Coursework Declaration and Feedback Form

The Student should complete and sign this part

Student Number: 2815452Z	Student Name: Ke Zhao
Programme of Study : MSc in Computer Systems Engineering	
Course Code: ENG5059P	Course Name: MSc Project
Name of **First** Supervisor: Umer Ijaz	Name of **Second** Supervisor: Oana Dobre
Title of Project: Exploration of zeta diversities and species interaction in gut microbial profile of Crohn's disease patients when treated with Exclusive Enteral Nutrition	

Declaration of Originality and Submission Information

I affirm that this submission is all my own work in accordance with the University of Glasgow Regulations and the School of Engineering requirements Signed (Student) : Ke Zhao	
Date of Submission : August 14th 2023	

Feedback from Lecturer to Student – to be completed by Lecturer or Demonstrator	
Grade Awarded: Feedback (as appropriate to the coursework which was assessed):	
Lecturer/Demonstrator:	Date returned to the Teaching Office:

Exploration of zeta diversities and species interaction in gut microbial profile of Crohn's disease patients when treated with Exclusive Enteral Nutrition

Student Name: KE ZHAO
Student ID: 2815452Z
Supervisor: Dr Umer Zeeshan Ijaz

August 14th 2023

ACKNOWLEDGEMENT

I would like to express my sincere thanks to my supervisor, Dr Umer ljaz, Uzma, Kelly and Aqsa. Thanks for all the guidance and suggestions toward the project with extremely patience. And the same sincerely thanks for the group mates.

How time goes fly, I've gained many "first times" in this year, the first time I came to the UK, the first time I came to a completely strange environment by myself, and so on. Fortunately, my family, girlfriend, and friends are always by my side, helping me solve problems and cheering me up, making my life very enjoyable this year. For this Msc project, I would like to express my heartfelt thanks again to my supervisor Dr Umer ljaz and Aqsa, who have provided great help to my project. Umer ljaz can still provide me with precious opinions by email even he went to take care of his family who suffered from the disease. Even after receiving my file on Friday night, Aqsa will still take her time to propose amendments for me. Finally, I would like to thank myself who made this decision a year ago.

ABSTRACT

The etiology of Crohn's disease's pathological processes has garnered substantial scholarly attention. Among the prevailing hypotheses, a prominent perspective underscores the intricate connection between the ailment and the perturbation within the gut's microbial equilibrium. Within the spectrum of therapeutic interventions, Exclusive Enteral Nutrition (EEN) assumes a preeminent position as a primary modality for patients afflicted with Crohn's disease (CD). This encompasses even those in critically compromised states, as well as children and adolescents undergoing initial induction therapy.

Extant evidence derived from singular-center cohort investigations highlights that EEN can yield a remarkable remission rate of up to 80% in select CD patients. The utility of EEN transcends the mere pursuit of instigating abatement of inflammatory processes; it serves a dual role by concurrently providing vital nutritional sustenance. This therapeutic avenue also boasts economic viability. It presents an invaluable option for individuals who grapple with contraindications to corticosteroids and immunosuppressive regimens, while also remaining financially precluded from embracing costly biologics.

In instances where these multifaceted challenges converge, EEN might emerge as the singular efficacious recourse. Nonetheless, the precise mechanistic underpinning of EEN's efficacy remains shrouded in ambiguity. The elucidation of its modus operandi, specifically its capacity to ameliorate intestinal function and abate inflammatory cascades in the context of CD patients, continues to elude comprehensive explication.

This report uses previous data: a total of 117 fecal samples from 23 CD children (in different EEN treatment periods) and 21 healthy children were collected, and the data set was obtained by 16s rRNA gene sequencing and shotgun gene sequencing. And by using two relatively novel methods of analysis: network inference and zeta diversity, to analyse the interaction of gut microorganisms and the change in diversity during EEN treatment, to try to find the bacteria associated with the pathogenesis of CD and to explore the role of EEN in curing the specific mechanism of Crohn's disease.

Ultimately, microorganisms like TM7x have been revealed which might exert an influence on the development of Crohn's disease (CD) and exacerbate inflammatory manifestations. While bacteria such as Firmicutes may secrete proteins that play a crucial role in curing CD and improving intestinal barrier function substances, and inhibit other harmful bacteria. EEN may reduce the diversity of gut microbiota and increase the abundance of certain beneficial bacteria in the process of curing CD, and regulate the dysbiosis of intestinal microbiota.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Explanation
EEN	Exclusive Enteral Nutrition
CD	Crohn’s Disease
UC	Ulcerative Colitis
IBD	Inflammatory Bowel Disease
GI	gastrointestinal
ESPGHAN	European Society of Pediatric Gastroenterology, Hepatology and Nutrition
ECCO	European Crohn’s and Colitis Organization
NGS	Next Generation Sequencing
ALB	Albumin Levels
CDIA	Crohn's Disease Activity Index
DNA	Deoxyribonucleic acid
SparCC	Sparse Correlations for Compositional data
SPIEC-EASI	SParse InversE Covariance Estimation for Ecological Association Inference
OTU	Operational Taxonomic Unit
IVI	Integrated Value of Influence
SCFA	Short Chain Fatty Acid
C1	Formic acid
C2	Acetic Acid
C3	Propionic Acid
C4	Butyric Acid
C5	Valerate Acid
C6	Caproic Acid

CHAPTER 1

INTRODUCTION

1.1 Background

Within the human organism, a myriad of microorganisms coexists symbiotically with their host, predominantly occupying regions such as the gastrointestinal tract, skin, saliva, oral mucosa, conjunctiva, and vagina. Among these, the microorganisms inhabiting the gastrointestinal tract—collectively referred to as gut microbiota—amount to approximately 1×10^{14}, playing a pivotal role in maintaining intestinal equilibrium, fostering developmental processes, and fortifying the body's defenses against invasive pathogens. Moreover, their residence within the gut engenders crucial immunomodulatory responses and catalyses essential metabolic reactions.

The components comprising the gut microbiota encompass a rich assortment of microorganisms, encompassing bacteria, yeasts, and viruses. Of particular significance are the bacteria, boasting a repertoire of over a thousand distinct species. These microbial entities can be systematically classified into six prominent phyla: Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria, Fusobacteria, and Verrucomicrobia. Among these, Firmicutes and Bacteroidetes emerge as preeminent, collaboratively comprising a substantial portion, up to 90%, of the gut microbiota's composition.

Remarkably, a healthy individual's gut microbiota composition exhibits discernible variation across distinct segments of the gastrointestinal tract and evolves over time, reflecting changes associated with aging (including infancy and early development) as well as being influenced by environmental factors such as dietary habits, lifestyle choices, and antibiotic utilization. The composition of microbiota diverges significantly among individuals, with these variations attributed to factors like age, ethnic background, lifestyle, and dietary preferences. Diverse microbiota is further classified into three distinct enterotypes. These variations are considered physiological and harmonious with a state of robust microbial health. However, alterations in microbiota composition are often associated with diseases, a condition known as dysbiosis. Yet, the causal relationship between modified microbiota and various diseases frequently remains enigmatic.

1.1.1 Crohn’s disease

In the realm of worldwide health concerns, the prevalence of Inflammatory Bowel Disease (IBD) has exhibited a notable rise post-2000. Within Western nations, the impact of IBD has extended to approximately 1 in 200 individuals in current times [1]. This ailment comprises two distinct disorders, namely Crohn’s disease (CD) and Ulcerative Colitis (UC), each characterized by disparate pathophysiological underpinnings, affected segments of the gastrointestinal (GI) tract, symptomatology, complications, disease progression, and management strategies.

The exact etiology of Crohn's disease remains elusive. Current understanding points to a multifaceted genesis involving a confluence of genetic susceptibility, immune responses, and environmental and bacterial influences. In individuals predisposed by

their genetic makeup, these factors converge to initiate the condition. Crohn's disease (CD) is typified by non-continuous areas of inflammation within the gastrointestinal (GI) tract, colloquially termed "skip lesions." This ailment entails persistent, cyclic transmural inflammation that can precipitate chronic abdominal pain, diarrhea, obstructions, and/or perianal lesions.

Clinical manifestations of Crohn's disease commonly encompass abdominal discomfort, diarrhea (which may take on a bloody nature under severe inflammation), fever, abdominal distension, and weight loss. Beyond the confines of the gastrointestinal tract, complications may manifest, such as anemia, cutaneous eruptions, arthritis, ocular inflammation, and fatigue.

Over the past two decades, developed nations have witnessed a consistent rise in the occurrence of Crohn's disease (CD) among pediatric populations, a trend marked by not only a decrease in the age of onset but also heightened disease activity [2]. While the root cause of this phenomenon remains enigmatic, one plausible explanation emerges from the ongoing societal shifts witnessed in the contemporary era. These transformative changes invariably impact the microbiome, an entity of pivotal significance in the intricate functioning of the gastrointestinal system. Unquestionably, the microbiome assumes a pivotal role in upholding the appropriate operation of the intestinal milieu. This role encompasses vital functions such as facilitating gut functionality, orchestrating the production of essential vitamins, nurturing immune system development, maintaining epithelial equilibrium, and generating crucial metabolites—all collectively contributing to the fortification of the gut barrier.

Establishing a definitive archetype for a healthy gut microbiome proves challenging, as its composition exhibits variance contingent on factors like age, surroundings, and dietary choices. Nonetheless, the cardinal attributes of a thriving gut microbiome generally encompass diversity, stability, and the faithful fulfillment of its metabolic responsibilities.

1.1.2 EEN (Exclusive Enteral Nutrition)

A definitive cure for Crohn's disease remains elusive. However, an inadvertent revelation during the 1970s unveiled the potential of exclusive enteral nutrition (EEN) to elicit remission in CD patients. Since then, EEN has emerged as a groundbreaking strategy for inducing remission while concurrently optimizing nutritional status post-diagnosis. In 2014, the European Society of Pediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) and the European Crohn’s and Colitis Organization (ECCO) jointly published updated consensus guidelines. These guidelines not only endorse EEN's role but also propose it as the primary first-line induction therapy for pediatric patients afflicted with CD.

For the purpose of remission induction, patients underwent treatment with exclusive enteral nutrition (EEN) administered either orally or via nasogastric tube feeding over

a span of approximately 6–8 weeks. During this regimen, consumption was restricted to solely chewing gum and water, as stipulated [3]. The dietary protein sources encompassed elemental, semi-elemental, and polymeric diets. Interestingly, variations in protein or lipid levels appeared to exert minimal influence on efficacy. Notably, its effectiveness in inducing clinical remission in children and adolescents grappling with active CD is tantamount to that of corticosteroids. Furthermore, it outshines corticosteroids in its capacity to elicit mucosal healing.

Despite a substantial body of evidence supporting the efficacy of nutritional therapy in pediatric cases, the precise mechanisms underpinning EEN's remission-inducing properties remain inscrutable. The distinctive gut microbiota and metabolomic profiles exhibited by CD patients in comparison to healthy individuals underline a significant proposition—that disrupted co-metabolism within the gut microbial community could substantively contribute to the pathogenesis of CD.

1.2 Prior research

Past investigations have primarily centered their attention on scrutinizing alterations within the intestinal flora as well as the concomitant intestinal inflammation within patients afflicted by Crohn's disease and subjected to exclusive enteral nutrition (EEN) treatment. To illustrate this, a study was conducted (Dong Guo et, al, 2022) [4] that initiated with the collection of data from 16 recently diagnosed Crohn's disease patients, designated as the experimental group. Concomitantly, a control group consisting of 10 healthy volunteers was assembled. Within the cohort of Crohn's disease patients, the 16 individuals were randomly allocated to either the EEN group or the corticosteroid group. The focal point of this study was the subsequent analysis and comparison of the gut microbiota composition.

This study methodology encompassed a two-tier evaluation framework. Initially, it entailed an assessment of the clinical treatment effects based on indicators such as albumin (ALB) levels and Crohn's Disease Activity Index (CDAI) scores. Subsequently, a comprehensive exploration into shifts within microbial composition throughout the treatment process was undertaken. The ultimate findings served to validate that EEN effectively sustained Crohn's disease remission while simultaneously mitigating inflammation, resulting in improvements in nutritional indices.

EEN, as well as corticosteroids, emerged as influential agents capable of enhancing microbiome diversity, thereby facilitating the induction of Crohn's disease remission. Notably, these interventions exhibited divergent effects on the proportions of microbiome species. Seeking to illuminate EEN's precise role in reestablishing microbial equilibrium within the gut, Joseph Runde et al. embarked on the task of reconstructing genomes extracted from the gut metagenomes of 12 pediatric subjects. This endeavor culminated in the classification of each microbial population into distinct "phenotypes" or response patterns. This classification was predicated on variations in the relative abundance of each microorganism throughout the treatment course. The outcomes elucidated that children who achieved clinical remission throughout treatment showcased enriched microbiota. However, these microbial communities were either suppressed or underwent transient proliferation, attributable to the functional influence of EEN. A pivotal hallmark of EEN's efficacy manifested as the transient emergence of rare microbial populations akin to those observed in healthy

individuals. This phenomenon corresponded with a concurrent reduction in microbes typically associated with gut homeostasis [39].

1.3 Aims and Objectives

1. Employing network inference, an in-depth analysis was executed utilizing a dataset of 16S rRNA obtained from both Crohn's disease (CD) patients and individuals in good health. The primary objective centered around elucidating the intricate interactions within the gut microbiota. This pursuit aimed to offer discernment into the functions of distinct bacterial species and their associated metabolites, thereby shedding light on their contributions to the symptom of CD.

2. Utilize the Zeta diversity technique as a pivotal tool in the research endeavors to explore the dynamic changes in diversity elicited by EEN treatment. By applying this technique, delve into the intricate mechanisms underlying the action of EEN. The core aim is to elucidate how EEN operates to induce therapeutic effects, shedding light on the intricate processes governing its modulatory impact on the gut microbiota and associated diversity patterns.

CHAPTER 2

METHODOLOGY

Microbiomes stand as intricate microbial communities, characterized by intricate structures and functions heavily shaped by interactions among microorganisms and between microorganisms and the host. These interactions manifest through diverse mechanisms, encompassing direct cell-to-cell communication and interspecies signalling, as well as indirect sensing of metabolites. Collectively, these mechanisms have emerged as pivotal players in modulating disease progression and clinical outcomes [13]. An exemplar of intricate microbial interactions contributing to disease exacerbation is polymicrobial synergism, a phenomenon where infections involving multiple interplaying species of bacteria yield more severe outcomes compared to single-agent infections. Notably, polymicrobial synergism has been linked to heightened levels of antibiotic resistance, biofilm formation, tissue damage, and adaptive responses to the environment [14, 15].

Consequently, comprehending the microbiome in its entirety necessitates a holistic perspective that encompasses not only the intricate interplay among various microbial taxa but also the intricate interactions these microorganisms have with their host organisms. This comprehensive comprehension stands as a pivotal prerequisite for unravelling the multifaceted roles that microbiomes undertake in terms of host health, developmental processes, dysbiosis, and the intricacies of polymicrobial infections. Despite the profound impact of NGS technologies in vastly expanding the scope and scale of microbiome studies, it remains surprising that the analytical methodologies tailored to scrutinize microbe–microbe and host–microbe interactions remain surprisingly constrained [12].

2.1 Data description

The presented report draws its foundation from a prior research endeavor [5]. Fecal samples were meticulously gathered from a cohort comprising 23 children (males: n =13; aged between 6.9–14.7 years) diagnosed with active Crohn's disease (CD), classified according to the Montreal classification. The disease locations encompassed L2 (n =3), L2+L4 (n =4), L3 (n =3), and L3+L4 (n =13), while the disease behaviors were categorized as B1 (n =20), B2 (n =2), and B3 (n =1) [5]. Among these subjects, 15 (including 11 newly diagnosed cases) contributed at least two consecutive fecal samples during the span of 8 weeks of exclusive enteral nutrition (EEN), as detailed previously [5]. Each patient contributed a maximum of five serial samples, comprising the initial sample (A) collected before or within six days of EEN initiation (89% collected within four days), two samples during EEN (B: ~16 days and C: ~32 days), one near the culmination of treatment (D: ~54 days), and a final sample (E: ~63 days post EEN) collected after the return to habitual diet. The distribution of participants contributing 1, 2, 3, 4, and 5 samples was 5, 3, 3, 2, and 10, respectively [5]. For comparative purposes, 21 healthy children (males: n =12; aged between 4.6–16.9 years) devoid of any documented familial history of inflammatory bowel disease were included in the study, with two fecal samples collected at intervals of at least two months apart serving as the control group [5].

2.2 Genome primary analysis

16S rRNA gene sequencing (16Ss) and shotgun metagenomic sequencing (SMs) are the two main NGS tools implemented for microbial community profiling [10].

2.2.1 16S rRNA gene sequencing

The 16S rRNA gene has stood as a cornerstone in the realm of sequence-based bacterial analysis for an extended period. Following the emergence of high-throughput sequencing methodologies, PCR-amplified 16S sequences have commonly been clustered based on their similarity to generate operational taxonomic units (OTUs). These OTUs, in turn, have been juxtaposed with reference databases to deduce plausible taxonomic classifications. While this approach has demonstrated convenience and potency, it has also given rise to certain underlying assumptions. One illustrative example is the historical assumption that sequences displaying > 95% identity pertain to the same genus, whereas sequences demonstrating > 97% identity correspond to the same species [7].

In this present study, the procedural steps for conducting 16S rRNA gene sequencing, based on previously established research, are outlined as follows:

The initial phase encompassed the isolation of bacterial DNA through the chaotropic method [8]. Subsequently, 16S rRNA sequencing was executed targeting the V4 region utilizing the MiSeq platform (Illumina). This entailed employing 2 × 250 base pair paired-end reads. To facilitate this process, the V4 region was amplified, and the reverse strand was barcoded with fusion Golay adaptors. The forward 16S rRNA primer sequence employed was 515f (GTGNCAGCMGCCGCGGTAA). Furthermore, the reverse primers, barcodes, and adaptors mirrored those that had been previously described [9].

Following amplification, the amplicons were subjected to purification utilizing AMPure XP DNA purification beads (Beckman Coulter, Danvers, MA, USA), adhering to the manufacturer's stipulated guidelines. The resultant purified amplicons were eluted using 25 µl of Elution Buffer (Qiagen, 19086, UK). Subsequent to this, the quantification of amplicons was conducted through employment of the KAPA SYBR® FAST qPCR Kit (Kapa biosystems, KK4824, UK). In order to avert potential base-calling issues arising from limited base diversity, the amplicons were then diluted to a concentration of 40 pM, accompanied by the addition of 40 pM of genomic DNA as a spike-in measure.

2.2.2 Shotgun metagenome sequencing

High-throughput sequencing methodologies offer a transformative avenue for genomic analyses, enabling comprehensive investigations of all microorganisms within a given sample, thereby surpassing the limitations posed by cultivability. Among these methodologies, shotgun metagenomics stands out—an untargeted approach that involves sequencing all microbial genomes present within a sample. This encapsulates the comprehensive range of microbial genomes, collectively referred to as 'metagenomics', comprising the given sample. Shotgun metagenomics serves a dual-fold purpose: it not only enables the characterization of the taxonomic composition and functional capabilities of microbial communities but also enables the recovery of entire

genome sequences. It is noteworthy that methods like high-throughput 16S rRNA gene sequencing, which spotlight specific organisms or individual marker genes, are at times colloquially referred to as metagenomics [11].
In this presented study, the procedural steps for conducting the 16S rRNA gene sequencing method, as informed by previous research, unfold as follows:

To capture the genetic functional capacity, shotgun metagenomics samples were meticulously crafted employing the Nextera XT Prep Kit (Illumina, FC-131-1096, UK), in conjunction with the Illumina dual-barcoding Nextera XT Index kit (Illumina, FC-131-1002, UK). Subsequent to their preparation, the sequencing libraries were combined in equimolar concentrations and subjected to quantification utilizing the KAPA SYBR® FAST qPCR Kit (Kapa biosystems, KK4824, UK). Following quantification, these libraries were loaded onto both lanes of a rapid run flow cell at concentrations of 10 pM. The HiSeq 2500 (Illumina) instrument was instrumental in generating clusters on-board, with the subsequent sequencing process employing TruSeq Rapid SBS Kit reagents (Illumina, FC-402-4001, FC-402-4002). The sequencing procedure adhered to a paired-end 150-cycle protocol.

2.3 statistical analysis method

In this report, the analysis employs network and zeta diversity methodologies, implemented through the R programming language. The overarching objective is to scrutinize and juxtapose the gut microbiota communities between individuals afflicted with Crohn's disease and a healthy control group. Furthermore, the study aims to delve into the transformations of the intestinal microbiota occurring throughout the course of exclusive enteral nutrition (EEN) treatment. This endeavor encompasses an exploration of the potential significance held by specific pivotal flora in the genesis of Crohn's disease, as well as an investigation into the therapeutic mechanisms underpinning the healing effects of EEN.

2.3.1 R Language and R Studio

R serves as both a language and an environment tailored for statistical computing and graphics. It boasts a wide-ranging assortment of statistical tools, encompassing linear and nonlinear modelling, traditional statistical tests, time-series analysis, classification, clustering, and more. The robustness of R lies in its extensibility, enabling users to augment its functionalities according to their requirements.

R encompasses an integrated suite of software features designed for tasks such as data manipulation, calculations, and graphical visualization. Notably, it furnishes an effective mechanism for handling and storing data, an array of operators optimized for array-based calculations (especially matrices), a comprehensive set of intermediate tools for data analysis, and graphical utilities for presenting data either on-screen or in hardcopy form. Furthermore, R boasts a well-developed programming language characterized by its simplicity and efficacy. This programming language encompasses features like conditionals, loops, user-defined recursive functions, and mechanisms for input and output.

In parallel, RStudio emerges as an integrated development environment (IDE) catering to both R and Python programming. It amalgamates a console, a syntax-highlighting

editor that supports direct execution of code, and an array of tools that encompass plotting, history tracking, debugging, and workspace management.

2.3.2 Network inference

Utilizing network theory, it becomes possible to model and dissect the intricacies of a microbiome along with its intricate interactions, all within a single network framework. An intriguing facet of network theory is its ability to unveil universal architectural traits present in a myriad of complex systems. These characteristics span across various domains, encompassing microbiomes, molecular interaction networks, computer networks, microcircuits, and social networks, among others. This pervasive universality provides an avenue for harnessing the insights gained from extensively studied non-biological systems. These insights can be applied to untangle the interlinked associations that define microbial interactions within the context of a microbiome.

Various techniques, exhibiting diverse levels of efficiency and precision, have been employed to construct networks rooted in microbiome data. Among these methods, some are based on (dis)similarity or distance measurements, while others lean towards correlation-based approaches. Notably, the latter category involves the identification of significant pairwise relationships between operational taxonomic units (OTUs), which group organisms based on specific levels of DNA sequence similarity at a designated marker gene. This identification typically relies on correlation coefficients like Pearson's or Spearman's, serving as indicators. Nonetheless, the application of correlation coefficients for detecting interdependencies within a microbiome encounter limitation. These constraints include susceptibility to detecting spurious correlations due to compositional considerations [16] and encountering statistical power constraints due to the relatively limited sample size. The apprehensions associated with correlation -centered analyses have instigated the development of methodologies that exhibit resilience in the context of compositional data. One such innovation is SparCC (Sparse Correlations for Compositional data), which utilizes linear Pearson correlations between logarithmically transformed components to unveil associations within compositional datasets. Additionally, the SPIEC-EASI approach (SParse InversE Covariance Estimation for Ecological Association Inference) is a statistical technique employed to infer microbial ecological networks. It amalgamates specialized data transformations designed for the analysis of compositional data with a graphical model inference framework. This framework is founded upon the assumption that the inherent ecological association network is characterized by sparsity [17].
In this report, we used ϕ-statistic and proportionality to characterize the correlation of relative data based on previous research [18]. Sometimes, for data carrying only relative information, common analysis methods, including correlation, can be very misleading while CoDA(Compositional Data Analysis) could be an efficient alternative. That is, while statistically independent variables X, Y, and Z are not correlated, their ratios X/Z and Y/Z must be, because of their common divisor.

The CoDA theory introduces three fundamental principles:
1. Scale invariance: Analyses are required to consider vectors featuring proportionally positive constituents as indicative of the same composition.
2. Sub-compositional coherence: Inferences drawn concerning sub-compositions (subsets of constituents) should demonstrate consistency, irrespective of whether these inferences are derived from the sub-composition or the entirety of the composition.

3. Permutation invariance: Analytical conclusions must remain independent of the arrangement sequence of the individual components.

Proportionality obeys all three principles for analysing relative data. If relative abundances x and y are proportional across experimental conditions i, their absolute abundances must be in proportion:

$$\frac{x_i}{t_i} \propto \frac{y_i}{t_i} => x_i \propto y_i \quad (1)$$

where t_i is the total abundance in condition i.

A "goodness-of-fit to proportionality" statistic ϕ was proposed by researchers to assess the extent to which a pair of random variables (x, y) are proportional (David Lovell et al.) [17]. ϕ is related to log-ratio variance, var(log(x/y)) as following equation 2, and is zero when x and y behave perfectly proportionally. However, when x and y are not proportional, ϕ has both a clear geometric interpretation and a meaningful scale, addressing concerns raised about log-ratio variance: the closer ϕ is to zero, the stronger the proportionality. Considering "Strength" of proportionality (goodness-of-fit) rather than testing the hypothesis of proportionality because it allows us to compare relationships between different pairs of mRNAs.

$$\begin{aligned} var\left(\log\left(\frac{x}{y}\right)\right) &= var(log^x - log^y) \quad (2)\\ &= var(log^x) + var(log^y) - 2cov(log^x, log^y)\\ &= var(log^x) \cdot \left(1 + \frac{var(log^y)}{var(log^x)} - 2\sqrt{\frac{var(log^y)}{var(log^x)}}\frac{cov(log^x, log^y)}{\sqrt{var(log^x)var(log^y)}}\right)\\ &= var(log^x) \cdot (1 + \beta^2 - 2\beta|r|)\\ &\triangleq var(log^x) \cdot \phi(log^x, log^y) \end{aligned}$$

In practice, one can use the relationship $\phi(log^x, log^y) = var\left(\log\left(\frac{x}{y}\right)\right)/var(log^x)$

Within a microbiome network, distinct entities often emerge, including clusters (referred to as modules) of closely interconnected microorganisms, alongside individual microbes that hold pivotal positions within the network's framework (referred to as keystone taxa). These clusters not only embody localized interaction patterns but also wield influence over the network's overall structure and interconnections. They can carry significance from biological, taxonomic, evolutionary, or functional perspectives. To identify these clusters, topological clustering methods are widely employed, serving as the preferred approach for detecting and delineating clusters within networks.

2.3.3 Zeta diversity

Spatial variability in species presence or absence within ecological communities, which forms the basis of biodiversity analysis, plays a critical role in ecological research. However, one notable challenge in comprehending the relationships among diverse biodiversity theories lies in the divergence of mathematical terminologies and the absence of a unified system of equations. Presently, there exists no singular metric that effectively links the array of patterns generated from species presence-absence, or

"incidence," data. This deficiency hampers the formulation of mathematical connections between these patterns.

The absence of a universal measure obstructs a comprehensive understanding of the interrelations between these patterns and hinders their quantitative formulation. Establishing a common metric would yield substantial benefits for modelling purposes and enhancing the comprehension of the underlying mechanisms driving spatial diversity patterns. The amalgamation of diverse biodiversity models into a coherent framework represents a pivotal objective within the field of ecology.

The exploration of why and how biodiversity varies across different sites and habitats, along with the implications of this variation, often involves an examination of species richness and composition. Metrics to quantify spatial differences in compositional similarity among communities commonly rely on β diversity. These metrics are derived by partitioning overall γ diversity into α and β components, utilizing either multiplicative (β=γ/α) or additive (γ=α+β) partitioning approaches [21]. A variety of community patterns, including species-area relationships, interspecific range size distributions, as well as patterns of rarity and endemism, are also frequently employed for analysis purposes [21].

In cases involving comparisons among three or more ecological assemblages, the approach often employed is the calculation of pairwise similarities and subsequent determination of their average. Regrettably, none of the metrics designed for assessing the turnover of species across sites based on presence-absence (incidence) data has the capacity to comprehensively calculate all components of diversity. This limitation implies that the diversity components for three or more distinct assemblages cannot be accurately represented solely using α and β diversity.

For instance, when dealing with three assemblages, the species exclusively shared between pairs of these assemblages within the comparison cannot be accurately computed using only α and β diversity. Likewise, the species that are shared by all three of the assemblages also cannot be accurately assessed using this approach. Consequently, the use of pairwise metrics proves inadequate for effectively representing the degree of similarity between multiple assemblages across various sites.

The notion of Zeta (ξ) diversity, serving as a comprehensive metric encompassing all facets of diversity stemming from assemblage partitioning, was introduced by researchers Hui and McGeoch in 2014 [21]. In this framework, let ξ_i denote the average number of species shared by i sites (as depicted in Figure 1). Notably, ξ_i (where $i = 1$) corresponds to the mean number of species across all sites. Given that species shared by i sites necessarily include those shared by $i - 1$ sites, the count of shared species ξ_i diminishes in a monotonic manner as i increases [21]. All pairwise β diversity metrics reliant on incidences can be represented utilizing ξ_1 and ξ_2. However, for scenarios involving three or more sites, the diversity components (e.g., A–G in the inset of Figure 1, where a component or partition represents the subset of species shared by a specific set of sites, following the concept outlined by Lande in 1996) cannot be solely estimated using pairwise components such as ξ_1 and ξ_2 [21]. The inclusion of higher-order ξ components become necessary to accomplish this task.

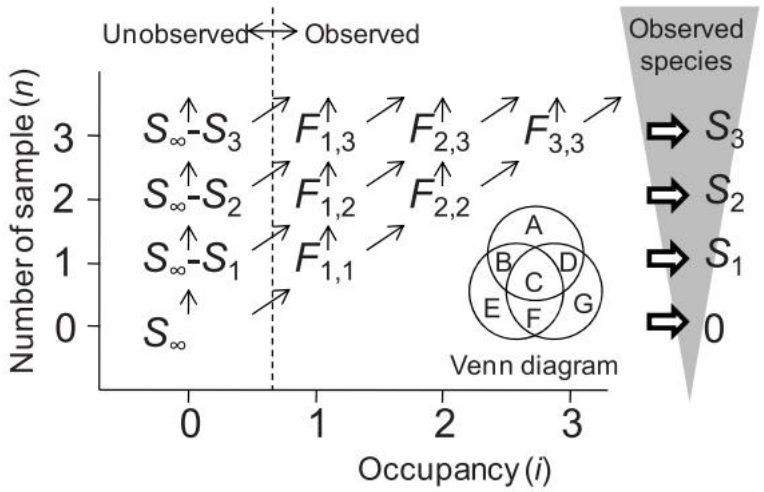

Figure 1: Recursive diagram of diversity partitioning across multiple assemblages (sites or samples) [21].

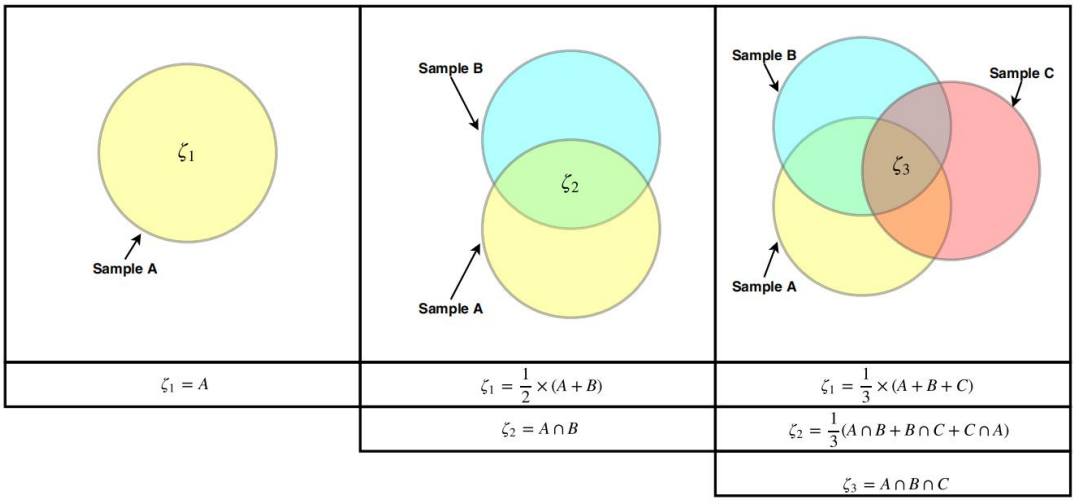

Figure 2: An illustration of the first three orders of ξ diversity [20].

The methodology for partitioning diversity can be visually depicted through a cumulative and recursive approach, as illustrated in Figure 1. Let S_n denote the total count of species present across n surveyed sites, and let $F_{i,n}$ represent the count of species found in exactly i out of the total n sites [21]. By applying the inclusion-exclusion principle, the subsequent general formulas can be systematically derived through the utilization of ξ components:

$$S_n = \sum_{k=1}^{n} (-1)^{k+1} \cdot C_n^k \cdot \xi_k \,, \tag{3}$$

$$F_{i,n} = C_n^i \cdot \sum_{k=1}^{n-i+1} (-1)^{k+1} \cdot C_{n-i}^{k-1} \cdot \xi_{i+k-1} \tag{4}$$

Where C_n^i is the number of combinations of choosing i from n sites and k is the standard index of summation.

CHAPTER 3

RESULTS

3.1 Network inference

Complex biological systems are intricately composed of networks, and unraveling the functional significance of individual components within these networks is pivotal for comprehending both healthy and diseased states. A network, often represented as N = (n, e), is characterized by its nodes (also termed vertices) and edges. Nodes represent distinct elements within the network that are interlinked by edges. Among these nodes, influential ones form the core of the entire network, and their identification is facilitated by evaluating network centralities. These centralities, which quantify the importance of nodes within a network, are derived from a comprehensive analysis of the network's overall structure. Notably, the fundamental properties of networks and measures of centrality possess a universal nature, extending across diverse domains, including the realm of biological networks.

Degree centrality, ClusterRank, neighborhood connectivity, local H index, betweenness centrality, and collective influence stand out as key determinants for pinpointing a network's most significant nodes, as illustrated in Table 1. To surmount the challenges associated with prior centrality measurement approaches developed for diverse networks, which included issues like adverse impacts stemming from centrality measurements between nodes and nonlinear centrality's positional discrepancies, a novel technique termed the "Integrated Value of Influence" (IVI) was introduced (Abbas Salavaty et al., 2020) [22]. IVI represents a pioneering approach that genuinely amalgamates the influence garnered from these six critical network centrality indicators. This method addresses the complexities posed by integrating these distinct dimensions of influence within networks.

Centrality Measure	Topological Scale	% Normality[a]
Degree centrality	Local	100% non-normally distributed
ClusterRank	local	95% non-normally distributed
LH index	semi-local	96% non-normally distributed
Neighborhood connectivity	semi-local	84.5% non-normally distributed
Betweenness centrality	global	97% non-normally distributed
Collective influence	global	97% non-normally distributed

[a]The normality percentage of each centrality measure among 200 networks studied.

Table 1: Six centrality measures with their Characteristics [22]

The analysis of the interdependence between ClusterRank and neighborhood connectivity, executed through four statistical methods (CANOVA, Hoeffding, MIC, and NNS), unveiled a substantial degree of interrelation between these two metrics across numerous networks (Abbas Salavaty et al., 2020) [22]. Furthermore, a significant negative correlation was observed between betweenness centrality and collective influence in a considerable subset of networks, in comparison to ClusterRank. Hence, leveraging the mathematical principles outlined earlier, researchers proposed utilizing the additive product of ClusterRank and neighborhood connectivity as a means of mitigating the potential bias arising from the additive product of betweenness centrality and collective influence.

Considering the distinct scales of various centrality measures, their integration sans normalization could introduce bias toward the index with the broader range. To address this, the researchers applied the Min-Max feature scaling technique to normalize all centrality measures to a uniform range while preserving their relative weight proportions. Building upon the foundations of these four measurements and the topological dimensions they encapsulate, the researchers named their ultimate amalgamation the "Spreading score." This composite metric is indicative of the vertices' potential for information propagation within a network.

$$Spreading_{score_i} = (NC_i^{'} + CR_i^{'})(BC_i^{'} + CI_i^{'}) \quad (5)$$

Where $NC_i^{'}$, $CR_i^{'}$, $BC_i^{'}$, $CI_i^{'}$ are range normalized neighborhood connectivity, ClusterRank, betweenness centrality, and collective influence of node i, respectively.

Among six selected centrality measures, the association between local H index and degree centrality was demonstrated by dependence analyses. According to the significance analysis of the dependence of local H index and degree centrality using CANOVA and Hoeffding's independence tests, these two centrality indices were significantly dependent on each other in the majority of the networks studied. Likewise, the analysis of dependence level using Hoeffding, MIC, and NNS indicated that local H index and degree centrality were dependent on each other with a noticeably high dependence value across the majority of networks. Altogether, using the same mathematical rules and normalization methods explained above, the Addition function was used to combine the effect of local H index and degree centrality [22]. Moreover, considering the same rationale used for the denomination of Spreading score, the additive product of local H index and degree centrality was named as the Hubness score, which could be reflective of the sovereignty of a vertex in it surrounding local territory [22].

$$Hubness_{score_i} = DC_i^{'} + LH_{index_i}^{'} \quad (6)$$

where are $DC_i^{'}$, $LH_{index_i}^{'}$ range normalized degree centrality and local H index of node i, respectively.

Each of the previously described scores, including the Spreading and Hubness scores, embodies a distinct and significant characteristic of each node within the network. The

Hubness score captures a vertex's dominance within its immediate vicinity, whereas the Spreading score is an indicator of its potential to disseminate information. Notably, when the multiplicative product of the Spreading and Hubness scores is higher, it signifies a greater influence exerted by the vertex across the entire network.

To combine the Spreading and Hubness scores and assess their combined impact, the Multiplication function was employed to generate the Integrated Value of Influence (IVI). In essence, the IVI represents the synergistic outcome of essential local centrality measures (degree centrality and ClusterRank), semi-local measures (neighborhood connectivity and local H index), as well as global measures (betweenness centrality and collective influence). Importantly, the IVI calculation method concurrently eliminates positional biases, thus providing a comprehensive perspective on a vertex's influence within the network.

$$IVI_i = (Spreading_{score_i}) \cdot (Hubness_{score_i}) \tag{7}$$

In this report, based on R language programming, computing various centrality measure values among nodes (microbial strains) are implemented, so that the interaction between gut microorganisms can be further studied. The following is the roughly calculation process based on R language programming against the OTU tables sequenced in previous research by 16s rRNA sequencing and shotgun sequencing methods:

Firstly, perform data cleaning, such as removing single samples, removing contaminants, adjusting sample size to match metadata, and adjusting data according to research hypotheses. Next, the microbiome data are aggregated by the specified taxonomic level (e.g., phylum, class, order, family, genus, or species). To analyse the correlation between microorganisms, a method based on symmetric regression and phi statistics was used. Finally, after identifying and visualizing the nodes of microbes with colours at Phyla level and attached name list at Genus level, the results are shown in the figures below:

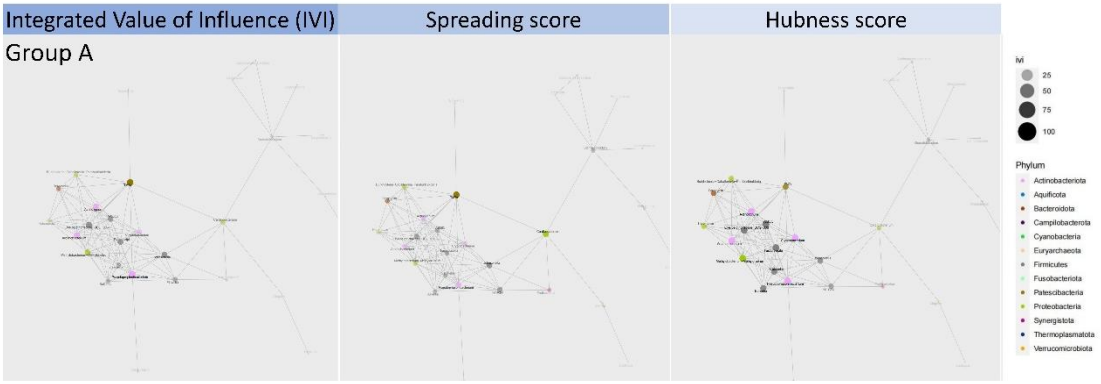

Figure 3: network visualization of three main factors for the gut microbiota of Crohn's disease patients within 6 days EEN treatment (Group A)

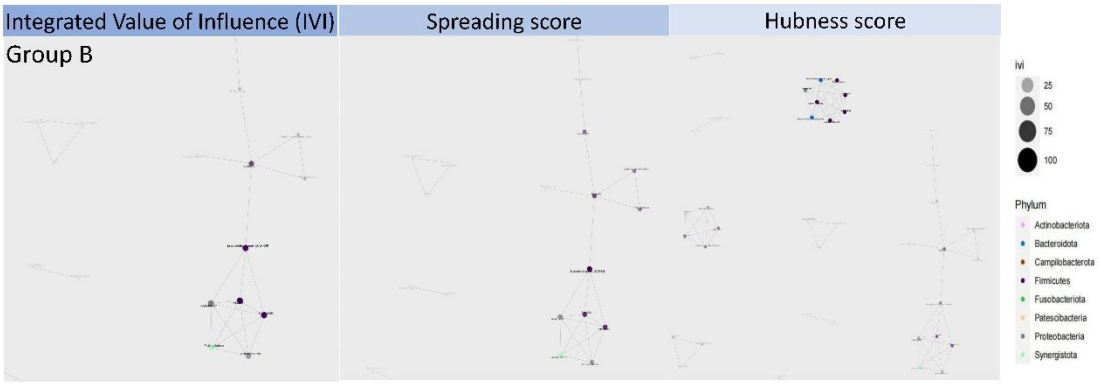

Figure 4: network visualization of three main factors for the gut microbiota of Crohn's disease patients within 16 days EEN treatment (Group B)

Figure 5: network visualization of three main factors for the gut microbiota of Crohn's disease patients within 32 days EEN treatment (Group C)

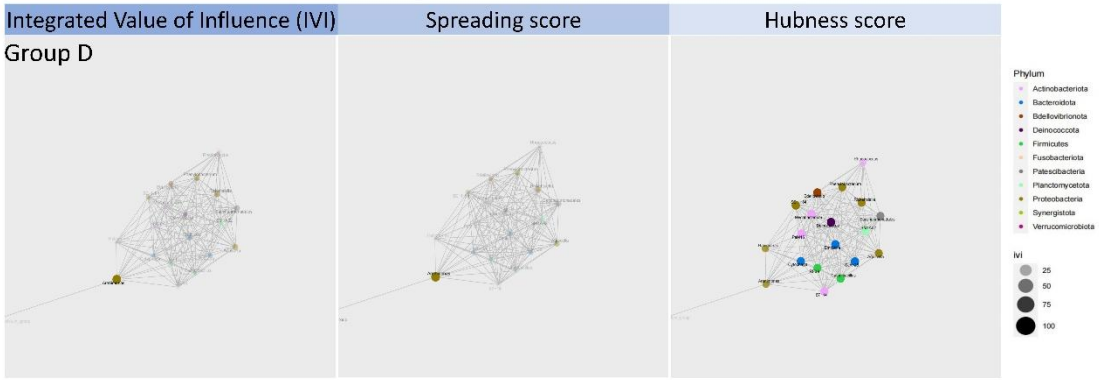

Figure 6: network visualization of three main factors for the gut microbiota of Crohn's disease patients within 54 days EEN treatment (Group D)

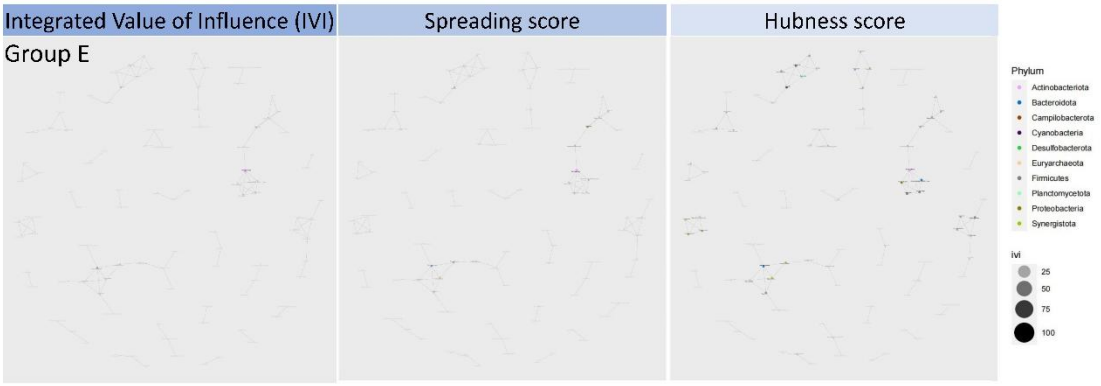

Figure 7: network visualization of three main factors for the gut microbiota of Crohn's disease patients within 63 days EEN treatment (Group E)

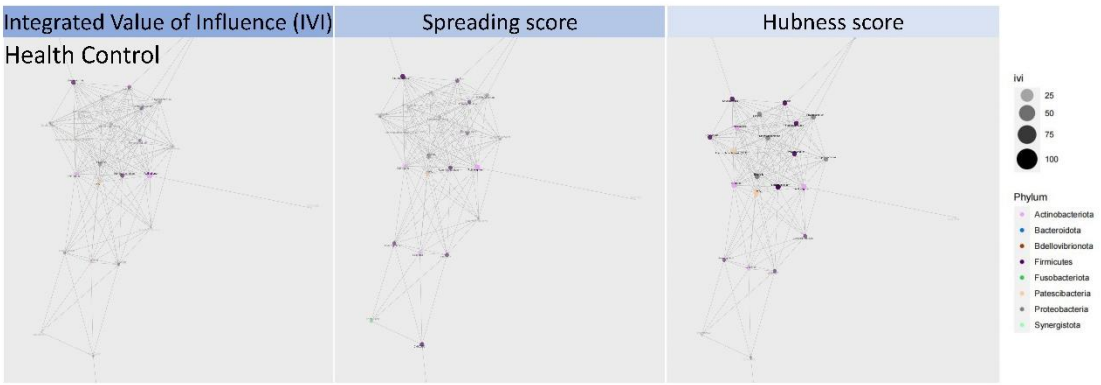

Figure 8: network visualization of three main factors for the gut microbiota of Health Control Group which with no known family history of inflammatory bowel disease

From Figure 1 (Crohn's disease patients within 6 days EEN treatment), we could find two highly influential nodes with the bigger shape than any rest of the nodes in the network, indicating that they have the extremely high IVI score (over 90 marks of 100) than any other nodes, which are as following:

Group A: 1. Genus TM7x (Phylum Patescibacteria; Class Saccharimonadia; Order Saccharimonadales; Family Saccharimonadaceae).

2. Genus Pseudopropionibacterium (Phylum Actinobacteriota; Class Actinobacteria; Order Propionibacteriales; Family Propionibacteriaceae). And for the whole the groups as well as health control group we could reveal nodes with high IVI scores, which are contained in the following table:

Group	Order	IVI score	Kingdom	Phylum	Class	Order	Family	Genus
A	1	100	Bacteria	Patescibacteria	Saccharimonadia	Saccharimonadales	Saccharimonadaceae	TM7x
	2	94.455		Actinobacteriota	Actinobacteria	Propionibacteriales	Propionibacteriaceae	Pseudopropionibacterium
B	1	100		Firmicutes	Clostridia	Christensenellales	Christensenellaceae	Catabacter
	2	100		Firmicutes	Clostridia	Peptostreptococcales-Tissierellales	Peptostreptococcales-Tissierellales	Ezakiella
	3	100		Alphaproteobacteria	Acetobacterales	Acetobacteraceae	Christensenellaceae	Acetobacter
	4	97.117		Firmicutes	Bacilli	Erysipelotrichales	Erysipelotrichaceae	Erysipelotrichaceae_UCG-006
C	1	100		Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Moraxella
D	1	100		Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Arenimonas
E	1	100		Actinobacteriota	Actinobacteria	Micrococcales	Microbacteriaceae	Microbacterium
HC	1	100		Actinobacteriota	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinotignum

Table 2: Influential nodes (microbes) with high IVI score in each group

The dynamic shifts in the predominant microbial composition of the gut in Crohn's disease (CD) patients were distinctly evident as the duration of enteral nutrition (EEN) treatment increased. The table above illustrates these changes at the genus level. Within the initial 6 days of EEN treatment, the central gut microbiota of CD patients comprised TM7x and Pseudopropionibacterium. Subsequently, after 16 days of treatment, the dominant gut microbiota encompassed Catabacter, Ezakiella, Acetobacter, and Erysipelotrichaceae_UCG-006. Remarkably, Catabacter, Ezakiella, and Erysipelotrichaceae_UCG-006 are constituents of the Firmicutes Phylum.

By day 32 of EEN treatment, the core gut microbiota transitioned to Moraxella, belonging to the Proteobacteria Phylum and the Gammaproteobacteria Class. Notably, this transformation echoed the dominant gut microbiome, Arenimonas, observed after 54 days of treatment. Ultimately, after 63 days of EEN treatment, the central microbiome in the CD patients' intestines shifted to Microbacterium. This genus shares the same Actinobacteriota Phylum and Actinobacteria Class with the predominant microbiota found in the intestines of healthy control individuals. These observations highlight the dynamic nature of the gut microbiome during the course of EEN treatment and underscore the significance of these microbial shifts in CD patient recovery.

3.2 Zeta diversity

Curves related to Zeta diversity was drawn based on R language programming for each group including the healthy control group, as shown in the following figures:

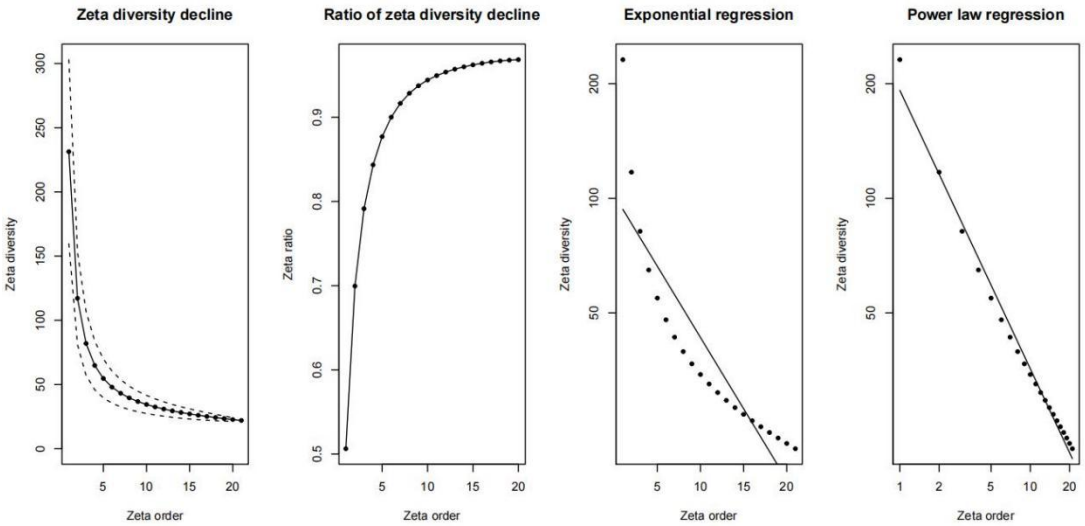

Figure 9: Zeta diversity decline curves and regression for Group A (within 6 days EEN treatment)

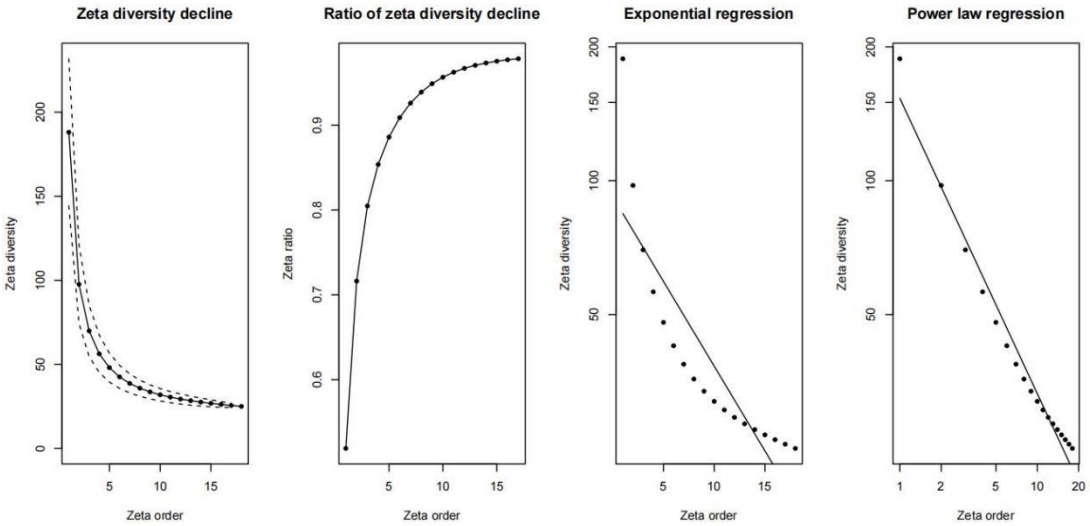

Figure 10: Zeta diversity decline curves and regression for Group B (within 16 days EEN treatment)

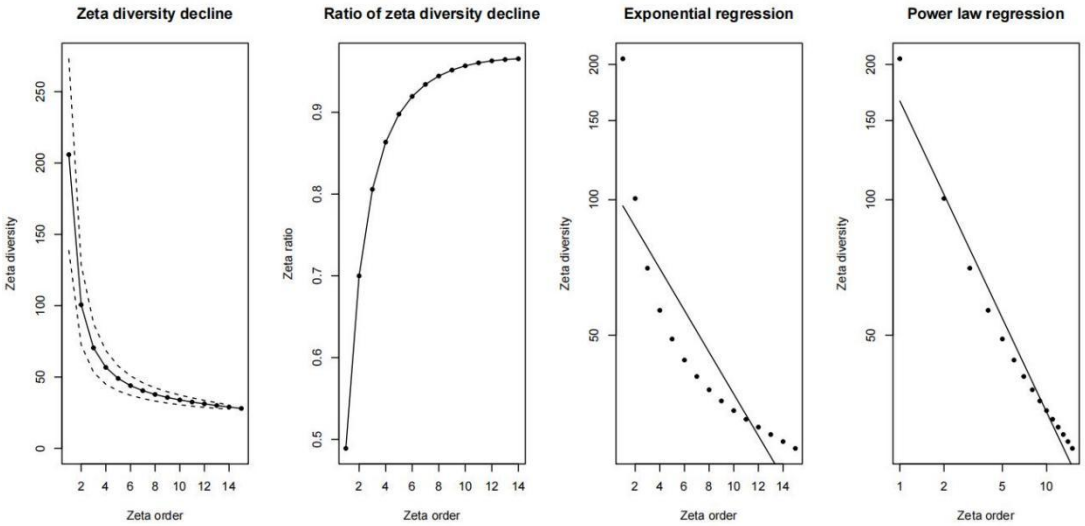

Figure 11: Zeta diversity decline curves and regression for Group C (within 32 days EEN treatment)

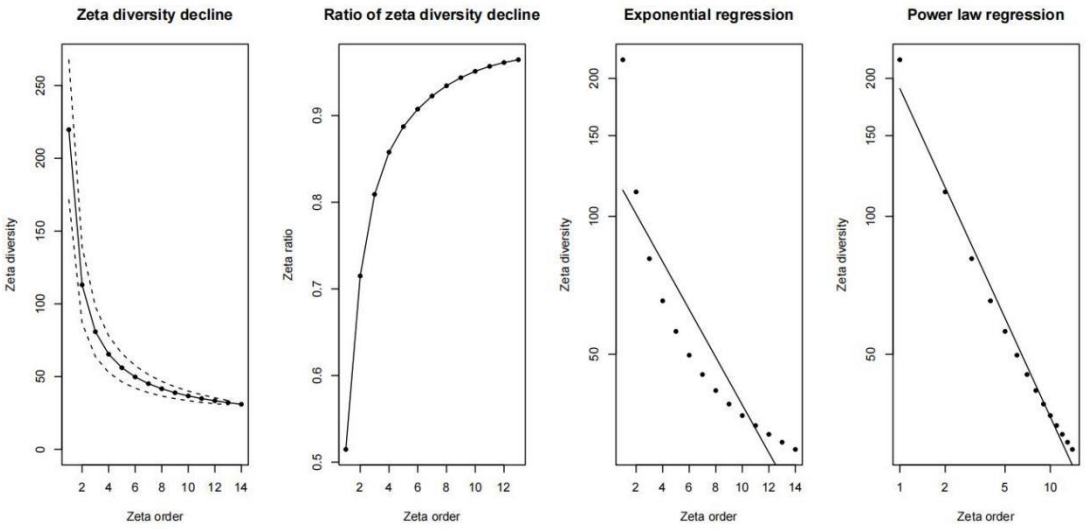

Figure 12: Zeta diversity decline curves and regression for Group D (within 54 days EEN treatment)

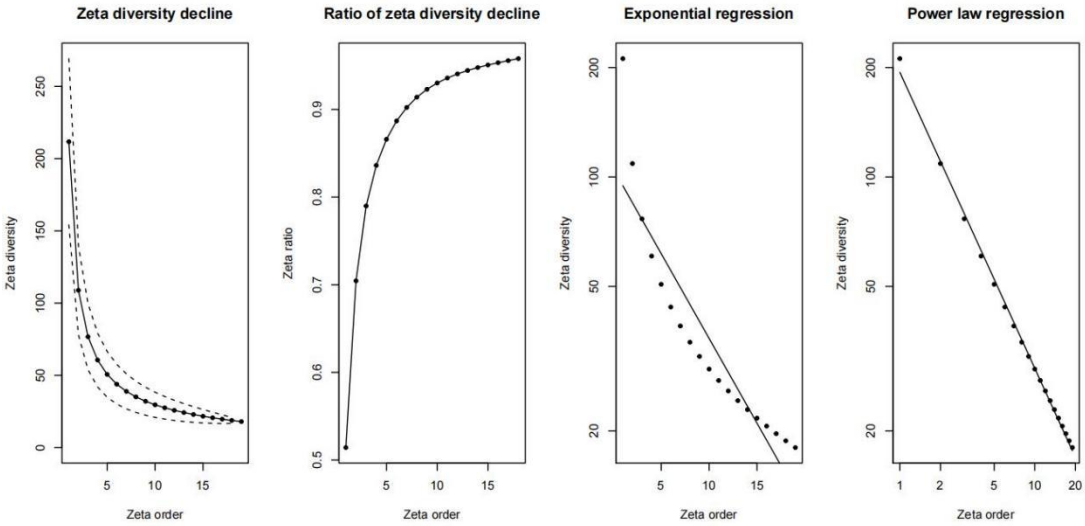

Figure 13: Zeta diversity decline curves and regression for Group E (within 63 days EEN treatment)

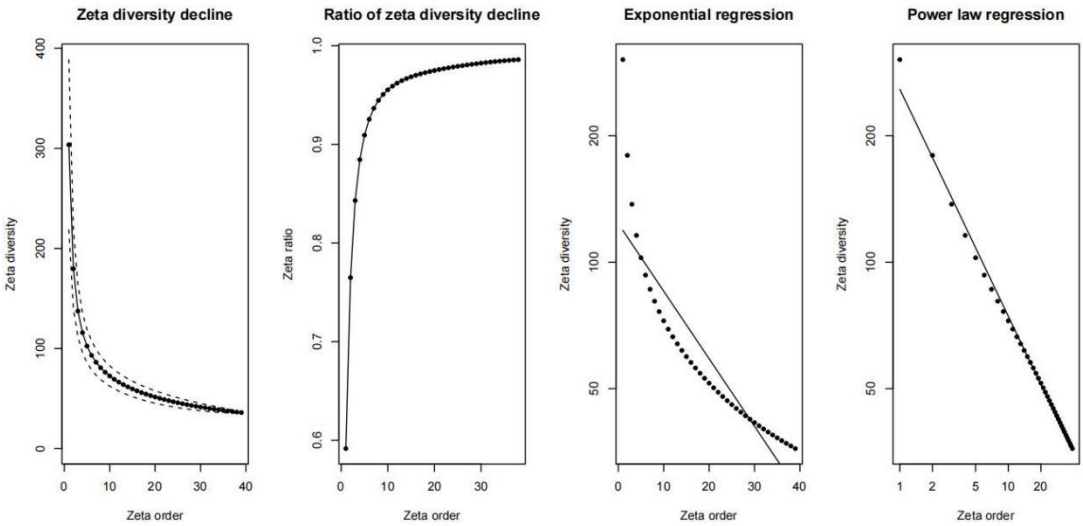

Figure 14: Zeta diversity decline curves and regression for Health Control Group which with no known family history of inflammatory bowel disease

CHAPTER 4

DISSCUSSION

4.1 Bacteria correlated with Crohn's Disease

Serving as the primary gateway to the human body, the oral cavity boasts a highly intricate environment that houses diverse microbial communities encompassing bacteria, fungi, viruses, protozoa, and archaea collectively known as the oral microbiota [25]. These commensal inhabitants of the oral microbiota play a vital role in safeguarding oral health by impeding the colonization of opportunistic pathogens and regulating inflammatory reactions. Nonetheless, any disruption to the equilibrium of this microenvironment can lead to the translocation of certain microorganisms. These microorganisms might find their way to the gastrointestinal tract through mechanisms such as bloodstream circulation or enteral migration. This translocation phenomenon can subsequently contribute to the onset or exacerbation of inflammatory bowel disease (IBD), and in more critical cases, it may even pose a threat to overall health [25].

Despite the distinct environmental conditions that set apart the oral and gut environments, which encompass factors like gastric acidity and antimicrobial bile acids in the duodenum, the findings from the report indicate a noteworthy phenomenon [25]. Specifically, more than half of the microbial species, such as Streptococcus, Actinomyces, and Veillonella Haemophilus, which are commonly identified in both the oral and gut sites, suggest the potential transfer of oral microbes to the intestine, even among individuals in a healthy state [25,26]. This intriguing observation suggests the possibility of oral-gut translocation and ectopic colonization of oral microbes within a healthy human, potentially contributing to the establishment and upkeep of intestinal immunity [25]. However, it's important to note that under specific circumstances, the ectopic colonization of certain oral microorganisms within the intestine could play a role in the development of gastrointestinal disorders, including inflammatory bowel disease (IBD) [27]. Conversely, individuals with IBD also exhibit oral manifestations that can influence the composition of the oral microbiota.

At the Genus level, TM7x and Pseudopropionibacterium were identified in Group A as core microbiomes within the intestines of Crohn's disease (CD) patients undergoing exclusive enteral nutrition (EEN) treatment within a span of 6 days. While TM7x constitutes a minor component of the oral microbiota, it has been observed in human subgingival plaque [25]. Previous studies have underscored the significant role played by the TM7 division in the initial phases of inflammatory mucosal diseases [25]. Comparative analyses of mucosal microbes in inflamed sites of patients with active CD and ulcerative colitis (UC), in comparison to non-inflammatory bowel disease (non -IBD) controls, have revealed that CD patients exhibit a greater diversity of TM7 phylotypes compared to both UC patients and non-IBD controls. Interestingly, the diversity of TM7 in UC patients was approximately on par with that of controls.

It is essential to emphasize that TM7 does not directly trigger inflammation [25]. However, its presence may influence the structure and function of the oral microbial community by impacting the physiology of the bacterial host [25]. This influence

could manifest in the inhibition of host growth or even direct elimination, leading to alterations in its relative abundance and, consequently, contributing to the onset of certain oral diseases [28].

Within Group A, the collection of fecal samples pertains to patients who either received fewer than six days of exclusive enteral nutrition (EEN) treatment or did not undergo any treatment at all. The nodes displaying remarkably high Integrated Value of Influence (IVI) scores identified within the Group A network are strongly implicated in the development of inflammatory symptoms in individuals afflicted with Crohn's disease. This correlation indirectly underscores the potential association of TM7 with Crohn's disease.

On the whole, the available evidence, as gleaned from various studies, strongly suggests a plausible link between TM7 and inflammatory bowel disease (IBD). Nevertheless, the precise mechanisms underlying the role of TM7 as a bacterial trigger in IBD, particularly Crohn's disease, necessitate further comprehensive investigation. It is important to note that while the pathophysiological interactions between TM7 and IBD may differ between ulcerative colitis (UC) and Crohn's disease, the intricate mechanisms driving this association warrant further elucidation.

As the course of treatment progressed, a noticeable shift in the dominant microbial composition of the gut towards Firmicutes was observed, particularly highlighted by the top two microorganisms with the highest Integrated Value of Influence (IVI) scores within the Group B network. This observation aligns with earlier research [32-35], where dietary interventions were separately administered to individuals with Crohn's disease (CD), resulting in elevated Firmicutes concentrations that closely approached levels seen in healthy control subjects.

Firmicutes, being one of the fundamental bacterial phyla present in the human body, holds significant importance. It is recognized as a probiotic that wields a direct influence on the equilibrium of the human gut. The concentration of Firmicutes directly impacts the gut's homeostasis. Its probiotic nature enables it to enhance the intestinal mucosal barrier and bolster immune system functionality. Additionally, it stimulates the secretion of anti-inflammatory factors, which, in turn, restricts the proliferation of detrimental bacteria within the intestinal tract. On the flip side, pathogens can compromise the integrity of the intestinal mucosal barrier, allowing their passage through or causing damage to the mucosal lining. Probiotics, however, can play a pivotal role in preventing or repairing such damage [36].

Within Group B, the progressive administration of exclusive enteral nutrition (EEN) treatment contributed to a transformation in the intestinal microbial landscape of the CD patients. As the dominance of Firmicutes intensified, it began to exert an inhibitory effect on the growth of other potentially harmful microbial species. This gradual shift in the microbial ecosystem was attributed to the growing prevalence of Firmicutes, aiding in the enhancement of the intestinal barrier function.

Earlier investigations have explored interventions involving the administration of Nuciferine and Protopine to individuals diagnosed with Crohn's disease (CD), as documented in previous studies [29-31]. Remarkably, both of these interventions exhibited the ability to ameliorate the composition of the gut microbiota as well as the

functionality of the gut's immune system, resulting in a notable improvement in inflammatory conditions. Notably, during the course of these interventions, specific bacterial groups demonstrated a significant decrease in their concentration, Actinobacteriota being one such example.

These findings underline the potential therapeutic impact of Nuciferine and Protopine in the context of CD management. The observed improvements in gut microbiota composition and enhanced immune function align with previous research, underscoring the importance of exploring alternative treatments to mitigate the inflammatory aspects of Crohn's disease. Furthermore, the decrease in the abundance of Actinobacteriota is indicative of the intricate interactions between specific bacterial groups and their role in modulating inflammatory responses within the gut.

4.2 Variations of Zeta diversity between CD children and Health Controls

In the analysis conducted in Section 3.2, the Zeta diversity correlation curve offers insights into the dynamics of microbial populations across different groups. Notably, the Zeta diversity curve for each group exhibits a consistent monotonic decrease as the number of samples or the zeta order increases. This pattern illustrates that as the sample size or zeta order expands within each group, the count of shared microbial populations among fecal samples diminishes. However, it becomes evident that after a specific zeta order is attained within each group, the decline in the number of shared species levels off.

Upon examining the curves derived from the initial five CD patients subjected to varying durations of EEN treatment, a distinct observation emerges. The Zeta decline curve for these patients appears to stabilize around a zeta order of 15, displaying a power-law decay pattern. Throughout this range, the count of shared species generally remains below 50. Moreover, the trend indicates that as the zeta order surpasses 15, the rate of species retention based on Zeta diversity decline tends to balance. This suggests that common species tend to persist across a larger number of samples compared to rare species.

Contrasting results are observed in the Zeta diversity curve of the healthy control group. Notably, this curve also follows a power-law decline, yet the slope is more gradual compared to that of the CD patient group. Intriguingly, for the healthy control group, the curve exhibits a tendency to flatten after reaching a zeta order of 40, with the number of shared species ranging between 50 and 100. This count is nearly twice that observed for the CD patient group's Zeta diversity decline curve.

Furthermore, it is pertinent to highlight that during the course of EEN treatment, a decreasing trend in the number of shared species becomes evident when the Zeta diversity decline curves flatten across different CD patient groups. This trend underscores a decline in species diversity. Consistent with prior research, successful outcomes of EEN are characterized by the temporary emergence of rare microbial populations commonly found in healthy individuals. This occurrence coincides with a decrease in microbes typically associated with gut homeostasis, ultimately resulting in an overall reduction in microbiota diversity [39]. This aligns with the observed reduction in species diversity depicted by the Zeta diversity curve.

4.3 Mechanism of action of EEN

The Phylum types of all groups are represented graphically, as shown in the following figures:

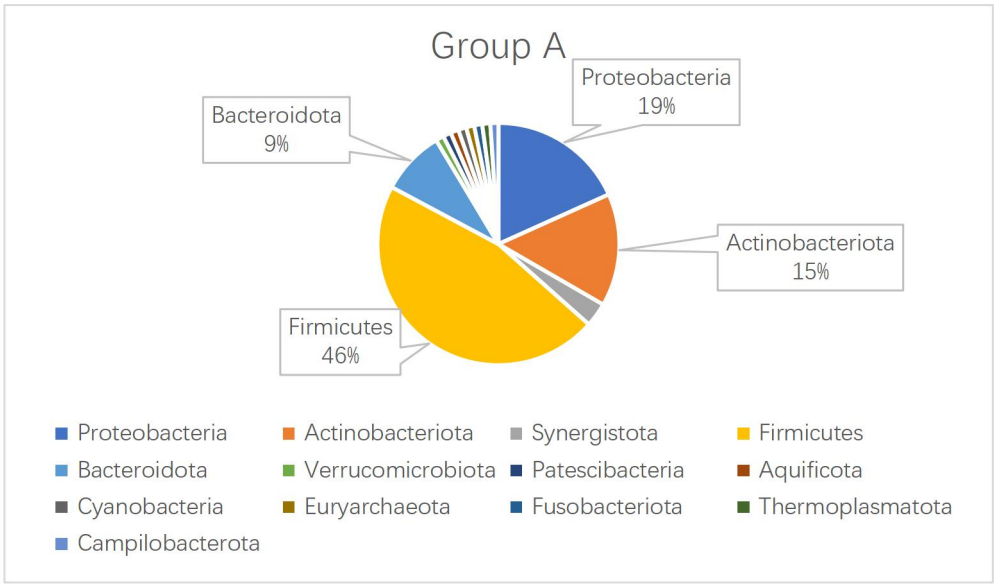

Figure 15: Microbial species composition pie chart for group A at Phylum level (receiving EEN treatment within six days)

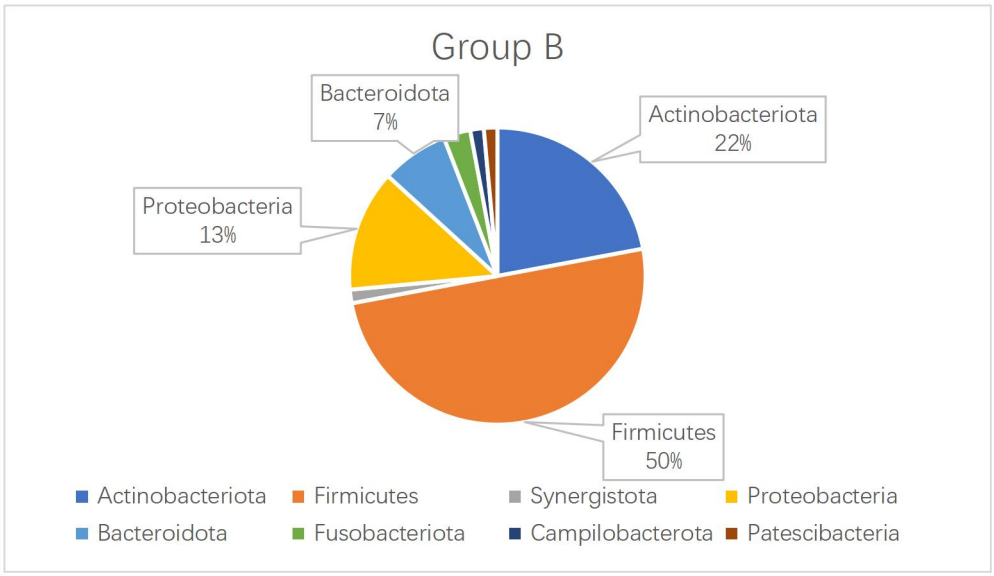

Figure 15: Microbial species composition pie chart at Phylum level for group B (receiving EEN treatment within 16 days)

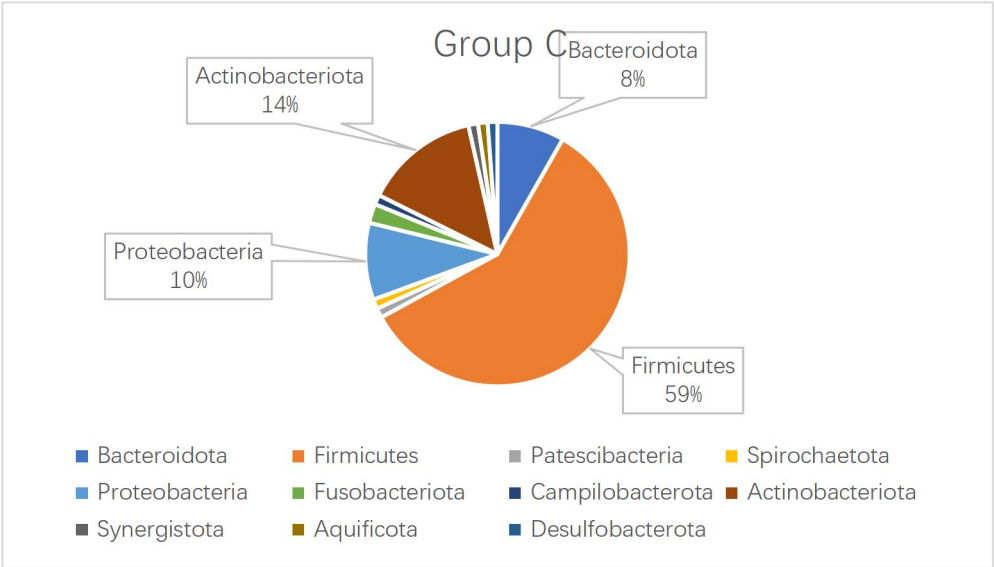

Figure 16: Microbial species composition pie chart at Phylum level for group C (receiving EEN treatment within 32 days)

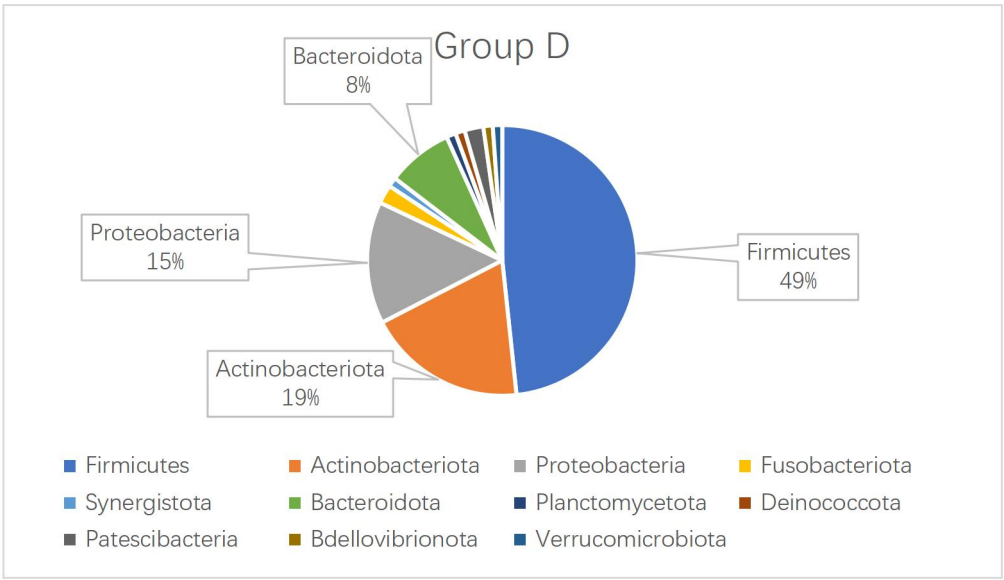

Figure 17: Microbial species composition pie chart at Phylum level for group D (receiving EEN treatment within 54 days)

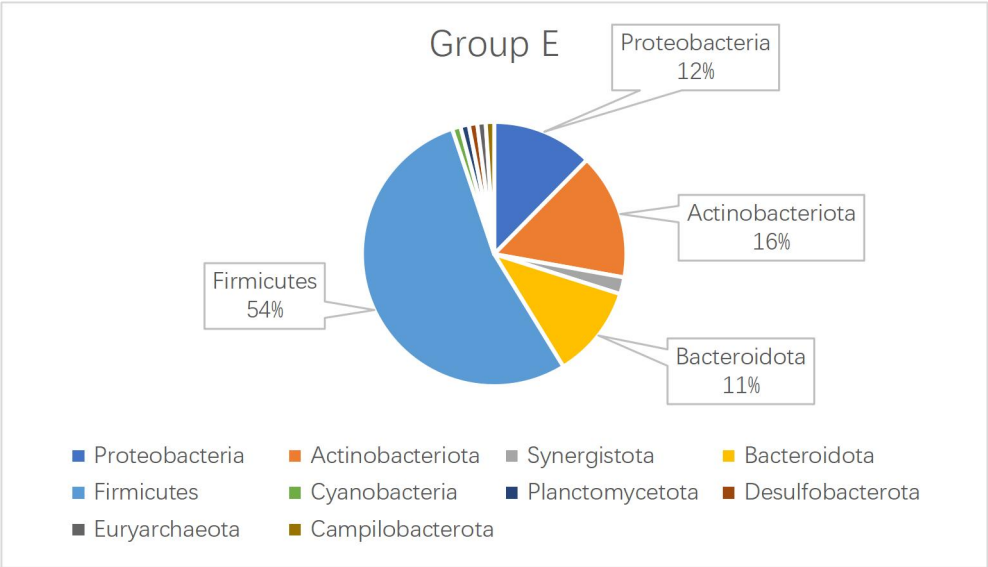

Figure 18: Microbial species composition pie chart at Phylum level for group E (receiving EEN treatment within 63 days)

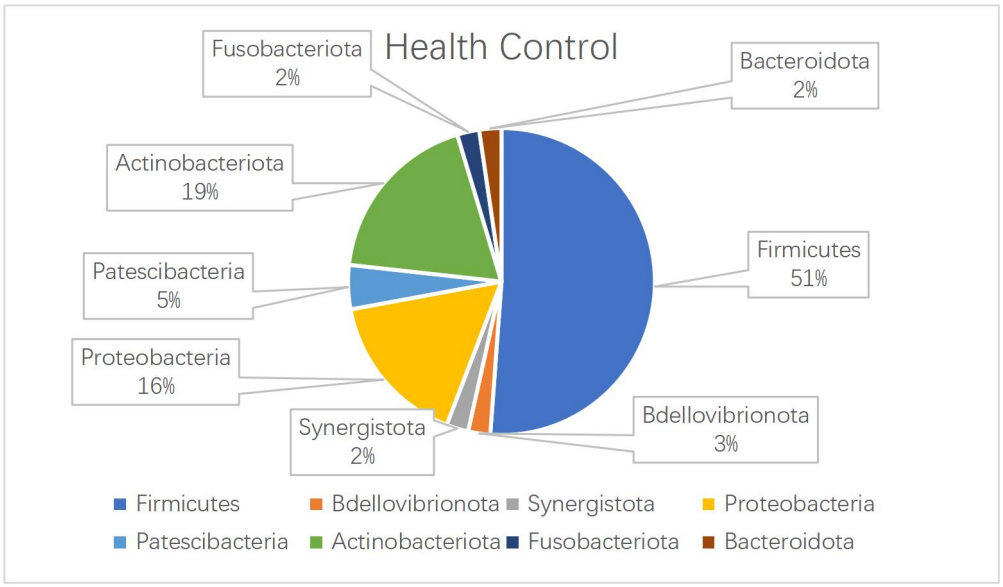

Figure 19: Microbial species composition pie chart at Phylum level for Health Control Group

The furnished diagrams furnish valuable insights into the alterations unfolding during the course of EEN treatment. Remarkably, a discernible pattern emerges, signifying a rise in the relative abundance of bacteria affiliated with the Firmicutes phylum over the duration of EEN treatment. This discernment concurs with precedent investigations (Jingjing Jiang et al., 2022) [35], which demonstrated a favorable association between the prevalence of this bacterial phylum and the concentrations of short-chain fatty acids (SCFA).

Short-chain fatty acids, commonly abbreviated as SCFAs, represent metabolites originating from the microbial fermentation of non-digestible carbohydrates within the gut ecosystem (Muhammad Akhtar et, al, 2021) [37]. These SCFAs hold significant promise as potential therapeutic bioactive molecules for addressing gut inflammatory

diseases. Acting as the outcome of dietary fiber fermentation by gut microbiota, SCFAs possess the capacity to remodel intestinal ecology. Moreover, they play a role in eliciting immunomodulation and exerting antibiotic effects. This suggests their potential to mediate host immune regulation and effectively modulate inflammatory responses during instances of intestinal inflammation.

Short chain fatty acids (SCFAs) are characterized by their saturated nature and their carbon atom content, ranging from 1 to 6 carbon atoms. The specific SCFAs within this group include formate (C1), acetate (C2), propionate (C3), butyrate (C4), valerate (C5), and hexanoate (C6). These SCFAs are synthesized by the gut microbiota through the fermentation of non-digestible carbohydrates, resulting in major end products, primarily acetate, butyrate, and propionate.

It is well-established that SCFA production is reliant on the gut microbiota, and these bioactive compounds play a pivotal role in maintaining host gut homeostasis. The immune system's proper functioning is intricately connected to the gut microbiota and its associated SCFAs. For instance, butyrate has been shown to regulate the production, trafficking, and effective function of both innate and adaptive immune cells in inflammatory bowel disease (IBD), implying its involvement in mitigating intestinal inflammation [38].

Conversely, gut dysbiosis, which represents a disruption in the structure and function of the gut microbiota, exacerbates the development of IBD and disrupts the production and normal function of SCFAs. The therapeutic mechanism of exclusive enteral nutrition (EEN) likely involves enhancing intestinal barrier function and reducing inflammation by exerting regulatory effects on the gut flora and influencing the expression levels of SCFAs.

4.4 Further work suggestions

The current research presented in this report has identified specific core microorganisms that could potentially be associated with the dysbiosis induced by Crohn's disease (CD). However, whether these bacteria hold the key to effectively treating the gut dysbiosis and inflammation caused by CD requires further investigation and in-depth research. Additionally, there is a gap in our understanding of the underlying pathogenic mechanisms of these bacteria in biological contexts, which warrants further exploration.

Similarly, the beneficial bacteria highlighted in the network analysis, which could potentially contribute to enhancing the intestinal environment in patients, need to undergo thorough experimental validation to determine their efficacy in practical treatment scenarios. For instance, the study demonstrated an increase in the abundance of Firmicutes during exclusive enteral nutrition (EEN) treatment, which correlated with the presence of short-chain fatty acids (SCFAs), known to regulate gut flora and alleviate intestinal inflammation. Investigating effective strategies to stimulate the secretion of these substances in patients’ guts is an avenue for further exploration.

In relation to the Zeta diversity analysis, it was observed that EEN appeared to diminish the diversity of the gut microbial population while simultaneously improving symptoms in CD patients. The proposed mechanism behind EEN's effects is based on the hypothesis that EEN may increase the abundance of beneficial intestinal flora

while reducing overall microbial diversity, thereby counteracting the dysbiosis associated with CD and promoting improved intestinal health. In practice, it is important to closely monitor bacterial community diversity and the shifts in abundance of each bacterial group before and after EEN, while taking other influencing factors into account. Furthermore, conducting a thorough analysis of the interplays within these microbial communities is imperative for attaining a more profound comprehension of the observed effects.

CHAPTER 5

CONCLUSION

This report initiates the exploration of interconnections among microorganisms within groups A, B, C, D, E, and the healthy control (HC) group by constructing and analysing a network. The Integrated Value of Influence (IVI) is calculated for each node in the network, ultimately leading to a visual representation of the network. Through the comparative examination of these depicted networks, a number of bacterial species have been discerned, potentially linked to the disruption of the intestinal microbiota balance and the inflammatory processes associated with Crohn's disease (CD). For instance, at the Genus level, TM7x emerges as a potentially significant player in the early stages of inflammatory intestinal mucosal diseases. Similarly, at the Phylum level, the prevalence of Firmicutes as a dominant gut microbiota for a certain duration emerges as noteworthy, given its positive correlation with short-chain fatty acids that hold potential in enhancing and regulating gut barrier function.

Furthermore, insights gained from the Zeta diversity curve and regression analyses indicate a general decline in the microbiota diversity within the intestines of CD patients undergoing exclusive enteral nutrition (EEN) treatment. Significantly, it's worth noting that the Zeta diversity within the CD patient cohort appears to be lower compared to that of the healthy control (HC) group. This intriguing observation suggests that EEN treatment could potentially modulate the symptoms of CD by concomitantly diminishing the overall diversity of the gut microbiota while promoting the proliferation of select bacterial populations. These bacteria could produce beneficial metabolites, enhancing overall human health, while suppressing the growth of potentially harmful bacteria that hinder intestinal function recovery. This dual approach of boosting beneficial bacteria while inhibiting the proliferation of detrimental ones might be central to the efficacy of EEN in ameliorating CD symptoms.

However, to comprehend the intricate mechanisms underlying EEN's effects, a more in -depth investigation is crucial. This involves closely monitoring shifts in bacterial community diversity and the abundance alterations in each bacterial community before and after EEN, while considering various influencing factors. Analysing the intricate interplay among these microbial communities is essential to unravel the precise mechanisms by which EEN leads to its observed outcomes in CD patients.

REFERENCES

[1] Yu, Y., Chen, K. & Chen, J. 2019, "Exclusive enteral nutrition versus corticosteroids for treatment of pediatric Crohn's disease: a meta-analysis", World journal of pediatrics : WJP, vol. 15, no. 1, pp. 26-36.
[2] Diederen, K., Li, J.V., Donachie, G.E., de Meij, T.G., de Waart, D.R., Hakvoort, T.B.M., Kindermann, A., Wagner, J., Auyeung, V., te Velde, A.A., Heinsbroek, S.E.M., Benninga, M.A., Kinross, J., Walker, A.W., de Jonge, W.J. & Seppen, J. 2020, "Exclusive enteral nutrition mediates gut microbial and metabolic changes that are associated with remission in children with Crohn's disease", Scientific reports, vol. 10, no. 1, pp. 18879.
[3] Kowalska-Duplaga, K., Gosiewski, T., Kapusta, P., Sroka-Oleksiak, A., Wędrychowicz, A., Pieczarkowski, S., Ludwig-Słomczyńska, A.H., Wołkow, P.P. & Fyderek, K. 2019, "Differences in the intestinal microbiome of healthy children and patients with newly diagnosed Crohn's disease", Scientific reports, vol. 9, no. 1, pp. 18880-11.
[4] Guo, D., Fang, L., Liu, R., Li, Y., Lv, L., Niu, Z., Chen, D., Zhou, Y. & Zhu, W. 2022, "Exploring Different Effects of Exclusive Enteral Nutrition (EEN) and Corticosteroids on the Gut Microbiome in Crohn's Disease Based on a Three-Stage Strategy", Gastroenterology research and practice, vol. 2022, pp. 1-8.
[5] Quince, C., Ijaz, U.Z., Loman, N., Eren, A.M., Saulnier, D., Russell, J., Haig, S.J., Calus, S.T., Quick, J., Barclay, A., Bertz, M., Blaut, M., Hansen, R., McGrogan, P., Russell, R.K., Edwards, C.A. & Gerasimidis, K. 2015, "Extensive Modulation of the Fecal Metagenome in Children With Crohn's Disease During Exclusive Enteral Nutrition", The American journal of gastroenterology, vol. 110, no. 12, pp. 1718-1729.
[6] Ding, N.S., McDonald, J.A.K., Perdones-Montero, A., Rees, D.N., Adegbola, S.O., Misra, R., Hendy, P., Penez, L., Marchesi, J.R., Holmes, E., Sarafian, M.H. & Hart, A.L. 2020, "Metabonomics and the Gut Microbiome Associated With Primary Response to Anti-TNF Therapy in Crohn's Disease", Journal of Crohn's and colitis, vol. 14, no. 8, pp. 1090-1102.
[7] Johnson, J.S., Spakowicz, D.J., Hong, B., Petersen, L.M., Demkowicz, P., Chen, L., Leopold, S.R., Hanson, B.M., Agresta, H.O., Gerstein, M., Sodergren, E. & Weinstock, G.M. 2019, "Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis", Nature communications, vol. 10, no. 1, pp. 5029-11.
[8] Gerasimidis, K., Bertz, M., Hanske, L., Junick, J., Biskou, O., Aguilera, M., Garrick, V., Russell, R.K., Blaut, M., McGrogan, P. & Edwards, C.A. 2014, "Decline in presumptively protective gut bacterial species and metabolites are paradoxically associated with disease improvement in pediatric Crohn's disease during enteral nutrition", Inflammatory bowel diseases, vol. 20, no. 5, pp. 861-871.
[9] Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N. & Knight, R. 2011, "Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample", Proceedings of the National Academy of Sciences - PNAS, vol. 108, no. Supplement 1, pp. 4516-4522.
[10] Han, D., Gao, P., Li, R., Tan, P., Xie, J., Zhang, R. & Li, J. 2020, "Multicenter assessment of microbial community profiling using 16S rRNA gene sequencing and shotgun metagenomic sequencing", Journal of advanced research, vol. 26, pp. 111-121.
[11] Quince, C., Walker, A.W., Simpson, J.T., Loman, N.J. & Segata, N. 2017, "Shotgun metagenomics, from sampling to analysis", Nature biotechnology, vol. 35, no. 9, pp. 833-844.

[12] Beiko, R.G., Hsiao, W., Parkinson, J. & SpringerLink (Online service) 2018, Microbiome Analysis: Methods and Protocols, Springer New York, New York, NY.
[13] Magalhães, A.P., Azevedo, N.F., Pereira, M.O. & Lopes, S.P. 2016, "cystic fibrosis microbiome in an ecological perspective and its impact in antibiotic therapy", Applied microbiology and biotechnology, vol. 100, no. 3, pp. 1163-1181.
[14] Dalton, T., Dowd, S.E., Wolcott, R.D., Sun, Y., Watters, C., Griswold, J.A. & Rumbaugh, K.P. 2011, "An in vivo polymicrobial biofilm wound infection model to study interspecies interactions", PloS one, vol. 6, no. 11, pp. e27317.
[15] Murray, J.L., Connell, J.L., Stacy, A., Turner, K.H. & Whiteley, M. 2014, "Mechanisms of synergy in polymicrobial infections", The journal of microbiology, vol. 52, no. 3, pp. 188-199.
[16] Chen, E.Z. & Li, H. 2016, "A two-part mixed-effects model for analyzing longitudinal microbiome compositional data", Bioinformatics (Oxford, England), vol. 32, no. 17, pp. 2611-2617.
[17] Kurtz, Z.D., Müller, C.L., Miraldi, E.R., Littman, D.R., Blaser, M.J. & Bonneau, R.A. 2015, "Sparse and Compositionally Robust Inference of Microbial Ecological Networks: e1004226", PLoS computational biology, vol. 11, no. 5.
[18] Lovell, D., Pawlowsky-Glahn, V., Egozcue, J.J., Marguerat, S. & Bähler, J. 2015, "Proportionality: A Valid Alternative to Correlation for Relative Data: e1004075", PLoS computational biology, vol. 11, no. 3.
[19] McGeoch, M.A., Latombe, G., Andrew, N.R., Nakagawa, S., Nipperess, D.A., Roigé, M., Marzinelli, E.M., Campbell, A.H., Vergés, A., Thomas, T., Steinberg, P.D., Selwood, K.E., Henriksen, M.V. & Hui, C. 2019, "Measuring continuous compositional change using decline and decay in zeta diversity", Ecology (Durham), vol. 100, no. 11, pp. 1-18.
[20] Simons, A.L., Mazor, R., Stein, E.D. & Nuzhdin, S. 2019, "Using alpha, beta, and zeta diversity in describing the health of stream-based benthic macroinvertebrate communities", Ecological applications, vol. 29, no. 4, pp. e01896-n/a.
[21] Hui, C. & McGeoch, M.A. 2014, "Zeta Diversity as a Concept and Metric That Unifies Incidence-Based Biodiversity Patterns", The American naturalist, vol. 184, no. 5, pp. 684-694.
[22] Salavaty, A., Ramialison, M. & Currie, P.D. 2020, "Integrated Value of Influence: An Integrative Method for the Identification of the Most Influential Nodes within Networks", Patterns (New York, N.Y.), vol. 1, no. 5, pp. 100052-100052.
[23] Volland, J. 2023, "Small cells with big secrets", Nature reviews. Microbiology, vol. 21, no. 7, pp. 414-414.
[24] Abdelbary, M.M.H., Hatting, M., Bott, A., Dahlhausen, A., Keller, D., Trautwein, C. & Conrads, G. 2022, "The oral-gut axis: Salivary and fecal microbiome dysbiosis in patients with inflammatory bowel disease", Frontiers in cellular and infection microbiology, vol. 12, pp. 1010853.
[25] Han, Y., Wang, B., Gao, H., He, C., Hua, R., Liang, C., Xin, S., Wang, Y. & Xu, J. 2022, "Insight into the Relationship between Oral Microbiota and the Inflammatory Bowel Disease", Microorganisms (Basel), vol. 10, no. 9, pp. 1868.
[26] Schmidt, T.S., Hayward, M.R., Coelho, L.P., Li, S.S., Costea, P.I., Voigt, A.Y., Wirbel, J., Maistrenko, O.M., Alves, R.J., Bergsten, E., de Beaufort, C., Sobhani, I., Heintz-Buschart, A., Sunagawa, S., Zeller, G., Wilmes, P. & Bork, P. 2019, "Extensive transmission of microbes along the gastrointestinal tract", eLife, vol. 8.
[27] Kitamoto, S., Nagao-Kitamoto, H., Hein, R., Schmidt, T.M. & Kamada, N. 2020, The Bacterial Connection between the Oral Cavity and the Gut Diseases, SAGE Publications, Los Angeles, CA.

[28] Bor, B., Bedree, J.K., Shi, W., McLean, J.S. & He, X. 2019, Saccharibacteria (TM7) in the Human Oral Microbiome, SAGE Publications, Los Angeles, CA.
[29] Ponce-Alonso, M., García-Hoz, C., Halperin, A., Nuño, J., Nicolás, P., Martínez-Pérez, A., Ocaña, J., García-Pérez, J.C., Guerrero, A., López-Sanromán, A., Cantón, R., Roy, G. & del Campo, R. 2021, "An Immunologic Compatibility Testing Was Not Useful for Donor Selection in Fecal Microbiota Transplantation for Ulcerative Colitis", Frontiers in immunology, vol. 12, pp. 683387-683387.
[30] Yue, M., Huang, J., Ma, X., Huang, P., Liu, Y. & Zeng, J. 2023, "Protopine Alleviates Dextran Sodium Sulfate-Induced Ulcerative Colitis by Improving Intestinal Barrier Function and Regulating Intestinal Microbiota", Molecules (Basel, Switzerland), vol. 28, no. 13, pp. 5277.
[31] Zhu, Y., Zhao, Q., Huang, Q., Li, Y., Yu, J., Zhang, R., Liu, J., Yan, P., Xia, J., Guo, L., Liu, G., Yang, X. & Zeng, J. 2022, "Nuciferine Regulates Immune Function and Gut Microbiota in DSS-Induced Ulcerative Colitis", Frontiers in veterinary science, vol. 9, pp. 939377-939377.
[32] Verburgt, C.M., Dunn, K.A., Ghiboub, M., Lewis, J.D., Wine, E., Sigall Boneh, R., Gerasimidis, K., Shamir, R., Penny, S., Pinto, D.M., Cohen, A., Bjorndahl, P., Svolos, V., Bielawski, J.P., Benninga, M.A., de Jonge, W.J. & van Limbergen, J.E. 2023, "Successful Dietary Therapy in Paediatric Crohn's Disease is Associated with Shifts in Bacterial Dysbiosis and Inflammatory Metabotype Towards Healthy Controls", Journal of Crohn's and colitis, vol. 17, no. 1, pp. 61-72.
[33] Starz, E., Wzorek, K., Folwarski, M., Kaźmierczak-Siedlecka, K., Stachowska, L., Przewłócka, K., Stachowska, E. & Skonieczna-Żydecka, K. 2021, "The Modification of the Gut Microbiota via Selected Specific Diets in Patients with Crohn’s Disease", Nutrients, vol. 13, no. 7, pp. 2125.
[34] Van Limbergen, J., Dunn, K., Wine, E., Sigall Boneh, R., Bielawski, J. & Levine, A. 2020, "OP22 Crohn’s disease exclusion diet reduces bacterial dysbiosis towards healthy controls in paediatric Crohn’s disease", Journal of Crohn's and colitis, vol. 14, no. Supplement_1, pp. S019-S020.
[35] Jiang, J., Chen, L., Chen, Y. & Chen, H. 2022, "Exclusive enteral nutrition remodels the intestinal flora in patients with active Crohn's disease", BMC gastroenterology, vol. 22, no. 1, pp. 212-212.
[36] Shen, Z., Zhu, C., Quan, Y., Yang, Z., Wu, S., Luo, W., Tan, B. & Wang, X. 2018, "Relationship between intestinal microbiota and ulcerative colitis: Mechanisms and clinical application of probiotics and fecal microbiota transplantation", World journal of gastroenterology : WJG, vol. 24, no. 1, pp. 5-14.
[37] Akhtar, M., Chen, Y., Ma, Z., Zhang, X., Shi, D., Khan, J.A. & Liu, H. 2022, "Gut microbiota-derived short chain fatty acids are potential mediators in gut inflammation", Animal Nutrition, vol. 8, pp. 350-360.
[38] Gonçalves, P., Araújo, J.R. & Di Santo, J.P. 2018, "A Cross-Talk Between Microbiota-Derived Short-Chain Fatty Acids and the Host Mucosal Immune System Regulates Intestinal Homeostasis and Inflammatory Bowel Disease", Inflammatory bowel diseases, vol. 24, no. 3, pp. 558-572.
[39] Runde, J., Veseli, I., Fogarty, E.C., Watson, A.R., Clayssen, Q., Yosef, M., Shaiber, A., Verma, R., Quince, C., Gerasimidis, K., Rubin, D.T. & Eren, A.M. 2023, "Transient Suppression of Bacterial Populations Associated with Gut Health is Critical in Success of Exclusive Enteral Nutrition for Children with Crohn's Disease", Journal of Crohn's and colitis, vol. 17, no. 7, pp. 1103-1113.