 shifted from mild fermentative to pronounced putrefactive dysbiosis (from 1.47 to 0.22), and S4 moved from mild fermentative toward mild putrefactive dysbiosis (from 1.38 to 0.71). S23 remained with mild putrefactive dysbiosis, with only minor variation (from 0.48 to 0.67). S10, however, followed the opposite trajectory. Starting from already severe fermentative dysbiosis at T0 (15.49), this patient worsened markedly to (64.73) at T2, the highest value recorded in the entire cohort across all timepoints.

Regarding the MDR colonization status, no consistent pattern was observed between COL and Non-COL patients in terms of B:B trajectory, with both groups showing substantial individual variability across timepoints.

4.5 Comparisons and Statistical Analyses

4.5.1 MicrobiomeAnalyst Comparison 1: Full Cohort Analysis at the Onset of Sepsis vs. Recovery

To assess microbial community changes between the onset of sepsis and recovery, and to compare GN and GP patients, MicrobiomeAnalyst was applied to all 14 patients using two metadata variables: time (onset of sepsis *vs* T2) and sepsis type (GN/GP). We considered the T0 withdrawal for patients who were septic at hospital admission and T1a for patients that developed sepsis during hospitalization. For simplification, the timepoints at the onset are labelled as T1 in the following figures and charts. A minimum count threshold of 15 was applied during data filtering. Five phyla, 27 genera, and 35 species were identified across all samples (Figure 4A – C).

Across the full cohort, five phyla were identified: *Bacillota*, *Bacteroidota*, *Pseudomonadota*, *Actinomycetota*, and *unclassified_Bacteria*. *Bacillota* and *Bacteroidota* were the two dominant classified phyla, with median relative abundances of 27.35% and 27.12%, respectively. *Pseudomonadota* was the third most abundant phylum (median 11.41%), followed by *Actinomycetota* (Figure 4A).

At the genus level, the most abundant genera were *unclassified_Bacteroidales*, followed by *unclassified Enterobacteriaceae*, *Phocaeicola*, *unclassified Oscillospiraceae*, *unclassified Lachnospiraceae*, *Finegoldia*, *Bacteroides*, *unclassified Bacillota*, *Parabacteroides*, *unclassified Prevotellaceae*, *unclassified Bacteroidaceae*, *unclassified Eubacteriales*, and *unclassified Enterobacterales*. Genera with a median relative abundance below 1% were grouped as “Others” (Figure 4B). The species-level composition across all four subgroups is shown in Figure 4C.

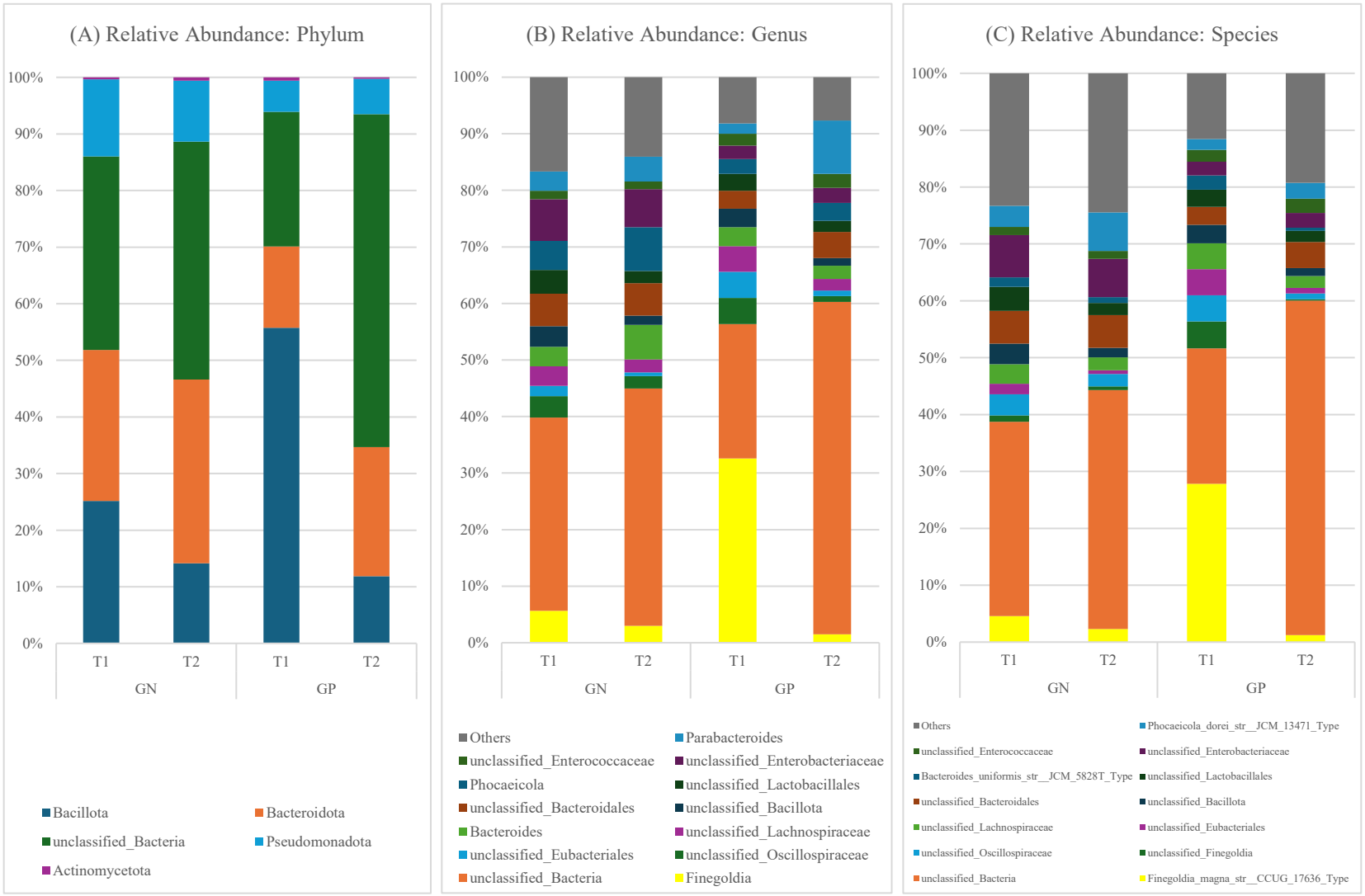


Figure 4: Relative abundance at the (A) phylum and (B) genera and (C) species. Stacked bar plot.

By considering the type of bacteria as an experimental factor, we found significant results. *Experimental factor, time:* Alpha-diversity analysis (Mann-Whitney/Kruskal-Wallis test) showed no significant differences between the onset of sepsis and recovery at any taxonomic rank for the Observed OTU, Shannon, Simpson or the Chao1 index.

Experimental factor, Sepsis Type: When comparing GN and GP subgroups, Observed OTUs and Chao1 index were significantly lower in GP than in GN patients at the class, order, family, and genus level (Figure 5A – B). Beta-diversity analysis (PERMANOVA) revealed no significant inter-sample compositional differences for either metadata variable at any taxonomic rank (*p*-values shown in Table 5).

Rank	Alpha diversity				Beta diversity		
	Observed OTUs	**Shannon**	**Simpson**	**Chao1**	**F-value**	**R-value**	**PCoA**
Phylum	0.70695	0.55543	0.46418	0.70695	1.6747	0.060515	0.166
Class	**0.017787**	0.52411	0.38164	**0.018914**	1.2368	0.045409	0.31
Order	**0.015109**	0.49367	0.3562	**0.013603**	1.2447	0.045687	0.299
Family	**0.022417**	0.19087	0.24492	**0.034754**	1.0674	0.039436	0.376
Genus	**0.029007**	0.098664	0.13274	**0.033008**	1.0935	0.040361	0.366
Species	0.073467	0.20785	0.17488	0.14442	1.0474	0.038726	0.376

Table 5: P-values for the sepsis type (GP or GN) as experimental factor.

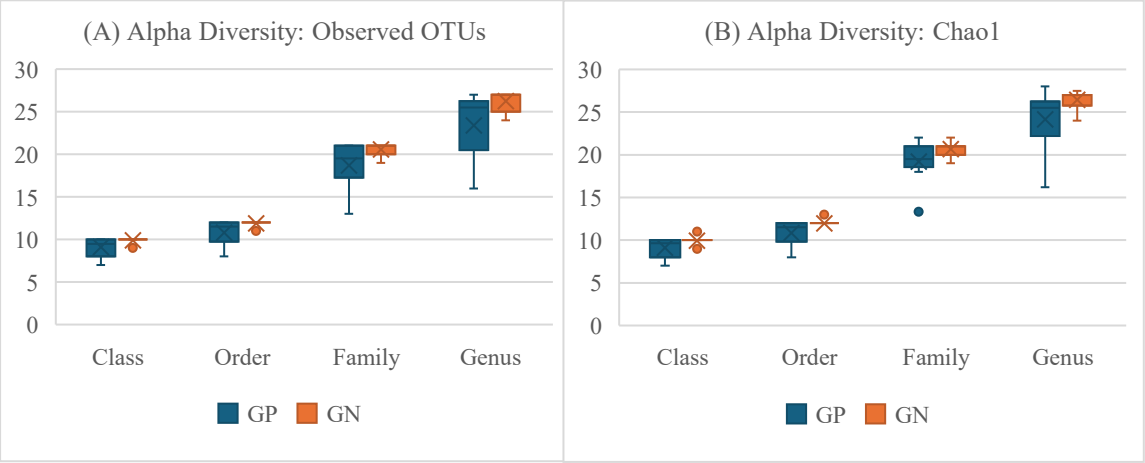
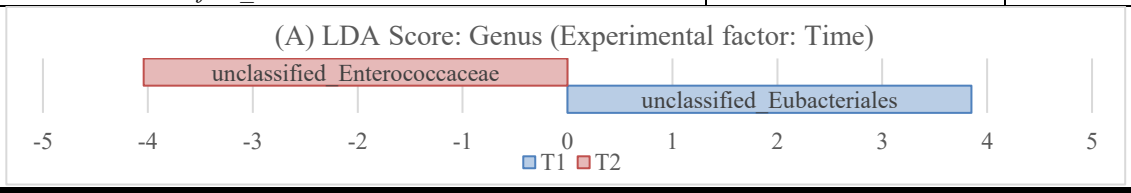


Figure 5: Alpha diversity analysis at various taxonomic ranks. (A) Observed index; (B) Chao1 index.

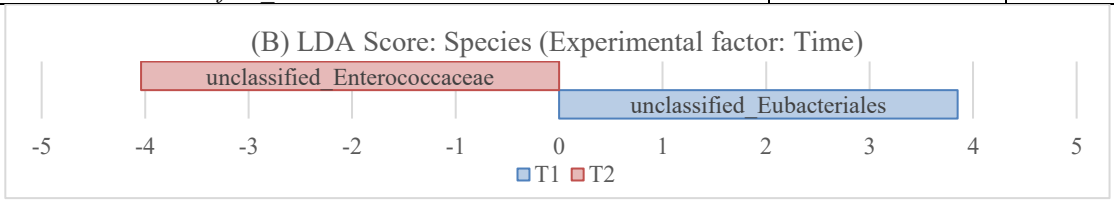
LEfSe analysis[60] using time as the experimental factor identified no FDR-adjusted significant features. At the nominal significance level ($p < 0.05$), eight features showed differential abundance across five taxonomic ranks: one class *(Clostridia, Bacillota)*, one order *(Eubacteriales, Bacillota)*, two families *(unclassified_Eubacteriales and Enterococcaceae, both Bacillota)*, two genera, and two species. At the genus level, *unclassified_Eubacteriales (Bacillota)* was nominally enriched at the onset of sepsis ($p = 0.013$, LDA score = 3.85), while *unclassified_Enterococcaceae (Bacillota)* was nominally enriched at recovery ($p = 0.027$, LDA

score = -4.04). These findings were reflected at the species level with identical significance values (Figure 6).

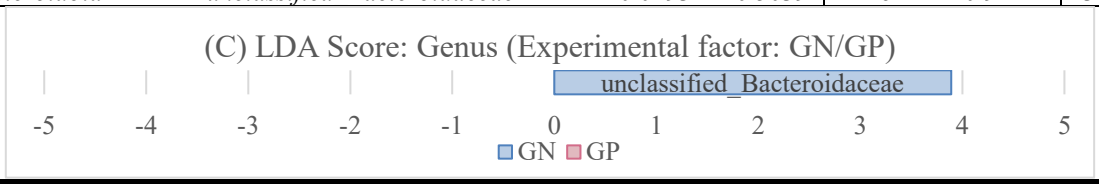
Phylum	Genus (Time)	*p* values	FDR	Rel Abundance (Median of Timepoints %)		LDA Score (A)
				T1	T2	
Bacillota	*unclassified_Eubacteriales*	0.013095	0.35357	3.21%	0.81%	3.85
	unclassified_Enterococcaceae	0.026998	0.36447	1.78%	1.93%	-4.04



Phylum	Species (Time)	*p* values	FDR	Rel Abundance (Median of Timepoints %)		LDA Score (B)
				T1	T2	
Bacillota	*unclassified_Eubacteriales*	0.013095	0.45833	3.21%	0.81%	3.85
	unclassified_Enterococcaceae	0.026998	0.47246	1.78%	1.93%	-4.04



Phylum	Genus (GP.GN)	*p* values	FDR	Rel Abundance (Median of Sepsis Type %)		LDA Score (C)
				GN	GP	
Bacteroidota	*unclassified_Bacteroidaceae*	0.04932	0.5039	2.46%	0.94%	3.89



Phylum	Species (GP.GN)	*p* values	FDR	Rel Abundance (Median of Sepsis Type %)		LDA Score (D)
				GN	GP	
Bacteroidota	*Parabacteroides_merdae_st__JCM_9497*	0.04929	0.4666	3.21%	0.81%	3.81
	unclassified_Bacteroidaceae	0.04932	0.4666	2.46%	0.94%	3.89

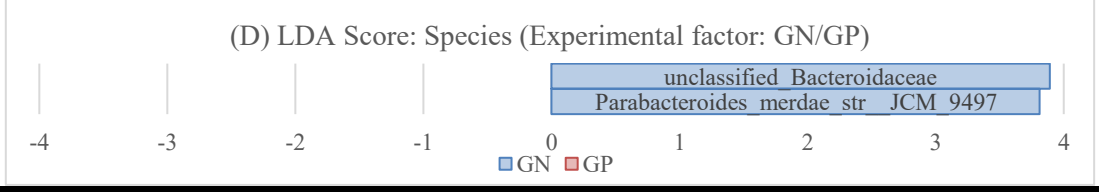


Figure 6: Tables for the p values accompanied by their respective histograms of the LDA scores computed for differential abundant taxa at T1 / GN (Blue) and T2 / GP (red) at the (A) Genus and (B) Species levels considering time as the experimental factor, (C) Genus and (D) Species considering sepsis type as the experimental factor. Abbreviations: GP, Gram positive; GN, Gram negative; Rel, Relative; str, strain.

Using sepsis type as the experimental factor, LEfSe identified no FDR-adjusted significant features. Three nominally significant features were detected: at the genus level,

unclassified_Bacteroidaceae (Bacteroidota) was enriched in GN compared to GP patients (p=0.049, LDA score = 3.89); at the species level, *Parabacteroides merdae (Bacteroidota)* and *unclassified_Bacteroidaceae (Bacteroidota)* were both enriched in GN patients (p = 0.049, LDA scores 3.81 and 3.89, respectively) (Figure 6).

Noticeably, relative abundances at the phylum level in both subgroups showed a consistent directional shift from the onset of sepsis to recovery. *Bacillota* decreased and *Bacteroidota* increased in both GN and GP patients, with the change more pronounced in the GP group (*Bacillota*: 55.75% to 11.85%; *Bacteroidota*: 14.37% to 22.82%) than in the GN group (*Bacillota*: 25.15% to 14.15%; *Bacteroidota*: 26.68% to 32.48%). These trends are consistent with the individual B:B ratio trajectories reported in section 4.4.4. At the genus level, *Finegoldia* showed a marked reduction in relative abundance from the onset of sepsis to recovery in both subgroups, particularly in GP patients (32.57% to 1.47%) compared to GN patients (5.66% to 2.98%). This change did not reach statistical significance in LEfSe analysis but was consistent across both groups and may be of biological interest given the known association of *Finegoldia magna* with opportunistic infections[61,62]. At the species level, this reduction was driven primarily by *Finegoldia magna (str. CCUG 17636)*, which decreased from 27.85% to 1.22% in GP patients and from 4.56% to 2.31% in GN patients (Figure 4C).

4.5.2 MicrobiomeAnalyst Comparison 2: Non-Septic patients at Admission (T0 vs. T1 vs. T2)

A three-timepoint longitudinal analysis was performed on the six patients who were not septic at hospital admission and had samples available at T0 (admission), T1 (onset of sepsis), and T2 (recovery): S5, S8, S14, S34, S35, and S45_NO, all of whom had GN sepsis. A minimum count threshold of 15 was applied during data filtering. Five phyla, 25 genera, and 35 species were identified across these samples.

Neither alpha-diversity (Observed OTUs and Chao1, Mann-Whitney/Kruskal-Wallis test) nor beta-diversity (PERMANOVA) analysis showed statistically significant differences across the three timepoints at any taxonomic rank. LEfSe analysis identified no significant features at either the FDR-adjusted or nominal significance level. The absence of statistically detectable microbial shifts in this subgroup is likely related in part to the small sample size (n = 6) and the variable time interval between hospital admission (T0) and the onset of sepsis (T1) across patients, and the implications of these conditions will be addressed in the Discussion section. It is worth noting that all six patients in this comparison had GN sepsis, meaning the

GP subgroup in comparison 1 was composed mostly of patients who were already septic at admission.

4.5.3 MicrobiomeAnalyst Comparison 3: Septic Patients at Admission (T0 vs. T2)

For the eight patients who were already septic at hospital admission (S4, S9, S10, S13, S16, S23, S28, and S45_AL), a two-timepoint comparison between T0 (septic at admission) and T2 (recovery) was performed. This subgroup included five GP and three GN patients. A minimum count threshold of 15 was applied during data filtering. Five phyla, 23 genera, and 31 species were identified across these samples. The relative abundances at the phylum, genus, and species levels for this subgroup are shown in Figure 7 A – C.

At the phylum level, five phyla were detected in this subgroup: *Bacillota*, *Bacteroidota*, *Pseudomonadota*, *Actinomycetota*, and a considerable fraction of *Unclassified_Bacteria* (Figure 7). At admission (T0), *Bacillota* was the predominant phylum in most patients, in line with the mainly GP composition of this subgroup, whereas at recovery (T2) its relative abundance decreased, accompanied by an increase in *Bacteroidota* and in the unclassified fraction. This phylum-level shift from *Bacillota* toward *Bacteroidota* between admission and recovery mirrors the trajectory observed in the full cohort (Comparison 1) and is consistent with the B:B ratio dynamics described in section 4.4.4.

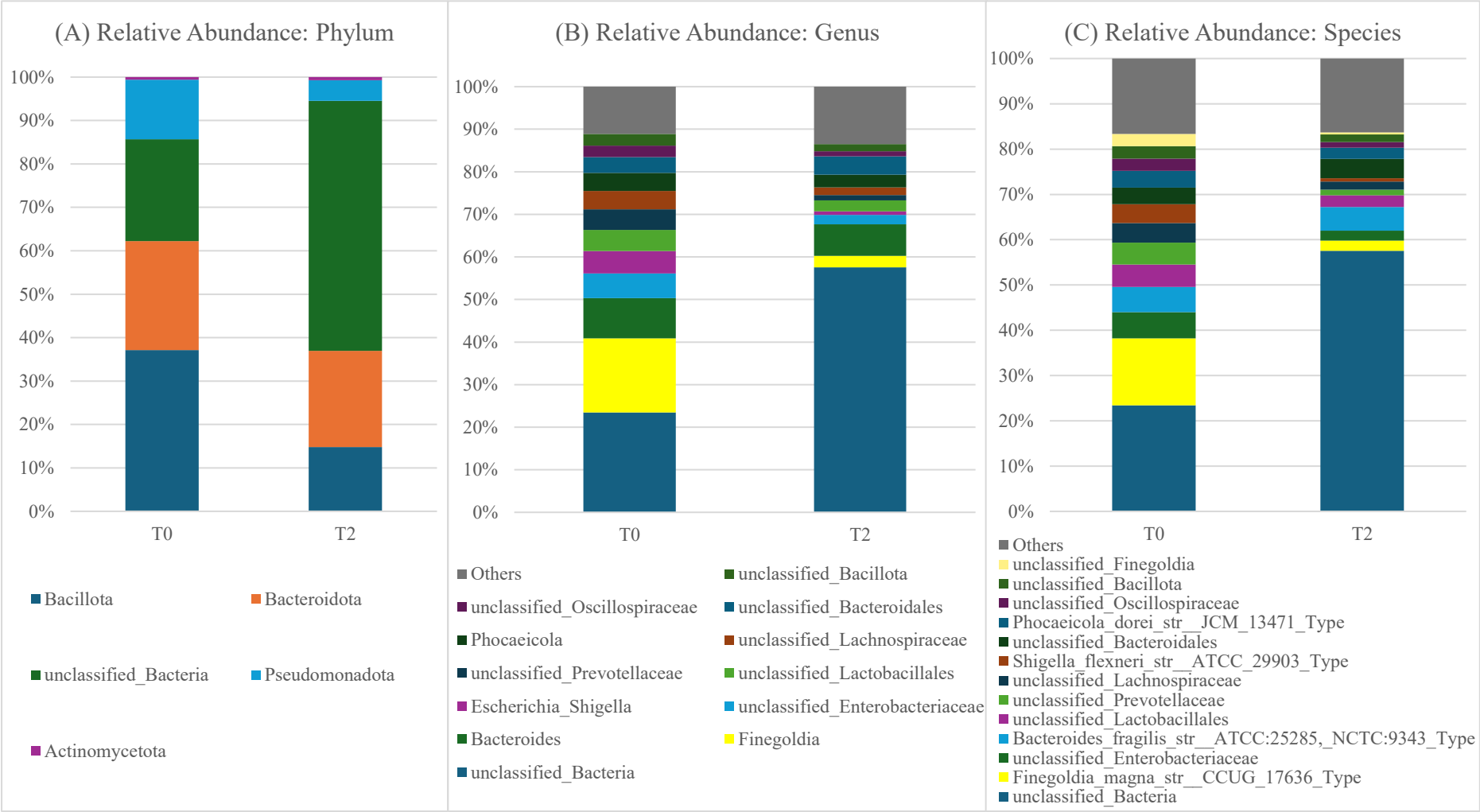
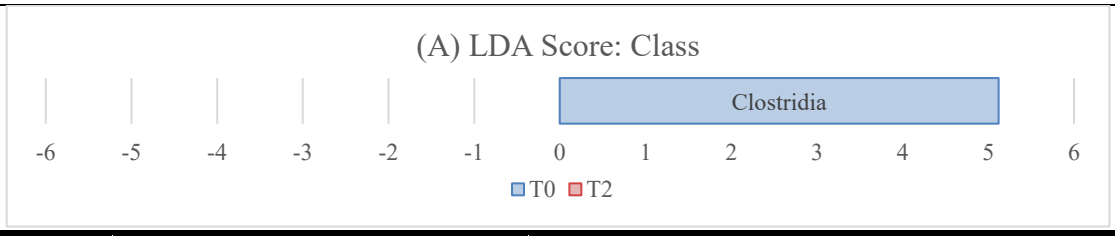


Figure 7: Relative abundance at the (A) phylum, (B) genus, and (C) species levels. Stacked bar plot.

Alpha-diversity (Observed OTUs and Chao1) and beta-diversity (PERMANOVA) analyses showed no statistically significant differences between T0 and T2 at any taxonomic rank. LEfSe analysis identified no FDR-adjusted significant features. At the nominal significance level, two features showed differential abundance: *Clostridia* (class, *Bacillota*) was enriched at T0 relative to T2 (p = 0.012, LDA score = 5.12), and *Eubacteriales* (order, *Bacillota*) showed the same pattern at the order rank (p = 0.012, LDA score = 5.12) (Figure 8). These two features represent the same biological signal at different taxonomic levels, pointing to a nominally higher relative abundance of organisms of the class *Clostridia*, order *Eubacteriales* at admission compared to recovery. This finding is consistent with the B:B ratio trajectories described in section 4.4.4 for this subgroup.

Phylum	Class	Pvalues	FDR	Rel Abundance (%)		LDA Score (A)
				T0	T2	
Bacillota	*Clostridia*	0.011719	0.10547	26.98%	6.62%	5.12



Phylum	Order	Pvalues	FDR	Rel Abundance (%)		LDA Score (B)
				T0	T2	
Bacillota	*Eubacteriales*	0.011719	0.11719	26.98%	6.62%	5.12

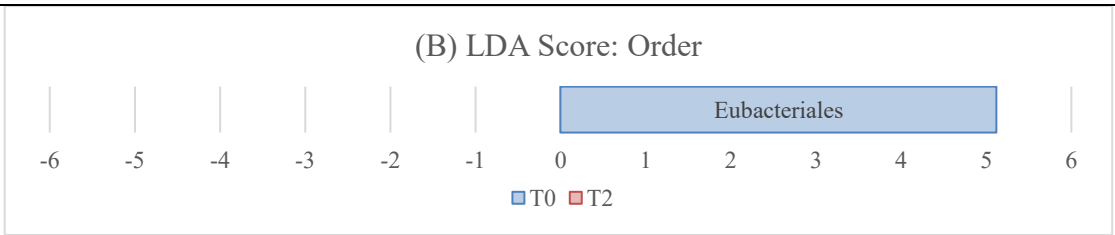


Figure 8: Tables for the p-values accompanied by their respective histograms of the LDA scores computed for differential abundant taxa at T0 (Blue) at (A) the Class and (B) Order levels. Abbreviations: Rel, Relative.

At the genus and species levels, this signal corresponded to a decrease in the class *Clostridia* and order *Eubacteriales* from 26.98% to 6.62% between admission and recovery, driven mainly by the genus *Finegoldia* (from 17.44% to 2.73%), represented largely by *Finegoldia magna str. CCUG 17636* with a small fraction of *unclassified_Finegoldia* (Figure 7B & 7C, yellow fraction). Because *Finegoldia* is classified within the order *Eubacteriales* and class *Clostridia*, its reduction is reflected at these higher ranks. The same trend was observed in the full-cohort analysis (Comparison 1), indicating that this decline from the acute phase to recovery was recurrent across groups, although it reached only nominal significance and did not survive the FDR adjustment.

5. Discussion

This study represents the metagenomic component of the SURVEIL project, a prospective, longitudinal, multi-omics investigation of the gut microbiome in critically ill adults, whose overall aim was to identify gut microbiota features associated with, and potentially predictive of, the development of sepsis, including colonization by multidrug-resistant (MDR) organisms[55]. Within this framework, the specific aim of the present work was to describe the composition and the longitudinal dynamics of the gut microbiome in septic patients by 16S rRNA gene sequencing and to identify possible relationships with the bacterial type responsible for sepsis and with the colonization status with respect to MDR bacteria, across the course of the disease from admission to recovery. The findings discussed below are therefore interpreted both on their own terms and against the background of the larger SURVEIL cohort, so that the focused, longitudinal picture obtained here can be placed in the context of the broader study from which it derives.

The samples analyzed in this study were a deliberately selected subset of the larger SURVEIL study cohort, which comprised more than one hundred and thirty critically ill adults consecutively enrolled at two University Hospitals of Eastern Piedmont, of whom forty-three developed microbiologically confirmed sepsis[55]. From these septic patients, sixteen of the most informative cases were selected for a parallel transcriptomic study based on three criteria: the availability of complete biological sample withdrawals at all three timepoints (admission, sepsis onset or confirmation, and discharge), the type of microorganism responsible for sepsis (GN *vs* GP), and the MDR rectal colonization status at admission. Two of these cases were subsequently excluded from the bioinformatic analysis of this study, one in which sepsis was attributed to *Staphylococcus epidermidis*, a skin contaminant, and one with a polymicrobial bloodstream infection sustained by both GN and GP organisms, leaving fourteen septic patients (nine GN and five GP). For these patients, a total of 34 rectal-swab samples spanning admission, sepsis onset, and recovery were sequenced on the Illumina MiSeq platform across the V3, V4, and V6 hypervariable regions of the 16S rRNA gene; all 34 samples passed primary and secondary bioinformatic analysis with optimal sequencing metrics (median 50,793 reads per sample), providing a robust basis for the comparisons discussed below.

Before turning to the microbiome data, it is useful to recapitulate the microbiological basis of sepsis in this cohort. Sepsis was defined according to the Sepsis-3 consensus criteria, and bacteremia was confirmed at the onset of sepsis by direct microbiological identification on blood, using Gram staining, FilmArray, direct MALDI-TOF mass spectrometry, and the T2Dx

system, after exclusion of organisms regarded as contaminants[55]. The lung was the most frequent primary site of infection. Among the fourteen septic patients, bacteremia was exclusively monomicrobial: GN organisms accounted for nine cases (64.3%) and GP organisms for five cases (35.7%). The GN infections were caused most often by *Klebsiella* species (five cases), followed by *Escherichia coli* (three cases) and a single case of *Serratia marcescens*, while the GP infections were due to *Enterococcus faecalis* (three cases) and *Staphylococcus aureus* (two cases). Almost all strains were antibiotic-susceptible, the two exceptions being carbapenemase-producing *Klebsiella pneumoniae*, one of the KPC type and one of the NDM type.

The choice of rectal swabs as the specimen for the metagenomic profiling followed directly from the design of the study: rectal swabs are routinely collected twice weekly for the microbiological surveillance of MDR organisms in hospitalized patients, and were therefore available at all three timepoints, which made them the natural and most consistent specimen for a longitudinal analysis of the gut microbiota.

Taken together, the metagenomics results describe a gut community that is markedly disturbed during the acute phase of sepsis and that undergoes only a partial and often incomplete reorganization by the time of discharge.

At the most general level, the septic gut community was built mainly on two phyla, *Bacillota* and *Bacteroidota*, which together dominated the classified fraction at comparable median abundances, with smaller contributions from *Pseudomonadota* and *Actinomycetota* and a variable, often considerable fraction of reads remains unclassified. Across the cohort, the phylum-level composition showed a consistent directional change from the onset of sepsis to recovery, with *Bacillota* decreasing and *Bacteroidota* increasing in both sepsis types, a shift that was more pronounced in the GP patients and was echoed in the B:B ratio. This overall trajectory, however, concealed substantial differences between individuals: within-patient diversity and the B:B ratio varied widely from one patient to another, and clinical recovery was not always accompanied by a return toward a more balanced gut community, with some patients remaining, or becoming, more imbalanced. Against this heterogeneous background, one signal proved consistent across all three comparisons and at every taxonomic level examined, and it is considered in detail in the following paragraphs.

The most consistent finding across all three comparisons is the expansion of the GP anaerobic coccus *Finegoldia magna* during the acute phase, followed by its pronounced reduction at recovery, a pattern that was particularly evident in patients with GP sepsis and that

was detectable at several taxonomic levels. This central observation is discussed below and, where possible, compared with the corresponding observations in the full SURVEIL cohort.

The reduction of *Finegoldia magna* from the onset of sepsis to recovery was captured by the different comparisons at different taxonomic resolutions. In the full-cohort analysis (Comparison 1), the genus *Finegoldia* decreased from 32.57% to 1.47% in GP patients and from 5.66% to 2.98% in GN patients, a change driven almost entirely by *Finegoldia magna*, which fell from 27.85% to 1.22% in the GP group and from 4.56% to 2.31% in the GN group. Although this shift did not survive correction for multiple testing across the comparison, it was consistent in both subgroups and dominated the genus- and species-level composition. The same signal reached nominal significance when the patients who were already septic at admission were analyzed separately (Comparison 3): in this GP subgroup, the class *Clostridia* and the order *Eubacteriales* were enriched at admission relative to recovery (26.98% versus 6.62% for both taxa, LDA score 5.12), and, because the genus *Finegoldia* fell from 17.44% to 2.73% in the same subgroup and is classified within these higher ranks, it was by far the largest contributor to this class- and order-level signal. Since no significant features were found in the species level for this comparison, it could indicate that grouping data into broader taxonomic ranks like genus or family levels aggregates biological signals from the species level, making it easier to detect significant trends by filtering out noise from rare taxa. This observation is in close agreement with the findings of the larger SURVEIL cohort, in which sepsis was associated with a profound dysbiosis characterized by the enrichment of opportunistic, mostly anaerobic GP genera, including *Finegoldia*, *Anaerococcus*, and *Peptoniphilus*, and in which the enrichment of *Finegoldia* in COL but non-septic patients at admission was proposed as an early marker of microbial destabilization and sepsis risk[55]. This analysis adds to the previous observation a species- and strain-level identification and a longitudinal trajectory of *Finegoldia magna str. CCUG 17636*, accompanied by a tiny fraction of unclassified *Finegoldia*, showing not only the bloom of this organism during the acute phase but also its clearance upon discharge from hospital.

This trajectory is biologically plausible. *Finegoldia magna* is a GP anaerobic coccus and a normal commensal of the skin and of the mucosal surfaces of the human body, including the gastrointestinal tract, but it is also a recognized opportunistic pathogen that can cause soft tissue, bone, joint, and device-associated infections, particularly in wounded or immunocompromised hosts[62], and it has occasionally been associated with severe presentations such as toxic shock syndrome[61]. The expansion of such an organism during the

acute phase of sepsis, followed by its contraction during recovery, is consistent with the broader view of the gut as a reservoir of opportunistic resident microbes that can bloom under the selective pressures of critical illness, antibiotic exposure, and altered host physiology[40,41,42,43]. Because the present analysis is observational and based on relative abundances, it cannot establish whether the rise in *Finegoldia magna* contributes to the pathophysiology of sepsis or simply reflects the disturbed intestinal environment of acutely ill patients. Nevertheless, the consistency of the pattern across taxonomic levels and across both sepsis types, together with the independent identification of *Finegoldia* as an early dysbiotic signal in the full cohort[55], makes *Finegoldia magna* a credible candidate marker of the acute phase of gut dysbiosis in sepsis.

The phylum-level composition and the (B:B) ratio describe the same phenomenon from a complementary angle. From the onset of sepsis to recovery, *Bacillota* decreased and *Bacteroidota* increased in both subgroups, with a much larger change in GP patients (*Bacillota* from 55.75% to 11.85%; *Bacteroidota* from 14.37% to 22.82%) than in GN patients (*Bacillota* from 25.15% to 14.15%; *Bacteroidota* from 26.68% to 32.48%). Because *Finegoldia* belongs to the phylum *Bacillota*, the high *Bacillota* content and the elevated B:B ratio seen in several GP patients at the septic phase are largely explained by the *Finegoldia magna* increase, while the fall in *Bacillota* and in the B:B ratio at recovery reflects its clearance. At the level of individual patients, however, the B:B trajectories were highly variable and moved in both directions, in agreement with the large interindividual variation reported for the gut microbiota of critically ill patients[44]. Patient S9, for example, moved from severe fermentative dysbiosis at admission (B:B = 52.41) to a *Bacteroidota*-dominant putrefactive state at recovery (0.36), the largest improvement in the cohort, whereas patient S10 followed the opposite course, worsening from an already high value (15.49) to the highest value recorded in the entire dataset (64.73) when discharged from hospital. This dissociation indicates that clinical recovery does not necessarily coincide with restoration of the gut microbial balance, and that the intestinal community can remain or even become more dysbiotic while the patient improves clinically. A shift toward *Bacteroidota* predominance at recovery, associated with a low B:B ratio and a putrefactive type of dysbiosis[58], should therefore not be read as a full return to a healthy configuration, but rather as a change in the type of imbalance.

The four non-septic patients described earlier, who were hospitalized but never developed sepsis, also allowed the role of *Finegoldia* to be examined outside the septic condition. In these patients the genus *Finegoldia* remained low and stable between admission

(T0) and discharge (T2) (2.97% and 2.93%), and *Finegoldia magna* stayed close to 2.0% at the species level. This contrasts with the septic patients, in whom *Finegoldia magna* expanded strongly during the acute phase before declining at recovery, indicating that the bloom of this organism is tied to the septic episode itself rather than to hospitalization alone. Consistently, their B:B ratio was almost unchanged across the two timepoints (0.85 and 0.84), whereas the septic patients showed a marked shift over a comparable interval, reinforcing that the changes seen in sepsis are not simply a consequence of being in hospital. Since these patients were few and not the focus of this work, this comparison is reported descriptively rather than as a separate figure.

With respect to within-sample diversity, the comparison between sepsis types showed that the number of observed taxa and the Chao1 richness estimate are significantly lower in GP than in GN patients at the class, order, family, and genus levels, while the Shannon and Simpson indices, which also account for evenness, do not differ significantly at any rank. GP communities therefore contained fewer taxa overall, a difference that is at least partly mechanical, since the dominance of a single organism such as *Finegoldia magna* necessarily reduces the apparent richness of the community. This difference must, however, be interpreted with caution, because of a structural feature of the cohort: all six patients who were not yet septic at admission, and who therefore provided the only true pre-sepsis baseline, had GN sepsis after their admission to the hospital, so that the GP subgroup consisted almost entirely of patients who entered the hospital already septic at admission and were sampled at a more acute point in their illness. The lower richness observed in GP patients thus partly reflects the overlap between GP sepsis and a more acute clinical presentation, rather than an effect of the infecting organism alone. A further possibility, which is more speculative and cannot be tested with the present data, is that some of the GP patients were experiencing a more complicated or recurrent septic status rather than a first episode. The markedly depleted and opportunistic microbe-dominated communities observed in this subgroup are characteristic of the loss of colonization resistance that follows repeated infection and prolonged antibiotic exposure, a state that has itself been associated with an increased risk of subsequent, and potentially recurrent, bloodstream infection[47,48,49]. Because the clinical history of previous septic episodes was not examined in this work, this explanation remains a hypothesis to be addressed in future studies. A loss of microbial diversity is nonetheless a well-documented hallmark of critical illness and of the antibiotic exposure that accompanies it[43,44,49,51], and the same direction of effect was reported in the full SURVEIL cohort, in which COL patients, and particularly those

harboring GP taxa, showed a significantly reduced alpha diversity compared with Non-Col individuals[55]. By contrast, neither the time factor (onset versus discharge) nor beta-diversity analysis revealed significant differences at any taxonomic rank in the present subset, which suggests that the changes observed here were driven by a small number of taxa rather than by a whole restructuring of the community, and that they were partly masked by the high interindividual variability and the limited sample size.

A specific consideration concerns the patients with GN sepsis, who made up nine of the fourteen cases and whose infections were caused mainly by *Enterobacterales* such as *Klebsiella pneumoniae* and *Escherichia coli*. On clinical grounds, GN sepsis is often expected to carry a heavy inflammatory burden, since the lipopolysaccharide of the GN cell wall is a potent trigger of the innate immune response. The gut profiling obtained here, however, did not translate this expectation into a more severely disrupted community: the GN subgroup retained a more diverse and commensal-rich gut than the GP subgroup, with higher observed richness, preservation of *Bacteroidota* members such as *Parabacteroides merdae* and *unclassified_Bacteroidaceae*, a less pronounced shift from *Bacillota* to *Bacteroidota*, and only a modest *Finegoldia magna* signal in place of the marked bloom seen in GP patients. The dominant acute dysbiotic signal in this sepsis cohort was therefore GP, in contrast to the GN, *Enterobacteriaceae*-dominated dysbiosis that characterizes many general critical-illness cohorts[49]. This apparent paradox is most plausibly explained by the structure of the cohort rather than by a true biological difference, since the GN patients included all of those who were not yet septic at admission and were therefore sampled at an earlier and less acute point in their illness, whereas the GP patients were almost all sampled when already septic at admission, meaning that the gut data alone cannot confirm the greater clinical severity expected of GN sepsis, and leading to whether the GN host mounts a stronger systemic response despite a comparatively preserved gut community is precisely the kind of question that the parallel transcriptomic study of the SURVEIL project is positioned to answer.

The longitudinal analysis of the six patients who were not septic at admission (Comparison 2) did not reveal any statistically significant change in diversity or in differential abundance across the three timepoints. This null result is most likely explained by the small size of this subgroup and by the variable interval between hospital admission and the onset of sepsis, which makes the septic timepoint less homogeneous across patients. It is nonetheless informative, because this is the only subgroup for which a sample was available before the clinical onset of sepsis, and it is precisely in this pre-septic, COL window that the full

SURVEIL cohort located the early *Finegoldia* signal[55]. Several of these patients showed a transient increase in the Shannon and Chao1 indices at the acute phase that did not persist to recovery, a pattern compatible with a brief reorganization of the community during active infection, although the small number of patients means that this observation should be regarded as exploratory.

The comparison between sepsis types also identified taxa that were relatively preserved in GN patients. At the genus level, unclassified *Bacteroidaceae* was enriched in GN compared with GP patients ($p = 0.049$, LDA score 3.89), and at the species level *Parabacteroides merdae* and *unclassified_Bacteroidaceae* were both enriched in GN patients (both $p = 0.049$, LDA scores 3.81 and 3.89, respectively) between T1 and T2.

Members of the genera *Bacteroides* and *Parabacteroides* are common commensals of the healthy gut that contribute to fermentation and to colonization resistance against incoming pathogens[25,26,28], and *Parabacteroides* was likewise among the discriminant taxa identified in the full SURVEIL cohort[55]. Their relative preservation in GN patients, set against the *Bacillota*- and *Finegoldia*-dominated communities of GP patients, is consistent with the interpretation that the GP subgroup, sampled at a more acute phase, was the more severely dysbiotic of the two.

Two features of the recovery samples deserve particular comment. First, a large proportion of the reads at recovery could not be assigned beyond the *Unclassified_Bacteria* domain (approximately 42% in GN and 59% in GP patients at the phylum level). This unclassified fraction may indicate a genuine shift toward taxa that are poorly represented in the reference database, but it may also reflect the well-known limitations of 16S rRNA gene sequencing, as discussed in the introduction, and it should therefore be considered when interpreting the recovery communities. Second, unclassified *Enterococcaceae* showed a nominal enrichment at recovery in the full cohort. Although the magnitude of this change was small, its direction is consistent both with the well-documented expansion of *Enterococcus* and related organisms in the gut of patients exposed to broad-spectrum antibiotics during hospitalization[46,47,48] and, more specifically, with the full SURVEIL cohort, in which *Enterococcus hirae* and *Enterococcus faecium* were among the taxa enriched at the recovery timepoint in COL septic patients[55]. The convergence of the two analyses on an *Enterococcus* signal at recovery is notable, even if the effect size in the present subset was modest.

A further observation concerns the starting condition of the patients. A subset entered the hospital with an already compromised gut community, reflected in low Shannon and Chao1

values, independently of the type of sepsis, and across the cohort the patients who started with a low species richness tended to show the largest gains by recovery, whereas those who started from a richer community more often showed declines. This pattern is partly a statistical effect, since extreme baseline values tend to move toward the average over time, but it is also compatible with a partial restoration of previously depleted communities in some patients and with the disruptive effect of antibiotics and hospitalization on richer communities in others; a depleted gut microbiota at baseline has itself been associated with increased susceptibility to infection and with worse outcomes in hospitalized patients[48,49].

With respect to MDR colonization, no consistent relationship was found between the colonization status at admission (and during hospitalization) and the microbial diversity or B:B ratio trajectories, with both COL and Non-COL patients showing substantial individual variability, a result that is likely related to the small number of patients in each group and to the predominance of the acute-sepsis signal. This does not diminish the clinical importance of MDR gut colonization, which the full SURVEIL cohort linked to bloodstream infection: in that study, a direct correspondence between an MDR strain carried in the gut and the pathogen isolated from the blood was documented in only two cases, both carbapenemase-producing *Klebsiella pneumoniae*, pointing to occasional translocation of organisms from the intestine to the bloodstream[55], in line with the long-standing view of the gut as a source of infection in critical illness[42]. In both of these cases, the carbapenemase-producing *Klebsiella pneumoniae* found in the blood had already been detected in the gut by the twice-weekly surveillance swabs. This suggests that the gut was the source of the infection. In the other patients, the organism in the blood was not the same as the one in the gut, which fits with the lung being the main site of infection. Carrying a resistant strain in the gut did not always mean it moved into the blood: one patient who was colonized by NDM-producing *Klebsiella pneumoniae* developed an *Escherichia coli* infection in the blood instead.

The fact that gut and bloodstream organisms matched in only two cases fits a broader rethinking of how the gut drives infection in critical illness. In a prospective, longitudinal study combining serial 16S rRNA gene sequencing of rectal swabs with single-cell profiling of systemic immunity, Schlechte and colleagues described the gut microbiota and the immune system as a single interacting metasystem, in which gut dysbiosis during the acute phase was linked to a weakened, poorly functioning innate immune response and to more frequent nosocomial infections[49]. Importantly, in that study the increase of opportunistic gut bacteria was not associated with infection by those same organisms, suggesting that dysbiosis raises the

risk of infection more through its effect on host defense than through direct translocation. This view is consistent with the present findings: the main signal in the acute phase was not a bloodstream pathogen carried over from the gut, but the expansion of *Finegoldia magna* together with reduced richness and weaker colonization resistance, the kind of community-level damage that the metasystem framework links to impaired immunity rather than to translocation alone. The gut profile described in this thesis therefore reaches its full meaning only when read together with the host response in the same SURVEIL patients, which is exactly the reason for the project's multi-omics design.

Antibiotic exposure is an important factor to consider when interpreting these findings. All patients received broad-spectrum antibiotics as part of the standard treatment for sepsis. Some of the changes described here, the bloom of *Finegoldia magna*, the rise of *Enterococcus* at recovery, and the overall loss of diversity, are also known effects of antibiotics on the gut microbiota. Because the antibiotics given to each patient were not systematically recorded and included in this analysis, the effect of sepsis cannot be fully separated from the effect of its treatment. The dysbiosis seen across the cohort should therefore be read as the combined result of sepsis and the antibiotics used to treat it, and not as an effect of sepsis alone.

A further point concerns what "recovery" means in this study. The recovery timepoint followed the clinical discharge schedule and fell at a two-week window, rather than at a clinically confirmed return to a pre-sepsis state, and no later samples were collected and partially followed-up. The partial recovery of the gut community seen at this timepoint should therefore be read as a snapshot taken at an early, fixed point after sepsis. Whether the patients whose communities had not yet recovered would have improved later, or stayed altered for longer, cannot be answered with these data.

6. Conclusion

This thesis characterized the gut microbiome of critically ill patients across the course of sepsis by 16S rRNA gene sequencing on an Illumina MiSeq platform, as the metagenomic component of the SURVEIL project. The gut microbiome of fourteen septic patients was profiled across 34 rectal-swab samples collected from admission to recovery. Samples have been successfully sequenced with optimal quality metrics and processed through primary and secondary bioinformatic analysis. The gut microbial community was markedly disturbed during the acute phase and showed only partial and heterogeneous bacterial remodeling by recovery. The most consistent finding was the bloom of the opportunistic GP anaerobic coccus *Finegoldia magna* at the acute phase and its pronounced reduction at recovery, a signal detectable from the species level up to the class and phylum levels, mirrored in the B:B ratio, and most marked in patients with GP sepsis. These patients showed lower microbial richness, partly explained by their earlier and more acute sampling, and their individual trajectories were highly variable, occasionally worsening despite clinical improvement. *Finegoldia* and the recovery-associated *Enterococcus* signals were concordant with those of the full SURVEIL cohort[55].

These results should be regarded as exploratory and hypothesis-generating. The analyzed subset was small (14 septic patients), so most changes reached only nominal significance, and sepsis type was partly confounded with sampling timing. Recovery was assessed at a fixed two-week window rather than at a true endpoint of resolution, and no longer-term follow-up was available. The unclassified fraction may limit species-level interpretation, and the observational, relative-abundance design cannot establish causal relationships or resolve bacterial function, which would require shotgun metagenomic sequencing.

Within the SURVEIL project, this metagenomic profiling provides complementary information with a multi-omics perspective to the concurrent analysis of the host transcriptional response, and the integration of the two layers through machine learning is the natural subsequent step. Combined host-microbe signatures, analyzed with classifiers suited to the high-dimensional and compositional nature of microbiome data, have recently achieved reliable diagnosis of sepsis[63,64], and the *Finegoldia magna* trajectory identified here, already an early dysbiotic signal in the full cohort, is a concrete candidate feature for such models. Their validation will require larger, multicenter cohorts, more homogeneous sampling schedules, and higher-resolution sequencing, so that signals such as this can be translated into tools of genuine clinical value.

7. References

1. Singer, M. et al. (2016) The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 315, 801–810. doi: https://doi.org/10.1001/jama.2016.0287
2. Gray, A.P. et al. (2025) Global, regional, and national sepsis incidence and mortality, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet Glob. Health*. doi: https://doi.org/10.1016/S2214-109X(25)00356-0
3. Rudd, K.E. et al. (2020) Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. *Lancet* 395, 200–211. doi: https://doi.org/10.1016/S0140-6736(19)32989-7
4. Fleischmann-Struzek, C. et al. (2020) Incidence and mortality of hospital- and ICU-treated sepsis: results from an updated and expanded systematic review and meta-analysis. *Intensive Care Med.* 46, 1552–1562. doi: https://doi.org/10.1007/s00134-020-06151-x
5. Hotchkiss, R.S. et al. (2016) Sepsis and septic shock. *Nat. Rev. Dis. Primers* 2, 16045. doi: https://doi.org/10.1038/nrdp.2016.45
6. Wiersinga, W.J. and van der Poll, T. (2022) Immunopathophysiology of human sepsis. *EBioMedicine* 86, 104363. doi: https://doi.org/10.1016/j.ebiom.2022.104363
7. Liu, D. et al. (2022) Sepsis-induced immunosuppression: mechanisms, diagnosis and current treatment options. *Mil. Med. Res.* 9, 56. doi: https://doi.org/10.1186/s40779-022-00422-y
8. Rhee, C. and Klompas, M. (2017) New sepsis and septic shock definitions: clinical implications and controversies. *Infect. Dis. Clin. North Am.* 31, 397–413. doi: https://doi.org/10.1016/j.idc.2017.05.001
9. Qiu, X. et al. (2023) SIRS, SOFA, qSOFA, and NEWS in the diagnosis of sepsis and prediction of adverse outcomes: a systematic review and meta-analysis. *Expert Rev. Anti Infect. Ther.* 21, 891–900. doi: https://doi.org/10.1080/14787210.2023.2237192
10. Lamy, B. et al. (2020) Bloodstream infections: standard and progress in pathogen diagnostics. *Clin. Microbiol. Infect.* 26, 142–150. doi: https://doi.org/10.1016/j.cmi.2019.11.017

11. Riedel, S. et al. (2008) Timing of specimen collection for blood cultures from febrile patients with bacteremia. *J. Clin. Microbiol.* 46, 1381–1385. doi: https://doi.org/10.1128/JCM.02033-07

12. Agnello, L. et al. (2021) The value of a complete blood count (CBC) for sepsis diagnosis and prognosis. *Diagnostics* 11, 1881. doi: https://doi.org/10.3390/diagnostics11101881

13. Barichello, T. et al. (2022) Biomarkers for sepsis: more than just fever and leukocytosis, a narrative review. *Crit. Care* 26, 14. doi: https://doi.org/10.1186/s13054-021-03862-5

14. Larsson, A. et al. (2020) Calprotectin is superior to procalcitonin as a sepsis marker and predictor of 30-day mortality in intensive care patients. *Scand. J. Clin. Lab. Invest.* 80, 156–161. doi: https://doi.org/10.1080/00365513.2019.1703216

15. Crouser, E.D. et al. (2019) Monocyte distribution width: a novel indicator of sepsis-2 and sepsis-3 in high-risk emergency department patients. *Crit. Care Med.* 47, 1018–1025. doi: https://doi.org/10.1097/CCM.0000000000003799

16. Evans, L. et al. (2021) Surviving Sepsis Campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Med.* 47, 1181–1247. doi: https://doi.org/10.1007/s00134-021-06506-y

17. World Health Organization (2024) WHO Bacterial Priority Pathogens List, 2024: Bacterial Pathogens of Public Health Importance to Guide Research, Development and Strategies to Prevent and Control Antimicrobial Resistance. *World Health Organization*

18. Magill, S.S. et al. (2014) Multistate point-prevalence survey of health care-associated infections. *N. Engl. J. Med.* 370, 1198–1208. doi: https://doi.org/10.1056/NEJMoa1306801

19. Peleg, A.Y. and Hooper, D.C. (2010) Hospital-acquired infections due to gram-negative bacteria. *N. Engl. J. Med.* 362, 1804–1813. doi: https://doi.org/10.1056/NEJMra0904124

20. Nordmann, P. et al. (2011) Global spread of carbapenemase-producing *Enterobacteriaceae*. *Emerg. Infect. Dis.* 17, 1791–1798. doi: https://doi.org/10.3201/eid1710.110655

21. Holmes, C.L. et al. (2021) Pathogenesis of gram-negative bacteremia. *Clin. Microbiol. Rev.* 34, e00234-20. doi: https://doi.org/10.1128/CMR.00234-20

22. Human Microbiome Project Consortium (2012) Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. doi: https://doi.org/10.1038/nature11234

23. Qin, J. et al. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. doi: https://doi.org/10.1038/nature08821

24. Sommer, F. and Bäckhed, F. (2013) The gut microbiota: masters of host development and physiology. *Nat. Rev. Microbiol.* 11, 227–238. doi: https://doi.org/10.1038/nrmicro2974

25. Belkaid, Y. and Hand, T.W. (2014) Role of the microbiota in immunity and inflammation. *Cell* 157, 121–141. doi: https://doi.org/10.1016/j.cell.2014.03.011

26. Louis, P. and Flint, H.J. (2017) Formation of propionate and butyrate by the human colonic microbiota. *Environ. Microbiol.* 19, 29–41. doi: https://doi.org/10.1111/1462-2920.13589

27. Schulthess, J. et al. (2019) The short chain fatty acid butyrate imprints an antimicrobial program in macrophages. *Immunity* 50, 432–445.e7. doi: https://doi.org/10.1016/j.immuni.2018.12.018

28. Buffie, C.G. and Pamer, E.G. (2013) Microbiota-mediated colonization resistance against intestinal pathogens. *Nat. Rev. Immunol.* 13, 790–801. doi: https://doi.org/10.1038/nri3535

29. Lozupone, C.A. et al. (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* 489, 220–230. doi: https://doi.org/10.1038/nature11550

30. Falony, G. et al. (2016) Population-level analysis of gut microbiome variation. *Science* 352, 560–564. doi: https://doi.org/10.1126/science.aad3503

31. Yatsunenko, T. et al. (2012) Human gut microbiome viewed across age and geography. *Nature* 486, 222–227. doi: https://doi.org/10.1038/nature11053

32. Claesson, M.J. et al. (2012) Gut microbiota composition correlates with diet and health in the elderly. *Nature* 488, 178–184. doi: https://doi.org/10.1038/nature11319

33. Woese, C.R. and Fox, G.E. (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U. S. A.* 74, 5088–5090. doi: https://doi.org/10.1073/pnas.74.11.5088

34. Klindworth, A. et al. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 41, e1. doi: https://doi.org/10.1093/nar/gks808

35. Fukuda, K. et al. (2016) Molecular approaches to studying microbial communities: targeting the 16S ribosomal RNA gene. *J. UOEH* 38, 223–232. doi: https://doi.org/10.7888/juoeh.38.223

36. Ewing, B. et al. (1998) Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res.* 8, 175–185. doi: https://doi.org/10.1101/gr.8.3.175

37. Callahan, B.J. et al. (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: https://doi.org/10.1038/nmeth.3869

38. Cole, J.R. et al. (2014) Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42, D633–D642. doi: https://doi.org/10.1093/nar/gkt1244

39. Bolyen, E. et al. (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: https://doi.org/10.1038/s41587-019-0209-9

40. Haak, B.W. and Wiersinga, W.J. (2017) The role of the gut microbiota in sepsis. *Lancet Gastroenterol. Hepatol.* 2, 135–143. doi: https://doi.org/10.1016/S2468-1253(16)30119-4

41. Dickson, R.P. (2016) The microbiome and critical illness. *Lancet Respir. Med.* 4, 59–72. doi: https://doi.org/10.1016/S2213-2600(15)00427-0

42. Alverdy, J.C. and Krezalek, M.A. (2017) Collapse of the microbiome, emergence of the pathobiome, and the immunopathology of sepsis. *Crit. Care Med.* 45, 337–347. doi: https://doi.org/10.1097/CCM.0000000000002007

43. Zaborin, A. et al. (2014) Membership and behavior of ultra-low-diversity pathogen communities present in the gut of humans during prolonged critical illness. *mBio* 5, e01361-14. doi: https://doi.org/10.1128/mBio.01361-14

44. Lankelma, J.M. et al. (2017) Critically ill patients demonstrate large interpersonal variation in intestinal microbiota dysregulation: a pilot study. *Intensive Care Med.* 43, 59–68. doi: https://doi.org/10.1007/s00134-016-4613-z

45. Shimizu, K. et al. (2011) Altered gut flora are associated with septic complications and death in critically ill patients with systemic inflammatory response syndrome. *Dig. Dis. Sci.* 56, 1171–1177. doi: https://doi.org/10.1007/s10620-010-1298-8

46. Donskey, C.J. (2004) The role of the intestinal tract as a reservoir and source for transmission of nosocomial pathogens. *Clin. Infect. Dis.* 39, 219–226. doi: https://doi.org/10.1086/422505

47. Taur, Y. et al. (2012) Intestinal domination and the risk of bacteremia in patients undergoing allogeneic hematopoietic stem cell transplantation. *Clin. Infect. Dis.* 55, 905–914. doi: https://doi.org/10.1093/cid/cis580

48. Freedberg, D.E. et al. (2018) Pathogen colonization of the gastrointestinal microbiome at intensive care unit admission and risk for subsequent death or infection. *Intensive Care Med.* 44, 1203–1211. doi: https://doi.org/10.1007/s00134-018-5268-8

49. Schlechte, J. et al. (2023) Dysbiosis of a microbiota-immune metasystem in critical illness is associated with nosocomial infections. *Nat. Med.* 29, 1017–1027. doi: https://doi.org/10.1038/s41591-023-02243-5

50. Schuijt, T.J. et al. (2016) The gut microbiota plays a protective role in the host defence against pneumococcal pneumonia. *Gut* 65, 575–583. doi: https://doi.org/10.1136/gutjnl-2015-309728

51. Wang, D. et al. (2025) Intestinal dysbiosis in critically ill patients: a case-control study of *Enterobacteriaceae* enrichment and reduced microbial diversity. *Front. Med.* 12, 1680262. doi: https://doi.org/10.3389/fmed.2025.1680262

52. Cho, N.A. et al. (2024) Pathogenesis and therapeutic opportunities of gut microbiome dysbiosis in critical illness. *Gut Microbes* 16, 2351478. doi: https://doi.org/10.1080/19490976.2024.2351478

53. Adelman, M.W. et al. (2020) The gut microbiome's role in the development, maintenance, and outcomes of sepsis. *Crit. Care* 24, 278. doi: https://doi.org/10.1186/s13054-020-02989-1

54. Pickard, J.M. et al. (2017) Gut microbiota: role in pathogen colonization, immune responses, and inflammatory disease. *Immunol. Rev.* 279, 70–89. doi: https://doi.org/10.1111/imr.12567

55. Bazzano, C. et al. (2026) Gut microbiota as a key modulator in the pathophysiology of sepsis: SURVEIL project. *MicrobiologyOpen* 15, e70301. doi: https://doi.org/10.1002/mbo3.70301

56. Lu, Y. et al. (2023) MicrobiomeAnalyst 2.0: comprehensive statistical, functional and integrative analysis of microbiome data. *Nucleic Acids Res.* 51, W310–W318. doi: https://doi.org/10.1093/nar/gkad407

57. McMurdie, P.J. and Holmes, S. (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* 8, e61217. doi: https://doi.org/10.1371/journal.pone.0061217

58. Di Pierro, F. (2021) Gut microbiota parameters potentially useful in clinical perspective. *Microorganisms* 9, 2402. doi: https://doi.org/10.3390/microorganisms9112402

59. Programma Nazionale sul Trapianto Fecale del Ministero della Salute-DGPRE 0018191-P-15/06/2018. Available online: https://www.salute.gov.it/imgs/C_17_pubblicazioni_2800_allegato.pdf (accessed on 28 May 2026)

60. Segata, N. et al. (2011) Metagenomic biomarker discovery and explanation. *Genome Biol.* 12, R60. doi: https://doi.org/10.1186/gb-2011-12-6-r60

61. Rosenthal, M.E. et al. (2012) Finegoldia magna (formerly Peptostreptococcus magnus): an overlooked etiology for toxic shock syndrome? *Med. Hypotheses* 79, 138–140. doi: https://doi.org/10.1016/j.mehy.2012.04.013

62. Murphy, E.C. and Frick, I.M. (2013) Gram-positive anaerobic cocci – commensals and opportunistic pathogens. FEMS Microbiol. Rev. 37, 520–553. doi: https://doi.org/10.1111/1574-6976.12005

63. Kalantar, K.L. et al. (2022) Integrated host-microbe plasma metagenomics for sepsis diagnosis in a prospective cohort of critically ill adults. Nat. Microbiol. 7, 1805–1816. doi: https://doi.org/10.1038/s41564-022-01237-2

64. Marcos-Zambrano, L.J. et al. (2021) Applications of machine learning in human microbiome studies: a review on feature selection, biomarker identification, disease prediction and treatment. Front. Microbiol. 12, 634511. doi: https://doi.org/10.3389/fmicb.2021.634511

Acknowledgements

*I would like to thank my supervisor, **Prof. Sandra D'Alfonso**, and my co-supervisor, **Dr. Marta Mellai**, for the opportunity to carry out this study and for their guidance throughout it. I am especially grateful to **Dr. Mellai**, who followed this project day by day and from whom I learned most of what I now know about working at the bench, for her constant guidance, her patience, and her support through every stage of the work. I am grateful for the trust placed in me, for holding this project to a high scientific standard, and for everything it taught me about doing research carefully and honestly with scientific rigor and intellectual integrity. The time spent within several projects at the CAAD Integrative Genomics Core Facility has shaped the way I think as a scientist, and I will carry that with me well beyond this thesis. I would also like to thank **Prof. Elisa Bona**, who produced the 16S rRNA gene sequencing data for the entire set of SURVEIL samples, and to extend my gratitude to **all the participants and members** of the SURVEIL project, whose work made this research possible.*

*To my father, **Nidal**, my mother, **Lina**, and my brother, **Ammar**, thank you for everything that could never fit onto these pages. You have been my foundation across every distance, and I owe this achievement to your patience, your sacrifices, and your belief in me even when the road was long and far from home. Every page of this work rests on what you have given me.*

*To my colleagues and lab mates at CAAD, **Samar** and **Martina**, thank you for the daily company at the bench, for the problems we puzzled through together, and for the small victories that made the long days worthwhile. Your help, both scientific and human, made this experience far richer than it would otherwise have been, and a great deal of what I learned, I learned alongside you.*

*My time in Italy would not have been the same without the friends I found here, above all **Alaa**, without whose help I would not have been in Italy at all, and **Ilenia**, who has stood by me since we met on this master's and accepted me exactly as I am. Thank you for turning this country into a home, for the meals and the laughter and the parties and the moments away from the lab that kept me grounded through the most demanding stretches of this journey. You gave me a place to belong.*

*And to my friends abroad, my other family, **Samer**, **Al-Abbass**, **Majd**, **Rami** and **Basel**, spread across borders and time zones, thank you for staying close despite the distance. Your messages and calls and your constant presence were a reminder that the friendships that matter are not measured in kilometers. Knowing I could always count on you carried me more than you know.*

To all of you, from the deepest of my heart, thank you.

-AlaaJ