T.R.
İSTANBUL UNIVERSITY
INSTITUTE OF SCIENCES

Ph.D. THESIS

APPLICATION OF METAGENOMICS APPROACHES IN ULCERATIVE COLITIS AND CROHN´S DISEASES

Omar SAJER NASER

Department of Molecular Biology and Genetics

Programme of Biotechnology

ADVISOR
Prof. Dr. Ercan ARICAN

Ağustos, 2023

İSTANBUL

This work is accepted as a Ph.D. Thesis in **Department of Molecular Biology and Genetics**
Department **Programme of Biotechnology** Program by the following jury in 15.08.2023

Thesis Jury

Prof. Dr. Ercan Arican(Advisor)
İstanbul University
Faculty of Science

Prof. Dr. Veysel Sabri HANÇER
Istınye University
Faculty of Medicine

Assoc.Prof. Murat PEKMEZ
İstanbul University
Faculty of Sciences

Assoc.Prof. Emre YÖRÜK
İstanbul Yeni Yüzyıl Üniversitesi
Faculty of Sciences

Assoc.Prof. Cenk KIĞ
İstanbul University
Faculty of Sciences

Ethic Declaration

According to the 9/2 and 22/2 articles of the Graduate Education Regulation, which was published in the Official Gazette on April 20, 2016, this graduate thesis is reported as meeting the criteria established by the Institute of Sciences using the plagiarism software to which Istanbul University is a subscriber..

Research Fund

This thesis was funded by the Istanbul University Scientific Research Projects Executive Secreteriat project with the number 36460..

Publications Produced From Thesis

Applying metagenomics techniques to treat Crohn's disease

Tarih ve Sayı: 03.06.2021-228333

T.C.
İSTANBUL ÜNİVERİTESİ
İstanbul Tıp Fakültesi Dekanlığı

Sayı :E-29624016-050.99-228333
Konu :Prof. Dr. Ercan ARICAN hk.

Sayın Prof. Dr. Ercan ARICAN
Fen Fakültesi

İlgi : Fen Fakültesi Dekanlığının 17/11/2020 gün ve 198962 sayılı yazısı

Sorumlu araştırıcılığını üstlendiğiniz ve Omar Sajer Naser NASER' in yürüteceği 2020/1762 dosya numaralı "Ülsretif Kolit ve Crohn Hastalıklarında Metagenomik Yaklaşmaların Uygulanması" başlıklı çalışma, kurulumuzun 28/05/2021 tarih ve 11 sayılı toplantısında görüşülerek etik yönden uygun bulunmuş olup, tutanaklar ekte sunulmuştur.

Bilgilerinizi rica ederim.

Prof. Dr. Ali Yağız ÜRESİN
Kurul Başkanı

PREFACE

I would like to express my gratitude to everyone who contributed to and assisted me in finishing this research effort by offering suggestions or recommendations.

My supervisor Prof. Dr. Ercan ARICAN provided guidance, support, and inspiration for my effort. I would like to express my gratitude to the educational staff for their help throughout fieldwork and lectures.

Since this work was supported by the Scientific research projects Coordination Unit of Istanbul University, we would like to thank Istanbul University and the Bilimsel Araştrma Projeleri Süreçleri Yönetim Sistemi (BASS) for their assistance. The project is 36460.

Great gratitude and appreciation to my FAMILY for their unwavering support during my years of study away home and for their tolerance, patience, and sincerity.

Ağustos 2023 Omar SAJER NASER

TABLE OF CONTENTS

Page Number

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

Symbols **Explanations**

a : Aaaa

b : Aaaa

c : Aaaa

ds : Aaaa

es : Aaaa

F : Aaaa

Abbreviations **Explanations**

BCL3: B cell lymphoma 3 protein

BSG: British society of gastroenterology

CD: Crohn’s disease

CRP : C-reactive protein

CSSC: Clinical services and standards committee

EC : Endometrial cancer

EIMs: Extraintestinal manifestations

GI: Gastrointestinal

GO: Gene ontology

GWAS: Genome-wide association studies

hASCs : Human adipose stem cells

HLA : Human leukocyte

IBD: Inflammatory bowel disease

mNGS : Metagenomics of next-generation sequencing

NGS : Next-generation sequencing

NOD2: Nucleotide-binding oligomerization domain 2

OSM: Oncostatin-M

PBCs : Peripheral blood cells

PBLs: Peripheral blood leukocytes

PBMCs : Peripheral blood mononuclear cells

PCR : Polymerase chain reaction

PPR: Pathogen Pattern Recognition

QOL: Quality of life

SCFA: Short-chain fatty acids

STAT3: Signal transducer and activator of transcription 3

STAT5: Signal transducer and activator of transcription 5

TGS : Third-generation sequencing

TWAS: Transcriptome-wide association studies

UC: Ulcerative colitis

UCEIS : Ulcerative colitis endoscopic index of severity

ABSTRACT

APPLICATION OF METAGENOMICS APPROACHES IN ULCERATIVE COLITIS AND CROHN´S DISEASES

THESIS Ph.D.

Omar SAJER NASER

İstanbul University

Institute of Sciences

Department of Molecular Biology and Genetics

ADVISOR
Prof. Dr. Ercan ARICAN

No official treatment exists for the complex, polygenic disease known as Inflammatory Bowel Disease (IBD). Patients are more likely to develop cancer and other diseases that require lifelong monitoring, expensive pharmaceuticals, and difficult procedures. Crohn's disease (CD) and ulcerative colitis (UC) are both types of IBD. IBD is unclear about its specific origin, however, individuals with a genetic predisposition to have an anomalous response to environmental triggers in the intestinal flora (such as microbiota). Since the early detection of colon cancer is difficult, the majority of patients will receive their diagnosis later on. Despite

this, they will experience serious damage to the colon and will have a hard time receiving treatment early on.

The purpose of this thesis was to investigate biomarkers that could be used to assess the risk level of IBD (specifically Ulcerative colitis and Crohn's diseases) in patients, and identify those who are at risk of the disease's severity increasing in advance. Initially, mutations in human patients with IBD were studied at locations that were prospective, including the Nucleotide-binding oligomerization domain-containing protein 2 (NOD2), the innate immunity gene, and the IL-23R gene that is associated with immunity. Next, NGS and PCR were employed to discover how patient's microbiome is altered. Statistical methods were employed to determine the microbiota as well as to assess the differences between the disease states and the control, these methods were then compared to previous studies. This research demonstrated a decrease in bacterial diversity in ulcerative colitis, Crohn's Disease, and a control group. At the genus level, the CD, UC and the control groups in the bacteria (100%) of Proteobacteria. Additionally, the Alphapeoteobacteria demonstrated other significant discrepancies with the levels in UC8 (17.4%) being the greatest and CD3 (0.7%) having the lowest total volume (3.6%). Additionally, CD7 was documented in Bacteroidia (0.2%), Clostridia (0.1%), and Negativicutes (0.4%) which were not documented in the other classes in the taxonomy. Additionally, the variation in UC8 is the greatest of any other group, but it isn't the highest of all bacteria. It demonstrates a percentage of Gammaproteobacteria (Pseudomonadales) (44.5%), Betaproteobacteria (Burkholderiales) (38.1%), Alphaproteobacteria (Sphingomonadales) (16.2%), and Alphaproteobacteria (Hyphomicrobiales) (1.2%). The higher increase in Gammaproteobacteria (Pseudomonadales) was in CD3 (88.6%), and the lowest was in UC8 (44.5%) and the total (75.6%). Ultimately, converting significant findings into a molecular test kit can serve as a molecular framework that helps gastroenterologists plan individualized treatments for IBD patients.

Ağustos 2023, 117 pages.

Keywords: Ulcerative colitis, Crohn’s Disease, Metagenomics, Gastrointestinal.

SUMMARY

APPLICATION OF METAGENOMICS APPROACHES IN ULCERATIVE COLITIS AND CROHN'S DISEASES

Ph.D. THESIS

Omar SAJER NASER

İstanbul University

Institute of Sciences

Department of Molecular Biology and Genetics

Supervisor : Prof. Dr. Ercan ARICAN

The intestinal microbial community is acknowledged to have a significant impact on the equilibrium of gut homeostasis. Variations in the composition of the microbiota groups, which are often referred to as dysbiosis, can negatively affect health and increase the chance of disease. Despite the fact that alterations to the microbiome of the gut have been associated with various diseases, it is still not clear if these changes are caused by or simply the result of the disorders.

A significant health threat for people is IBD, a chronic condition of the GI tract that triggers both active and passive inflammation. It's segregated into two classes: Ulcerative Colitis (UC)

and Crohn's Disease (CD), the latter of which has a similar symptomatology, but a different presentation in the clinical sphere. IBD is frequently linked to a alteration in the composition of the microbiota in the gut. On the microbiological composition of the disorders, there is, however, not a universal agreement. The primary differences between UC and CD are what cause the discrepancies in research. Previous studies have demonstrated that the microbial composition of UC is different than that of CD, and is similar to the composition of healthy individuals. However, some research studies consider both populations as separate entities and have found significant differences in the microbial patterns of UC, CD, and healthy individuals.

In order to determine whether the composition and function of the microbiome in the stomach could be distinguished between CD and UC patients, and whether or not data on the microbiome could be used as a means of diagnosis and predictions, the purpose of this thesis was to discover dysbiosis in a cohort of Turkish IBD patients. DNA extraction, 16sRNA gene PCR amplification that identifies the microbiota, and the analysis of the amplified regions via sequencing techniques.

We discovered dysbiosis in IBD patients, as expected. Additionally, we found that CD and UC had different microbial compositions, with UC having a larger dysbiosis than CD. Additionally, it was found that UC has more dysbiosis than CD and that UC has more genes associated with metabolic pathways and immunological problems than CD and healthy individuals.

August 2023, 117 pages.

Keywords: Ulcerative colitis, Crohn’s Disease, Metagenomics, Gastrointestinal.

1. INTRODUCTION

inflammatory bowel disease (IBD), which is a condition that affects the gastrointestinal tract (GI) system, has a damaging effect on the intestinal system. This debilitating disease, that affects both the structure and function of the GI tract and intestinal lining, necessitates long-term treatment. The cause of IBD has been attributed to multiple factors, including urbanization, westernization, dietary changes, increased antibiotic utilization, and disregards to the equilibrium between host and microbes. Based on the symptoms of disease and the way it is pathologically characterized, IBD can be classified as Ulcerative Colitis (UC) or Crohn's Disease (CD). This division suggests that UC and CD are separate medical entities that require different treatment methods (Elhag D.A, et al, 2022).

Additionally, while the two main types may be considered separate diseases, they have similar patterns. UC primarily harms the rectum and the surrounding tissue of the colon, whereas CD can affect any part of the digestive system, including the mouth, perineo-spincterian region, and the rear end. Several factors contribute to the development of IBD, including genetic inheritance and the microbiome. Notably, variations in SNPs associated with both innate and adaptive immunity have a significant impact on the predisposition for IBD (Candan, G, et al, 2019).

Inflammatory bowel disease is considered by many doctors to have a significant genetic component. Recent studies of the genome have identified genes that are associated with IBD. CD was the first disease associated with the NOD2 gene, but additional IBD-related genes have since been identified. The probability of having IL23R in IBD (2.01), NOD2 in CD (3.01), and HLA in UC (1.44) was the greatest of the genetic variants. Despite a few exceptions, the majority of the gene loci associated with Crohn's disease and ulcerative colitis had a similar effect. An example of this is seen with NOD2 and PTPN22, which have been observed to promote a strong defense against ulcerative colitis, but they also have the potential to contribute to Crohn's disease. Our comprehension of genetics and their role in the development and progression of IBD has been greatly enhanced by GWAS. It is important to recognize that certain IBD risk regions are also present in other conditions with autoimmune or chronic inflammation, these include multiple sclerosis, vitiligo, and psoriasis. Certain loci

are associated with either Crohn's disease or ulcerative colitis, while others are associated with both types of intestinal disease. These locations are particularly important for genes involved in cellular innate immunity (for Crohn's disease) and the epithelial barrier (for ulcerative colitis).. (Marcin W and Michael S, 2018).

1.1 HISTORY OF INFLAMMATORY BOWIL DISEASES

The nineteenth and twentieth centuries focused on diagnostic standards and distinguishing IBD from gastrointestinal diseases. Extensive therapeutic research in the twentieth century led to the utilization of biological medicines in medical practice. However, since the beginning of the new century, holistic criteria have emerged to challenge conventional definitions and search for an elusive cure. The cause of chronic intestinal mucosal inflammation in IBD remains unknown. In search of a global solution for IBD, scientists have spent centuries researching its causes and developing this groundbreaking method. Throughout the past 70 years, a worldwide focus on inflammation patterns has emerged, replacing the traditional organ-specific symptoms with more subtle indicators of mild inflammation. This novel method distinguishes the discomfort, swelling, and redness related to IBD. The substantial changes that impacted the digestive tract are what led to the epidemic-like appearance of Ulcerative Colitis (UC) and Crohn's Disease (CD). These two kinds of inflammatory bowel illnesses are the most prevalent ones. In their knowledge of the difficulties brought on by the growth of IBD, particularly in Western regions, researchers have distinguished three key periods.:

1. An explanation of diagnostic standards.
2. The pursuit of therapy, and
3. An effort to adopt a holistic perspective. (Actis, G C, et al., 2019).

Medical research has advanced significantly over the past century, particularly in regards to developing effective treatments for ulcerative colitis and Crohn's disease. As a direct result of these developments, the British Society of Gastroenterology's Clinical Services and Standards Committee (CSSC) recently approved a novel strategy for treating adult patients with inflammatory bowel disease (IBD). It was anticipated that the Society's 2011 directions on the subject would be superseded by this newly discovered recommendation. (Lamb CA, et al., 2021).

The prevalence of IBD remained low up until the latter part of the 20th century, but following World War II, the frequency of UC and CD skyrocketed impressively. The surge in cases suggests that factors beyond industrialization are playing a role in the changing prevalence of IBD. A recent study has linked various environmental influences, such as diet, obesity, how the body responds to sugar, medications, and gut bacteria, to an uptick in the occurrence of IBD. It is estimated that over 5 million individuals across the globe are affected by UC and

CD. Antibiotic use in the early years of life can raise the risk of IBD. Additionally, restricting a child's childhood exposure to environmental germs increases this risk. Although the scientific understanding of the etiology of IBD has advanced significantly, there is still a long way to go before this information can be applied to a tailored treatment strategy. (Kumar M, et al., 2019).

The COVID-19 pandemic has led to a rise in inflammatory bowel disease prevalence in various populations, with developing countries experiencing initial appearances and more countries experiencing the increased prevalence of IBD. Western countries have seen a significant increase in IBD cases, potentially reaching the Prevalence Equilibrium stage due to increased mortality during the pandemic. To reduce the global burden of IBD, effective strategies and technological innovations are needed to address changing demographics. The first phase of IBD's rise occurred in North America and Western Europe, with the industrial revolution impacting manufacturing, urbanization, agriculture, and transportation. (Kaplan and Windsor, 2021).

The planet's environmental science and geological timeline were both affected by human activity, this led to the creation of the Anthropocene. The increase in IBD in the West paralleled the industrialization process during this time. A significant increase in human population growth has been noted as well. After the Second World War, analysis of data showed a significant increase in the population of the world. This increase, called the ' Great acceleration', occurred after 1950. The rapid increase in acceleration is affected by multiple factors, including technological advances, healthcare enhancements, energy consumption, farming and dietary choices, the growth of cities, and the interconditional of global economies and societies. (Max Roser, et al., 2019).

Throughout the latter half of the 20th century, a number of population-based epidemiological studies in the Western world that examined the prevalence of ulcerative colitis and Crohn's disease. Recent studies have consistently documented a growing number of IBD. The earliest and original epidemiological investigation of how frequently Crohn's disease occurs dates back to 1932, when the first report of regional ileitis was published in Cardiff, Wales. The figure 1.1 below depicts the initial incidence as a constant and low rate, with records of 0.27 and 0.09 per 100,000 in 1933 and 1943, respectively. However, from the 1980s onward, the frequency of occurrence increased gradually, with an average of 0.98 per 100,000 in 1953, 2.4

per 100,000 in 1963, 5.2 per 100,000 in 1973, and 9.1 per 100,000 in 1983. Later, the frequency decreased steadily from 6.0 per 100,000 in 1993 to 7.0 per 100,000 in 2003 (Coward S., 2019).

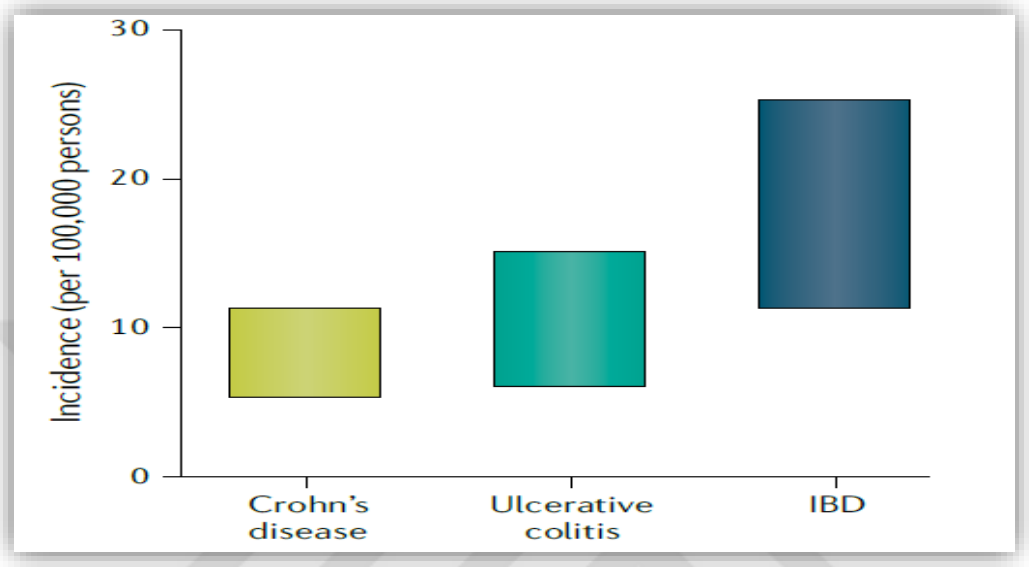


Figure 1.1 The figure describes Inflammatory Bowel Disease (IBD) Incidence Ranges in Western Regions are Converging (Coward S., 2019).

The third level of epidemiological analysis is called compounding prevalence, this was first described in an American article published in the year 1980. They discovered that prevalence is the result of both illness prevalence and the average degree of severity in a consistent population. When the duration of disease is short due to the high mortality rates or rapid recoveries, the method determines the prevalence of chronic conditions in a population that is aging. To understand the effects of acute and chronic diseases, healthcare providers and managers may combine information on the occurrence and duration of diseases. However, this equation is unable to adequately describe the prevalence of IBD in communities. Although it can occur in any age group, the condition is most commonly diagnosed in adolescents and young adults (ages 18-35). Despite the lack of a treatment and the low mortality rate, those who are diagnosed with IBD early on often have a long life. As a result, it's impossible to calculate the typical duration of disorders in IBD over the course of a lifetime; this is true of the prevalence of IBD as well. (Jones G R, et al., 2019).

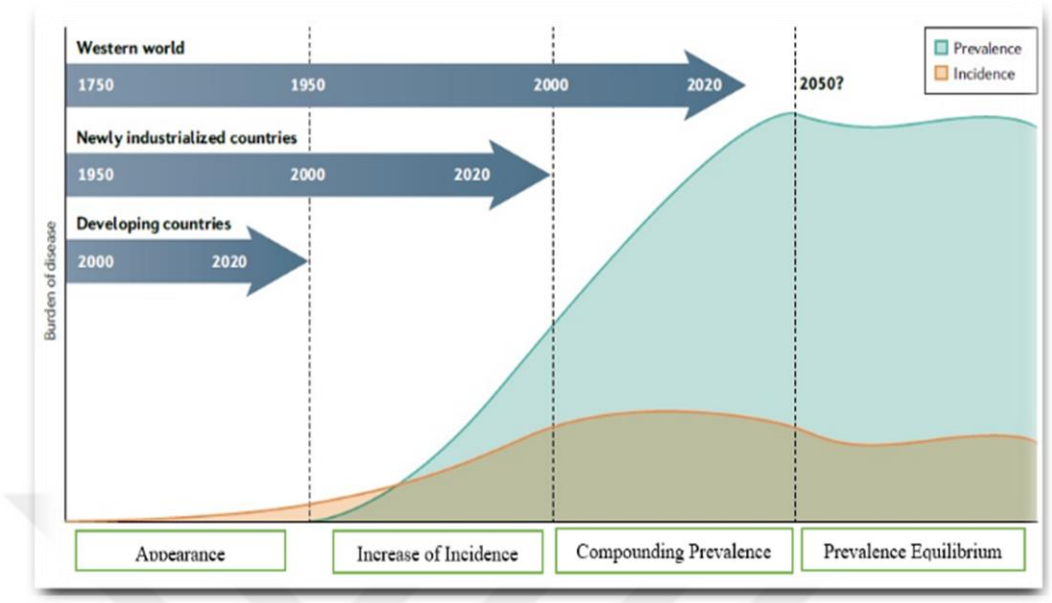


Figure 1.2 The figure describes the four epidemiological stages of Inflammatory Bowel Diseases evolution (Nguyen, G C, et al., 2019).

Between 1990 and 2020, the number of IBD patients has increased significantly. In Ontario, Canada, the prevalence of IBD among older adults increased by 5.2% per year between 1999 and 2008. Therefore, a patient with IBD in 1990 was more likely to live into 2020 than a patient diagnosed in 2020. Compared with patients without IBD, patients with IBD have a slightly higher risk of death from age-related cancers and cardiovascular disease and a lower life expectancy. The prevalence of IBD is not expected to reach 2%. The four epidemiological stages of IBD defined by Nguyen, et al, 2919 the situation as shown in Figure 1.2 explains the relevant years.

Research on inflammatory bowel disease presents a puzzling picture in which multiple factors including genetics, immune system, environment and gut microbiota all play a role. Thanks to next-generation sequencing, high-throughput omics data generation, and molecular networking, technological advances have enabled deeper profiling of IBD. In particular, the combination of artificial intelligence (especially machine learning) and molecular genetics opens the door to efficiently analyze large datasets and discover valuable treatment-relevant information (Seyed Tabib NS., et al., 2019).

The human gastrointestinal system is inhabited by a variety of microorganisms, including archaea, eukaryotes, viruses, and bacteria. In this complex ecosystem, the gut microbiota is mainly composed of different types of bacteria. The microbiome contains a large number of genes, many times the number in the human genome. This difference greatly affects the phenotypic characteristics and overall health of the host. In particular, there is extensive communication between gut microbes and human hosts. Extensive research has identified three main roles of the gut microbiome.

The biochemical activity of the microbiota has various beneficial effects on calcium and iron absorption, nutritional and metabolic processes such as short-chain fatty acid (SCFA) production and vitamin synthesis in the colon. Examples of protective measures include preventing the spread of infectious diseases or dangerous invasive species.

A study by Aldars-García et al. (2021) found that nutritional effects have implications for growth and regulation of neuroendocrine and immune responses. Furthermore, they found that these effects also play a role in the proliferation and specification of intestinal epithelial cells.

Due to genetic defects or environmental stress, ecological diseases can develop in the human gut, affecting the trillions of commensal microorganisms that reside there. The immune system's fight against invading gut bacteria may lead to persistent inflammation, leading to increased cases of inflammatory bowel disease (IBD). It has been observed that the disease mainly affects young adults and has spread from developed Western countries to newly industrialized countries. To combat the two types of IBD, ulcerative colitis (UC) and Crohn's disease (CD), the standard pharmaceutical approach is to take high daily doses of immunosuppressive or anti-inflammatory drugs. Unfortunately, these drugs come with unwanted side effects and potential toxicity. However, thanks to nanotechnology, a new method of drug delivery has emerged. The goal is to improve drug packaging and enable precise delivery while bypassing the limitations of traditional drug delivery methods (Jacob EM., et al., 2020).

Chronic abdominal discomfort, weight loss, and an imbalance of neutrophils and macrophages that lead to inflammation and ulceration of the intestinal lining can be signs of inflammatory bowel disease. These disorders usually appear early in life in both sexes. While the cause of IBD remains a unknown, recent research advances have shed light on how the

disease progresses. Studies have found that IBD can be triggered by a variety of factors, including genetic predisposition, changes in gut bacteria, environmental exposures and abnormalities in the immune system. Analysis of IBD-associated genes and loci has revealed important roles in various processes that maintain intestinal homeostasis. They exert mechanisms such as epithelial barrier function, innate mucosal defense, immune regulation, cell migration, autophagy, and adaptation. The porosity of the epithelial barrier facilitates microbial entry and triggers the innate immune system to secrete extracellular mediators in response. These mediators attract additional cells, including adaptive immune cells, that help trigger appropriate tolerance, inflammation, and necessary responses for healing (Qingdongguan, 2019).

An interesting observation is that the immune response in IBD patients is unbalanced, with irregularities between inflammatory and anti-inflammatory responses. Of particular interest, the development of IBD has been linked to certain genes that produce proteins involved in innate immunity. An example of such a gene is nucleotide-binding oligomerization domain 2, which encodes a pathogen pattern recognition receptor (PPR) that plays a key role in bacterial detection. This is the first susceptibility gene associated with IBD identified. Thus, individuals carrying the NOD2 mutant allele have reduced defenses against invading pathogens, which is associated with a higher likelihood of developing celiac disease, especially in the ileum, as shown in Table 1.1. In IBD, chronic inflammation is associated with an imbalance between the pro-inflammatory and anti-inflammatory responses of the adaptive immune system. This was also affected by two genes, Autophagy related 16 like 1 (ATG16L1) and Immunity Related GTPase M (IRGM), that play a role in pathogen-induced autophagy. The risk loci lie between genes encoding cytokines and cytokine receptors, with IL23R having the greatest impact. IL23R activates pro-inflammatory cytokine production through JAK2 and STAT3 in response to ligand binding. IL23R also plays a critical role in the growth of Th17 cells, further promoting inflammation. Studies have found that an overactive gene mutation in the IL23R-Th17 signaling pathway increases the risk of IBD. Instead, studies suggest that protection may be associated with loss-of-function of the IL23R allele. Furthermore, the IL23R-Th17 pathway is involved in IBD through its downstream effectors JAK2, STAT3 and IL12. Recognition of the importance of signaling pathways in disease has led to the development of new drugs to treat IBD (Zhao M and Burisch J, 2019).

Table 1.1 The table describes the overview of functional genes related to IBD.

Gene	Polymorphism(s)	Function	Effect	Type
NOD2	rs2066845 (G908R)	Bacterial sensing	Risk of disease	CD
	rs2066844 (R702W)		Stricturing disease; Penetrating disease	
	rs2066847 (p.L1007fsX)		Surgery	
			Reoperation	
ATG16L1	rs2241880 (T300A)	Autophagy	Risk of disease	CD
			Risk perianal disease	
			Response to steroids, immunosuppressants, Biologics	
IL23R	rs11209026 (R381Q)	Proinflammatory	Protection against disease	CD, UC
	rs76418789			
	rs76575803			
	rs41313262	IL17-IL23R pathway		
	rs2201841			
	rs10489629			
	rs10889677			
JAK2	rs744166		Risk of disease	CD

There are approximately 100 trillion different types of microorganisms in the gut microbiome, including bacteria, fungi, viruses and protozoa. Furthermore, the gut microbiome has more than 100 times as many genes as the host genome. Firmicutes, Bacteroides, Proteobacteria and Actinobacteria are the four genus of bacteria that make up the majority of gut bacteria, with Firmicutes and Bacteroidetes being the most common in healthy people. Also, the number of bacteria varies depending on where you are in your digestive system, the large intestine always contains more different species than the small intestine. Many aspects of host homeostasis, including nutrient absorption, immune system growth, metabolism, and pathogen defense, depending on every gut microbiota. Figure 3 illustrates how the interplay between environmental, microbial, and immune-mediating variables can lead to IBD in a genetically susceptible host.

Risk factors for developing IBD are shown in Figure 1.3. A complex set of links exists between the environment, microbiota, and immune system of a genetically vulnerable host (Kerri L. Glassner et al., 2019).

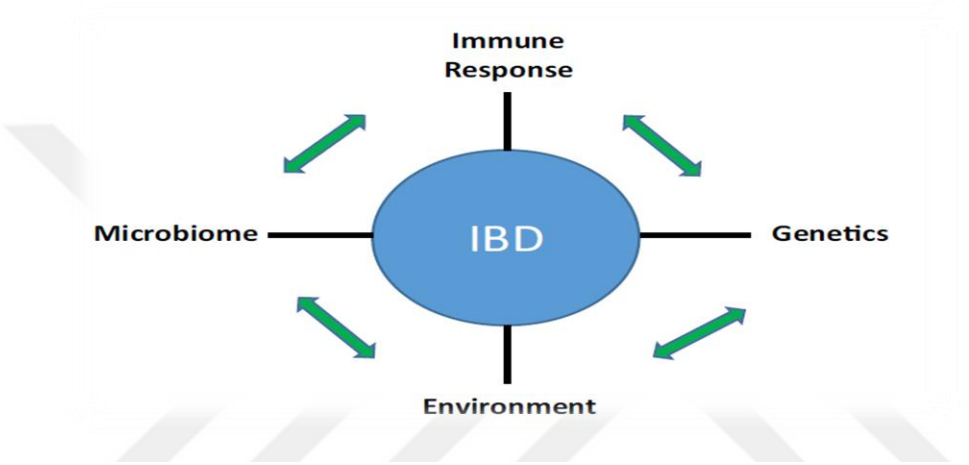


Figure 1.3 The figure describes the factors that increase the risk of IBD developing. (Kerri L. Glassner, et al., 2019).

The fact that disease activity is more readily observed in regions with high bacterial populations (colon) and a lack of fecal material stasis are two further signs from human research indicating the microbiome may play a role in the pathophysiology of inflammatory bowel diseases.

1. Fecal conversion is a successful Crohn's disease treatment strategy, with recovery taking place in the portion of the intestine that is not converted.
2. After the gastrointestinal system and feces have returned to normal, disease recurrence occurs.
3. Treatment with antibiotics can be beneficial for certain IBD patients.
4. The immune system's interaction with the microbiota is linked to genetic markers for IBD, and
5. Microbes have been found to have a role in either causing or regulating inflammation. (Kerri L. Glassner, et al., 2019).

Additionally, each patient's gene expression profile is different due to the complexity and polygenic nature of IBD. Furthermore, more than 200 genes have been linked to an increased chance of developing CD, UC, or both disease categories in multiple genome-wide

association studies and in silico meta-analyses of genetic data (CNAs). In general, genetics has led to the identification of a number of archetypical biological concepts that play a variety of roles in the pathophysiology of disease (such as intestinal epithelial cells, B cells, and T cells). The guiding ideas consist of those that are stated below.:

1. Impaired gut barrier performance.
2. Improper autophagy.
3. innate immune system disruption with decreased phagocytic activity.
4. Improperly activated and differentiated T- and B-cells (Aden K and Reindl W, 2019.

Researchers have employed a variety of methodologies to examine the gut microbiota, and it is the synthesis of various methods that permits a comprehensive understanding of the gut microbiota and health. Feces collection is the most extensively used, non-invasive, and highly populated technique for identifying the gut microbiome in human investigations, as shown in Figure 1.4. DNA exists in the cell as a complex with some proteins (histones, non-histone proteins, High mobility group (HMG) proteins) and RNA. In microorganisms like viruses, DNA is found in a protein coat. DNA isolation can be performed basically in three steps:

1. Destruction of cell-wall.
2. Separation and Disruption of DNA-protein complexes, and
3. Separation of DNA from other molecules in the matrices

Moreover, and for the Lysis of Cell Wall:

1. First stage is the weakening of the cell wall:
   a. Physical (freezing-thawing)
   b. Chemical (lyzozyme, EDTA)
2. Destruction stage:
   a. Ionic detergents (sodium dodecyl sulphate, SDS)
   b. Non-ionic detergents (Triton X-100)
   c. Chemical treatment times differs according to the organism (https://acikders.ankara.edu.tr 2023).

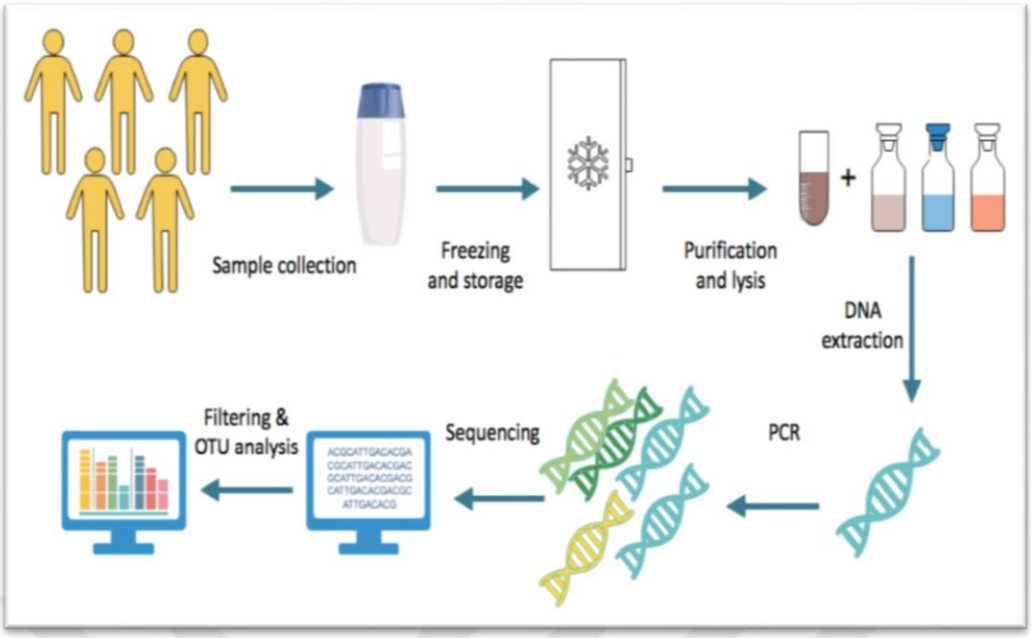


Figure 1.4 The figure describes the general steps undertaken in stool microbiome analysis (Zahraa Al Bander, et al., 2020).

Inflammatory bowel illnesses generally present their symptoms in a variety of digestive system locations. Nearly identical symptoms are present in both Crohn disease (CD), which can affect any portion of the gastrointestinal system from the mouth to the anus, and ulcerative colitis (UC), which only affects the colonic mucosa. Additionally, this illness is not the only one with these characteristics. The indications include ebbing:

1. Stomach cramps
2. Unusual bowel movements, the passing of mucous devoid of blood or pus
3. Slimming down
4. Sweating and fever
5. Illness and weariness
6. Arthralgias
7. Delayed or uncompleted sexual maturation in children, as well as growth retardation Extraintestinal symptoms (10%–20%) such as liver disease, uveitis, or arthritis
8. Extremely bloody stools, occasionally accompanied by tenesmus: Frequent in UC, less so in CD
9. Perianal disease: 50% of CD patients have it.

The following signs of digestive system inflammation are associated with it, according to the World Gastroenterology Organization.

1. Diarrhea: Nighttime occurrence, blood or mucous in the stool, incontinence
2. Constipation: In UC confined to the rectum, constipation may be the primary symptom, followed by obstipation and bowel obstruction.

3. Tenesmus, sharp urgency, and discomfort or rectal bleeding are examples of abnormal bowel motions.
4. In moderate to severe UC, they can occur in the periumbilical or left lower quadrant; in CD, they most frequently occur in the right lower quadrant.
5. CD patients experience nausea and vomiting more frequently than UC patients. (William A Rowe, et al., 2020).

A vague disorder of the intestinal mucosa known as ulcerative colitis or Crohn's disease, inflammatory bowel disease is defined by ongoing inflammation. Since neither the clinical symptoms nor the morphologic indicators of IBDs are specific enough, the diagnosis of this condition necessitates a thorough analysis of the clinical, endoscopic, and microscopic data. the growth of the lamina propria lymphoplasmacytic infiltration, basal plasmacytosis, and architectural deformation, all of which are signs of chronic inflammation, are the long-standing histologic criteria for diagnosis. In order to measure the severity of the sickness, Kvári B. and colleagues (2022) emphasize the significance of evaluating neutrophil inflammation as well as the harm to crypts and epithelial cells. However, it can be challenging to validate these markers, particularly in the early phases of the illness or when individuals have only partially recovered.

Inflammatory bowel diseases (IBD) are a class of persistent, chronic, and even crippling gastrointestinal conditions. IBD is becoming more and more common, especially in Africa, the Middle East, and the Asia-Pacific area. IBD is thought to arise as a result of aberrant immune responses in people who have a genetic vulnerability to particular environmental triggers. Due to major improvements in our knowledge of the causes of the condition and the development of new therapeutic methods, such as biological and small molecule therapies, the goals of IBD treatment have altered. As a result, the present strategy tries to lessen the condition's impairments while also addressing symptoms and achieving mucosal repair and preventing intestinal damage. (Alkhatry M, et al., 2020).

The identification of monogenic IBD at an early stage is crucial for determining the appropriate prognosis and treatment plan, which may involve bone marrow transplantation. This is because more than 50 genes related to various immunological processes, inflammatory homeostasis, and intestinal barrier functions have been associated with causative genetic variants. Monogenic IBD, regardless of age at onset, has a worse prognosis compared to polygenic IBD, particularly in terms of gastrointestinal involvement and extraintestinal symptoms frequency (Lega S, et al, 2020).

In recent years, numerous genome-wide association studies (GWAS) have been conducted with an emphasis on genetic analysis of chronic medical conditions such inflammatory bowel diseases. This intervention began a new age of personalized medicine with pharmaceuticals that precisely target genes or gene products implicated in the disease's pathogenesis, leading to the creation of novel therapy approaches for controlling IBD. As a result, genetic roots are being discovered and getting more attention.. Only a small number of the numerous IBD-related genes that show a high link with the condition have been pinpointed by extensive investigation.

Identifying the pathogen, the first step in the process, is crucial.

3. Pathogen clearance by innate and cell-mediated immunity occurs following obstacle entry via the gut mucosal barrier.

Despite the significance of these genetic findings, it is conceivable that a person without any symptoms or signs of IBD could still show a positive test result for an IBD-related mutation. This emphasizes the role of additional factors in the emergence of the condition, such as social, nutritional, and environmental influences (Younis N, et al., 2020).

Table 1.2 The thable describes the inflammatory bowel disease: between genetics and microbiome, Molecular Biology Reports; Younis N. et al., 2020. Key IBD risk loci identified in multiple WGS and WES among Crohn's Disease and Ulcerative Colitis patients

Genes	Locus	IBD variant	Type of disease
NOD2	16q12.1	rs2066844	CD
		rs2066845	CD
		rs41450053	CD
ATG16L1	2q37.1	rs2241880	CD
IRGM	5q33.1	rs13361189	CD
		rs4958847	CD
LRRK2	12q12	rs33995883	CD
PTPN2	18p11.21	rs2542151	CD and UC
		rs7234029	CD and UC
HNF4α	20q13.12	rs6017342	UC
Il10RA	11q23.3	rs3135932	Pediatric CD and UC

1.2 CROHN'S DISEASE

Crohn's disease (CD), sometimes referred to as chronic inflammation of the intestines, typically affects the terminal segment of the small intestine, which is situated at the farthest point from the stomach. Regional enteritis or regional ileitis are other names for this illness. A Polish surgeon named Antoni Leniowski was the first to describe Crohn's disease in 1904. Burrill Bernard Crohn, Leon Ginzburg, and Gordon D. Oppenheimer were pioneering American doctors who discovered the intestinal ailment known as regional ileitis in 1932.. They gave this illness Crohn's disease as a name. These days, this disorder is referred to as Crohn's disease, which is a kind of IBD (inflammatory bowel disease). Researchers have discovered links between Crohn's disease and immune system performance as well as genetic variants. It's interesting to note that this illness has also been linked to aberrant microbial

pathogen alterations in the intestine. A couple of examples include Mycobacterium avium paratuberculosis (MAP), which has been detected in the blood of a few patients, and the digestive systems of ruminants with Johne disease, a condition that is similar to CD in humans and causes lower levels of the normal resident in the intestine called Faecalibacterium prausnitzii whcih is confirmed both in fecal samples in inflammatory bowel disease (IBD) patients. Numerous observational studies have suspected dysbiosis, an imbalance between protective and harmful bacteria to be relevant to the etiology and pathogenesis of IBD. However, it is still unknown what causes Crohn's disease (Barbara A S, et al., 2022).

Crohn's disease can appear in various ways including the stricturing type caused by fibrosis and the penetrating type caused by fistulas connecting the stomach and other tissues. There is also the non-stricturing, nonpenetrating type and the stricturing, penetrating type. The disease might change from an inflammatory form to a stricturing, penetrating condition due to repeated inflammation resulting in intestinal damage (Imai T, et al, 2019).

An ongoing condition called Crohn's disease, which affects the gastrointestinal tract, is marked by chronic inflammation. This illness is becoming more prevalent worldwide and has a slow, debilitating progression. CD's specific cause is unknown, however it is thought to be influenced by a malfunctioning immune system, altered gut flora, genetics, and environmental factors. In the early stages of the disease, prompt and thorough therapy is frequently necessary to stop exacerbations and other intestinal problems. In people with CD, transmural inflammation results in blockage, perianal lesions, diarrhea, and/or ongoing stomach discomfort. Skip intestinal lesions in any area of the GI system can be used to diagnose CD. On the other hand, UC only affects the colon and causes erosions, ulcers, and bloody diarrhea. The inflammation is only superficial, and the lesions are ongoing. Ten years after diagnosis, half of all CD patients have intestinal problems including strictures or fistulae, illustrating the disease's persistent and debilitating character. Extraintestinal manifestations (EIMs), which impact 21–47% of patients, have serious negative effects on patients' quality of life. Surgery, hospitalization, and complications are all feasible long-term effects (Roda G., et al., 2020).

Inflammatory bowel illnesses occur more frequently in North America, northern and western Europe, and Oceania than in other locations. Between 0.1 and 58 instances per 100,000 persons are reported each year. The prevalence of CD varies between 0.3 and 12.7 cases per 100,000 people per year in Europe, compared to 0 to 20.2 cases per 100,000 people per year

in North America. The highest documented incidence of CD are in Canada, where there are 319 instances per 100,000 individuals (Siew C Ng, et al., 2018).

Since the turn of the millennium, inflammatory bowel illness has become more common, particularly in newly industrialized nations in Asia, Africa, and South America. The rate of CD cases varies by area in mainland China, with the north and west having a higher concentration. According to research, Korea had 1.68 instances per 100,000 people annually in 2005, and the rate then stabilized. Although IBD is less common in Asia than in Western countries, the incidence of CD has increased more quickly than UC (Park S H, et al., 2019).

The fast changes in Crohn's disease epidemiology pose a substantial challenge to disease prevention and early sickness diagnosis. The rising incidence of CD in recently industrialized nations like Asia is a clear indication of the dangers linked with aspects of the Western lifestyle, such as food, urbanization, and industry. Furthermore, studies on migrants have shown that children of immigrants who move from low- to high-prevalence areas have a higher risk of CD than non-migrants do (Piovani D, et al., 2019).

Numerous environmental factors can affect how Crohn's disease develops and progresses in a genetically predisposed host. These elements are very significant. Smoking stands out as the primary risk factor for CD in Western nations. It makes the condition more likely to develop, especially in women and at various ages. Smoking is also linked to early sickness start, the demand for immunosuppressive therapy, the need for surgery, and greater rates of disease recurrence following surgery. Additional environmental risk factors for Crohn's disease are briefly described in Table 1.3. Depending on one's ethnicity, there has been evidence linking passive smoking to an increased chance of getting CD in Japan (Kondo K, et al., 2019).

Table 1.3 The table describes the lists some of the environmental risk factors for Crohn's disease. Kondo, K. et al.2019.

Environmental factors	Association
Smoking	There is a strong link between the beginning of illness and the progression of the disease.
Appendectomy	There is a positive correlation with illness initiation but no correlation with disease progression.
Low dietary vitamin D	There is a link between the start and progression of the condition.
Oral contraceptive use	There is a strong positive correlation with disease beginning but no relationship with disease progression.
Postmenopausal hormone use	There is no link between the development of the disease and the progression of the illness.
NSAID use	Positive correlation with illness start and a substantial relationship with disease progression.
Antibiotic use	Positive relationship with illness start and progression.
Depression and psychosocial stress	There is a positive correlation with illness initiation but no correlation with disease progression.
Low dietary fibre	There is a negative correlation with illness initiation but no correlation with disease progression.
High dietary fat	There is no link between the beginning of the disease and the progression of the illness.
High dietary protein	There is no link between the beginning of the disease and the progression of the illness.

Gastrointestinal dysbiosis is a symptom of Crohn's disease, and among the environmental elements that have changed over the last 10 years, food is most likely to have an impact on the microbiota in the gut. Changes in dietary composition, as well as a shift away from processed foods with food additives and toward high-fiber, low-fat meals, have an effect on the connection between the host and gut flora. Low dietary fiber intake and frequent nutritional swings between meals high in and low in fiber are linked to decreased gut microbial diversity and the beginning of Crohn's disease. Because nutrition has a temporary effect on microbial composition, it is still unclear what influence dietary alterations played in CD's altered

microbial biodiversity. Higher Mediterranean diet adherence was linked to a noticeably decreased prevalence of later-onset diabetes in two prospective studies carried out in Sweden. CD (Khalili H, et al., 2020).

For more than a generation, no one has been able to determine the cause of Crohn's disease. Additionally, researchers have made significant progress in understanding the underlying mechanisms that predispose to the condition by combining molecular biology, molecular genetics, and gene sequencing with first-hand patient experience. The aetiology of CD is heavily influenced by genetics.. A CD patient's sibling is about twice as likely to get the condition than the general population is, and first-, second-, and third-degree relatives are at particularly high risk (incidence rate ratios of 7.8, 7.8, and 7.8). Monozygotic twins have a higher illness phenotypic concordance (37%) than dizygotic twins (7%), according to a study including more than 300 twin pairs. Unexpectedly, gastrointestinal dysfunction affects nearly one-third of first-degree relatives, including increased permeability or signs of inflammation like the neutrophil protein S100A12. It has proven challenging to identify the genetic mutations that result in these significant genetic impacts (Segal A w, 2019).

In meta-analyses, the findings of genome-wide association studies (GWAS) on CD are increasingly being evaluated. Not a single penetrant mutation that might be the only factor in the development of the disease has been found. With the exception of the nucleotide binding oligomerization domain containing 2 (NOD2), the most of the related genes have small impact sizes despite being extremely significant and are believed to work together to produce a tendency. Their data therefore supported a negligible influence.. IBD cases and controls total roughly 70 000, yet the cumulative impact of these loci only explains 14% of the disease. These GWAS studies focused on IBD rather than CD or UC separately. These two distinct illnesses have been combined because they both have an effect on the lower intestine and are connected to inflammation. Even though it was challenging to collect precise phenotypic data on the tens of thousands of people investigated in this GWAS, the influence of the genes that cause one of these disorders will be decreased by the absence of effect on the other. For instance, the two most prominent risk genes for CD, NOD2 and PTPN2, had strong protective benefits in UC (Wainberg M, 2019).

To find gene-trait connections, transcriptome-wide association studies (TWAS), which combine GWAS and gene expression datasets, are currently being developed. IBD loci are

also overrepresented in genes related to early immunodeficiencies, particularly those conferring Mendelian vulnerability to leprosy and mycobacterial illness. Gene ontology (GO) terms related to cytokine synthesis, lymphocyte activation, and response to bacterial compounds were most highly enriched. In addition to NOD2, the molecules IRGM and ATG16L1, which are involved in membrane movements, and molecules of the IL-23/Il-17 axis, which activate T cells in the adaptive immune system, have the strongest associations to CD (Denson L, et al., 2019).

Pediatric onset IBD, which makes up around 25% of all IBD diagnoses, frequently manifests as gastrointestinal stricture, perianal disease, and a failure to respond to traditional medication. The first and most frequently repeated locus connected to adult IBD to date is NOD2. On the other hand, it's uncertain what impact it has for pediatric IBD. (Crowley E, et al., 2020).

A population of 1,183 IBD patients with juvenile onset (aged 0 to 18.5 years) was examined using whole-exome sequencing. Biallelic uncommon and low frequency nucleotide binding oligomerization domain containing 2 (NOD2) mutations were found in 92 probands, or around 8% of the population, in the study, suggesting a Mendelian disease inheritance pattern. Researchers also looked at the importance of recessive NOD2 allele inheritance in adult IBD patients using a huge clinical community sample. By demonstrating that 7% of cases in this adult IBD cohort, including 10% of CD patients, are brought on by recessive inheritance of NOD2 mutations, the study supported the conclusions of our pediatric IBD group. Some of these adult IBD patients got IBD diagnoses even before the age of 18, showing an early onset of illness, according to an examination of EHR data. 40 unexpected mutations were found in 31 probands after an initial screening of exome sequence data for pathogenic and predicted pathogenic alterations in genes known to cause monogenic types of IBD in the juvenile population. A proband-based analytical methodology was used to analyze 492 complete trios in order to find all pertinent recessive X-linked and de novo mutations (DNM) is any mutation or alteration in the genome of an individual organism (human, animal, plant, or microbe) that was not inherited from its parents. in the afflicted probands. The initial 10 families (2 percent MAF) with recessive rare mutations in NOD2 were all determined to have CD. According to Table 1.4, two people have the p.G908R mutation in trans along with the less common p.N852S NOD2 CD threat variation in one case and the new truncating index

p.S506Vfs*73 in the other. This includes the p.G908R missense variant, one of the uncommon mutations in these probands. The identification of a CD-associated NOD2 allele in trans from other novel or unique genes prompted the examination of all probands, including singletons and those in incomplete trios, for recessive inheritance of NOD2 variations, whether homozygous or compound heterozygous, while extending the allelic range to low-frequency variants (% MAF 5%) (Horowitz J E, et al., 2021).

Table 1.4 The figure describes the common NOD2 variations are the three main low-frequency risk variants for Crohn's disease (p.R702W, p.G908R, and p.L1007fs); uncommon NOD2 variations are other low-frequency variants (MAF 5%).. ((Horowitz J E, et al., 2021).

NOD2 variant	# EO-IBD probands	Mean age (range)	% CD Dx	Tissue involvement		
				Colon (%)	Ileum (%)	Perianal (%)
Compound heterozygous						
Rare/rare	4 T; 1S	12.5 (9.0–14.6)	88.9	80.0	60.0	40.0
Common/rare	1Q; 9 T; 6D; 14S	12.6 (5.5–16.5)	93.1	57.1	82.1	25.0
Common/common	17 T; 5D; 10S	11.8 (2.1–18.5)	90.6	56.3	90.6	25.0
Homozygous						
Rare	1D; 2S	9.6 (4.2–15.1)	66.7	33.3	33.3	33.3
p.R702W	1Q; 2 T; 3D; 1S	11.7 (5.8–13.8)	100	85.7	57.1	42.9
p.G908R	2 T; 1D; 2S	10.6 (7.7–13.7	80.0	40.0	40.0	0.0
p.L1007fs	4 T; 2D; 4S	12.1 (5.9–13.4)	100	60.0	90.0	20.0

Epigenetics has become a promising field of study in the quest to enhance the detection, prediction, and reaction to inflammatory bowel disease. An uptick in the number of individuals afflicted with IBD, particularly those affected by Crohn's disease, has been observed. It is believed that the intricate interplay between inherited inclinations, external influences, and modified gut bacteria is responsible for the disrupted function of the body's native defense mechanisms in Crohn's disease. Although a wide array of genetic variations have been linked to Crohn's disease, the exact role of epigenetics in its development remains largely unknown. On the flip side, the disease's genetic vulnerability poses a significant health hazard. Additionally, various lifestyle factors such as diet, smoking, early antibiotic use, and the composition of gut bacteria also contribute significantly to its development.. In countries

of South America, Africa, and Asia that have become industrialized in recent years, there is an increasing occurrence of CD, indicating the significance of lifestyle. The boost in CD cases can be attributed, at least in part, to environmental factors that may cause alterations in DNA methylation or histone modification. Recognizing the potential of studying these epigenetic changes, such as methylation of DNA or modification of histones, provides a valuable approach to identifying new biomarkers or targets for treatment, ultimately leading to personalized therapies for CD. Colorectal cancer, allergies, asthma, and cardiovascular disease are just a few of the conditions that can be diagnosed, prognosticated, and treated using epigenetic tests. By comparing the DNA methylation changes in CD patients and healthy controls, we can uncover biomarkers and enhance IBD care. (Hornschuh M, et al., 2021).

The DNA methylation pattern of individuals with CD has been the main focus of recent epigenome-wide association studies (EWAS). These studies have looked at various samples such as whole blood, peripheral blood cells, and serum samples. In some cases, DNA methylation investigations have also been conducted on specific cell types like B cells, T cells, and monocytes. While a few papers have looked at methylation in pediatric donors, the majority of studies have focused on adults. Comparing these studies poses a challenge due to the influence of aging and external factors such as nutrition and medication on DNA methylation. Interestingly, trials involving long-term IBD patients outperformed those involving pediatric patients, who benefit from less medication. This suggests that epigenetic changes are more indicative of illness. Methylation status is affected by both age and gender, leading some studies to only include participants of the same gender. As a result, the comparisons between patient and control groups are more reliable. (Moret-Tatay I, et al., 2019).

In the context of IBD, certain loci that have been methylated regulate or send downstream signals to the Interleukin-23 (IL-23) signaling pathway. These loci include B cell lymphoma 3 protein (BCL3), signal transducer and activator of transcription 3 (STAT3), oncostatin-M (OSM), and signal transducer and activator of transcription 5 (STAT5). The immune system's IL-23 cytokine, which forms a heterodimer and triggers autoimmune inflammation, could potentially contribute to autoimmune inflammatory diseases like CD. Additionally, examination of blood samples from CD patients has revealed DNA methylation changes in

genes related to immune-related pathways such as the immune response and defense against bacteria, as depicted in Figure 1.5. The limited sample size and small effect sizes identified mean that many of the findings will have to be verified in bigger cohorts before they can be considered effective biomarkers. Biomarkers, in terms of their helpfulness, could be quite significant. (Li Yim, et al., 2020).

Figure 1.5 schematically illustrates the DNA methylation studies conducted on various cell types in individuals with Crohn's disease. Genes that are both hypo- and hypermethylated can be observed in B cells, monocytes, and T cells. Human adipose stem cells (recovered from subcutaneous adipose tissue of CD patients), isolated intestinal epithelial cells (IECs), and isolated fibroblasts from biopsies of CD patients, in addition to the hypo- and hypermethylated genes discovered in bowel biopsies, were all examined. The genes highlighted in bold have been identified in multiple investigations.. (Li Yim, et al., 2020).

The diverse appearances and vague symptoms of Crohn's disease often lead to a delay in its diagnosis. In Europe, CD is typically diagnosed within 5-9 months, but in certain countries it can take up to 2 years. Early detection of CD is considered when there is a diagnosis within 18 months of symptom onset, without prior treatment using thiopurines, methotrexate, or biologics, and no irregularities present. A comprehensive understanding of early CD is crucial in investigating and characterizing the impact of early therapy on long-term outcomes. Studies consistently highlight the importance of initiating therapy promptly in individuals with severe disease phenotype (Loy L, et al., 2019).

The introduction of new medications for Crohn's disease has sparked a change in therapy approaches, as healthcare professionals recognize that certain clinical characteristics can indicate a heightened likelihood of experiencing a worsening and debilitating form of CD. Furthermore, it is now increasingly evident that the recovery of the gut lining and the fading or disappearance of internal lesions, known as mucosal recovery, is linked to positive outcomes. These include a reduced risk of relapse, fewer hospitalizations, a remission without the need for steroids in follow-up examinations, and longer periods between surgical interventions. Notably, mucosal repair also lowers the incidence of severe ulcers and the necessity for surgery in CD patients compared to those with more extreme ulceration.. In the pursuit of a complete recovery, medical experts highly recommend achieving both clinical and endoscopic functional recuperation. Consequently, mucosal restoration has become an

increasingly imperative objective in treatment. Moreover, in the near future, the introduction of cross-sectional imaging techniques and histology will supplement mucosal repair by highlighting the significance of transmural recovery. Conversely, in the realm of research, the majority of randomized controlled studies primarily evaluate symptomatic response or remission as outcomes, with only recent inclusion of endoscopic outcomes in medical studies, particularly within the last few years. Starting treatment quickly after diagnosis can lead to higher rates of success and improved mucosal repair. Considering the patient's perspective is essential prior to initiating therapy (Roda G, et al., 2020).

Mucosal healing is the main therapeutic objective for managing Crohn's disease. The STRIDE program identified two possible treatment options for CD: clinical and patient-reported achievement recovery. Endoscopic recovery, defined as the disappearance of ulcers or infection resolution on imaging, is one aspect of achievement recovery. Additionally, remission of biomarkers was a secondary goal. As our understanding of Crohn's disease continues to grow, therapeutic goals and endpoints evolve.. Mucosal recovery has received positive feedback from society, research, and meta-analyses. The reduction in surgical procedures, recurrence rates, and improvements in quality of life make it a promising avenue. Close monitoring, using serum, C-reactive protein (CRP), and faecal calprotectin as biological markers, has shown higher rates of mucosal healing in early CD patients who tried corticosteroid treatment compared to traditional treatments. This approach has also been linked to fewer flare-ups, hospital stays, and overall improvement in quality of life.. (Colombel J.-F, et al., 2018).

In the treatment of Crohn's disease, biologic therapy plays a vital role. It offers a unique approach that conventional treatments cannot achieve, with the ability to induce recovery and response rates. Over the past twenty years, CD therapy has undergone significant advancements thanks to the introduction of anti-TNF drugs like infliximab, adalimumab, and certolizumab. Thanks to the introduction of infliximab and adalimumab biosimilars, these medicines are now available worldwide. Additionally, ustekinumab and vedolizumab, which have recently been approved in the United States and Europe for the treatment of moderate-to-severe CD, provide effective relief. (Roda G, et al., 2020).

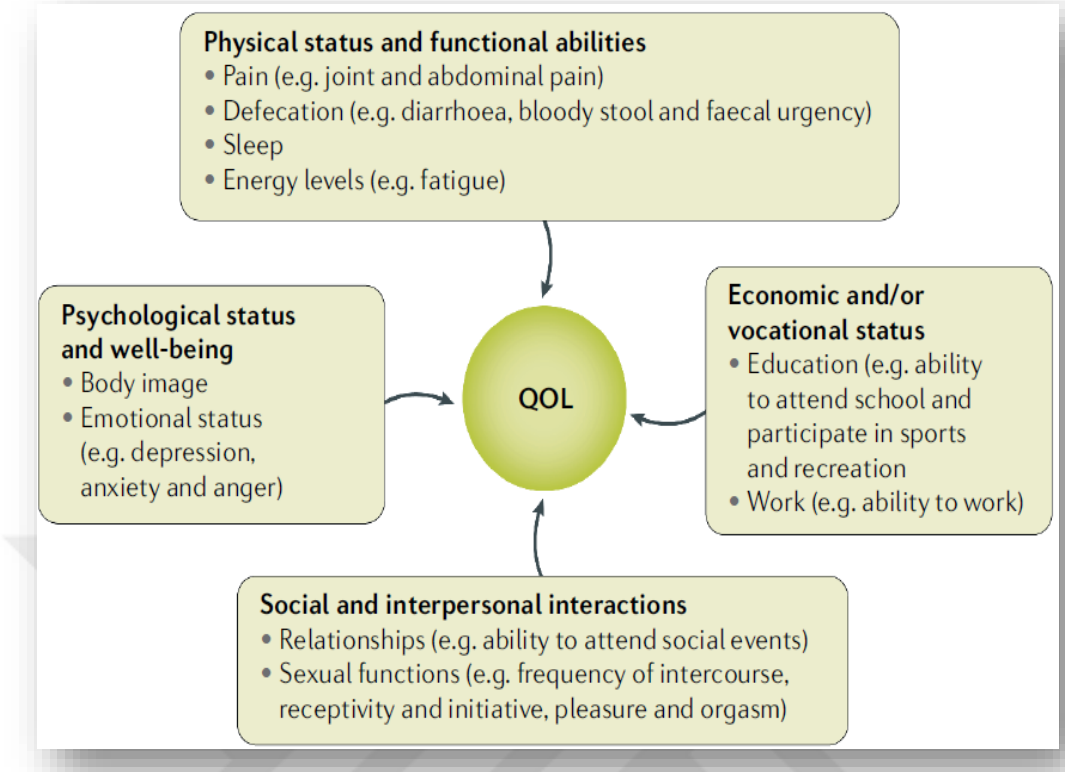


Figure 1.5 The figure describes the impact of certain factors on Crohn's illness with regards to QOL domains. Assessing patients' standards of living is dependent on their physical, psychological, cultural, and economic backgrounds. The symptoms of their condition, challenges they face, extraintestinal manifestations, and the therapies they receive all contribute to the alteration of patients' quality of life with Crohn's disease. (Colombel J.-F, et al., 2018).

In the year 2019, a total of 955 individuals diagnosed with Crohn's disease participated in a research study that included the administration of infliximab. Results from the study revealed that a significant factor leading to lack of response to the treatment was the presence of low levels of infliximab at week 14. In cases where the concentration of infliximab reached a maximum of 7 g/ml at week 14, positive recovery outcomes were observed both at week 14 and week 54. It was also observed that inadequate dosage levels at week 14 contributed to the development of immunogenicity and the appearance of detectable anti-infliximab antibodies. Moreover, a study conducted in France identified a connection between higher levels of adalimumab in the blood and improved outcomes in patients with IBD. The average blood levels of adalimumab for clinical improvement and endoscopic remission in CD were 6.02 g/ml and 6.5 g/ml, respectively. Moreover, Vedolizumab, an approved treatment for

moderate-to-severe active CD in patients with insufficient response to conventional therapy, targets 47 integrin on lymphocytes. This information can be found in Table 1.5. It has been suggested that the dosage of medication and its effectiveness may have a connection, even though the reliability of results is not as strong as with anti-TNF medications.. (Argollo M, et al., 2020).

Table 1.5 The table describes the displays the results of numerous studies looking into infliximab and adalimumab serum concentrations in Crohn's disease. (Argollo M, et al., 2020).

Disease	Number of patients	Dose	Outcomes
Crohn's Disease	105	Episodic vs maintenance IFX	Remission 82% vs 6% (detectable TL vs undetectable); TL higher in maintenance group
Crohn's Disease	72	IFX 5–10 mg/kg	IFX TL >8 μg/ml in maintenance associated with EH
Crohn's Disease	40	ADL 40 mg EOW	ADA maintenance TL >6 μg/ml associated with clinical remission/EH
Crohn's Disease	66	ADL 40 mg EOW	Mean random ADL concentrations >7.5 μg/ml associated with improved EH
Crohn's Disease	23	ADL 40 mg EOW	Increased ADL serum levels at week 48 (>10 μg/ml) in patients in remission
Crohn's Disease	52	IFX 5 mg/kg	Increase of IFX concentration >0.5 μg/ml associated with EH ($P = 0.001$)
Crohn's Disease	110	IFX 5 mg/kg	IFX ≥9.7 μg/ml associated with EH ($P = 0.006$)

1.3 ULCERATIVE COLITIS DISEASE

Samuel Wilks first described ulcerative colitis (UC) in his book Morbid Appearances in the Intestine of Miss Bankes back in 1859. This chronic and incurable inflammatory bowel disease affects the colon and rectum, but its origin remains a mystery. UC is one of the two primary forms of IBD, and it is characterized by mucosal inflammation that progresses predictably from the rectum upwards in the colon. The development of UC is thought to be influenced by various factors, including a person's genetic background, environmental factors, luminal factors, and dysregulation of the mucosal immune system. Bloody diarrhea is the

most common symptom used to diagnose UC. It has become a global issue with its increasing prevalence in affluent countries and higher incidence in underdeveloped ones. UC diagnosis relies on a range of symptoms, endoscopy, and histology (Kobayashi T, et al., 2020).

Ulcerative colitis in the United Kingdom is reported to have an incidence rate of 12.6 per 100,000 person years. Recent data from the Lothian region reveals a point prevalence of 432 per 100,000. Moreover, the incidence is increasing steadily. The occurrence of UC follows a bimodal age distribution, peaking between 50 and 80 years and again in the second and third decades of life. Common symptoms of ulcerative colitis include frequent stomach discomfort, pain, fatigue, and difficulties with bowel control (Jonathan P Segal, et al., 2021).

Sufferers of this illness often find that it worsens over time, with periods of spontaneous remission followed by relapses. Nonsteroidal anti-inflammatory drugs and quitting smoking are often two culprits that make ulcerative colitis worse. Additionally, around 10% to 30% of UC patients experience various extraintestinal symptoms (EIMs). Disease severity is often associated with symptoms such as episcleritis, scleritis, uveitis, peripheral arthropathies, and pyoderma gangrenosum. Primary sclerosing cholangitis, a severe hepatic extraintestinal manifestation of UC, is associated with a greater risk of colorectal cancer. Axial arthropathies, sacroiliitis, and ankylosing spondylitis occur as extraintestinal symptoms independently of UC activity (Whitney D. Lynch and Ronald Hsu, 2021).

A multitude of mysterious causes trigger IBDs that are yet to be identified. A strong genetic tie is evident as multiple factors contribute to a history of illness. The chances of developing UC are quadrupled if a first-degree relative is a sufferer. While there isn't solid evidence to back this theory, it has been suggested that deficiencies in mucosal immunity and changes in gut bacteria are the main culprits behind UC. According to some studies, smoking may prevent the emergence of colitis, but no one has yet been able to demonstrate a clear, meaningful connection between the two. It is possible that the autoimmune disease could have contributed to its development (Cambrian Y. Liu and D. Brent Polk, 2018).

The initial step in correctly diagnosing inflammatory bowel disease, such as UC, involves thoroughly documenting the medical history of the patient. It is important to take into account the longevity of the symptoms to determine their significance. Infections and intestinal ischaemia are commonly diagnosed instead of IBD, particularly if the symptoms have only

recently appeared. Therefore, if someone experiences acute-onset symptoms of ulcerative colitis within a specific timeframe, it should raise concern. Typically, adults between the ages of 20 and 50 are diagnosed with UC. Prompt treatment is advised when people present with severe bloody diarrhea, anemia, tachycardia, and fever in other settings. In these cases, acute severe colitis should be considered as a possibility. Despite technological advancements, the patient's medical history remains vital in determining a diagnosis and cannot be substituted (Kobayashi T, et al., 2020).

To ensure prompt and accurate treatment, it is crucial to promptly diagnose ulcerative colitis. The current diagnostic criteria for UC haven't changed drastically and involve assessing clinical signs, endoscopic appearance, histological examination, and ruling out other possible causes like infections and different types of colitis. Nonetheless, the use of non-invasive methods such as intestinal ultrasonography and biomarkers like faecal calprotectin or faecal lactoferrin for diagnosis and monitoring is gaining popularity. Finding the right diagnosis for ulcerative colitis is vital to begin treatment promptly. Table 1.6 displays the various possible diagnoses for UC because the diagnostic criteria for UC have not had significant changes recently and rely on a combination of clinical signs, endoscopic appearance, histological examination, as well as eliminating other possible causes of colitis such as infections or other types. However, there is an increasing trend of using non-invasive methods and biomarkers like intestinal ultrasonography and faecal calprotectin or faecal lactoferrin for diagnosis and monitoring. The integration of patient history, endoscopy, histopathology, laboratory testing, and imaging modalities may potentially benefit from the usage of artificial intelligence, resulting in future diagnosis performance improvements. Diagnosing UC requires a comprehensive approach due to the absence of a single gold-standard methodology (Matsuoka K, et al., 2018).

Table 1.6 The table describes the different ulcerative colitis diagnoses (Matsuoka K, et al., 2018).

Differential diagnosis	Diagnostic tool	Distinguishing features
Infection		
Bacteria (Salmonella spp. or Shigella spp., Escherichia coli, Clostridioides difficile, Campylobacter spp., Mycobacteriaceae)	History plus stool culture	Acute onset, positive stool culture
Protozoa (amoebic colitis and strongyloidiasis)	Stool ova and protozoa study	Positive stool test
Virus (cytomegalovirus)	Colonic histopathology with IHC	Positive immunohistochemistry (IHC)
Immune related		
Crohn's disease	Endoscopic picture	Skipped lesions, longitudinal ulcers
Graft versus host disease	History plus endoscopic picture	History of transplantation
Eosinophilic colitis	Colonic histopathology	Eosinophilic infiltration
Vascular		
Ischaemic colitis	History plus endoscopic picture	Acute onset, ischaemic change in the watershed area
Vasculitis	History plus other lab data	Anti-neutrophil cytoplasmic antibody

The appropriate sickness identification and evaluation of the degree and severity of disease in people who present the primary symptoms depend on early endoscopic assessment. Treatment options (local versus systemic treatment) and prognosis are influenced by the infection's severity. Some recommendations urge a complete colonoscopy with ileal intubation for anyone displaying clinical IBD symptoms, as shown in Figure 1.6. When someone exhibits the fundamental symptoms, early endoscopic assessment is also a key tool for identifying the appropriate illness and evaluating the degree and severity of sickness.. Treatment options (local versus systemic treatment) and prognosis are influenced by the infection's severity. Some recommendations urge a complete colonoscopy with ileal intubation for anyone displaying clinical IBD symptoms, as shown in Figure 7. (Kobayashi T, et al., 2020).

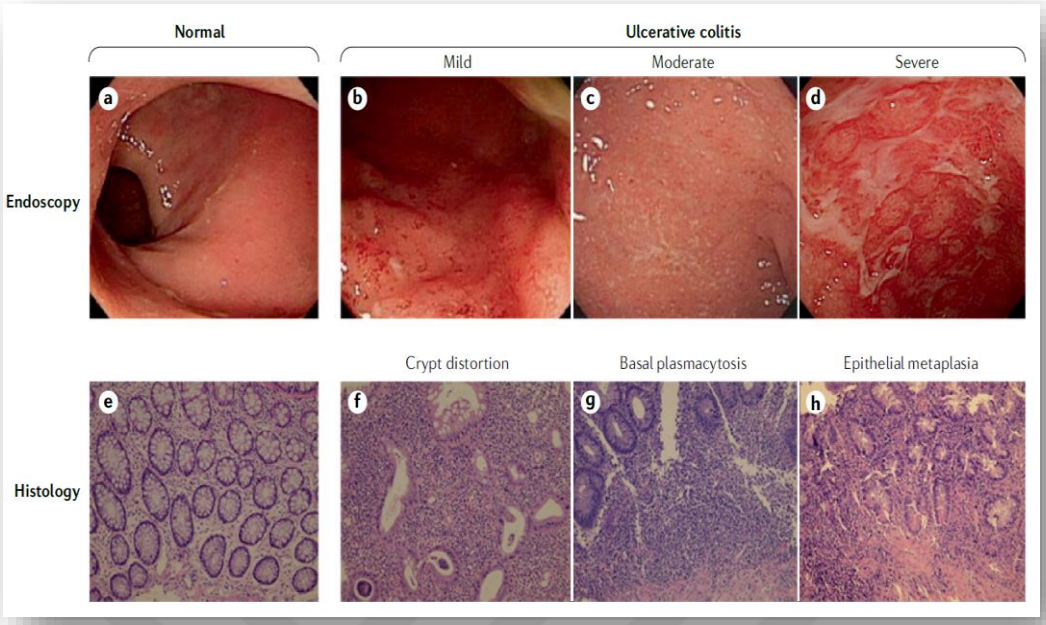


Figure 1.6 The figure describes the endoscopic and histological features of ulcerative colitis. Endoscopic and histological information are essential for determining the differential diagnosis of ulcerative colitis (UC). A healthy colon as seen through an endoscope (Kobayashi T, et al., 2020).

Clinical and laboratory markers are used to determine the disease's severity. The Ulcerative Colitis Endoscopic Index of Severity (UCEIS) is the only accepted endoscopic grading system for determining severity. The Mayo score is also widely utilized in medical treatment due to its simplicity of use (Lamb CA, et al., 2021).

According to a widely accepted idea, activities that damage the mucosal barrier, alter the normal balance of the intestinal microbiota, and abnormally stimulate gut immune responses are what cause ulcerative colitis to develop and raise a person's risk of developing the condition. Genetic research has identified 260 susceptibility loci for inflammatory bowel disease, including both widespread and uncommon genetic variants. The majority of genetic variables that affect both UC and CD are shared, which is only one of the interesting discoveries that the research has produced. An early examination of 15 genome-wide association datasets revealed that 110 of 163 susceptibility loci (67%) were related to both illnesses.. These common genes, which also encode the innate and adaptive immune systems, cytokine signaling, and immunological detection, include IL23-R, IL-12, JAK2, CARD9,

TNFSF18, and IL-10. Additionally, the highest genetic signaling inside UC-specific loci is connected to the HLA region of the chromosome. In a more thorough genomic analysis, 16 HLA allelic correlations with UC (mainly class II) were discovered, including HLA DRB1*01*03 for IBD colonic involvement.. More research reveals that they are associated with both UC and CD colonic involvement. Despite the identification of numerous susceptibility loci, genetics only contributes to 19% of ulcerative colitis heredity. When a polygenic risk score that takes into account all susceptibility loci is determined, many people are found to be genetically free (Porter R J, et al., 2020).

Environmental factors are anticipated to play a role in the rapid development of ulcerative colitis in nations with recently developed economies. This is similar to the fashions that were popular in the West around the start of the 20th century. Particularly in urban areas, UC first appears with an abrupt rise in frequency and a subsequent fall. There are many factors that are thought to contribute to westernization, including modern urban lifestyles, exposure to pollution, dietary changes, access to antibiotics, improved hygiene, and a drop in disease rates. However, more specific environmental factors linked to UC have long been found. The most obvious example is the protective benefit of smoking cigarettes, as well as the crucial finding that individuals who cease smoking experience their first case of UC (Zhu F, et al., 2019).

Ulcerative colitis must be treated first before undergoing any procedures that may be required. The key variables that influence how UC patients are treated include the severity of the condition, the degree of the infection, and how it varies over time. As a result, the therapy objective must be the same for all patients: reduction of mucosal friability and ulceration at lower endoscopy, as well as remission of diarrhea and rectal bleeding within three months after medication initiation. Patients with acute severe colitis must be admitted to the hospital right away.. In situations involving a toxic megacolon, a perforation, or substantial bleeding, immediate emergency surgery may be advised. Although it's important to rule out alternate diagnosis like C. difficile infection, serious diseases shouldn't wait to start therapy.. (Kobayashi T, et al., 2020).

In addition, screening for hepatitis B and latent tuberculosis is indicated in the majority of patients at the time of admission to avoid postponing potential rescue therapy with an anti-TNF drug. For the majority of patients, intravenous corticosteroids are the first step in serious

medical care. Since they are more likely to have thrombosis, patients with active ulcerative colitis should begin taking anticoagulants such low-molecular-weight heparin in addition to receiving adequate fluid and electrolyte replacement. Patients in hospitals should routinely be evaluated by a multidisciplinary team.. (Higashiyama M, et al, 2019).

A diagnostic step-up strategy that combines 5-ASA, corticosteroids, and thiopurines like azathioprine and 6-mercaptopurine can be used to effectively control and treat many ulcerative colitis patients in an outpatient setting. For instance, a conventional regimen that includes oral 5-ASA, rectal 5-ASA, or steroids will work best for individuals with left-sided, mild-to-moderate, or severe colitis. Oral corticosteroids are advised if the patient's condition does not improve after attempting this initial course of therapy. Due to their higher safety profile, topical drugs like budesonide MMX and beclomethasone dipropionate are recommended to systemic steroids. Even though corticosteroids must be reduced after complete remission to avoid side effects, 5-ASA therapy should be continued for recovery management. (Nielsen O H, et al., 2020).

1.4 METAGENOMICS

Analyzing genetic material that has been directly collected from environmental components is known as metagenomics. Metagenomics, which uses screening tools based on sequencing and operation, reveals details about the microbiomes of species that cannot be cultivated in the natural world (Datta S, et al., 2020).

Metagenomics is a revolutionary method for analyzing microorganisms from a particular habitat using functional gene screening or sequencing analysis. A wide range of topics are covered by metagenomics analysis, such as microbial diversity, community composition, genetic antecedents, biological functions, interactions with the environment, etc. Shotgun sequencing has been supplanted by high-throughput, next-generation sequencing (NGS), and third-generation sequencing (TGS) technologies. In 1998, the phrase "metagenome," also known as the "microbial environmental genome," was first used to describe "the genome of the whole microbiota observed in nature." The word "metagenome" now refers to the collection of genomes that can be found in samples of naturally occurring bacteria and fungi (Zhang, et al., 2021).

Following recent advancements in high-throughput sequencing technology and computational methodologies, genome-resolved metagenomics adds a new level to metagenomic data management. In specifically, it entails the recovery of high-quality or draft microbial genomes as well as their taxonomic classification and functional annotation. As illustrated in Figure 1.7, numerous studies have employed genome-resolved metagenome analysis to reveal new microbial communities in human and ecological metagenomes (Kayani, et al., 2021).

Metagenomics is also very beneficial for analyzing soil microbiology. More alternative microorganisms are anticipated in one gram of soil than in all the cultivated microbial species to yet. Therefore, it appears that metagenomics is the most effective culture-independent technique for figuring out soil microbial diversity and studying how this biodiversity is impacted by changing environments (Wael N. Hozzein, 2020).

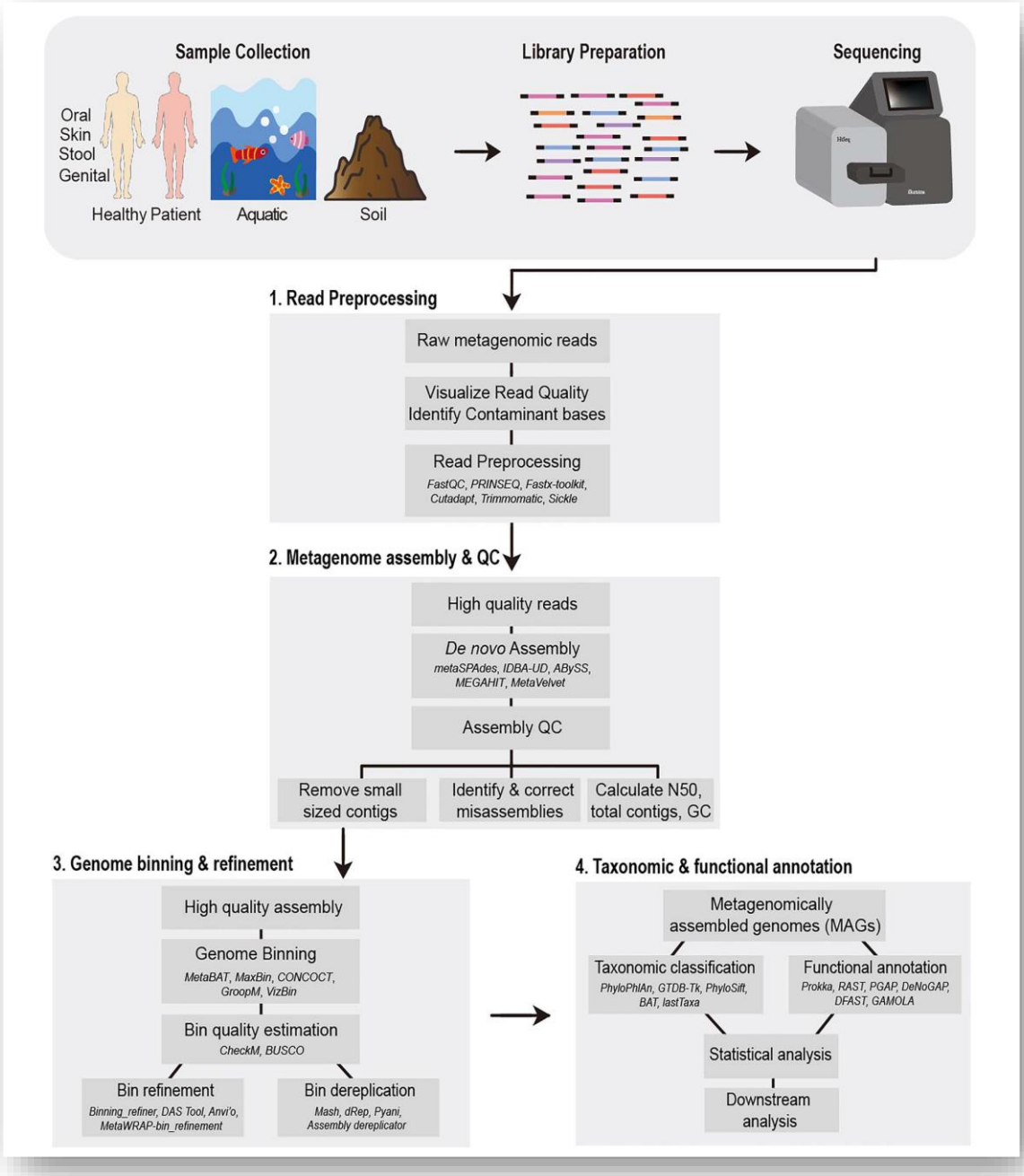


Figure 1.7 The figure describes the pipeline for genome-resolved metagenomics analysis is depicted in this figure. The four stages of a typical genome-resolved analysis of metagenomes from any source are read preprocessing, metagenome assembly and quality control, genome binning and refinement, and taxonomic and functional annotation of the returning MAGs. For each stage, the image also provides samples of numerous tools Colombel J.-F, et al., 2018.

Metagenomics has also gained appeal as a method for identifying and examining the range of human viruses found in aquatic systems, including both their role as waterborne pathogens and as indicators of water quality. Recent developments in environmental viral metagenomics have made it possible to discover, recognize, and completely sequence viruses in a variety of environmental and clinical materials. Traditional molecular methods, such as polymerase chain reaction (PCR) and sequencing, were developed before metagenomics and were predominantly used to investigate and identify enteric virus species in various natural habitats. Metagenomics has allowed for greater in-depth study of the major waterborne viruses found in wastewater treatment plant effluents and surface water (Osunmakinde C O, et al., 2018.

Genomic DNA must first be extracted from any region of the soil sample that comprises the whole population in order to do a metagenomic study. After the DNA has been isolated, a DNA library is created and utilized to search for a certain gene. A suitable DNA extraction technique must be described in order to capture the variety of the entire microbial population in the targeted system.. This is still one of the trickier steps in the metagenomic analysis process. As a result, the creation of a standard method for DNA extraction from soils is difficult since the chemical and physical features of soils vary substantially and extensively depending on the type of soil studied. However, the analysis of the massive amount of sequencing data that was produced from the constructed library is by far the most difficult stage in the metagenomics screening process. Over time, a variety of bioinformatic tools have been developed to help with the analysis of metagenomic data and comparison with databases (Mazziotto M, et al., 2018).

1.5 APPLICATIONS OF METAGENOMICS

Targeted metagenomics and shotgun metagenomics are two frequently used metagenomic applications. Targeted metagenomics or microbiomics are frequently used to examine the phylogenetic diversity and relative abundance of a particular gene in a sample. This is accomplished by searching an environment for an individual gene's whole complement of sequences. With this method, the diversity of small-subunit rRNA (16S/18S rRNA) sequences in a sample is commonly evaluated. Small subunit rRNA sequencing is frequently used by microbial ecologists to assess the taxonomic diversity of an ecosystem. It can also be used to look into how environmental toxins affect the make-up of microbial communities. Shotgun metagenomics analyzes the entire genetic composition of an environmental population using

genomic sequencing. Although it is an effective method for determining a microbial community's functional potential, its application is limited by the depth of sequencing. To undertake a full examination of the community's functional potential, however, adequate coverage of every member's entire genetic make-up is necessary. Shotgun metagenomics often favors the dominant bacteria in a population by only examining the genetic material of the community's low-abundance members. Furthermore, the correct annotation of a range of gene sequences, the majority of which lack homologs in the present sequence databases, is required in order to assess metagenomic sequencing data (Ram N. Bharagava, et al., 2019).

The genetically modified particles known as viruses can infect not only human and animal cells but also bacteria, archaea, and fungi. They are present everywhere. Given the expanding interest in figuring out how the structure and function of viruses relate to human health and disease states. The area of viral metagenomics applications, which will continue development and be utilized in a number of areas from discovery to surveillance, is also anticipated to be significantly impacted by the return of the development of worry materials and databases (Tasha M. Santiago-Rodriguez and Emily B. Hollistera, 2020).

A few examples of clinical metagenomics applications to date include the identification of tumor-associated viruses and their genomic integration sites, infectious disease diagnostic tests for a variety of syndromes and sample types, microbiome analysis in both healthy and diseased conditions, gene-expression characterization of the human host response to disorder, and clinical metagenomics. Apart from the detection of infectious diseases, the uptake of metagenomics and next-generation sequencing (mNGS) in clinical laboratories has been slow, with the majority of technologies still not being used in ordinary clinical practice, as shown in Figure 1.8. But these applications will likely alter diagnostic microbiology in the near future given their diversity and significant promise for clinical benefit. (Charles Y. Chiu and Steven A. Miller, 2019).

Several approaches for assessing metagenomics of cancers, involving both the experimental and the computational strategies, have indeed been advanced during last years with the evolution of sequencing methods. Moreover, there are still obstacles to overcome in researching metagenomics-related tumors. Both computational and experimental approaches for metagenomics are still in their infancy, and greater attention should be paid to this promising field. As an examples of that, Endometrial cancer (EC) is considering among the

most prevalent malignant cancer in women. The existence of poor prognostic variables promoting tumor recurrence contributed significantly to EC patient death. Using bioinformatics technologies to investigate the involvement of the E2F family in EC. In EC tissues, scientists discovered that E2F1, E2F2, E2F3, E2F7, and E2F8 expression was dramatically elevated whereas E2F4 expression was downregulated. This discovery suggested that the E2F family had therapeutic promise in preventing and treating EC (Huang G, et al., 2021).

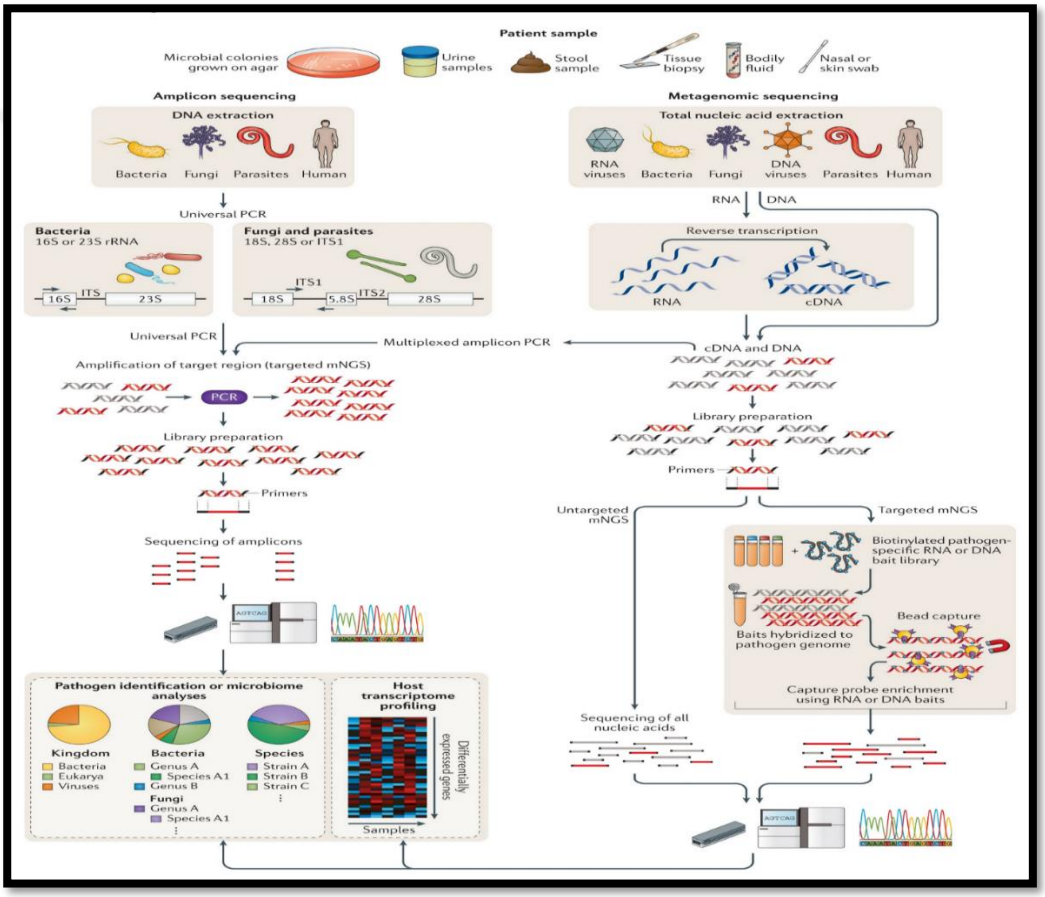


Figure 1.8 The figure describes the clinical applications of metagenomic sequencing, descript the whole process of the application of metagenomics approaches human and other animal diseases from collection samples and through the analysis, Targeted mNGS, AND sequencing (Charles Y. Chiu and Steven A. Miller, Clinical Metagenomics, Natural Review (Genetics), Volume 20, 2019).

1.6 CORRELATION OF METAGENOMICS WITH CROHN'S DISEASE

IBD, also referred to as Crohn's Disease or Ulcerative Colitis, is a condition that involves 163 different gene loci. Interestingly, 30 of these genes are specifically associated with Crohn's Disease, while the other 110 are more commonly found in both Crohn's Disease and Ulcerative Colitis. These genes play a variety of roles in different metabolic processes, such as autophagy, maintaining the epithelial barrier, and the body's immune responses. It is important to note that IBD cannot be traced back to just one single gene.. Provided by epithelial cells, the initial line of defense against invading pathogenic organisms involves the recognition of bacterial components through receptors for TLR and NOD2. Receptors such as TLR and NOD2/CARD15, which identify extracellular bacterial peptides, play a crucial role. The synthesis of antimicrobial peptides (HBD-defensins) and the expression of MHC molecules are then initiated to kickstart the adaptive immune response of the mucosa. Unfortunately, individuals with Crohn's disease experience a compromise in these signaling pathways leading to changes in microbial peptide synthesis and the innate immune response. (Ahlostrom, 2022)..

The initial gene to be found associated with the susceptibility of Crohn's disease is NOD2. NOD2 brings ATG16L1 into the area where pathogens are present, thus stimulating autophagy. It has been identified that three noteworthy mutations (R702W, G908R, and 1007fs) in the NOD2 gene are connected to CD. These mutations, impacting the amino acid composition, are linked to an escalated chance of developing Crohn's disease. Furthermore, genes involved in autophagy (ATG16L1), innate immunity (NOD2), and adaptive immunity (IL-23R) are also relevant factors.. (Feki, et al., 2018).

The ATG16L1 gene plays an important role in maintaining the functioning of intestinal epithelial stem cells by protecting Paneth cells. These specialized cells contain lysozyme enzymes that break down the cell walls of bacteria and defend the intestinal flora through phagocytosis. Abnormalities in the lysozyme granules of Paneth cells have been associated with the ATG16L1 (T300A) polymorphism in patients with Crohn's disease. The autophagy mechanism is altered when NOD2 variations make it difficult to attract ATG16L1 to the site of a bacterial infection. This hinders the body's defense mechanism against pathogens by

reducing NOD2 activation, which typically attracts ATG16L1 to the microbes' location and triggers auto-phagosome production.. (Gammoh, 2020).

Increases in IL-23 are beneficial for controlling infections and initiating inflammatory responses through IL-17 release from neutrophils and T cells. The presence of IL-1, TGF-, IL-6, and IL-23 promotes the differentiation of Th17 cells from CD4+ T cells, resulting in the production of IL-17A and IL-23R. Subsequently, IL-23 facilitates the production of IL-17A, IL-17F, IL-21, and IL-22. Furthermore, IL-12 and IL-6 secreting dendritic cells are increased in CD (Álvarez-Salamero et al., 2020).

1.7 CORRELATION OF METAGENOMICS WITH ULECRATIVE COLITIS

Notably, UC is associated with 23 distinctive genes, while CD and UC share a staggering 110 genes (Ahlstrom, 2022). In simple terms, mucus is deemed a vital barrier between the epithelium's surface and the microbiota. Its composition primarily consists of mucins, which are secreted by goblet cells and play a crucial role in preserving and safeguarding the intestinal mucosa. Nevertheless, the mucosal homeostasis may be disrupted in UC patients due to aberrant mucin M1/MUC5AC expression, alteration in MUC2 glycosylation patterns, and a decline in MUC2 secretion... Astonishingly, CD and UC have a whopping 110 genes in common, while UC is associated with a distinct 23 genes (Ahlstrom, 2022). Essentially, mucus acts as a crucial barrier, protecting the surface of the epithelium from the microbiota. It is primarily composed of mucins, which are secreted by goblet cells and have a vital role in preserving and safeguarding the mucosa of the intestines. However, UC patients may experience disruption in mucosal homeostasis due to abnormal expression of mucin M1/MUC5AC, changes in glycosylation patterns of MUC2, and reduced secretion of MUC2 (Nacimento et al,2020).

Crucial for maintaining the integrity of the intestinal epithelium and fending off infections, ILCs are positioned near the intestinal lining in UC. When stimulated by an altered microbiota and excessive antigens, ILCs may struggle to generate anti-inflammatory cytokines, potentially triggering or perpetuating colon inflammation. This can result in ILCs producing abundant levels of IL-17, IL-22, IL-23, and interferon gamma (IFN-) via tissue (Kunkl et al, 2020) states that gene expression or synthesis...

IL-23 plays a crucial role in both innate and adaptive immunity by encouraging the release of IL-17, which aids in infection control by neutrophils and inflammation response by T cells. Additionally, IL-1, TGF-, IL-6, and IL-23 work together to enhance the differentiation of CD4+ T cells into Th17 cells, ultimately leading to the production of IL-17A and IL-23R. This ultimately results in an increase in the production of IL-17A, IL-17F, IL-21, and IL-22, all thanks to IL-23's presence. The protective function of IL-17A is to safeguard barrier tissues such as the skin and intestinal mucosa.. The rise in IL-23 is caused by bacteria such as Escherichia coli and Enterococcus faecalis. This leads to the activation and phosphorylation of JAK2 and STAT3, primarily due to the binding of IL-23 to its receptor. As a result of this connection, JAK2 and STAT3 are activated and later induce the expression of the transcription factor RORt, which is crucial for the production of IL-17 (Álvarez-Salamero et al., 2020).

2. GENERAL PARTS

Comparing the gut microbiome of individuals with Crohn's disease and ulcerative colitis to healthy control samples was the main objective of this study. Specifically, the research focused on genes associated with immunity that increase the likelihood of developing inflammatory bowel disease. To assess the impact of mutations on microbial imbalances in both UC and CD patients, a combination of computational techniques and DNA sequencing methods were used. Through analyzing a diverse range of blood samples, the different components of the gut microbiota were identified. this study also analyze human blood samples in order to determine the specific microbial makeup and functions that differentiate individuals with IBD from those who are healthy. To validate the accuracy of our methods in identifying IBD, this study additionally compared the 16S ribosomal RNA data from disease samples and control samples separately.

3. MATERIALS AND METHODS

3.1. MATERIALS AND METHODS

3.1.1. Preparation of Blood Samples for Ulcerative Colitis, Crohn’s disease and control samples

This study was carried out at two hospitals and one clinic in Istanbul, Turkey, to collect human blood samples for Crohn's disease, Ulcerative colitis, and control samples. Istanbul University has accepted to participate in this experiment and has given their informed consent for the collection, storage, and analysis of human blood samples by the molecular biology and genetics department. Application of Metagenomic Approaches in Ulcerative Colitis and Crohn's Diseases" file number 2020/1762, which the supervisor of the project is the responsible researcher and will be conducted by Omar Sajer Naser NASER. The study titled was discussed in the meeting of our board dated 28/05/2021 and numbered 11, and was deemed appropriate in terms of ethics, and the minutes are attached. 10 blood control samples were collected from a group of participants as healthy individuals, 25 blood samples for Ulcerative Colitis patients, and 25 blood samples for Crohn’s patients from both genders, ranging in age from 20 to 65, have been obtained to compare the two diseases and also to compare them with the control samples. İn total, it was 50 samples for both diseases and 10 samples for the control. This study was applied to the patients to obtain their approval because the two hospitals and the clinic provided their informed consent for the study. A formal informed consent form with details about the study's aims was signed by each participant. Within a year of their initial consent, the research participants completed surveys that asked about general information about inflammatory bowel diseases (IBDs), as well as specific information about Crohn's and ulcerative colitis, including disease phenotypic characteristics, surgical histories, and treatment histories. Additionally, to learn about CD and UC's smoking history and family background, only individual surveys were used.

3.1.2. The DNA isolation from blood samples

The DNA from blood samples were applied following the instructions kit manual(s) from Hibrigen Biyotechnology R&D Industry and Trade Inc. It was added 400μl of blood sample into microtube and centrifuge for 2 minutes at 14,000 RPM. Removing 200μl of blood sample before begans the isolation. According to the kits instructions, it was added 800μl of DP

buffer then let rest for 3 minutes while mixing up-side-down slowly. Centrifuge for 5 minutes at 7500 RPM to let the vast amount DNA settle down of the Eppendorf, then we removed the up part of the blood and use the DNA at the bottom for continuing the isolation.

It was added 200μl of DA buffer and mixing slowly by the pipette then we mix by utilizing Vortex Mixer for 1 minute. It was added 20μl of Proteinase-K (the concentration of Proteinase-K was 50–100µg/ml) (already left in the freezer for the later use), after that, it was added directly 220μl of DB buffer and utilize the water bath for 20 minutes at 65°C. The Buffers are used in the DNA isolation kit because it keeps the pH stable. When cells are lysed open they release many types of compounds that can change pH which could alter the properties of the target molecule. Including a buffer prevents this and keeps the pH to something similar to that in the cell.. While the use of the WiseBath, mixing the samples up-side-down slowly every 5 minutes. After ending the WiseBath, we add 220μl Ethanol and utilize the Vortex mixer for less than a minute. We transfer the samples into filter and centrifuge for 2 minutes at 12,000 RPM, then we remove what remains under the filter then centrifuge again for 1 minute and then we remove what had remains down. We perform this part once more again then we take the Eppendorf which have the DNA and stored.

The results of the isolated DNA were quantified using spectrophotometer (thermo-NanoDrop-USA). The DNA concentration (ng/ml blood) and purity [absorption rate at 260/280 (A260/A280)] in tubes containing pure DNA were measured using the NanoDrop 2000 apparatus. Pure DNA has an A260/A280 ratio between 1.8 and 2.0...

3.1.3. Primer Preparation for 16S rRNA

The DNA sequences for total bacterial 16S rRNA-specific primer pairs are shown in Table 3.1. The specificity of the primers was evaluated using NCBI and BLAST (https://www.ncbi.nlm.nih.gov/tools/primerblast/). We purchased primers at Macrogen in Korea.

Table 3.1 The table describes the primers for PCR analysis.

Target microorganism	DNA sequences of the primers (5′-3′)		Annealing Temperature (ºC)	Product Size (bp)
Total bacteria	Forward	5'-GAGTTTGATCCTGGCTCAG-3'	57	20
	Reverse	5′-TACCTTGTTACGACTT-3	44	16

Table 3.2 The table describes the references of primers sequences

Oligonucleotide	Target site	Sequence (5′→3′)	Specificity	Source
8bF	8– 26	GAGTTTGATCCTGGCTCAG	Bacteria	Burggraf et al. (1992)
1512uR	1493–1513	TACCTTGTTACGACTT	Universal	Lane (1991)

3.1.4. Polymerase Chain Reaction (PCR) analysis

First, in accordance with usage manuals, 100 M (main stock) of lyophilized primers were diluted. After then, 10 M of intermediate stock for each pair of primers was created. FastStart High was used for PCR analysis with primers created for the designated gene areas. Purchased a dNTPack kit from Istanbul, Turkey's Hibrigen Biyotechnology R&D Industry and Trade Inc. The T100 thermal (Bio rad, France) cycler was used to complete entire processes. Before being added to the solutions, each reagent was vortexed and centrifuged. Samples were used in accordance with the company kit protocol. The reactions were carried out in a final volume of 25 HL. The PCR tubes contain 11.5 HL of DNA template, 12.5 l of *Taq* PCR Master Mix from Hibrigen Biyotechnology R&D Industry and Trade Inc, 1 HL of forward primer, and 1 HL of reverse primer. The GeneRuler DNA Ladder Mix was used to size and roughly quantify duplicate DNA fragments on the agarose gel

Table 3.3 The table describes the PCR reaction ingredients.

Components (Per reaction)	**Volume (µl)**
PCR-grade water (ddH2O)	1 µl
Forward Primer (intermediate stock 10 µM)	1 µl
Reverse Primer (intermediate stock 10 µM)	1 µl
PCR Grade Nucleotide Mix	12 µl
DNA sample	10 µl
Total Volume	25 µl

The T100 thermal cycler was used to do the PCR, and various modifications to the protocol listed in Table 3.4 were made depending on the difference in DNA concentrations. The capacity was 25 µl, and the TM were set between 55 and 68 °C.

Table 3.4 The table describes the PCR analysis procedure.

Program Name	Initial Denaturation	Amplification			Final Elongation	Cooling
		Denaturation	Annealing	Elongation		
Cycles	1	25-35			1	1
Temperature [°C]	94°C	94°C	61.3*°C	72°C	72°C	4°C
Time	3 min.	30 sec.	30 sec.	1 min.	10 min.	Unlimited time

- * The recommended temperature for our gradient PCR, which yielded results between 55 and 68 degrees Celsius, is 61.3 degrees Celsius.

3.1.5. Agarose Gel Electrophoresis

Gel electrophoresis had applied to the experiment to capture the DNA. İt was added 12.5 μl of Taq master into new tubes. Then added 10.5 μl of each sample in the tubes. After that, prepared the TAE (Tris/Acetate/EDTA) buffer by adding 20 ml of the TAE 1x buffer to 880 ml (ddH_2O) and mixing, then used it in preparing the gel. 1.5 gr agarose added to 100 ml buffer (which prepared) then applied into the microwave and heated for 30 seconds, it was added 2.8 μl of ethidium bromide with concentraction 1 µg/mL and mix well carefully.

The agarose then poured into the gel mold while still hot. The gel tray was put inside the casting device. The agarose was allowed to cool and set. (The agarose was allowed to cool on the work surface or by incubation in a water bath set at 65 °C. The gel tray that was inserted into the casting mechanism). The gel was taken out and put in the gel box by the comb. For every 5 HL of samples, 1 HL of the loading dye is applied. Then, for 45 minutes, the Gel electrophoresis instrument was set to V:65, mA:95.

3.1.6. Next generation sequencing (NGS) of PCR protocols

Next generation sequencing (NGS) was done using the results of the amplicon sequencing. In 60 blood DNA samples (25 samples for each of the UC, CD, and 10 (healthy) control groups), the 16S rRNA region was targeted and amplified using the primers 8BF(5'-AGAGTTTGATCCTGGCTCAG-3') and 1512uR(5′-TACCTTGTTACGACTT-3). attached

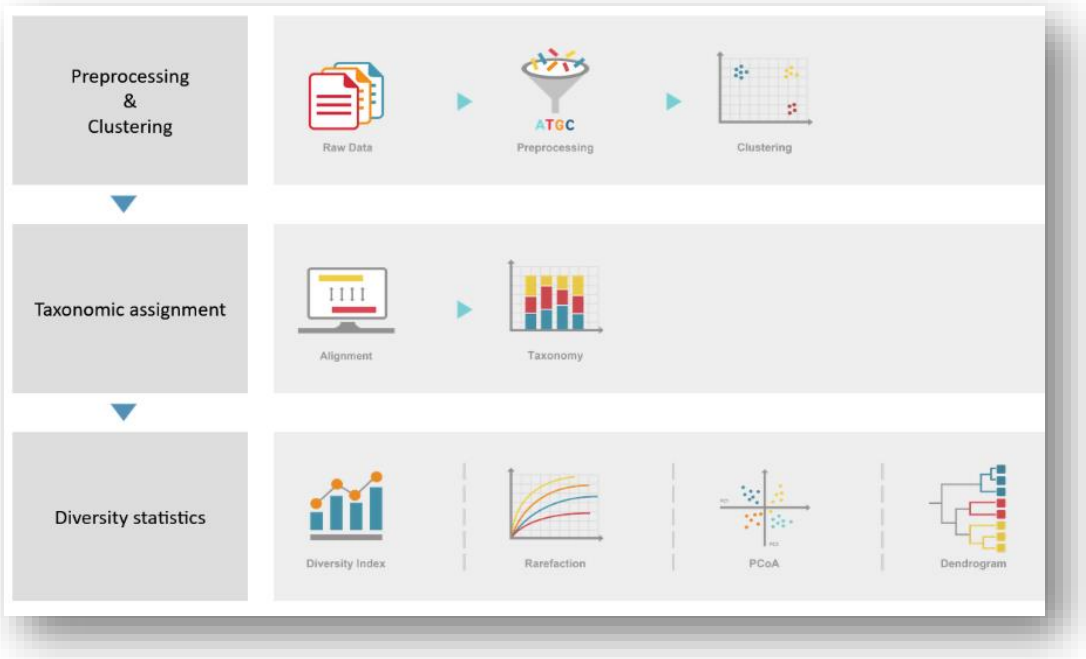


Figure 3.1 The figure describes the workflow and data analysis.

to an via PCR, an adapter sequence. After that, agarose gel electrophoresis was done to quantify and qualify the PCR results. On an agarose gel, duplicate DNA fragments were sized and approximatively quantified using the GeneRuler DNA Ladder Mix. Figure 3.1 depicts the data analysis procedure, which also includes the taxonomic assignment, grouping, and diversity statistics..

3.1.7 The preprocessing

- • Extra-long tails are clipped and short reads are removed.
- • CD-HIT-DUP is used to cluster filtered reads with 100% identity.
- Chimeric readings have been located.
- • Primary clusters recruit secondary clusters.
- • The elimination of noise sequences in clusters of size x or smaller. Here, x is calculated using statistics.
- At a user-specified OUT threshold (for instance, 97% ID at certain levels), remaining representative readings from non-chimeric clusters are sorted into OUTs using a greedy technique.

3.1.8 Clustering

- Uncertain readings are eliminated, and excessively long tails are trimmed.
- • • Recognize chimeras and exterminate them.
- Reference files for RDP can be found at http://www.monthur.org/wiki/RDP_reference_files.
- The creation of a distances matrix.
- Utilizing the average neighbor approach for clustering.

3.1.9 Taxonomic assignment

- To determine taxonomy, use representative sequences from each OUT.
- Create phylogenetic trees using representative OUT sequences that have been aligned and filtered.
- Utilize statistical analysis and visualization to produce graphics that are suitable for publication.

3.2. Experimental Methods and Workflow

3.2.1 Experiment Overview

As seen in Figure 3.2, the Illumina (NGS) workflow consists of 4 fundamental phases.:

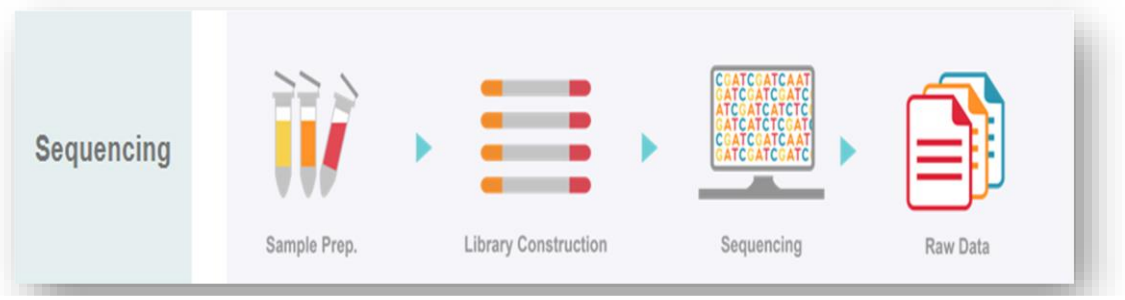


Figure 3.2 The figure describes the experiment overview.

3.2.2 Sample Preparation

DNA/RNA is taken out of a sample and used to build a library. Qualified samples move on to library creation when quality control (QC) has been completed..

3.2.3 Library Construction

The sequencing library is created by 5' and 3' adapter ligation after randomly fragmenting the DNA or cDNA sample. Instead, the library preparation process can be streamlined by combining the fragmentation and ligation reactions into a single step called "tagmentation (Illumina DNA Prep (M) Tagmentation Library Preparation for use on an Illumina MiSeq Sequencer)". Gel purify and amplify the adapter-ligated fragments using PCR afterwards.

3.2.4 Sequencing of NGS samples

Using bridge amplification, each fragment in the library is multiplied into unique, clonal clusters, after being caught on a lawn of surface-bound oligos that are complementary to the library adapters in a flow cell. The templates are then prepared for sequencing, following the completion of cluster generation.

Utilizing a one-of-a-kind, terminator-based reversible method, the Illumina SBS technology detects individual bases as they are appended to DNA templates. The presence of all 4 reversible, terminator-bound dNTPs during each sequencing cycle lessens raw error rates and prevents incorporation bias, offering a substantial advantage over alternative technologies. Consequently, sequencing of homopolymers and repetitive sections yields remarkable accuracy, eliminating errors specific to context. Additionally, the sequencing data undergoes

conversion into raw format for subsequent processing. Using an integrated primary analysis program known as RTA (Real Time Analysis), the base calling and system control software of the Illumina sequencer produces raw images. An essential step in the process is the conversion of the BCL (base calls) binary into FASTQ, which is accomplished by the illumina package bcl2fastq. It is worth noting that the readings do not undergo any adapter removal.

4 RESULTS

A study cohort of 60 individuals participated, comprising 10 healthy controls, 25 individuals with Crohn's disease, and 25 individuals with Ulcerative Colitis. None of the patients had first-degree kinship. Conducted was a longitudinal investigation into the gut's microbial populations. The patients shared essential details, such as age, the specific ailment, IBD family history, smoking habits related to IBD, and any medication being taken to facilitate recovery.

The bacterial population in blood samples from IBD (specifically UC and CD) and the control group were analyzed using gene amplicon sequence analysis. Data was cleaned by combining raw data with other data. To evaluate read quality, the FASTQ file format was utilized. Representative sequence OUT was performed using a greedy technique to cluster non-chimeric clusters into OUTs at a species level ID cutoff of 97%. From the left side to the right, Figure 4.1 show cases the gel electrophoresis outcomes for all samples (Loading, Control, and UC samples, then empty space to distinguish them from the other sections that went as Loading, Control, and the CD samples). To establish the phylum annotation at each taxonomic order (kingdom, phylum, class, order, family, genus, and species), a blast analysis with QIIME was conducted for each OTU.

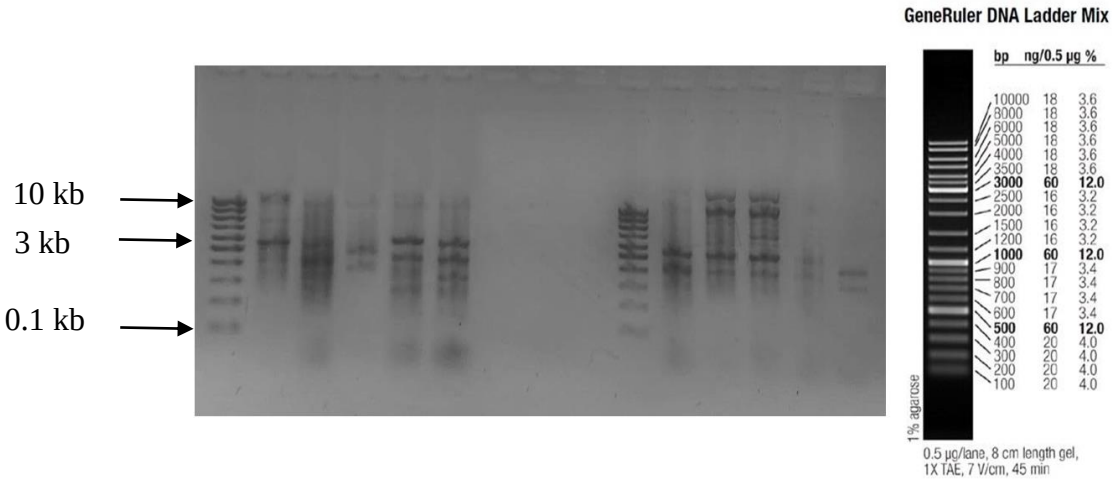


Figure 4.1 The figure describes the gel electrophoresis for whole samples.

4.1.1 Summary of Produced Data

In determining the outcome for each of the 10 samples, the following is calculated: reads, bases, GC (%) content, Q20 (%), and Q30 (%). For example, in the case of C1, a total of 118,302 reads are produced, with a cumulative base count of 35.6M bp. Referring to Table 4.1, it is noted that the GC content measures to be 52.56% while the Q30 ratio is 79.21%.

Table 4.1 The table describes the total number of bases, reads, GC (%), Q20 (%), and Q30 (%).

Sample ID	Total read bases (bp)	Total reads	GC(%)	AT(%)	Q20(%)	Q30(%)
C1	35,608,902	118,302	52.56	47.44	89.13	79.21
C2	53,534,656	177,856	52.38	47.62	90.98	82.31
CD_1	40,935,398	135,998	52.4	47.6	90.88	81.82
CD_3	55,149,220	183,220	51.8	48.2	91.67	83.02
CD_6	41,969,634	139,434	52.58	47.42	90.3	81.18
CD_7	55,181,728	183,328	51.78	48.22	90.5	81.57
UC_2	48,510,364	161,164	52.17	47.83	91.27	82.38
UC_4	45,794,742	152,142	52.44	47.56	90.82	81.67
UC_5	44,275,294	147,094	52.25	47.75	91.28	82.67
UC_8	69,422,640	230,640	53.02	46.98	91.55	83.01

Table 4.1 above presents data for 10 samples. Out of the 10, two are for healthy individual (C1 and C2), for samples are from patients suspected have crohn’s disease, while the rest four are from patients suspected to have ulcerative colitis. From the table each subject is presented along with the respective base pair, nitrogenous base pair percentages (AT for adenine-Thymine and GC for Guanine-Cytosine). Furthermore, there is also provision for nucleotides consistency (phred score) Q 20(%) and Q 30(%).

As can be observed in the chart, individuals diagnosed with ulcerative colitis and crohn's disease exhibit elevated amounts of genetic substances compared to healthy patients. The confirmation of this can be found in the presence of nitrogenous base pairs like AT and CG, which serves as evidence. Abundant research findings further support the notion of an increased quantity of genetic material. Italy's Vrakas, Mountlouris, Michalopoulous, Paramanchis, Papatheodorilis, and Trathas (2017) conducted a study examining the composition of intestinal bacteria and the movement of bacteria in inflammatory bowel

disease, affirming this theory. According to their findings, there are commensal gut bacteria present in the peripheral circulation that may potentially trigger inflammation..

Recently diagnosed adult patients with active or inactive Crohn's disease (CD) and active or inactive ulcerative colitis (UC) and healthy individuals had their blood collected for analysis. Antibiotics were not administered to any of the patients. The composition of their microbiota was determined, and the genotyping of NOD2/CARD15 was also examined...

According to the findings, IBD patients' blood samples had higher concentrations of total bacterial DNA than healthy controls in comparison to the instances of inactivity, the active cases of IBD exhibited higher concentrations of total bacterial DNA. Dysbiosis in IBD could be identified by the following species: an increase in *Bacteroides* spp. in both active and inactive samples, a decline in *Clostridium leptum group (IV),* and a decrease in *Faecalibacterium prausnitzi* in both active and inactive patients of IBD. The presence of NOD2/CARD15 mutations did not show any significant association with bacterial translocation. In summary, both active cases of CD and UC had notably higher quantities of blood bacterial DNA compared to healthy participants, as well as inactive cases of CD and UC. Active CD and UC cases displayed notably higher total bacterial DNA concentration compared to both inactive CD and UC samples and healthy controls. Notably, there were no noteworthy differences observed between CD and UC cases, as well as inactive cases, and healthy controls in terms of statistical significance.

Contrary to expectations, it can be observed from the figures above that the control subject 2 has a reasonably considerable increase in base pairs and total readings. It is clear from that subject that the base pair, total reads, and other pertinent characteristics along the row of that concerned subject are higher than those of the impacted cases CD1, CD6, UC2, UC4, and UC5. Although this is undoubtedly surprising, the same Italian study that was mentioned before supports the possibility of such a chemical paradox. That finding implied that bacterial DNA was present in the blood of 11 out of 20 healthy persons, or 55% of them..

There is also a significant statistical difference between the used CD and UC cases. According to the data in the above table, the mathematical mean average of all base pairs in UC is 52 million, while that in CD is 48.1 million. The results of the study mentioned above, however,

revealed that there had not been any statistically significant differences between CD and UC cases, as well as between inactive cases and healthy controls.

For a more comprehensive description:

The Phred quality score provides a numerical representation of each nucleotide's accuracy. A higher Q value indicates more accuracy. If Phred assigns a base a quality score of 30, for example, the probability of a base call error is 1 in 1000.

The Phred Quality Score Q is calculated using the probability of an inaccurate base call, P.

Sanger Quality (ASCII Character Code=Phred Quality Value + 33) is the encoding used.

Table 4.2 The table describes the percentage of the accuracy of each nucleotide.

Quality of phred score	Probability of incorrect base call	Base call accuracy	Characters
10	1 in 10	90%	!"#$%&'()*+
20	1 in 100	99%	,-./012345
30	1 in 1000	99.9%	6789:;h=i?
40	1 in 10000	99.99%	@ABCDEFGHIJ

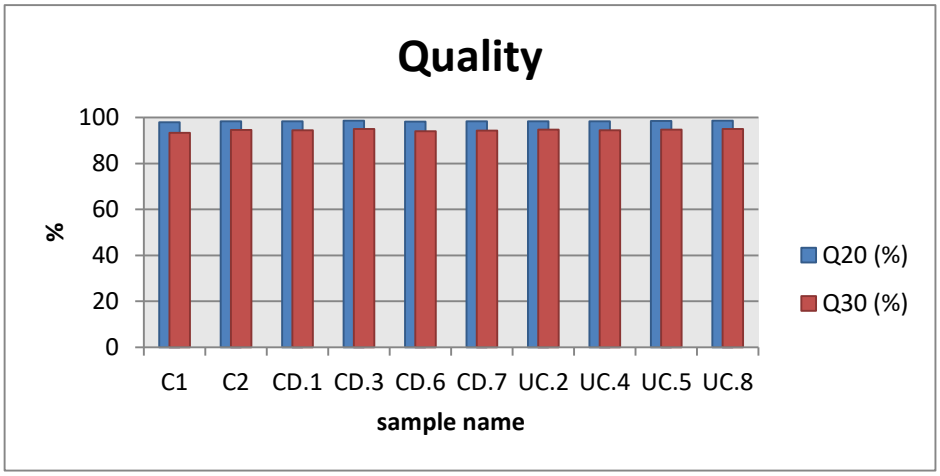


Figure 4.2 The figure describes the quality score for the disease and the control samples.

Table 4.3: The table describes the result of assembly

Sample Name	Total Bases	Read Count	N (%)	GC (%)	Q20 (%)	Q30 (%)
C1	26,378,952	56,883	0	52.08	97.97	93.26
C2	40,022,349	86,261	0	51.98	98.38	94.62
CD.1	30,777,392	66,304	0	51.96	98.39	94.46
CD.3	41,620,754	89,578	0	51.32	98.54	94.98
CD.6	31,311,397	67,542	0	52.17	98.2	94.07
CD.7	40,962,002	88,404	0	51.29	98.27	94.26
UC.2	36,243,640	78,174	0	51.7	98.37	94.66
UC.4	34,185,795	73,848	0	51.98	98.37	94.48
UC.5	33,288,136	71,758	0	51.82	98.42	94.7
UC.8	51,885,809	112,687	0	52.65	98.56	95.03

Details for the total results of sample sequencing are in table 4.3 above. The provides information on the percentage contents Guanine-Cytosine in all the samples used. After the analysis, it was revealed that sample UC 8 has the highest percentage composition of 52.65 followed by UC 5 with a percentage score of 51.82. Also from the table, UC 8 has the highest Q20(%) and Q30(%) of 98.56 and 95.03 in a respective manner.

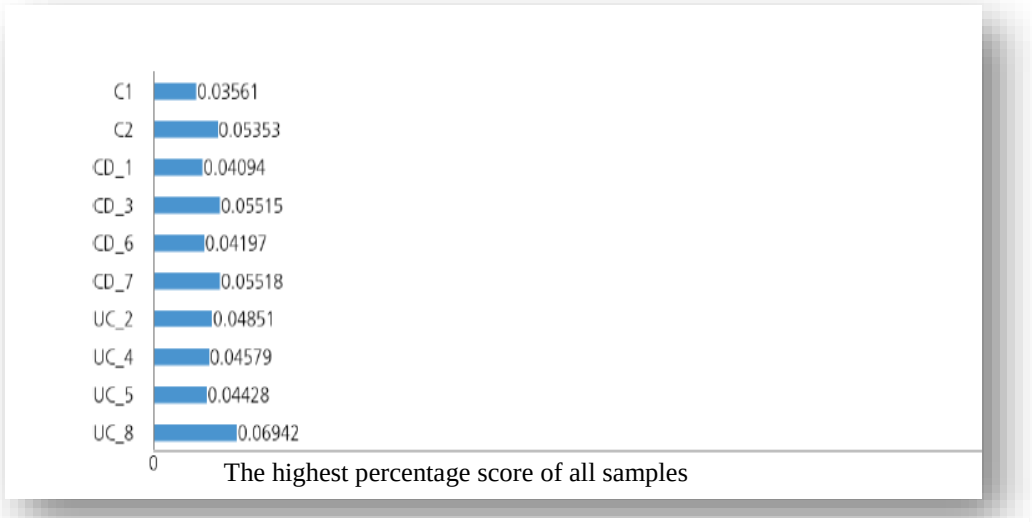


Figure 4.3 The figure describes the throughput of Raw data.

The Figure 4.3 above shows the throughput of all the raw data obtained from the study. I can be inferred that the sample titled UC8 is the highest with highest index of 0.06942, followed by CD 7, CD 3, the paradoxical C2, UC 2, UC 4, UC 5, CD 6, CD 1 and finally C1.

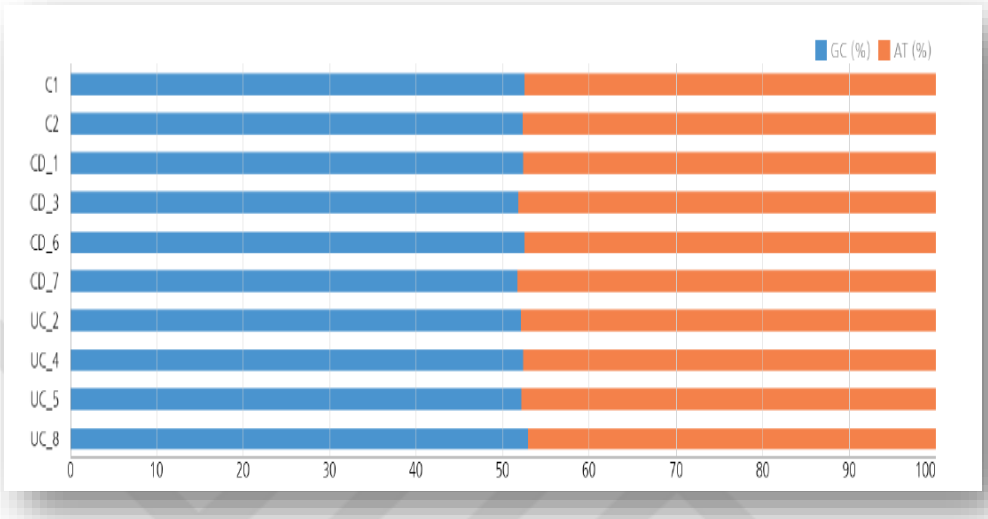
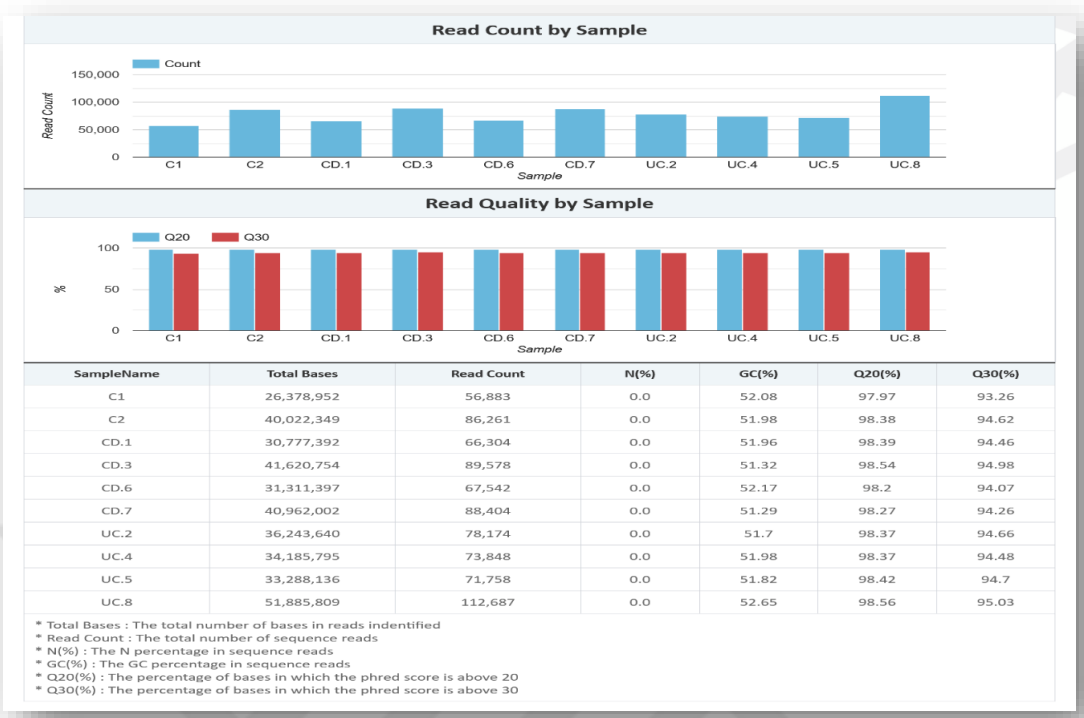


Figure 4.4 The figure describes the total read count of Raw data.

The read counts for all the samples (C1, C2, CD 1, CD 3, CD 6, CD 7, UC 2, UC 4, UC 5 and UC 8) are shown in the bar chart in Figure 4.4 above. The sample with the most reads is UC 8; it is followed in order by CD 3, CD 7, C 2, UC 2, UC 4 & 5, CD 6, CD 1, and C1. According to the ratings, UC 8 has the highest read count, while C 1 has the lowest read count. This is, of course, to be expected as healthy subjects, including control 1 (C 1), are supposed to have the least genetic material.



SampleName	Total Bases	Read Count	N(%)	GC(%)	Q20(%)	Q30(%)
C1	26,378,952	56,883	0.0	52.08	97.97	93.26
C2	40,022,349	86,261	0.0	51.98	98.38	94.62
CD.1	30,777,392	66,304	0.0	51.96	98.39	94.46
CD.3	41,620,754	89,578	0.0	51.32	98.54	94.98
CD.6	31,311,397	67,542	0.0	52.17	98.2	94.07
CD.7	40,962,002	88,404	0.0	51.29	98.27	94.26
UC.2	36,243,640	78,174	0.0	51.7	98.37	94.66
UC.4	34,185,795	73,848	0.0	51.98	98.37	94.48
UC.5	33,288,136	71,758	0.0	51.82	98.42	94.7
UC.8	51,885,809	112,687	0.0	52.65	98.56	95.03

* Total Bases : The total number of bases in reads indentified
* Read Count : The total number of sequence reads
* N(%) : The N percentage in sequence reads
* GC(%) : The GC percentage in sequence reads
* Q20(%) : The percentage of bases in which the phred score is above 20
* Q30(%) : The percentage of bases in which the phred score is above 30

Figure 4.5 The figure describes the read count of all samples DNA isolation.

Figure 4.5 shows the content of GC/AT by (%). As it shows the % GC is almost around 53%.

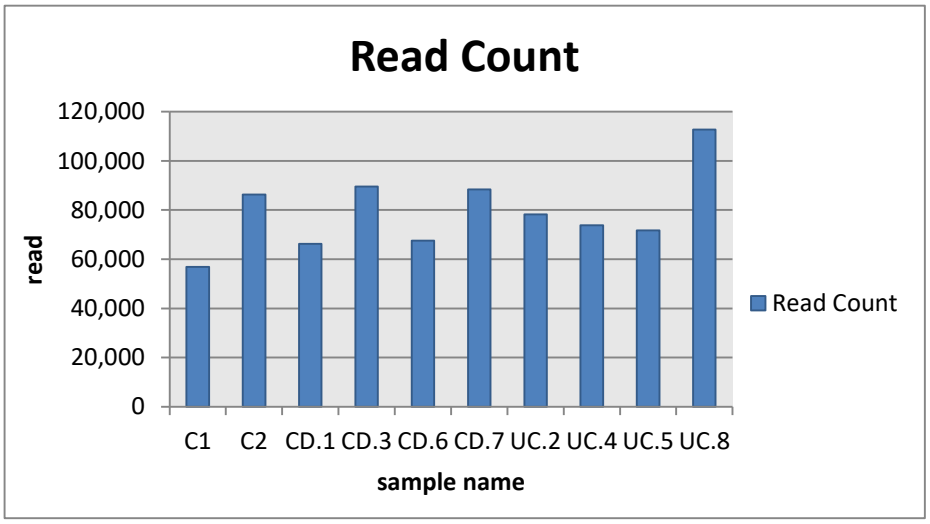


Figure 4.6 The figure describes the read count percentag of all samples.

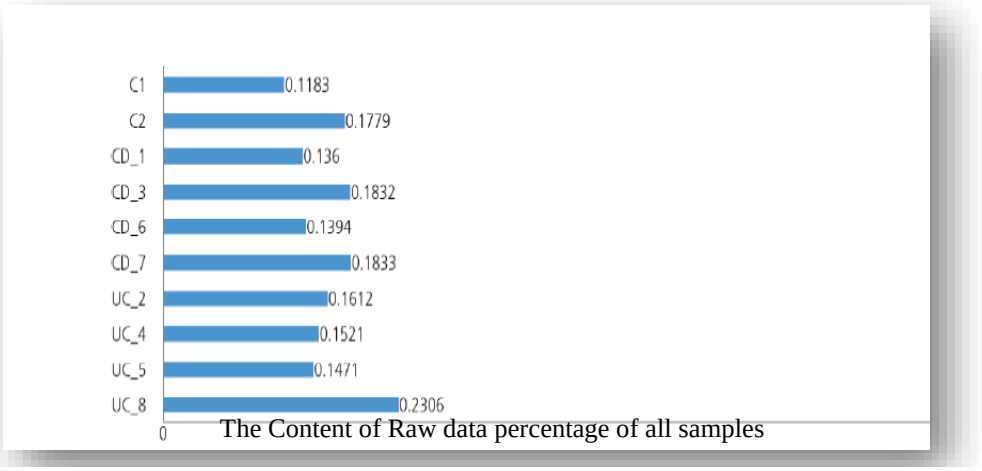


Figure 4.7 The figure describes the Data Content of the GC/AT.

The figure 4.7 above presents the nitrogenous base pair contents of the all the samples analyzed. It can be inferred that the sample titled UC 8 has the highest GC/AT ratio contents, with an index of 0.2306, followed by CD -7 and CD 3 with respective indices of 0.1833 and 0.1832. The least index is carried by C -1.

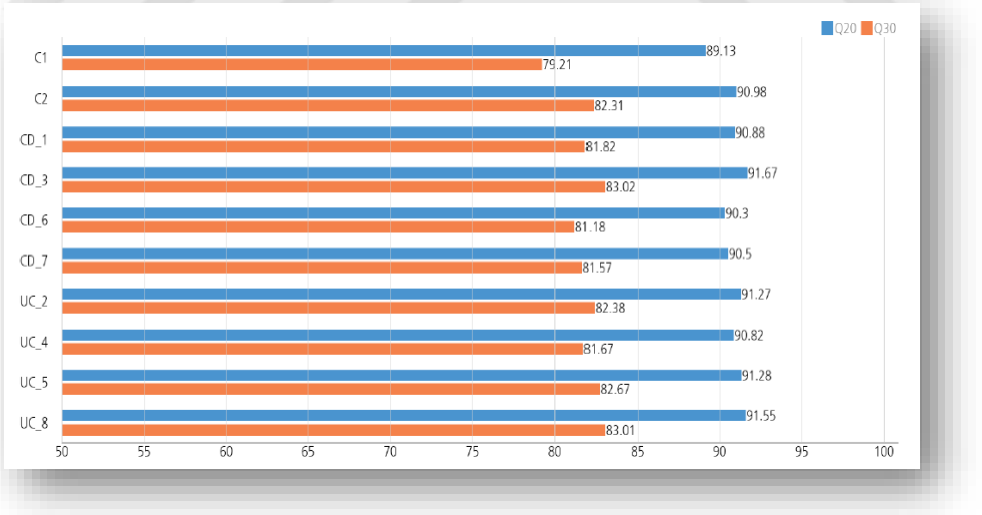
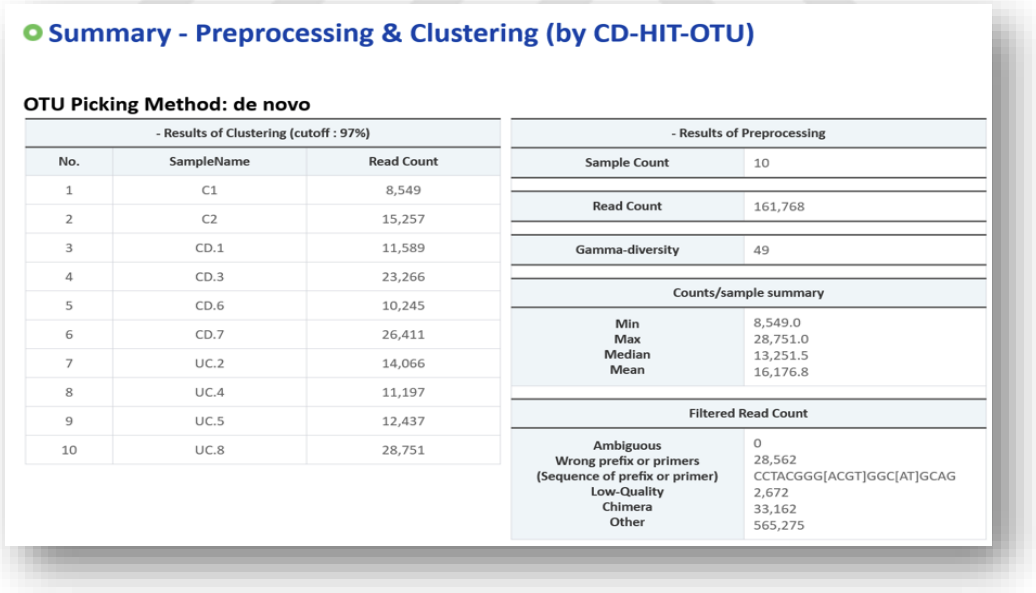


Figure 4.8 The figure describes the Q20/Q30 scores of raw data.

Figure 4.8, shows some summary information about preprocessing and clustering of sequences data the picture presents the phred score of all samples used. From the figure above, CD 3 has the highest Q20(%) and Q30(%) of 91.67%, and 83.02%, followed by UC 8

and The Q20(%) and Q30(%) figures for UC 5 are 91.55% & 83.01% and 91.28% & 82.27%, respectively. It suggests that the sample designated as CD 3 has the highest consistency of nitrogenous bases. It denotes an error rate of 1 in 100 in other wards and with reference to Q20(%), which means that each sequence read of 100 base pairs has a 99% call accuracy rate. In contrast, Q30(%) indicates an error rate of 1 in 1000, which means that there may be one error for every 1000 base pairs read during sequencing, with a corresponding call accuracy of 99.9%.

OTU (operational taxonomic unit) which is an operational definition of a species or group of species often used when only DNA sequence data is available. However, classification was utilized, as shown in the tables above. This was done to group the bacteria according to the relevant DNA sequence. Sequence can be determined by how similar two sequences are to one another, and operational taxonomic units are established using this similarity threshold. As a result, the criterion was attained at 97% (clustering findings).



Summary - Preprocessing & Clustering (by CD-HIT-OTU)

OTU Picking Method: de novo

- Results of Clustering (cutoff : 97%)

No.	SampleName	Read Count
1	C1	8,549
2	C2	15,257
3	CD.1	11,589
4	CD.3	23,266
5	CD.6	10,245
6	CD.7	26,411
7	UC.2	14,066
8	UC.4	11,197
9	UC.5	12,437
10	UC.8	28,751

- Results of Preprocessing

Sample Count	10
Read Count	161,768
Gamma-diversity	49

Counts/sample summary

Min	8,549.0
Max	28,751.0
Median	13,251.5
Mean	16,176.8

Filtered Read Count

Ambiguous	0
Wrong prefix or primers	28,562
(Sequence of prefix or primer)	CCTACGGG[ACGT]GGC[AT]GCAG
Low-Quality	2,672
Chimera	33,162
Other	565,275

Figure 4.9 The figure describes the result of clustering (cutoff: 97%), and the result of preprocessing.

On the other hand, the table also shows the variety of microorganisms found in all the samples taken into account. The aforementioned alpha-diversity relates to species diversity in localized places and ecosystems. In contrast, beta-diversity includes species diversity across

sites and habitats. In other words, it refers to diversity within diversity. Gamma diversity, which includes the two diversity types mentioned, stands for diversity throughout the entire terrain (alpha + beta diversity). It describes the total diversity of all the microorganism species found both within and outside of a sample. The value for this task is 49, as demonstrated above.

The table also shows the filtered read count in the manner described below:

Low quality reads had 2,672, chimeric reads had 33,162, and filtered sequences without any particular base cells had zero reads out of the overall count of 565,275.

4.1.2. Taxonomic assignment and diversity statistic by using QIIME

Each OUT's representative sequences were given a taxonomy by the QIIME. Additionally, as shown in table 4.4, phylogenetic trees were constructed using aligned, filtered sample sequences of the OUT and publication-quality graphic results were produced..

Table 4.4 The table describes The QIIME that used for representative sequences from each OUT.

Process		**Program**		**result**
Assembly-MiSeq PE		FLASH		Assembled reads (SE, fastq)
Pre-processing (denoising)		CD-HIT-OUT/ rDnaTools		Removed reads (SE. fastq)
Clustering		CD-HIT-OUT/ rDnaTools		OTU cluster
Diversity analysis	Alpha-diversity	QIIME	Alpha-diversity.py	OTUs, Chao1, Shannon, Simpson.
	Taxonomy		UCLSUT/ RDP(16S) or UNITE(ITS)	Taxonomy Composition
	Alpha-Rarefaction		Alpha-Rarefaction.py	Rarefaction curve graph
	PCoA		Make_2d_plots.py	PCoA graph _ 2D, 3D
	UPGMA Tree		Upgma_cluster.py	UPGMA Tree graph
	Heatmap		Make_otu_heatmap_html.py	OTU Heatmap

The table below 4.5 presents OTUs results using Chao1, Shannon and gini-simpson. In the column of Shannon, since it depicts rareness, UC 8 has the rarest and most abundant indices.

Table 4.5 The table describes the OTUs results using Chao1.

Sample Name	OTUs	Chao1	Shannon	Gini-Simpson	Good's Coverage
C1	16	16	1.0767711	0.412548246	0.999883027
C2	23	23.2	1.0617267	0.396738687	0.999868913
CD.1	19	21	1.0280415	0.389458224	0.999654845
CD.3	17	19	0.5805871	0.206066259	0.999785094
CD.6	24	26.5	1.2031449	0.463996936	0.999511957
CD.7	21	24	0.8397119	0.264745417	0.999886411
UC.2	20	21.5	0.8214982	0.276944653	0.99978672
UC.4	24	24.75	1.0203645	0.36215403	0.999732071
UC.5	20	23	0.896411	0.321433575	0.999758784
UC.8	20	20.5	1.5897642	0.631947908	0.999895656

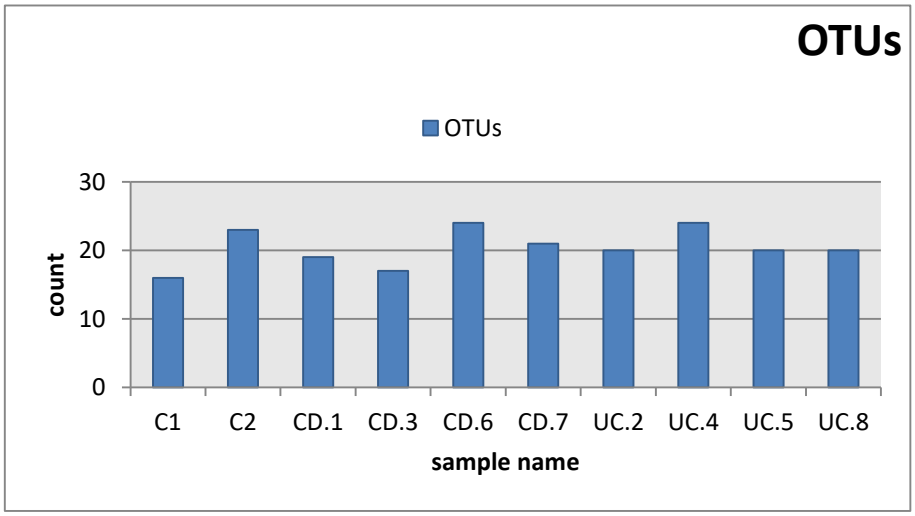
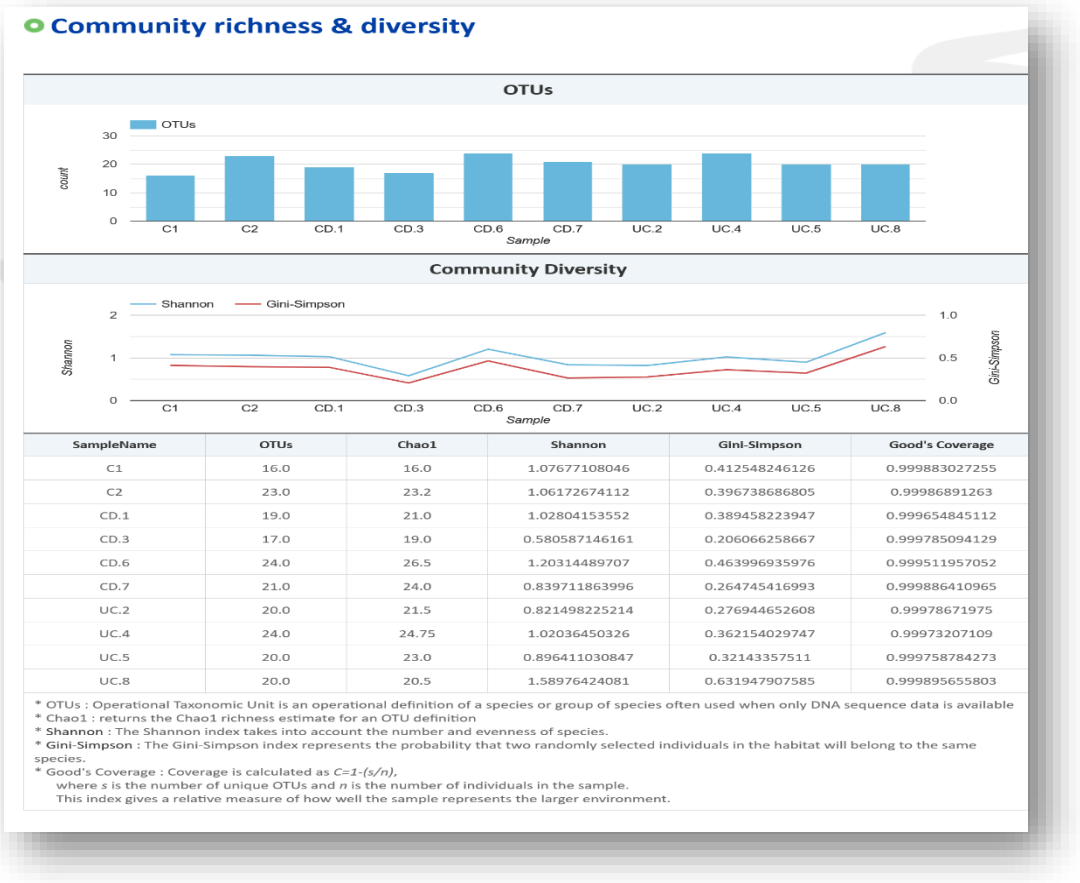


Figure 4.10 The figure describes the diversity index of every type of Ulcerative colitis, Crohn's disease amd the control samples.

Figure 4.10 above displays the diversity index of all the microbial species found in the samples taken during the data analysis. The diversity indexes Shannon and Gini-Simpson were used. Shannon describes relative rarity, while Simpson's index of diversity takes both the overall number of species present and their relative abundance into account.



SampleName	OTUs	Chao1	Shannon	Gini-Simpson	Good's Coverage
C1	16.0	16.0	1.07677108046	0.412548246126	0.999883027255
C2	23.0	23.2	1.06172674112	0.396738686805	0.99986891263
CD.1	19.0	21.0	1.02804153552	0.389458223947	0.999654845112
CD.3	17.0	19.0	0.580587146161	0.206066258667	0.999785094129
CD.6	24.0	26.5	1.20314489707	0.463996935976	0.999511957052
CD.7	21.0	24.0	0.839711863996	0.264745416993	0.999886410965
UC.2	20.0	21.5	0.821498225214	0.276944652608	0.99978671975
UC.4	24.0	24.75	1.02036450326	0.362154029747	0.99973207109
UC.5	20.0	23.0	0.896411030847	0.32143357511	0.999758784273
UC.8	20.0	20.5	1.58976424081	0.631947907585	0.999895655803

* OTUs : Operational Taxonomic Unit is an operational definition of a species or group of species often used when only DNA sequence data is available
* Chao1 : returns the Chao1 richness estimate for an OTU definition
* **Shannon** : The Shannon index takes into account the number and evenness of species.
* **Gini-Simpson** : The Gini-Simpson index represents the probability that two randomly selected individuals in the habitat will belong to the same species.
* Good's Coverage : Coverage is calculated as $C=1-(s/n)$,
where s is the number of unique OTUs and n is the number of individuals in the sample.
This index gives a relative measure of how well the sample represents the larger environment.

Figure 4.11 The figure describes the community richness and diversity of all samples..

The highest Gini-Simpson and Shannon values are found for UC 8 (see Figure 4.11 above), with values of 0.65 and 1.6, respectively. The following CD -6 has a gini-Simpson and Shannon index of 0.48 and 1.2, respectively. The CD-3 Gini-Simpson index (0.25) and Shannon index (0.45) have the two indices with the lowest values. However, error rate may be negatively connected with population heterozygosity, according to (Konopiski, 2020), even while associations and dissociations between study findings' data are true.

The alpha rarefaction graph demonstrates if the study's read count was sufficient to identify the species or OTU. If the curve flattens out to the right, more sequencing is not necessary because a respectable number of reads were used in the study. However, if the graph does not plateau, more OTUs will likely be discovered in the sample as a result of the additional reads. (Read number on x-axis; OTUs on y-axis).

The collectors curve above shows the number of operational taxonomic units (OTUs), given the quantity of sequences gathered and a pairwise sequence identity. More sequencing is unlikely to uncover additional OTUs at that level of percent identity, which are prevalent in the disease, but when a collectors curve reaches a plateau, it will be feasible to observe them in the controls. Below each plot is a table containing the average values for each measure of alpha diversity for each group of samples in the relevant category...

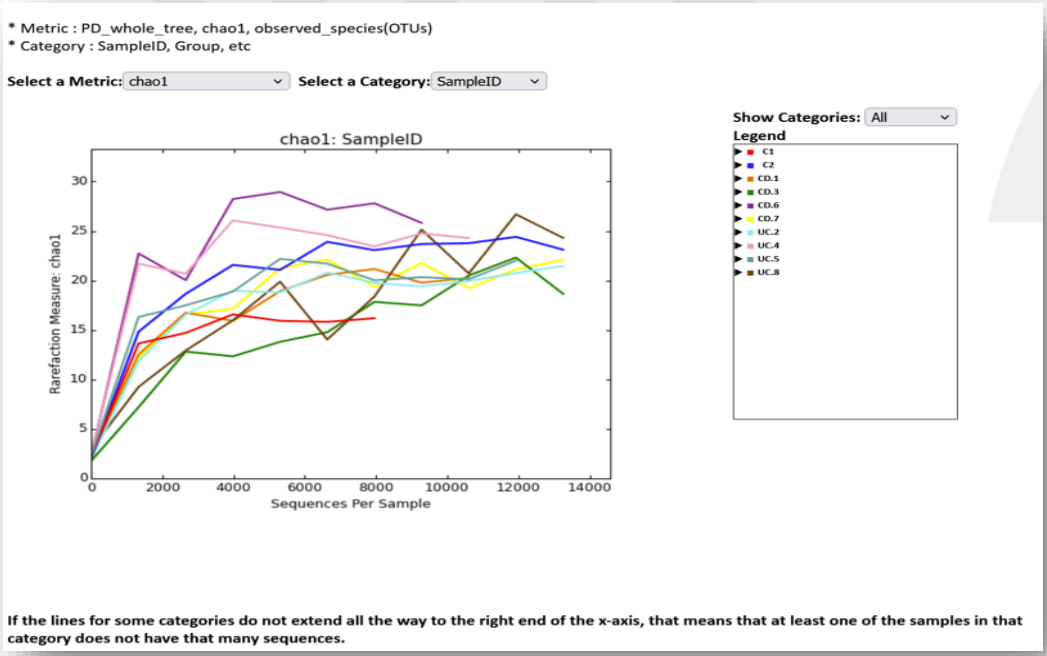


Figure 4.12 The figure describes the rarefaction Curve Graph for each samples.

Figure 4.12 above depicts the measurement of species richness and diversity. Richness and variety were contrasted in this case. The number of species can be determined from the number of samples. Every time, the curves take a plateauing pattern since only the rarest species could be tested later, despite the fact that the pattern initially increases swiftly due to the abundance of the majority of species. The pattern is presented in a color-coded manner.

The species in sample UC 4 had gone above their maximum range. The color-coded pattern that was noticed is unique from other patterns that were discovered in prior examinations. Sometimes the trend may not be as eviden (Nelson, 2011).

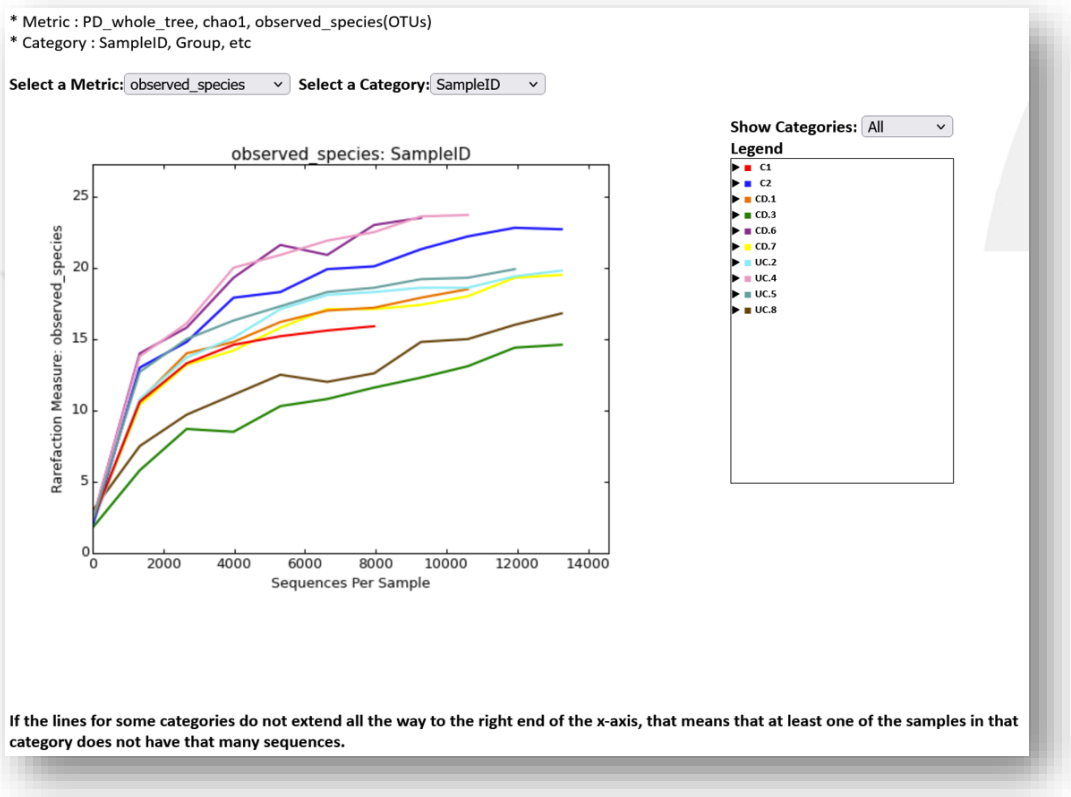


Figure 4.13 The figure describes the Phylogenetic Diversity (PD)-whole tree.

The Figure above 4.13 describes the phylogenetic diversity of all the species in the samples used. The pattern of refraction is colour coded and the graph shows the pattern assumed by each, starting from the zero origin. The pattern in all the cases the pattern grows fast starting from zero, later attain plateau because only the rarest species could be found.

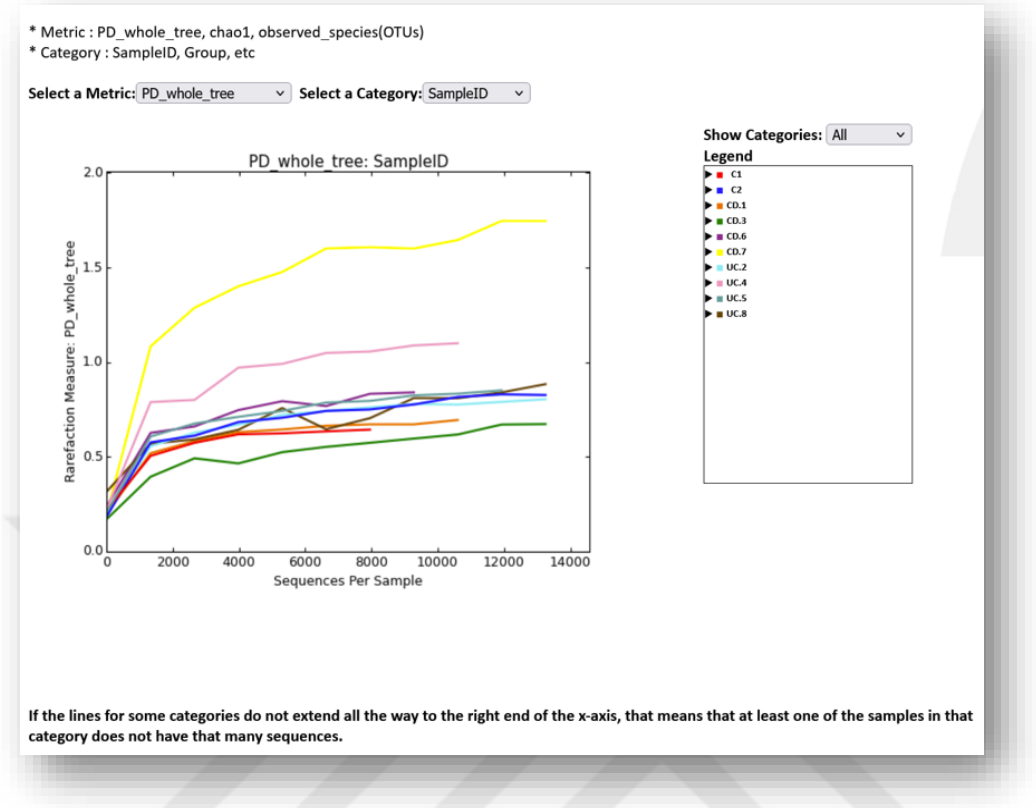


Figure 4.14 The figure describes the Phylogenetic Diversity (PD)-whole tree.

The Figure 4.14 above presence the pattern of phylogenetic diversity. The diversities are presented in different colour codes. The pattern of refraction is colour coded and the graph shows the pattern assumed by each, starting from the zero origin. The pattern in all the cases the pattern grows fast starting from zero, later attain plateau because only the rarest species could be found. From the figure sample UC 4 has the sharpest plateau, which means it contains the rarest species of microbial community. The least definitive plateau is seeing in CD 7, which means the rarest species are least in that sample.

Principle Coordinate Analysis of Samples (unweighted unifrac): This analysis, often known as PCoA, is a technique that helps extract and visualize a few particularly interesting components of variance from complex, multidimensional data. In order for the first principal coordinate, which displays a percent of 55.09% for the PC1, and the second principal coordinate, which displays a percent of 22.26%, to be able to explain the most variation, the samples in the distance matrix are transformed into a new set of orthogonal axes. The data

structure is clearly visible in the two-dimensional principal coordinates, which also enable comparisons between the samples.

The separations between samples are represented in this method by a low-dimensional space. In specifically, the maximum linear correlation between the distances in the distance matrix and the distances in a low-dimensional space is achieved..

The ordinates of objects that are nearer to one another are more similar than those that are farther apart. The sample distance is represented by the spatial distance between sample sites.

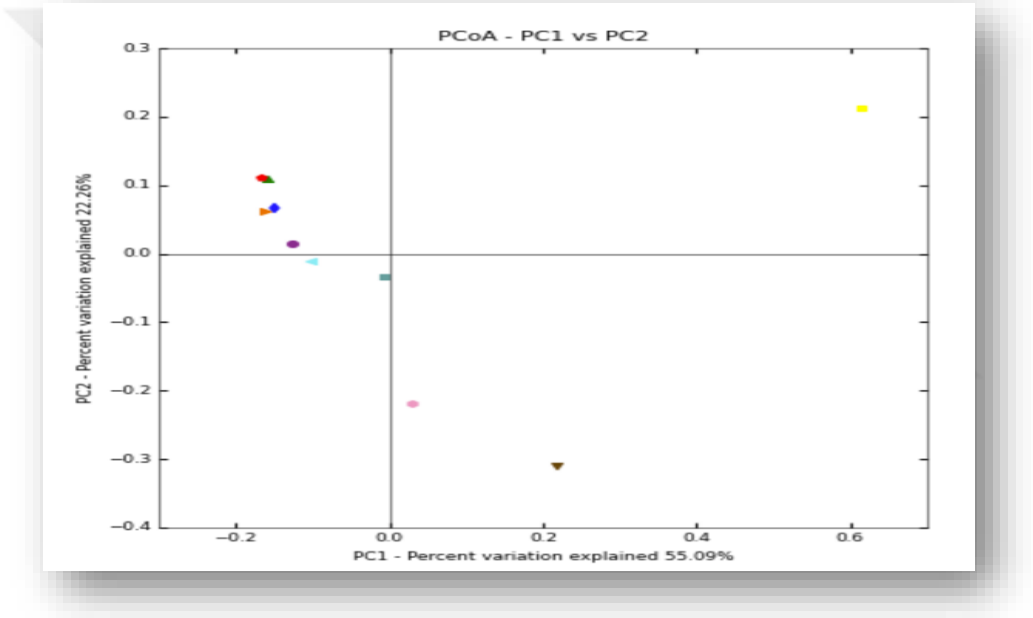


Figure 4.15 The figure describes the principle coordinates with percent variation 22.26% for PC2 and 55.09% for PC1.

In the diagram above, principle coordinate axis 2 is the principle coordinate that accounts for the majority of the remaining data, while principle coordinate axis 1 describes the data exchange. Objects are shown in the above graphic using color codes. Since the red is associated with the rest of the colors more closely than the other colors, their distribution and richness in the biocommunity are likely related.

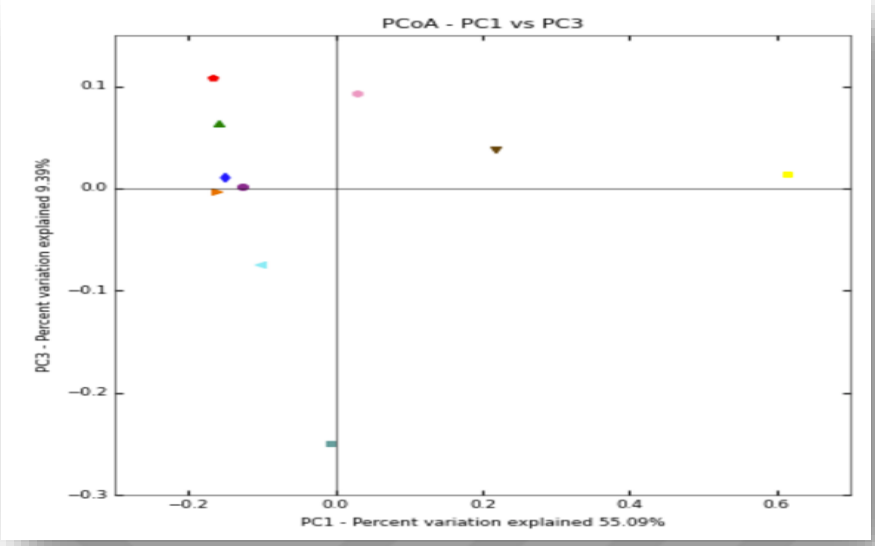


Figure 4.16 The figure describes the principle coordinates with percent variation 9.39% for PC2 and 55.9% for PC1.

According to the aforementioned figure, the second main coordinate, which displays a percent of 9.39% for the PC3 and a percent of 22.26% for the PC2, accounts for the second-largest amount of variation.

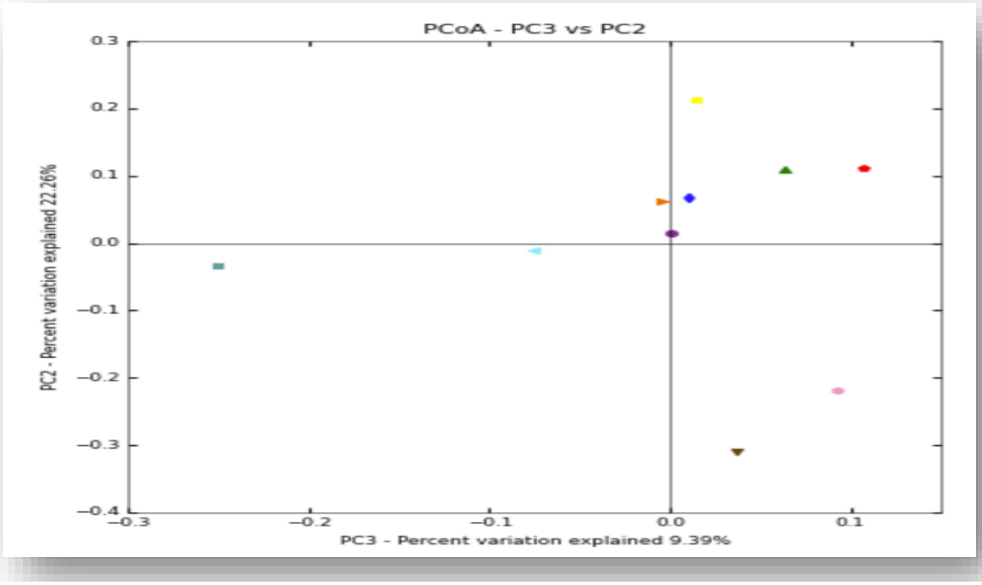


Figure 4.17 The figure describes the principle coordinates with percent variation 22.26% for PC2 and 9.39% FOR PC1.

Additionally, the 3D plot figure is displayed in the following top figures: The third main The third amount of variation is explained by the coordinate, which shows the percent of 55.09% for the PC1 and the percent of 9.39% for the PC3 from the figure.

Samples' principal coordinate analysis (weighted unifrac):

The largest amount of variation is accounted for by the first principal coordinate, which shows a percent of 89.24% for PC1 and a percent of 9.71% for PC2.

.

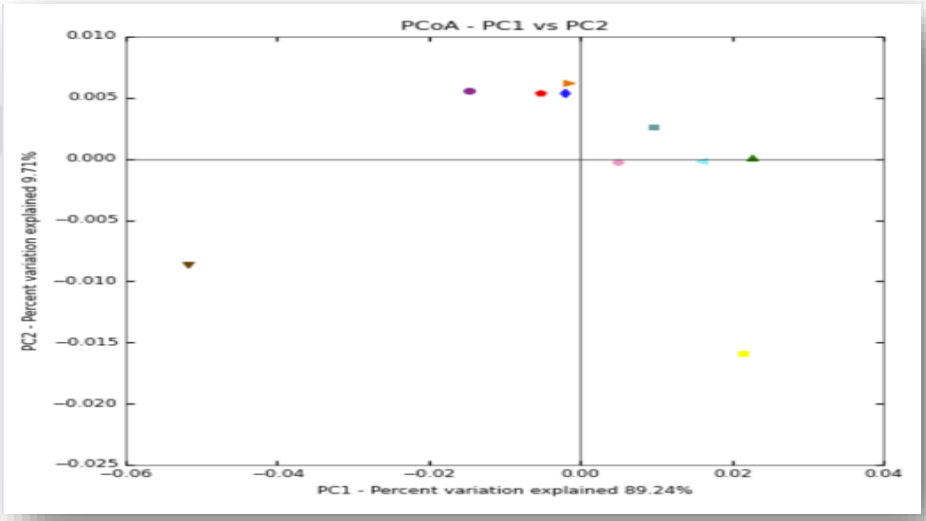


Figure 4.18 The figure describes the first primary correlation's widest possible range.

In the picture above 4.18, the second principal coordinate, which displays a percent of 0.61% for PC3 and a percent of 9.71% for PC2, is responsible for the second amount of variance.

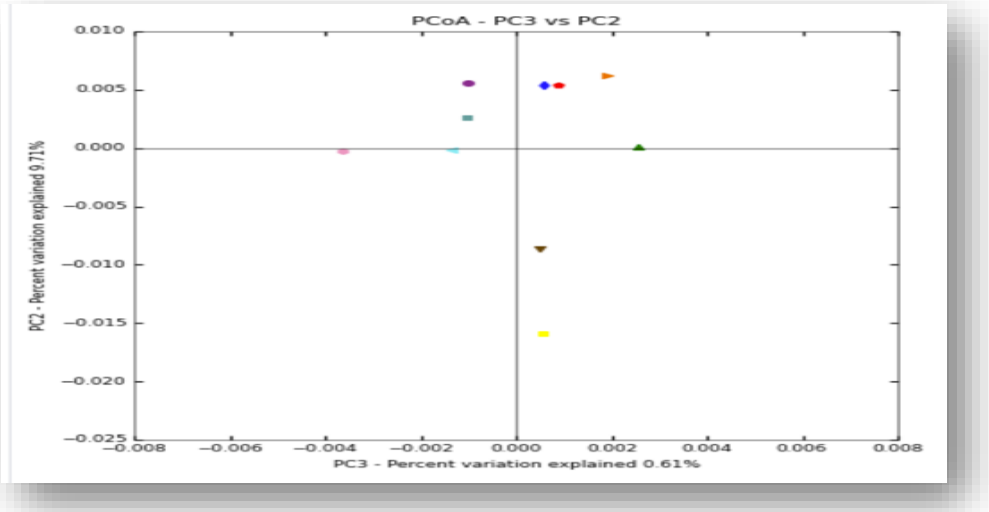


Figure 4.19 The figure describes the second primary coordinate's greatest possible variation.

The second major coordinate, which displays a percent of 0.61% for the PC3 and a percent of 9.71% for the PC2, explains the second amount of variation from the figure above 30.

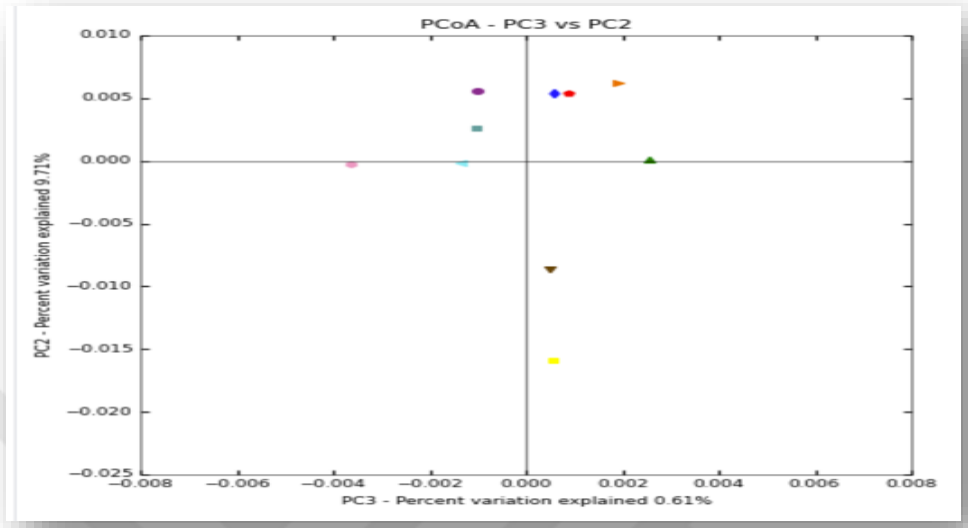


Figure 4.20 The figure describes the third primary coordinate's widest range of variation.

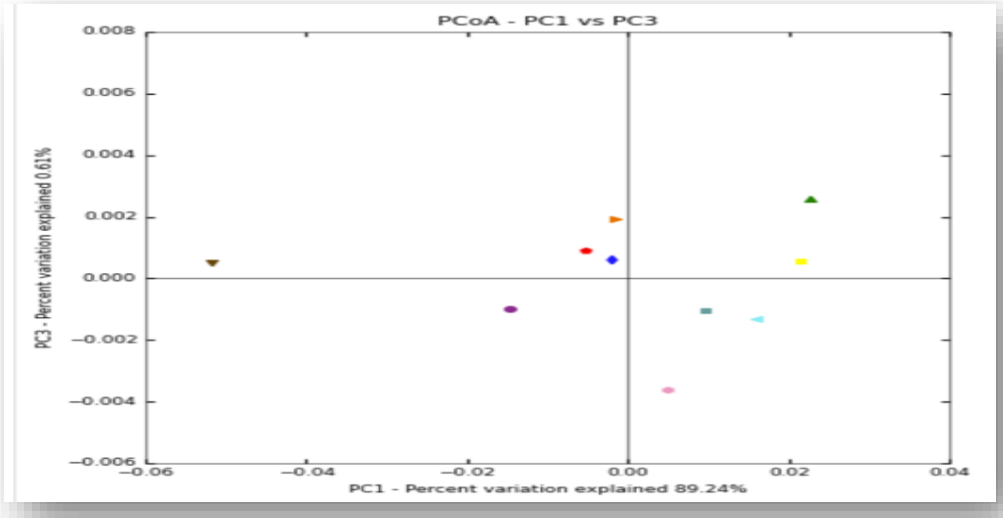


Figure 4.21 The figure describes the fourth primary coordinate's largest possible range of variation.

The third principal coordinate, which displays a percent of 0.61% for the PC3 and a percent of 89.24% for the PC1, provides an explanation for the third amount of variance..

Unweighted Pair Group Method with Arithmetic Mean (UPGMA): This method, often known as UPGMA, produces hierarchical clusters by using average linkage and can be used to interpret the distance matrix. A suitable tool can be used with the distance matrix file, as shown in figure32, to create a phylogenetic tree depending on the intended use

Method for Weighted Pair Groups with Arithmetic Mean: The distance matrix can be interpreted using this method, also known as UPGMA, which uses average linkage in hierarchical clustering (Weighted_unifrac_dm.txt). Figure 4.22 below illustrates and explains how the distance matrix file can be utilized, depending on the goal, to create a phylogenetic tree using an appropriate tool.

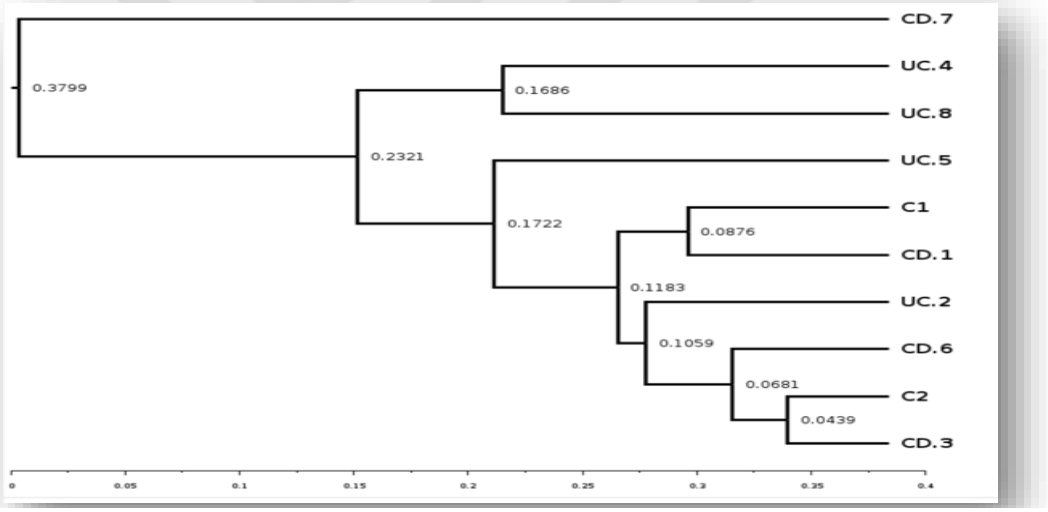


Figure 4.22 The figure describes the UPGMA tree comparing samples (unweighted unifrac).

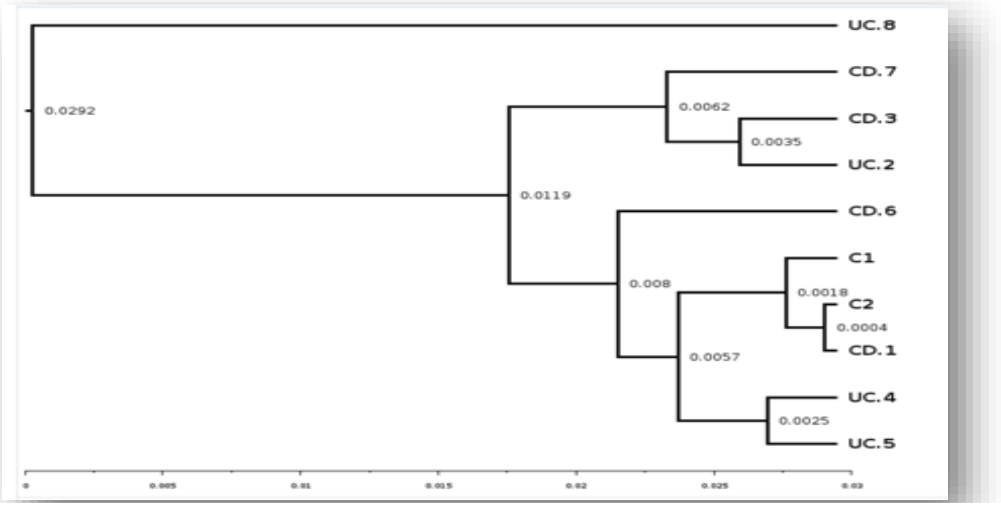


Figure 4.23 The figure discribes the UPGMA tree comparing samples (weighted unifrac).

From the figure above 4.23, phylogenetic tree was built with the origin from sample UC 8 from where the rest samples took their origin. From the figure, UC 4 and UC 5 has a common relationship. Also, CD 1 and C 2 share a common similarity by taking their origin from the same point as C 1.CD 1, C2 and C1 all originate from CD 6. Also UC 2, CD 3, share similarity and originate from the same point as CD 7.

Taxonomic analysis:

In follow up the Fasta analysis the OTU groups defined.

The data which got from sequencing was BLAST by 16s bacterial reference of NCBI.

BLAST results in the figure below 4.24 show the indication of all samples according to their Bar type plus the bacteria percentage found in each of them. Proteobacteria was the most indicated followed by firmicutes in both diseases and the healthy individual group.

Name	Q_Leng	Q_Best	Q_Start	Q_End	Q_Cove	Descrip	Accessi	S_Leng	S_Start	S_End	S_Cove	Bit	E-Value	I_Matc	I_Pct.(	G_Mat	G_Pct.I	Query	Subject	Organis	species	genus	family	order	class	phylum	kingdo	superk
denovo0	465	3	1	465	100	NR_15698	NR_15698	1428	316	780	32.56	848.7	0	463/465	99	0/465	0	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo1	465	3	1	465	100	NR_11412	NR_11412	1463	314	778	31.78	848.7	0	463/465	99	0/465	0	Plus	Plus	Bacteria; F	Ralstonia	Ralstonia	Burkholde	Burkholde	Betaprote	Proteobacteria		Bacteria
denovo2	440	2	1	440	100	NR_13418	NR_13418	1407	304	743	31.27	802.6	0	438/440	99	0/440	0	Plus	Plus	Bacteria; F	Sphingom	Sphingom	Sphingom	Sphingom	Alphapro	Proteobacteria		Bacteria
denovo3	440	5	1	440	100	NR_12563	NR_12563	1490	312	751	29.53	808.1	0	439/440	99	0/440	0	Plus	Plus	Bacteria; F	Bradyrhiz	Bradyrhiz	Bradyrhiz	Hyphomic	Alphapro	Proteobacteria		Bacteria
denovo4	440	1	1	440	100	NR_07426	NR_07426	1480	303	742	29.73	802.6	0	438/440	99	0/440	0	Plus	Plus	Bacteria; F	Agrobacte	Agrobacte	Rhizobiac	Hyphomic	Alphapro	Proteobacteria		Bacteria
denovo5	440	3	1	440	100	NR_11412	NR_11412	1463	314	753	30.08	636.4	0	411/443	92	6/443	1	Plus	Plus	Bacteria; F	Ralstonia	Ralstonia	Burkholde	Burkholde	Betaprote	Proteobacteria		Bacteria
denovo6	440	3	1	440	100	NR_15698	NR_15698	1428	316	755	30.81	636.4	0	411/443	92	6/443	1	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo7	440	3	1	440	100	NR_15698	NR_15698	1428	316	755	30.81	653	0	413/442	93	4/442	0	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo8	440	3	1	440	100	NR_15698	NR_15698	1428	316	755	30.81	741.6	0	428/441	97	2/441	0	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo9	440	5	1	440	100	NR_13572	NR_13572	1483	317	756	29.67	597.6	2E-170	404/443	91	6/443	1	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo10	440	4	1	440	100	NR_11422	NR_11422	1462	308	747	30.1	808.1	0	439/440	99	0/440	0	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo11	440	1	1	440	100	NR_04159	NR_04159	1342	140	579	32.79	763.8	0	431/440	97	0/440	0	Plus	Plus	Bacteria; F	Megamon	Megamon	Selenomo	Selenomo	Negativicu	Firmicutes		Bacteria
denovo12	440	5	1	440	100	NR_07425	NR_07425	1478	304	743	29.77	808.1	0	439/440	99	0/440	0	Plus	Plus	Bacteria; F	Methylor	Methylor	Methylob	Hyphomic	Alphapro	Proteobacteria		Bacteria
denovo13	440	3	1	440	100	NR_11412	NR_11412	1463	314	753	30.08	647.4	0	412/442	93	4/442	0	Plus	Plus	Bacteria; F	Ralstonia	Ralstonia	Burkholde	Burkholde	Betaprote	Proteobacteria		Bacteria
denovo14	440	3	1	440	100	NR_11412	NR_11412	1463	314	753	30.08	730.5	0	426/441	96	2/441	0	Plus	Plus	Bacteria; F	Ralstonia	Ralstonia	Burkholde	Burkholde	Betaprote	Proteobacteria		Bacteria
denovo15	440	3	1	440	100	NR_11412	NR_11412	1463	314	753	30.08	691.8	0	418/440	95	0/440	0	Plus	Plus	Bacteria; F	Ralstonia	Ralstonia	Burkholde	Burkholde	Betaprote	Proteobacteria		Bacteria
denovo16	439	5	1	439	100	NR_10958	NR_10958	1532	330	768	28.66	595.7	6E-170	404/443	91	8/443	1	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo17	439	2	1	439	100	NR_07451	NR_07451	1510	350	788	29.07	800.7	0	437/439	99	0/439	0	Plus	Plus	Bacteria; F	Phocaeico	Phocaeico	Bacteroid	Bacteroid	Bacteroid	Bacteroidetes		Bacteria
denovo18	439	5	1	439	100	NR_04211	NR_04211	1546	345	783	28.4	806.3	0	438/439	99	0/439	0	Plus	Plus	Bacteria; F	Lactobacil	Lactobacil	Lactobacil	Lactobacil	Bacilli	Firmicutes		Bacteria
denovo19	439	5	1	439	100	NR_12563	NR_12563	1490	312	725	27.79	566.2	5E-161	399/439	90	25/439	5	Plus	Plus	Bacteria; F	Bradyrhiz	Bradyrhiz	Bradyrhiz	Hyphomic	Alphapro	Proteobacteria		Bacteria
denovo20	439	3	1	437	99.54	NR_15194	NR_15194	1424	321	757	30.69	725	0	422/437	96	0/437	0	Plus	Plus	Bacteria; A	Cutibacte	Cutibacte	Propionib	Propionib	Actinomy	Actinobacteria		Bacteria
denovo21	439	2	1	439	100	NR_15608	NR_15608	1410	336	774	31.13	712.1	0	421/439	95	0/439	0	Plus	Plus	Bacteria; F	Kineothrix	Kineothrix	Lachnospi	Eubacteria	Clostridia	Firmicutes		Bacteria
denovo22	439	1	1	435	99.09	NR_11534	NR_11534	1489	338	771	29.15	603.1	4E-172	400/436	91	3/436	0	Plus	Plus	Bacteria; F	Acetivibri	Acetivibri	Oscillospir	Eubacteria	Clostridia	Firmicutes		Bacteria
denovo23	439	1	1	439	100	NR_02470	NR_02470	1497	322	760	29.33	806.3	0	438/439	99	0/439	0	Plus	Plus	Bacteria; F	[Pseudomonas] genii	Xanthomo	Xanthomo	Gammapr	Proteobacteria		Bacteria
denovo24	439	2	1	439	100	NR_14469	NR_14469	1478	333	771	29.7	767.5	0	431/439	98	0/439	0	Plus	Plus	Bacteria; F	Clostridiu	Clostridiu	Clostridia	Eubacteria	Clostridia	Firmicutes		Bacteria
denovo25	439	5	1	439	100	NR_11422	NR_11422	1462	308	771	31.74	595.7	6E-170	421/464	90	25/464	5	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo26	439	1	1	437	99.54	NR_11750	NR_11750	1477	311	747	29.59	802.6	0	436/437	99	0/437	0	Plus	Plus	Bacteria; A	Bifidobact	Bifidobact	Bifidobact	Bifidobact	Actinomy	Actinobacteria		Bacteria
denovo27	439	3	1	439	100	NR_11552	NR_11552	1445	283	721	30.38	712.1	0	421/439	95	0/439	0	Plus	Plus	Bacteria; F	Sphingom	Sphingom	Sphingom	Sphingom	Alphapro	Proteobacteria		Bacteria
denovo28	439	3	1	439	100	NR_15698	NR_15698	1428	316	754	30.74	617.9	1E-176	407/442	92	6/442	1	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo29	439	1	1	439	100	NR_13379	NR_13379	1432	312	775	32.4	606.8	3E-173	424/465	91	27/465	5	Plus	Plus	Bacteria; F	Noviherba	Noviherba	Oxalobact	Burkholde	Betaprote	Proteobacteria		Bacteria
denovo30	439	1	1	439	100	NR_13379	NR_13379	1432	312	775	32.4	584.7	1E-166	420/465	90	27/465	5	Plus	Plus	Bacteria; F	Noviherba	Noviherba	Oxalobact	Burkholde	Betaprote	Proteobacteria		Bacteria
denovo31	439	3	1	439	100	NR_15698	NR_15698	1428	316	754	30.74	573.6	3E-163	398/441	90	4/441	0	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo32	439	1	1	434	98.86	NR_11430	NR_11430	1338	206	640	32.51	497.9	2E-140	382/437	87	5/437	1	Plus	Plus	Bacteria; C	Aerosakk	Aerosakk	Oscillator	Oscillatoriales		Cyanobacteria		Bacteria
denovo33	439	5	1	439	100	NR_11411	NR_11411	1416	284	697	29.24	533	5E-151	394/440	89	27/440	6	Plus	Plus	Bacteria; F	Sphingom	Sphingom	Sphingom	Sphingom	Alphapro	Proteobacteria		Bacteria
denovo34	439	3	1	439	100	NR_11412	NR_11412	1463	314	752	30.01	507.1	3E-143	384/439	87	0/439	0	Plus	Plus	Bacteria; F	Ralstonia	Ralstonia	Burkholde	Burkholde	Betaprote	Proteobacteria		Bacteria
denovo35	439	1	1	439	100	NR_11814	NR_11814	1362	314	752	32.23	634.5	0	410/442	92	6/442	1	Plus	Plus	Bacteria; F	Haemoph	Haemoph	Pasteurell	Pasteurell	Gammapr	Proteobacteria		Bacteria
denovo36	439	5	1	439	100	NR_13418	NR_13418	1407	304	717	29.42	549.6	5E-156	397/440	90	27/440	6	Plus	Plus	Bacteria; F	Sphingom	Sphingom	Sphingom	Sphingom	Alphapro	Proteobacteria		Bacteria
denovo37	439	4	1	439	100	NR_15698	NR_15698	1428	316	754	30.74	573.6	3E-163	397/440	90	2/440	0	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo38	439	1	1	438	99.77	NR_02895	NR_02895	1470	296	733	29.8	804.4	0	437/438	99	0/438	0	Plus	Plus	Bacteria; A	Corynebac	Corynebac	Corynebac	Corynebac	Actinomy	Actinobacteria		Bacteria
denovo39	439	3	1	439	100	NR_15698	NR_15698	1428	316	754	30.74	695.5	0	418/439	95	0/439	0	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo40	439	1	1	439	100	NR_02498	NR_02498	1371	340	777	31.95	557	3E-158	397/443	89	9/443	2	Plus	Plus	Bacteria; F	Smithella	Smithella	Syntroph	Syntroph	Deltapro	Proteobacteria		Bacteria
denovo41	439	5	1	439	100	NR_10958	NR_10958	1532	330	768	28.66	540.3	3E-153	395/444	88	10/444	2	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo42	439	2	1	439	100	NR_11869	NR_11869	1423	338	776	30.85	795.2	0	436/439	99	0/439	0	Plus	Plus	Bacteria; F	Ruminoc	Mediterra	Lachnospi	Eubacteria	Clostridia	Firmicutes		Bacteria
denovo43	439	3	1	439	100	NR_15698	NR_15698	1428	316	754	30.74	656.7	0	412/440	93	2/440	0	Plus	Plus	Bacteria; F	Pseudom	Pseudom	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo44	439	5	1	439	100	NR_13476	NR_13476	1405	280	718	31.25	806.3	0	438/439	99	0/439	0	Plus	Plus	Bacteria; F	Methylop	Methylop	Methyloc	Hyphomic	Alphapro	Proteobacteria		Bacteria
denovo45	439	5	1	439	100	NR_14863	NR_14863	1525	330	768	28.79	590.2	3E-168	404/444	90	10/444	2	Plus	Plus	Bacteria; F	Glaciimon	Glaciimon	Oxalobact	Burkholde	Betaprote	Proteobacteria		Bacteria
denovo46	439	1	1	439	100	NR_13379	NR_13379	1432	312	775	32.4	562.5	6E-160	415/464	89	25/464	5	Plus	Plus	Bacteria; F	Noviherba	Noviherba	Oxalobact	Burkholde	Betaprote	Proteobacteria		Bacteria
denovo47	439	1	1	439	100	NR_04103	NR_04103	1481	319	757	29.64	562.5	6E-160	398/443	89	8/443	1	Plus	Plus	Bacteria; F	Azotobact	Azotobact	Pseudom	Pseudom	Gammapr	Proteobacteria		Bacteria
denovo48	439	2	1	438	99.77	NR_11336	NR_11336	1485	328	765	29.49	793.3	0	435/438	99	0/438	0	Plus	Plus	Bacteria; A	Cutibacte	Cutibacte	Propionib	Propionib	Actinomy	Actinobacteria		Bacteria

Figure 4.24 The figure describes the BLAST results for all samples.

Taxonomy abundance – Count

The table below 4.25 lists the phylum, order, family, genus, and species taxonomy of each microbial community identified in the sampled materials. The results show that all known microorganisms are bacteria, but that they vary in phylum, order, family, genus, and species. The problem in this case is that samples C 1 and C 2 have a bacterial colony in them. It might have another explanation because it wasn't expected. The possibility exists that the used controls weren't sufficiently screened...

In healthy people, those with inflammatory bowel disease, and those who have colon cancer, firmicutes, bacteroidetes, proteobacteria, and actinobateria—of which firmecutes and bacteroidetes are the most abundant—are the leading phyla of bacteria.

Furthermore, it can be inferred that pseudomonadellae are prevalent at the order level, gammaproteobacteria are prevalent at the class level, and proteobacteria are dominant at the phylum level. On the other hand, Ma, Y, et al. (2021) found this in their investigation.

Kingdo	Phylum	Class	Order	Family	Genus	Species	C1	C2	CD.1	CD.3	CD.6	CD.7	UC.2	UC.4	UC.5	UC.8
Bacteria	_Actinobacteria	_Actinomycetia	_Bifidobacteriales	_Bifidobacteriaceae	_Bifidobacterium	_Bifidobacterium longum	0	0	0	0	0	10	0	0	0	0
Bacteria	_Actinobacteria	_Actinomycetia	_Corynebacteriales	_Corynebacteriaceae	_Corynebacterium	_Corynebacterium appendicis	0	0	0	0	0	0	0	1	0	2
Bacteria	_Actinobacteria	_Actinomycetia	_Propionibacteriales	_Propionibacteriaceae	_Cutibacterium	_Cutibacterium granulosum	0	0	0	0	0	0	0	0	0	3
Bacteria	_Actinobacteria	_Actinomycetia	_Propionibacteriales	_Propionibacteriaceae	_Cutibacterium	_Cutibacterium namnetense	0	0	0	0	0	0	0	4	0	0
Bacteria	_Bacteroidetes	_Bacteroidia	_Bacteroidales	_Bacteroidaceae	_Phocaeicola	_Phocaeicola vulgatus	0	0	0	0	0	47	0	0	0	0
Bacteria	_Cyanobacteria	__	_Oscillatoriales	_Oscillatoriaceae	_Aerosakkonema	_Aerosakkonema funiforme	0	0	0	0	0	3	0	0	0	0
Bacteria	_Firmicutes	_Bacilli	_Lactobacillales	_Lactobacillaceae	_Lactobacillus	_Lactobacillus gallinarum	0	0	0	0	0	13	0	0	0	0
Bacteria	_Firmicutes	_Clostridia	_Eubacteriales	_Clostridiaceae	_Clostridium	_Clostridium saudiense	0	0	0	0	0	10	0	0	0	0
Bacteria	_Firmicutes	_Clostridia	_Eubacteriales	_Lachnospiraceae	_Kineothrix	_Kineothrix alysoides	0	0	0	0	0	4	0	0	0	0
Bacteria	_Firmicutes	_Clostridia	_Eubacteriales	_Lachnospiraceae	_Mediterraneibacter	_[Ruminococcus] gnavus	0	0	0	0	0	3	0	0	0	0
Bacteria	_Firmicutes	_Clostridia	_Eubacteriales	_Oscillospiraceae	_Acetivibrio	_Acetivibrio alkalicellulosi	0	0	0	0	0	7	0	0	0	0
Bacteria	_Firmicutes	_Negativicutes	_Selenomonadales	_Selenomonadaceae	_Megamonas	_Megamonas funiformis	0	0	0	0	0	100	0	0	0	0
Bacteria	_Proteobacteria	_Alphaproteobacteria	_Hyphomicrobiales	_Bradyrhizobiaceae	_Bradyrhizobium	_Bradyrhizobium valentinum	36	40	28	31	37	48	4	5	7	208
Bacteria	_Proteobacteria	_Alphaproteobacteria	_Hyphomicrobiales	_Methylobacteriaceae	_Methylorubrum	_Methylorubrum populi	0	0	0	0	0	8	4	1	3	20
Bacteria	_Proteobacteria	_Alphaproteobacteria	_Hyphomicrobiales	_Methyloligellaceae	_Methyloligella	_Methyloligella hexanensis	0	0	0	0	0	1	0	0	0	2
Bacteria	_Proteobacteria	_Alphaproteobacteria	_Hyphomicrobiales	_Rhizobiaceae	_Agrobacterium	_Agrobacterium fabrum	0	1	0	3	1	34	28	23	16	113
Bacteria	_Proteobacteria	_Alphaproteobacteria	_Sphingomonadales	_Sphingomonadaceae	_Sphingomonas	_Sphingomonas adhaesiva	0	2	0	0	0	0	0	0	0	2
Bacteria	_Proteobacteria	_Alphaproteobacteria	_Sphingomonadales	_Sphingomonadaceae	_Sphingomonas	_Sphingomonas kyeonggiensis	112	186	105	139	188	1,238	225	282	145	4,653
Bacteria	_Proteobacteria	_Alphaproteobacteria	_Sphingomonadales	_Sphingomonadaceae	_Sphingomonas	_Sphingomonas pituitosa	3	4	1	1	3	0	1	2	0	0
Bacteria	_Proteobacteria	_Betaproteobacteria	_Burkholderiales	_Burkholderiaceae	_Ralstonia	_Ralstonia pickettii	2,111	3,522	2,672	2,487	3,012	2,325	1,980	2,084	2,134	10,947
Bacteria	_Proteobacteria	_Betaproteobacteria	_Burkholderiales	_Oxalobacteraceae	_Glaciimonas	_Glaciimonas frigoris	0	2	0	0	1	0	0	0	0	0
Bacteria	_Proteobacteria	_Betaproteobacteria	_Burkholderiales	_Oxalobacteraceae	_Noviherbaspirillum	_Noviherbaspirillum suwonense	0	4	0	1	1	0	3	0	0	0
Bacteria	_Proteobacteria	_Deltaproteobacteria	_Syntrophobacterales	_Syntrophaceae	_Smithella	_Smithella propionica	0	0	0	0	0	3	0	0	0	0
Bacteria	_Proteobacteria	_Gammaproteobacteria	_Pasteurellales	_Pasteurellaceae	_Haemophilus	_Haemophilus parainfluenzae	0	0	0	0	0	0	0	0	3	0
Bacteria	_Proteobacteria	_Gammaproteobacteria	_Pseudomonadales	_Pseudomonadaceae	_Azotobacter	_Azotobacter armeniacus	0	0	0	0	2	0	0	0	0	1
Bacteria	_Proteobacteria	_Gammaproteobacteria	_Pseudomonadales	_Pseudomonadaceae	_Pseudomonas	_Pseudomonas guguanensis	2	11	7	5	6	0	10	7	7	0
Bacteria	_Proteobacteria	_Gammaproteobacteria	_Pseudomonadales	_Pseudomonadaceae	_Pseudomonas	_Pseudomonas monteilii	7	8	5	1	2	62	7	6	19	25
Bacteria	_Proteobacteria	_Gammaproteobacteria	_Pseudomonadales	_Pseudomonadaceae	_Pseudomonas	_Pseudomonas paralactis	6,274	11,463	8,768	20,598	6,980	22,493	11,896	8,776	10,098	12,769
Bacteria	_Proteobacteria	_Gammaproteobacteria	_Pseudomonadales	_Pseudomonadaceae	_Pseudomonas	_Pseudomonas punonensis	0	4	2	0	6	1	2	3	1	3
Bacteria	_Proteobacteria	_Gammaproteobacteria	_Pseudomonadales	_Pseudomonadaceae	_Pseudomonas	_Pseudomonas tolaasii	4	10	1	0	6	0	6	3	3	1
Bacteria	_Proteobacteria	_Gammaproteobacteria	_Xanthomonadales	_Xanthomonadaceae	__	_[Pseudomonas] geniculata	0	0	0	0	0	1	0	0	1	2

Figure 4.25 The figure describes the Taxonomy abundance – Count For each sample

Taxonomic assignment - Bar type

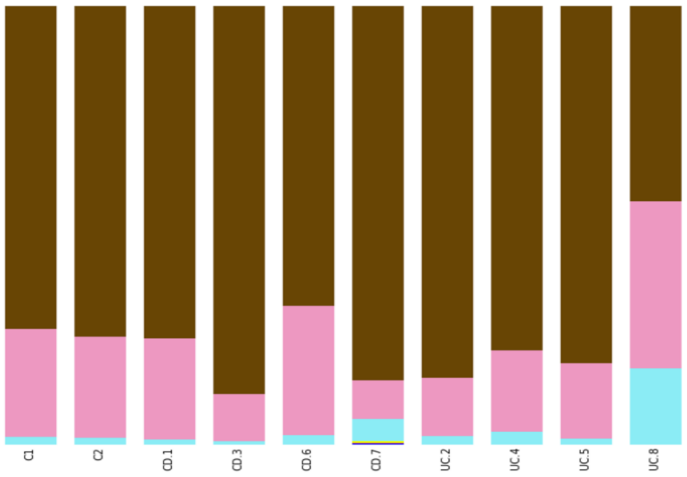
This thesis included Crohn's disease, the gut microbiota of patients with ulcerative colitis, and the control group in an extensive summary of inflammatory bowel disease. The following graphs (x-axis: sample name; y-axis: OTU proportions) show the taxonomic breakdown of each sample, from the Phylum to Genus levels. However, the data for the CD, UC, and control groups were shown in Figure 4.26 at the phylum level. At the phylum level, all of the bacteria in the CD, UC, and control groups belonged to the Proteobacteria class. 0% is used to indicate that the other bacterial categories are not significant. However, there have been changes in the proportions of Bacteroidetes (0.2%), Firmicutes (0.5%), and Proteobacteria (99.3%), as observed in CD7..

View Table (.txt)

Legend	Taxonomy	Total %	C1 %	C2 %	CD.1 %	CD.3 %	CD.6 %	CD.7 %	UC.2 %	UC.4 %	UC.5 %	UC.8 %
	Bacteria;__Actinobacteria	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Bacteroidetes	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.5%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria	99.9%	100.0%	100.0%	100.0%	100.0%	100.0%	99.3%	100.0%	100.0%	100.0%	100.0%

Figure 4.26 The figure describes the taxonomic composition for Phylum levels - Bar type

The following picture (with sample name on the x-axis and OTU proportions on the y-axis) displays the taxonomic composition of each sample from the Phylum to Genus levels.

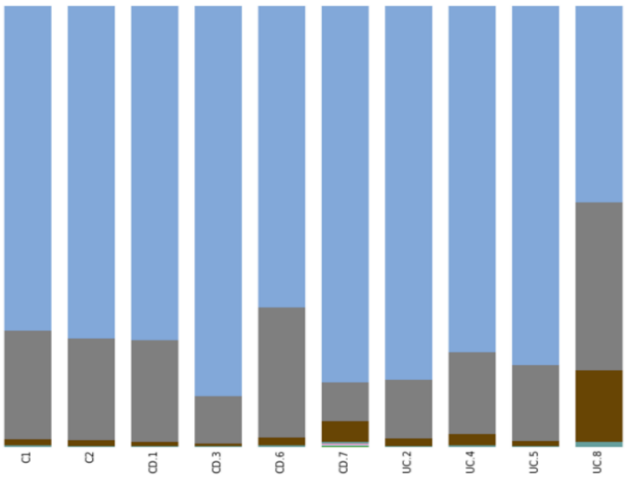


View Table (.txt)

Legend	Taxonomy	Total	C1	C2	CD.1	CD.3	CD.6	CD.7	UC.2	UC.4	UC.5	UC.8
		%	%	%	%	%	%	%	%	%	%	%
	Bacteria;__Actinobacteria;__Actinomycetia	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Bacteroidetes;__Bacteroidia	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria;__	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Bacilli	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Negativicutes	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria	3.6%	1.8%	1.5%	1.2%	0.7%	2.2%	5.0%	1.9%	2.8%	1.4%	17.4%
	Bacteria;__Proteobacteria;__Betaproteobacteria	20.7%	24.7%	23.1%	23.1%	10.7%	29.4%	8.8%	13.4%	18.6%	17.2%	38.1%
	Bacteria;__Proteobacteria;__Deltaproteobacteria	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria	75.6%	73.5%	75.3%	75.8%	88.6%	68.3%	85.4%	84.8%	78.5%	81.5%	44.5%

Figure 4.27 The figure describes the taxonomic composition for Class levels - Bar type.

The class levels that significantly differed between the study groups are shown in Figure 4.27 above. Additionally, CD3 revealed higher levels of Gammaproteobacteria (88.6%) and lowest levels in UC8 (44.5%), but it indicated (75.6%) in the aggregate of all groups. Additionally, an increase in UC8 (38.1%), the greatest level, and CD7 (8.8%), the lowest level, for a total of (20.7%), was seen in betaproteobacteria. The levels of the alpha-proteobacteria also revealed other notable variations, with the highest levels in the UC8 (17.4%) and the lowest levels in the CD3 (0.7%) and overall levels (3.6%). Furthermore, compared to the other groups, CD7 was only indicated in Bacteroidia (0.2%), Clostridia (0.1%), and Negativicutes (0.4%).

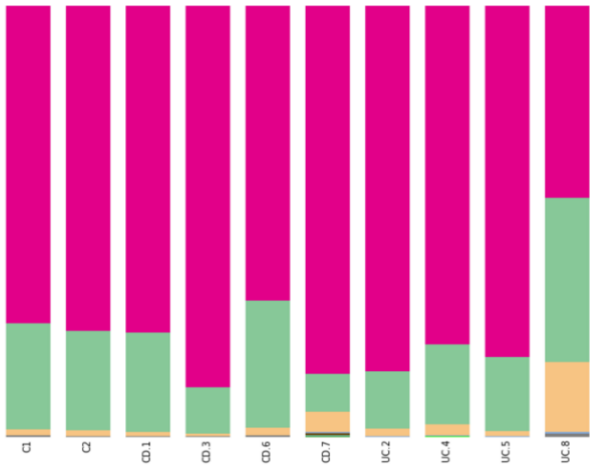


View Table (.txt)

Legend	Taxonomy	Total %	C1 %	C2 %	CD.1 %	CD.3 %	CD.6 %	CD.7 %	UC.2 %	UC.4 %	UC.5 %	UC.8 %
	Bacteria;__Actinobacteria;__Actinomycetia;__Bifidobacteriales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Corynebacteriales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Propionibacteriales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Bacteroidetes;__Bacteroidia;__Bacteroidales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria;__;__Oscillatoriales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Bacilli;__Lactobacillales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Negativicutes;__Selenomonadales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales	0.4%	0.4%	0.3%	0.2%	0.1%	0.4%	0.3%	0.3%	0.3%	0.2%	1.2%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales	3.2%	1.3%	1.3%	0.9%	0.6%	1.9%	4.7%	1.6%	2.5%	1.2%	16.2%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales	20.7%	24.7%	23.1%	23.1%	10.7%	29.4%	8.8%	13.4%	18.6%	17.2%	38.1%
	Bacteria;__Proteobacteria;__Deltaproteobacteria;__Syntrophobacterales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pasteurellales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales	75.6%	73.5%	75.3%	75.8%	88.6%	68.3%	85.4%	84.8%	78.5%	81.4%	44.5%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Xanthomonadales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 4.28 The figure describes the taxonomic composition for Order levels - Bar type.

With the exception of the UC8 group, Figure 4.28 of the order levels above demonstrates a substantial difference between the study groups. Gammaproteobacteria (Pseudomonadales) (44.5%), Betaproteobacteria (Burkholderiales) (38.1%), Alphaproteobacteria (Sphingomonadales) (16.2%), and Alphaproteobacteria (Hyphomicrobiales) (1.2%) are the groupings with the most variance among the other groups, but not the highest of all bacteria. Gammaproteobacteria (Pseudomonadales) increased most in CD3 (88.6%), and least in UC8 (44.5%) and overall (75.6%).

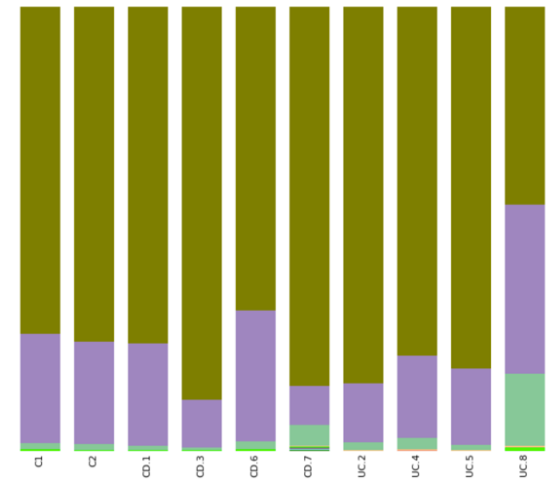


View Table (.txt)

Legend	Taxonomy	Total %	C1 %	C2 %	CD.1 %	CD.3 %	CD.6 %	CD.7 %	UC.2 %	UC.4 %	UC.5 %	UC.8 %
	Bacteria;__Actinobacteria;__Actinomycetia;__Bifidobacteriales;__Bifidobacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Corynebacteriales;__Corynebacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Propionibacteriales;__Propionibacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Bacteroidetes;__Bacteroidia;__Bacteroidales;__Bacteroidaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria;__;__Oscillatoriales;__Oscillatoriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Bacilli;__Lactobacillales;__Lactobacillaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Clostridiaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Lachnospiraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Oscillospiraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Negativicutes;__Selenomonadales;__Selenomonadaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Bradyrhizobiaceae	0.2%	0.4%	0.3%	0.2%	0.1%	0.4%	0.2%	0.0%	0.0%	0.1%	0.7%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylobacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylocystaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Rhizobiaceae	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.2%	0.1%	0.4%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales;__Sphingomonadaceae	3.2%	1.3%	1.3%	0.9%	0.6%	1.9%	4.7%	1.6%	2.5%	1.2%	16.2%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Burkholderiaceae	20.7%	24.7%	23.1%	23.1%	10.7%	29.4%	8.8%	13.4%	18.6%	17.2%	38.1%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Oxalobacteraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Deltaproteobacteria;__Syntrophobacterales;__Syntrophaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pasteurellales;__Pasteurellaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae	75.6%	73.5%	75.3%	75.8%	88.6%	68.3%	85.4%	84.8%	78.5%	81.4%	44.5%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Xanthomonadales;__Xanthomonadaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 4.29 The figure describes the taxonomic composition for family levels - Bar type.

Figure 4.29 reveals a substantial difference in family levels between the study groups, while group UC8 showed more variances than the other groups. UC8 shows a highest level of betaproteobacteria-Burkholderiales (Burkholderiaceae) (38.1%), Alphaproteobacteria-Sphingomonadales (Shpingomonadaceae) (16.2%), Alphaproteobacteria-Hyphomicrobiales (Rhizobiaceae) (0.4%), and Alphaproteobacteria-Hyphomicrobiales (Bradyrhizobiaceae) (0.7%), but the lowest in Gammaproteobacteria-Pseudomonadales (Pseudomonadaceae) (44.5%). Additionally, the levels of order suggest that the total Gammaproteobacteria (75.6%) in all groupings.

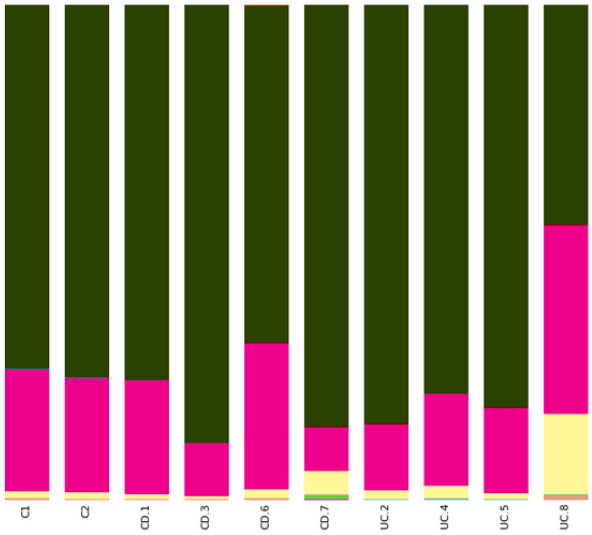


View Table (.txt)

Legend	Taxonomy	Total	C1	C2	CD.1	CD.3	CD.6	CD.7	UC.2	UC.4	UC.5	UC.8
		%	%	%	%	%	%	%	%	%	%	%
	Bacteria;__Actinobacteria;__Actinomycetia;__Bifidobacteriales;__Bifidobacteriaceae;__Bifidobacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Corynebacteriales;__Corynebacteriaceae;__Corynebacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Propionibacteriales;__Propionibacteriaceae;__Cutibacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Bacteroidetes;__Bacteroidia;__Bacteroidales;__Bacteroidaceae;__Phocaeicola	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria;__;__Oscillatoriales;__Oscillatoriaceae;__Aerosakkonema	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Bacilli;__Lactobacillales;__Lactobacillaceae;__Lactobacillus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Clostridiaceae;__Clostridium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Lachnospiraceae;__Kineothrix	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Lachnospiraceae;__Mediterraneibacter	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Oscillospiraceae;__Acetivibrio	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Negativicutes;__Selenomonadales;__Selenomonadaceae;__Megamonas	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Bradyrhizobiaceae;__Bradyrhizobium	0.2%	0.4%	0.3%	0.2%	0.1%	0.4%	0.2%	0.0%	0.0%	0.1%	0.7%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylobacteriaceae;__Methylorubrum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylocystaceae;__Methylopila	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Rhizobiaceae;__Agrobacterium	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.2%	0.1%	0.4%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales;__Sphingomonadaceae;__Sphingomonas	3.2%	1.3%	1.3%	0.9%	0.6%	1.9%	4.7%	1.6%	2.6%	1.2%	16.2%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Burkholderiaceae;__Ralstonia	20.7%	24.7%	23.1%	23.1%	10.7%	29.4%	8.8%	13.4%	18.6%	17.2%	38.1%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Oxalobacteraceae;__Glaciimonas	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Oxalobacteraceae;__Noviherbaspirillum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Deltaproteobacteria;__Syntrophobacterales;__Syntrophaceae;__Smithella	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pasteurellales;__Pasteurellaceae;__Haemophilus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Azotobacter	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas	75.6%	73.5%	75.3%	75.8%	88.6%	68.3%	85.4%	84.8%	78.5%	81.4%	44.5%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Xanthomonadales;__Xanthomonadaceae;__	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 4.30 The figure describes the taxonomic composition for genus levels - Bar type.

Figure 4.30 demonstrates a substantial difference in UC8 between the study groups based on genus levels. Betaproteobacteria (20.7%), Alphaproteobacteria (3.2%), and Gammaproteobacteria (75.6%) made up the majority of the population. Additionally, only Nagativicutes (Selenomonadales) Bacteroidia (Bacteroidales) (0.2%) showed CD7 indication, compared to the other categories.



View Table (.txt)

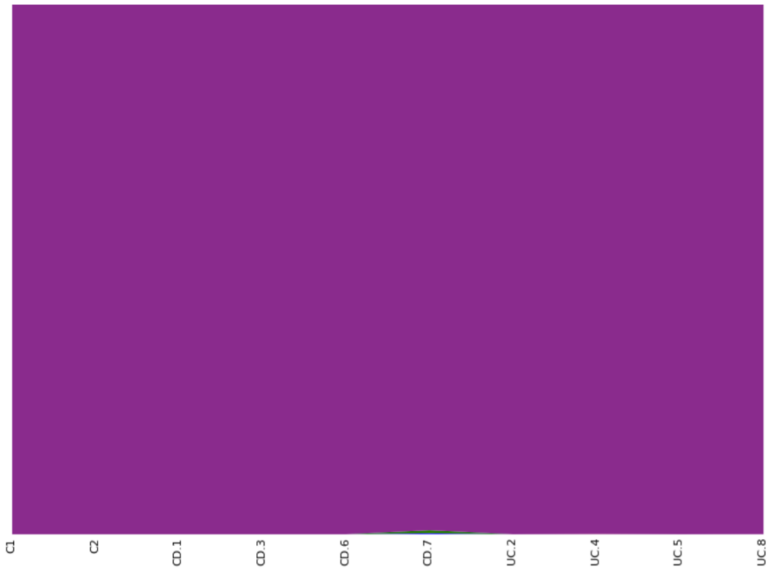
Legend	Taxonomy	Total %	C1 %	C2 %	CD.1 %	CD.3 %	CD.6 %	CD.7 %	UC.2 %	UC.4 %	UC.5 %	UC.8 %
	Bacteria;__Actinobacteria;__Actinomycetia;__Bifidobacteriales;__Bifidobacteriaceae;__Bifidobacterium; __Bifidobacterium longum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Corynebacteriales;__Corynebacteriaceae;__Corynebacterium; __Corynebacterium appendicis	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Propionibacteriales;__Propionibacteriaceae;__Cutibacterium; __Cutibacterium granulosum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Propionibacteriales;__Propionibacteriaceae;__Cutibacterium; __Cutibacterium namnetense	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Bacteroidetes;__Bacteroidia;__Bacteroidales;__Bacteroidaceae;__Phocaeicola; __Phocaeicola vulgatus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria;__;__Oscillatoriales;__Oscillatoriaceae;__Aerosakkonema; __Aerosakkonema funiforme	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Bacilli;__Lactobacillales;__Lactobacillaceae;__Lactobacillus; __Lactobacillus gallinarum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Clostridiaceae;__Clostridium; __Clostridium saudiense	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Lachnospiraceae;__Kineothrix; __Kineothrix alysoides	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Lachnospiraceae;__Mediterraneibacter; __[Ruminococcus] gnavus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Oscillospiraceae;__Acetivibrio; __Acetivibrio alkalicellulosi	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Negativicutes;__Selenomonadales;__Selenomonadaceae;__Megamonas; __Megamonas funiformis	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Bradyrhizobiaceae;__Bradyrhizobium; __Bradyrhizobium valentinum	0.2%	0.4%	0.3%	0.2%	0.1%	0.4%	0.2%	0.0%	0.0%	0.1%	0.7%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylobacteriaceae;__Methylorubrum; __Methylorubrum populi	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylocystaceae;__Methylopila; __Methylopila henanensis	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Rhizobiaceae;__Agrobacterium; __Agrobacterium fabrum	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.2%	0.1%	0.4%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales;__Sphingomonadaceae;__Sphingomonas; __Sphingomonas adhaesiva	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales;__Sphingomonadaceae;__Sphingomonas; __Sphingomonas kyeonggiensis	3.2%	1.3%	1.2%	0.9%	0.6%	1.8%	4.7%	1.6%	2.5%	1.2%	16.2%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales;__Sphingomonadaceae;__Sphingomonas; __Sphingomonas pituitosa	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Burkholderiaceae;__Ralstonia; __Ralstonia pickettii	20.7%	24.7%	23.1%	23.1%	10.7%	29.4%	8.8%	13.4%	18.6%	17.2%	58.1%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Oxalobacteraceae;__Glaciimonas; __Glaciimonas frigoris	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Oxalobacteraceae;__Noviherbaspirillum; __Noviherbaspirillum suwonense	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Deltaproteobacteria;__Syntrophobacterales;__Syntrophaceae;__Smithella; __Smithella propionica	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pasteurellales;__Pasteurellaceae;__Haemophilus; __Haemophilus paraphrohaemolyticus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Azotobacter; __Azotobacter armeniacus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas; __Pseudomonas guguanensis	0.0%	0.0%	0.1%	0.1%	0.0%	0.1%	0.0%	0.1%	0.1%	0.1%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas; __Pseudomonas monteilii	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.2%	0.0%	0.1%	0.2%	0.1%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas; __Pseudomonas parafulva	75.5%	73.4%	75.1%	75.7%	88.5%	68.1%	86.2%	84.6%	78.4%	81.2%	44.4%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas; __Pseudomonas punonensis	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas; __Pseudomonas tolaasii	0.0%	0.0%	0.1%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Xanthomonadales;__Xanthomonadaceae;__;__[Pseudomonas] geniculata	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 4.31 The figure describes the taxonomic composition for species levels - Bar type.

The Figure 4.31 above, the species levels shows a significant difference in UC8 between the study groups. Similar to the total genes levels of Gammaproteobacteria was (75.6%), Betaproteobacteria (20.7%), and *Alphaproteobacteria* (3.2%) were indicated in species levels. Moreover, CD7 was indicated in *Bacteroidia (Bacteroidales)* (0.2%) *Nagativicutes (Selenomonadales)*, which did not indicate with the other groups.

Taxonomic assignment - Area type

The UC and CD patients showed a greater microbial load at several taxonomic levels compared to the control group. Proteobacteria, Betaproteobacteria, and Alphaproteobacteria have been found to be more abundant in the class, family, genus, and species than in the control groups. The phylum, however, revealed the same outcome for the control, UC, and CD groups. This is probably a result of a lack of uniformity in microbiome technologies, or possibly as a result of disease-related microbiome variability. (Alam M, T, et al.,2020).

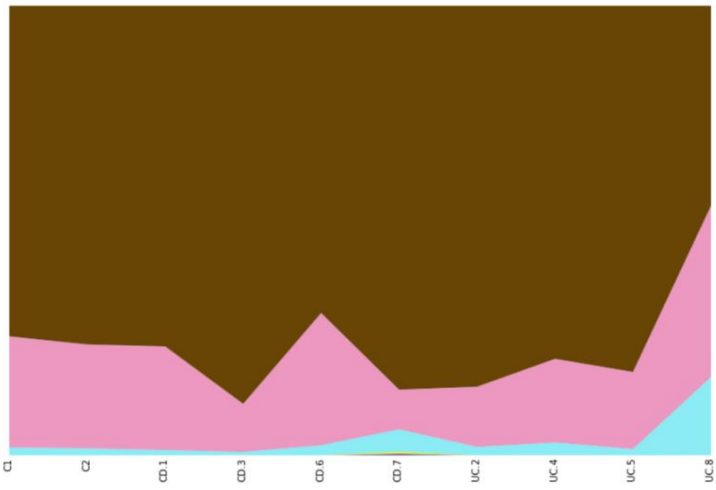


View Table (.txt)

Legend	Taxonomy	Total	C1	C2	CD.1	CD.3	CD.6	CD.7	UC.2	UC.4	UC.5	UC.8
		%	%	%	%	%	%	%	%	%	%	%
	Bacteria;__Actinobacteria	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Bacteroidetes	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.5%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria	99.9%	100.0%	100.0%	100.0%	100.0%	100.0%	99.3%	100.0%	100.0%	100.0%	100.0%

Figure 4.32 The figure describes the taxonomic assignment for phylum level -

The Figure above 4.32, which depicts the phylum level, CD, UC, and control groups in the bacteria, demonstrates that Proteobacteria made up all of the bacteria. The other bacterial groupings are represented as non-significant with (0%). However, there have been some value shifts in CD7 that have been seen in the protists (99.3%), firmicutes (0.5%), and bacteroidetes (0.2%).

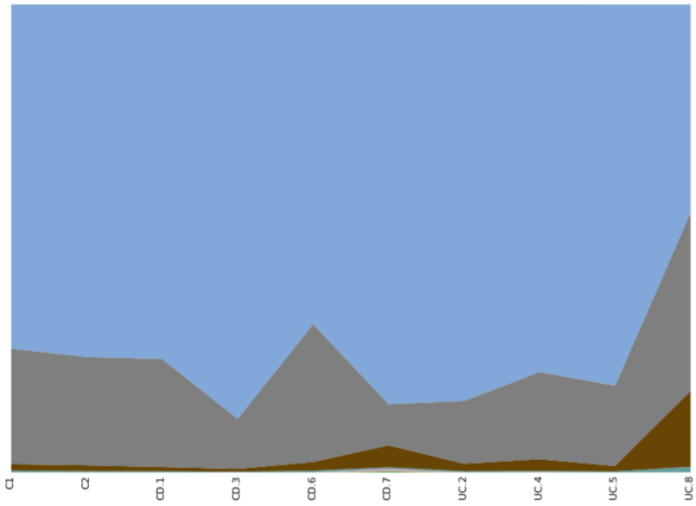


View Table (.txt)

Legend	Taxonomy	Total %	C1 %	C2 %	CD.1 %	CD.3 %	CD.6 %	CD.7 %	UC.2 %	UC.4 %	UC.5 %	UC.8 %
	Bacteria;__Actinobacteria;__Actinomycetia	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Bacteroidetes;__Bacteroidia	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria;__	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Bacilli	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Negativicutes	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria	3.6%	1.8%	1.5%	1.2%	0.7%	2.2%	5.0%	1.9%	2.8%	1.4%	17.4%
	Bacteria;__Proteobacteria;__Betaproteobacteria	20.7%	24.7%	23.1%	23.1%	10.7%	29.4%	8.8%	13.4%	18.6%	17.2%	38.1%
	Bacteria;__Proteobacteria;__Deltaproteobacteria	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria	75.6%	73.5%	75.3%	75.8%	88.6%	68.3%	85.4%	84.8%	78.5%	81.5%	44.5%

Figure 4.33 The figure describes the taxonomic assignment for class level - Area type.

The class levels that significantly differed between the study groups are shown in Figure 4.33 above. Additionally, CD3 revealed higher levels of Gammaproteobacteria (88.6%) and lowest levels in UC8 (44.5%), but it indicated (75.6%) in the aggregate of all groups. Additionally, an increase in UC8 (38.1%), the greatest level, and CD7 (8.8%), the lowest level, for a total of (20.7%), was seen in betaproteobacteria. The levels of the alpha-proteobacteria also revealed other notable variations, with the highest levels in the UC8 (17.4%) and the lowest levels in the CD3 (0.7%) and overall levels (3.6%). Furthermore, compared to the other groups, CD7 was only indicated in Bacteroidia (0.2%), Clostridia (0.1%), and Negativicutes (0.4%).

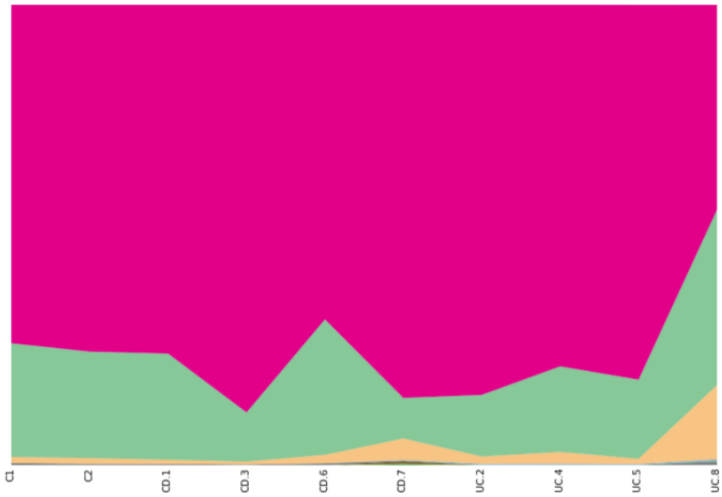


View Table (.txt)

Legend	Taxonomy	Total %	C1 %	C2 %	CD.1 %	CD.3 %	CD.6 %	CD.7 %	UC.2 %	UC.4 %	UC.5 %	UC.8 %
	Bacteria;__Actinobacteria;__Actinomycetia;__Bifidobacteriales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Corynebacteriales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Propionibacteriales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Bacteroidetes;__Bacteroidia;__Bacteroidales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria;__;__Oscillatoriales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Bacilli;__Lactobacillales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Negativicutes;__Selenomonadales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales	0.4%	0.4%	0.3%	0.2%	0.1%	0.4%	0.3%	0.3%	0.3%	0.2%	1.2%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales	3.2%	1.3%	1.3%	0.9%	0.6%	1.9%	4.7%	1.6%	2.5%	1.2%	16.2%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales	20.7%	24.7%	23.1%	23.1%	10.7%	29.4%	8.8%	13.4%	18.6%	17.2%	38.1%
	Bacteria;__Proteobacteria;__Deltaproteobacteria;__Syntrophobacterales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pasteurellales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales	75.6%	73.5%	75.3%	75.8%	88.6%	68.3%	85.4%	84.8%	78.5%	81.4%	44.5%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Xanthomonadales	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 4.34 The figure describes the taxonomic assignment for order level - Area type.

With the exception of the UC8 group, Figure 4.34 of the order levels above demonstrates a substantial difference between the study groups. Gammaproteobacteria (Pseudomonadales) (44.5%), Betaproteobacteria (Burkholderiales) (38.1%), Alphaproteobacteria (Sphingomonadales) (16.2%), and Alphaproteobacteria (Hyphomicrobiales) (1.2%) are the groupings with the most variance among the other groups, but not the highest of all bacteria. Gammaproteobacteria (Pseudomonadales) increased most in CD3 (88.6%), and least in UC8 (44.5%) and overall (75.6%).

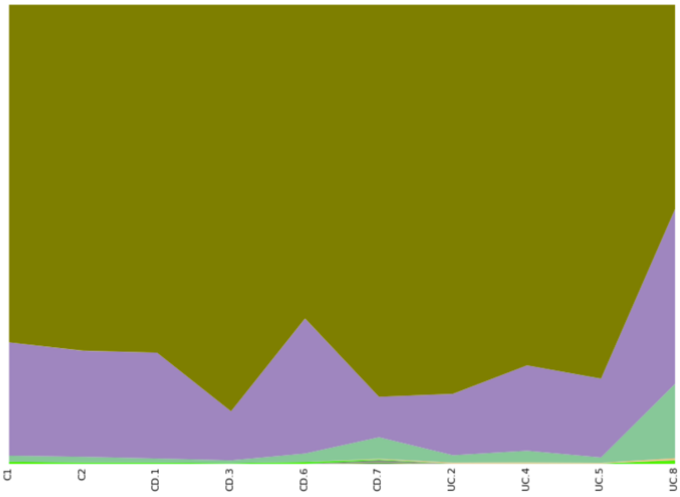


View Table (.txt)

Legend	Taxonomy	Total %	C1 %	C2 %	CD.1 %	CD.3 %	CD.6 %	CD.7 %	UC.2 %	UC.4 %	UC.5 %	UC.8 %
	Bacteria;__Actinobacteria;__Actinomycetia;__Bifidobacteriales;__Bifidobacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Corynebacteriales;__Corynebacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Propionibacteriales;__Propionibacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Bacteroidetes;__Bacteroidia;__Bacteroidales;__Bacteroidaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria;__;__Oscillatoriales;__Oscillatoriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Bacilli;__Lactobacillales;__Lactobacillaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Clostridiaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Lachnospiraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Oscillospiraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Negativicutes;__Selenomonadales;__Selenomonadaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Bradyrhizobiaceae	0.2%	0.4%	0.3%	0.2%	0.1%	0.4%	0.2%	0.0%	0.0%	0.1%	0.7%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylobacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylocystaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Rhizobiaceae	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.2%	0.1%	0.4%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales;__Sphingomonadaceae	3.2%	1.3%	1.3%	0.9%	0.6%	1.9%	4.7%	1.6%	2.5%	1.2%	16.2%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Burkholderiaceae	20.7%	24.7%	23.1%	23.1%	10.7%	29.4%	8.8%	13.4%	18.6%	17.2%	38.1%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Oxalobacteraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Deltaproteobacteria;__Syntrophobacterales;__Syntrophaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pasteurellales;__Pasteurellaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae	75.6%	73.5%	75.3%	75.8%	88.6%	68.3%	86.4%	84.8%	78.5%	81.4%	44.5%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Xanthomonadales;__Xanthomonadaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 4.35 The figure describes the taxonomic assignment for family level - Area type.

Figure 4.35 reveals a substantial difference in family levels between the study groups, however the group of UC8 showed more variances than the other groups. UC8 shows a highest level of betaproteobacteria-Burkholderiales (Burkholderiaceae) (38.1%), Alphaproteobacteria-Sphingomonadales (Shpingomonadaceae) (16.2%), Alphaproteobacteria-Hyphomicrobiales (Rhizobiaceae) (0.4%), and Alphaproteobacteria-Hyphomicrobiales (Bradyrhizobiaceae) (0.7%), but the lowest in Gammaproteobacteria-Pseudomonadales (Pseudomonadaceae) (44.5%). Additionally, the levels of order suggest that the total Gammaproteobacteria (75.6%) in all groupings.

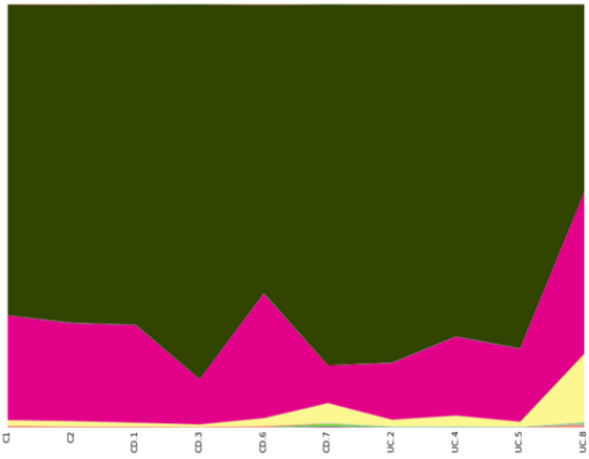


View Table (.txt)

Legend	Taxonomy	Total %	C1 %	C2 %	CD.1 %	CD.3 %	CD.6 %	CD.7 %	UC.2 %	UC.4 %	UC.5 %	UC.8 %
	Bacteria;__Actinobacteria;__Actinomycetia;__Bifidobacteriales;__Bifidobacteriaceae;__Bifidobacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Corynebacteriales;__Corynebacteriaceae;__Corynebacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Propionibacteriales;__Propionibacteriaceae;__Cutibacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Bacteroidetes;__Bacteroidia;__Bacteroidales;__Bacteroidaceae;__Phocaeicola	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria;__;__Oscillatoriales;__Oscillatoriaceae;__Aerosakkonema	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Bacilli;__Lactobacillales;__Lactobacillaceae;__Lactobacillus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Clostridiaceae;__Clostridium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Lachnospiraceae;__Kineothrix	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Lachnospiraceae;__Mediterraneibacter	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Oscillospiraceae;__Acetivibrio	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Negativicutes;__Selenomonadales;__Selenomonadaceae;__Megamonas	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Bradyrhizobiaceae;__Bradyrhizobium	0.2%	0.4%	0.3%	0.2%	0.1%	0.4%	0.2%	0.0%	0.0%	0.1%	0.7%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylobacteriaceae;__Methylorubrum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylocystaceae;__Methylopila	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Rhizobiaceae;__Agrobacterium	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.2%	0.1%	0.4%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales;__Sphingomonadaceae;__Sphingomonas	3.2%	1.3%	1.3%	0.9%	0.6%	1.9%	4.7%	1.6%	2.5%	1.2%	16.2%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Burkholderiaceae;__Ralstonia	20.7%	24.7%	23.1%	23.1%	10.7%	29.4%	8.8%	13.4%	18.6%	17.2%	38.1%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Oxalobacteraceae;__Glaciimonas	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Oxalobacteraceae;__Noviherbaspirillum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Deltaproteobacteria;__Syntrophobacterales;__Syntrophaceae;__Smithella	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pasteurellales;__Pasteurellaceae;__Haemophilus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Azotobacter	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas	75.6%	73.5%	75.3%	75.8%	88.6%	68.3%	85.4%	84.8%	78.5%	81.4%	44.5%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Xanthomonadales;__Xanthomonadaceae;__	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 4.36 The figure describes the taxonomic assignment for genus level - Area type.

Figure 4.36 demonstrates a substantial difference in UC8 between the study groups based on genus levels. Betaproteobacteria (20.7%), Alphaproteobacteria (3.2%), and Gammaproteobacteria (75.6%) made up the majority of the population. Additionally, only Nagativicutes (Selenomonadales) Bacteroidia (Bacteroidales) (0.2%) showed CD7 indication, compared to the other categories.



View Table (.txt)

Legend	Taxonomy	Total	C1	C2	CD.1	CD.3	CD.6	CD.7	UC.2	UC.4	UC.5	UC.8
		%	%	%	%	%	%	%	%	%	%	%
	Bacteria;__Actinobacteria;__Actinomycetia;__Bifidobacteriales;__Bifidobacteriaceae;__Bifidobacterium;__Bifidobacterium longum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Corynebacteriales;__Corynebacteriaceae;__Corynebacterium;__Corynebacterium appendicis	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Propionibacteriales;__Propionibacteriaceae;__Cutibacterium;__Cutibacterium granulosum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Actinobacteria;__Actinomycetia;__Propionibacteriales;__Propionibacteriaceae;__Cutibacterium;__Cutibacterium namnetense	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	Bacteria;__Bacteroidetes;__Bacteroidia;__Bacteroidales;__Bacteroidaceae;__Phocaeicola;__Phocaeicola vulgatus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Cyanobacteria;__;__Oscillatoriales;__Oscillatoriaceae;__Aerosakkonema;__Aerosakkonema funiforme	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Bacilli;__Lactobacillales;__Lactobacillaceae;__Lactobacillus;__Lactobacillus gallinarum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Clostridiaceae;__Clostridium;__Clostridium saudiense	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Lachnospiraceae;__Kineothrix;__Kineothrix alysoides	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Lachnospiraceae;__Mediterraneibacter;__[Ruminococcus] gnavus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Clostridia;__Eubacteriales;__Oscillospiraceae;__Acetivibrio;__Acetivibrio alkalicellulosi	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Firmicutes;__Negativicutes;__Selenomonadales;__Selenomonadaceae;__Megamonas;__Megamonas funiformis	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Bradyrhizobiaceae;__Bradyrhizobium;__Bradyrhizobium valentinum	0.2%	0.4%	0.3%	0.2%	0.1%	0.4%	0.2%	0.0%	0.0%	0.1%	0.7%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylobacteriaceae;__Methylorubrum;__Methylorubrum populi	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Methylocystaceae;__Methylopila;__Methylopila henanensis	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Hyphomicrobiales;__Rhizobiaceae;__Agrobacterium;__Agrobacterium fabrum	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.2%	0.1%	0.4%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales;__Sphingomonadaceae;__Sphingomonas;__Sphingomonas adhaesiva	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales;__Sphingomonadaceae;__Sphingomonas;__Sphingomonas kyeonggiensis	3.2%	1.3%	1.2%	0.8%	0.6%	1.8%	4.7%	1.6%	2.5%	1.2%	16.2%
	Bacteria;__Proteobacteria;__Alphaproteobacteria;__Sphingomonadales;__Sphingomonadaceae;__Sphingomonas;__Sphingomonas pituitosa	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Burkholderiaceae;__Ralstonia;__Ralstonia pickettii	20.7%	24.7%	23.1%	23.1%	10.7%	29.4%	8.8%	13.4%	18.6%	17.2%	38.1%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Oxalobacteraceae;__Glaciimonas;__Glaciimonas frigoris	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Betaproteobacteria;__Burkholderiales;__Oxalobacteraceae;__Noviherbaspirillum;__Noviherbaspirillum suwonense	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Deltaproteobacteria;__Syntrophobacterales;__Syntrophaceae;__Smithella;__Smithella propionica	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pasteurellales;__Pasteurellaceae;__Haemophilus;__Haemophilus paraphrohaemolyticus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Azotobacter;__Azotobacter armeniacus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas;__Pseudomonas guguanensis	0.0%	0.0%	0.1%	0.1%	0.0%	0.1%	0.0%	0.1%	0.1%	0.1%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas;__Pseudomonas monteilii	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.2%	0.0%	0.1%	0.2%	0.1%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas;__Pseudomonas paralactis	75.6%	73.4%	76.1%	76.7%	88.8%	68.1%	86.2%	84.6%	78.4%	81.2%	44.4%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas;__Pseudomonas punonensis	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Pseudomonadales;__Pseudomonadaceae;__Pseudomonas;__Pseudomonas tolaasii	0.0%	0.0%	0.1%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
	Bacteria;__Proteobacteria;__Gammaproteobacteria;__Xanthomonadales;__Xanthomonadaceae;__;__[Pseudomonas] geniculata	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 4.37 The figure describes the taxonomic assignment for species level - Area type.

The Figure 4.37 above, the species levels shows a significant difference in UC8 between the study groups. Similar to the total genes levels of Gammaproteobacteria was (75.6%), Betaproteobacteria (20.7%), and *Alphaproteobacteria* (3.2%) were indicated in species levels. Moreover, CD7 was indicated in *Bacteroidia (Bacteroidales)* (0.2%) *Nagativicutes (Selenomonadales)*, which did not indicate with the other groups.

2 DISCUSSION AND CONCLUSION

This study considered as one of the few investigating studies using human blood samples in the Turkish community and for that we found a difficulties in finding information, sources, and in convincing voluntaries for this work. Thus, even finding similar studies for comparing was rarely to find, so we tried to compered some the results with fecal studies.

In our investigation, alterations in the blood samples were discovered, suggesting a link to inflammatory bowel illness. Investigations that made use of fecal samples have verified our findings. We discovered a dysbiosis in the IBD samples and noted variations in ulcerative colitis and Crohn's disease. We have confirmed that UC and CD are separate subtypes at the microbiome level, as earlier research in adult populations have previously revealed, despite sharing similar traits and being examined together. (Forbes et al. 2018).

The PCR process uses a range of temperatures to get the optimal concentration; these temperatures are discussed in the materials and methods section. The preferred temperature, as indicated by our own PCR results, is 61.3 ^{o}C, while the designated temperature range is 55 to 68 ^{o}C. The results of our investigation's PCR analysis showed a drop in the amount of Faecalibacterium prausnitzii. When utilizing targeted quantitative PCR analysis, a study comparable to ours (Lopez- Siles, M., et al., 201) suggests that the existence or absence of F. prausnitzii phylogroups may be a disease predictor. We still don't fully understand the F. prausnitzii group's diversity or how that diversity affects disease. This is because most amplicon-based studies only sequence a small number of fleeting hypervariable areas, making it challenging to get a complete picture. Even after sequencing the full 16S rRNA gene, several strains of F. prausnitzii do not fit into any recognized phylogroups..

Raw data is converted from the initial raw pictures produced by the Illumina sequencing into raw data for processing. System control and base calling are performed using the sequencing control software in conjunction with the integrated primary analysis program, known as RTA. According to our research, sample UC 8 had the greatest compositional percentage (52.65), followed by sample UC 5 (51.82%). With percentages of 98.56 and 95.03 for Q20(%) and Q30(%), UC 8 has the highest values, indicating a change in the way samples were sequenced in earlier research. On the other hand, (Schirmer, M. et al., 2018) found a connection between

the disappearance of Clostridium cluster IV and XIV species and an increase in Enterobacteriaceae species and a decrease in Firmicutes.

Our study yielded different results compared to a study conducted by Shan Shah et al. in 2016. While they found higher diversity in CD, we found that UC has higher diversity compared to CD and healthy individuals. This discrepancy may be attributed to the samples we analyzed, which were distinct from those used in previous studies.

According to the study's findings, both those with ulcerative colitis and crohn's disease and those in the control group had lower levels of bacterial diversity. Participants in the control and CD groups especially displayed a predominance of microorganisms from the Proteobacteria phylum..

According to the study, people with IBD had more total bacterial DNA in their blood samples than people without the condition. Our results, however, are in contrast to those of Jessica D. Forbes et al., who observed that Clostridium abundances were significantly higher in both active and inactive illness groups of IBD. However, we noticed a drop notably in the group of Clostridium leptum...

Firmicutes and Bacteroidetes make up the majority of gut microorganisms. These two phyla make up the division of gut bacteria, together with Proteobacteria and Actinobacteria. In contrast to the research done by Sartor and Mazmanian (2012), it has been found that CD patients have less variety in their gut bacteria than healthy participants do. Several additional research have also produced similar findings. Families with a large number of members, like the Bacteroidaceae, Prevotellaceae, Clostridiaceae, Lachnospiraceae, Ruminococcaceae, Lryspelotrichaceae, Enterobacteriaceae, Fusobacteriaceae, Lactobacillaceae, Enterococcaceae, and Streptococcaceae, regularly experience alterations..

Metabolizing proteins and carbohydrates are tasks that the class Bacteroidetes excel at in the gut. They also generate butyrate and carry out important duties in preventing harmful bacteria from colonizing the gastrointestinal tract.. (Markelova M, et al, 2023).

Jessica D. Forbes et al. (2018) discovered in their study that each group of sick people included a unique variety of germs. Although our research showed that Proteobacteria were the most abundant in CD, CD contained a lot more Firmicutes at the phylum level. Our

findings supported the earlier study's conclusions that the bacteria types in IBD, CD, UC, and healthy controls varied from those in healthy controls. In addition, we found that disease activity and phenotype influence the quantity and variety of bacteria in IBD...

Few pathobionts (Escherichia and Fusobacterium), which could be seen in the differences in microbial composition between CD and UC, were enriched in Faecalibacterium. In contrast to healthy people, CD lacked any commensal symbionts. According to Gophna et al. (2006)'s findings, our data indicate that CD patients have more dysbiosis than UC patients do, while also showing that the microbial composition of UC patients and healthy people is similar, albeit with minor differences. In addition, our study identified decreased levels of Faecalibacterium prausnitzii, a regular resident of the human intestine, and Mycobacterium avium paratuberculosis, which is frequently detected in the digestive systems of ruminants suffering from Johne disease, in people with CD....

The bacterial population makeup in people with Crohn's disease and ulcerative colitis is different from that in healthy individuals. Notably, it was shown that patients with Crohn's and ulcerative colitis had higher quantities of genetic material in their blood than healthy people did. Surprisingly, the study's findings showed that nitrogenous base pairs (AT and CG) were present in the patients despite the fact that healthy controls had significantly more genetic material.. These results strongly suggest that microbes are involved in the onset of IBD, as evidenced by the prevalence of bacterial DNA with nitrogenous bases in the affected people. Unexpectedly, bacteria and genetic material were discovered in blood samples taken from healthy people. The limitations of the study prevented more thorough investigation, therefore the reason for this occurrence is still unknown. In addition to a potential genetic predisposition, the prevalence of foods high in purines and pyrimidines may also have a role in the development of IBD in otherwise healthy people. This study shows the crucial part bacteria play in the development of IBD by providing support for this claim..

Additionally, there are notable statistical differences between the CD and UC cases used, which is the opposite of what other studies' findings have shown. The number of total base pairs in UC is 52 million, but the number in CD is 48.1 million on a mathematical average. This indicates that UC patients have higher levels of microorganisms than CD patients do..

UC 8 had the greatest percentage composition of 52.65 based on the guanine-Cytosine content in the samples used, followed by UC 5 with a percentage score of 51.82. Additionally, UC 8 has the greatest Q20 (%) and Q30 (%) of 98.56 and 95.03, respectively, in terms of phred quality. According to the study, UC 8 has the highest read count, and C1 has the lowest read count. Naturally, this is expected because healthy persons, including the control 1(C 1) participant, are predicted to have the least genetic material.

The types of microorganisms that cause CD and UC vary depending on the samples that were used. The gamma diversity index's numerical value serves as evidence for this. This has many implications, but the most crucial one is that the microbes in all the samples are similar, suggesting that a single pathogen may be the source of both UC and CD. The study found that nitrogenous base pairs are present in large concentrations wherever genetic material is abundant. The sample with the most genetic material is UC 8, which also has the greatest AT/GC ratio. As a result, it will be determined that patient has been the most severely impacted by the condition.

Sample UC 8 has the greatest indices after taking Shannon and Gini Gimson into account. The most prevalent and rarest species of microorganisms, specifically bacteria, are reflected in this. It inevitably indicates that sample UC 8 has the highest levels of both alpha and beta diversity, which accounts for the presence of the most genetic material..

IBD is therefore a complex disease involving the host, microorganisms (in this case, bacteria), and environmental factors. A multidisciplinary approach connecting genetics, risk factors, and microorganisms would be required to fully understand the fundamental process of host-microorganism interaction in developing the disease. According to the results, in contrast to other studies, Crohn's disease and ulcerative colitis are always caused by bacteria, despite the fact that the biological taxonomy of these bacteria varies from phylum to phylum and species to species. Other research suggests that UC and CD can be caused by many bacteria, as well as by genetic predisposition or other connected factors. Therefore, despite differences in their taxonomic classification, bacteria were found to be the main cause of IBD in our analysis...

Although the current study used human blood samples to assess crohn's disease and ulcerative colitis, but the other studies used fecal samples, the findings of this inquiry have been compared to those of other investigations. Our research is distinctive among those that

analyze these illnesses using blood samples as a result. Additionally, when this study is compared to other investigations, it's likely that the findings or those pertaining to the largest bacterial community or perhaps not discovering the same bacterial.

Numerous studies have supported these conclusions, including those by (Xu, et al., 2022) and (R. Balfour Sartor, MD, and Sarkis K. Mazmanian, 2012)..

According to our findings, the initial study also found an increase in proteobacteria and actinobacteria. The present study demonstrated a decrease in this same bacteria, which displays different findings from the same study, where we discovered a rise in Firmicutes, particularly in the healthy groups. This difference in results may be related to the type of samples used. Additionally, the results of the second trial, which are identical to those of the current study.

REFERENCES

Actis G C., et al., 2019, History of Inflammatory Bowel Diseases, *journal of clinical medicine*, Volume 8.

Aden K and Reindl W, 2019, The Gut Microbiome in Inflammatory Bowel Diseases: Diagnostic and Therapeutic Implications, *Visceral medicine*, Volume 35.

Ahlostrom, G, 2022, Metagenomic analysis of crohn's disease, *UPTEC X 22031*

Alam M, T, 2020, Microbial imbalance in inflammatory bowel disease patients at different taxonomic levels, *Gut Pathogens*, Volume 12, Issue 1.

Aldars-García L., et al., 2021, The Interplay between Immune System and Microbiota in Inflammatory Bowel Disease: A Narrative Review, *International Journal of Molecular Sciences*, Volume 22.

Alkhatry M, et al., 2020, First United Arab Emirates consensus on diagnosis and management of inflammatory bowel diseases: A 2020 Delphi consensus, *World Journal of Gastroenterol*, Volume 26, Issue 43.

Álvarez-Salamero, et al, 2020, IL-23 signaling regulation of pro-inflammatory T-cell migration uncovered by phosphoproteomics, *PLoS biology*, Volume 18, Issue 3.

Argollo M, et al., 2020, Optimizing Biologic Therapy in IBD: How Essential Is Therapeutic Drug Monitoring?, *Nature Reviews Gastroenterology and Hepatology*. Volume 17.

Barbara A. S, et al., 2022, Crohns Disease, The Editors of Encyclopaedia Britannica, https://www.britannica.com/science/Crohn-disease.

Cambrian Y. Liu and D. Brent Polk, 2018, Microbiomes Through the Looking Glass: What Do UC?, *Cell Host Microbe*, Volume 24, Issue 4.

Candan, G, et al, 2019, Irak-4 rs4251481 gene variant: as a risk factor on inflammatory bowel disease, *Turkish Journal of Medical Sciences*, Volume 49, Issue 2.

Charles Y. Chiu and Steven A. Miller, 2019, Clinical Metagenomics, *Natural review (Genetics)*, Volume 20.

Colombel J.-F, et al., 2018, Effect of tight control management on Crohn's disease (CALM): a multicentre, randomised, controlled phase 3 trial, *The Lancet*, Volume 390, Issue 10114.

Coward S., 2019, Forecasting the Future: A Trek through the Changing Landscape of Inflammatory Bowel Disease, *Thesis, university of Calgary*.

Crowley E, et al., 2020, Prevalence and clinical features of inflammatory bowel diseases associated with monogenic variants, identified by whole-exome sequencing in 1000 children at a single center, *HHS Public Access Gastroenterology*, Volume 158, Issue 8.

Datta S, et al., 2020, Metagenomic Applications in Microbial Diversity, Bioremediation, Pollution Monitoring, Enzyme and Drug Discovery. A review, *Environmental Chemistry Letters*, Volume 18.

Denson L, et al., 2019, Genetic and transcriptomic variation linked to neutrophil granulocyte-macrophage colony-stimulating factor signaling in pediatric Crohn's disease. *Inflammatory Bowel Diseases, Research sponsored by the Crohn's & Colitis Foundation*, Volume 25, Number 3.

Elhag, D, A et al, 2022, Inflammatory Bowel Disease Treatments and Predictive Biomarkers of Therapeutic Response, *International Journal of Molecular Sciences*, Volume 23, Issue 13.

Examensarbete 30 hp.

Feki, S, et al, 2018, Is there any Association of NOD2/CARD15 Gene Polymorphism with Tunisian Pemphigus Foliaceus, *Journal of Dermatology and Plastic Surgery*, Volume 3, Issue 1.

Gammoh N, 2020, The multifaceted functions of ATG16L1 in autophagy and related processes, *Journal of Cell Science*, Volume 133, Issue 20.

Ghour A, Y, et al, 2020, Secondary causes of inflammatory bowel diseases, *World Journal of Gastroenterology*, Volume 26 Issue 28.

Gilaad G. Kaplan and Joseph W. Windsor, 2021, The four epidemiological stages in the global evolution of inflammatory bowel disease, *perspectives*, volume 18.

Gophna, Uri, et al, 2006, Differences between Tissue-Associated Intestinal Microfloras of Patients with Crohn's Disease and Ulcerative Colitis, *Journal of Clinical Microbiology* Volume 44, Issue 11.

Higashiyama M, et al., 2019, Management of Elderly Ulcerative Colitis in Japan, *Journal of Gastroenterology*, Volume54.

Hornschuh M, et al, the role of epigenetic modifications for the pathogenesis of Crohn's disease, *Clinical Epigenetics*, Volume 13, Article number 108.

Horowitz J E, et al., 2021, Mutation spectrum of NOD2 reveals recessive inheritance as a main driver of Early Onset Crohn's Disease, *Scientific Reports*, Volume 11, Article number: 5595.

Huang G, et al., 2021, Editorial: Applications of Metagenomics in Studying Human Cancer, *Frontiers in Genetics*, Volume 12.

Imai T, et al., 2019, Characterization of fungal dysbiosis in Japanese patients with inflammatory bowel disease, *Journal of Gastroenterol*, Volume 54.

Jacob E M., et al., 2020, Inflammatory Bowel Disease: The Emergence of New Trends in Lifestyle and Nanomedicine as the Modern tool for Pharmacotherapy, Nanomaterials, Volume 10.

Jessica D. Forbes, et al, 2018, A comparative study of the gut microbiota, *Microbiome*, Article number, 221, issue 7.

Jonathan P Segal, et al., Ulcerative Colitis: An Update, *Clinical Medicine (London, England)*, Volume 21.

Jones G R, et al., 2019, IBD prevalence in Lothian, Scotland, derived by capture-recapture methodology, *BMJ Journals, Gut*, Volume 68, Issue 11.

Kayani, et al., 2021, Genome-Resolved Metagenomics Using Environmental and Clinical Samples, *Briefings in Bioinformatics (Published by Oxford University Press)*, Volume 22, Issue 5.

Kerri L. Glassner, et al., 2019, The microbiome and inflammatory bowel disease, *American Academy of Allergy, Asthma & Immunology*, Volume 145, Number 1.

Khalili, H. et al., 2020, Adherence to a Mediterranean diet is associated with a lower risk of later- onset Crohn's disease: results from two large prospective cohort studies, *BMJ Journals, Gut*, Volume 69, Issue 9.

Kobayashi T, et al., 2020, Ulcerative Colitis, *Nature Reviews Disease Primers*, 6, Article number 74.

Kondo, K. et al., 2019, The association between environmental factors and the development of Crohn's disease with focusing on passive smoking: a multicenter case-control study in Japan. PLoS One, Volume 14.

Konopiński MK, 2020, Shannon diversity index: a call to replace the original Shannon's formula with unbiased estimator in the population genetics studies, *Bioinformatics and Genomics*, PeerJ 8:e9391 https://doi.org/10.7717/peerj.9391

Kővári B, et al., 2022, Histologic Diagnosis of Inflammatory Bowel Diseases, *Advances in Anatomic Pathology*, Volume 29, Number 1.

Kumar, M., et al., 2019, Integrating omics for a better understanding of Inflammatory Bowel Disease: a step towards personalized medicine, *Journal of Translational Medicine*, Article number: 419.

Kunk M, et al, 2020, T Helper Cells: The Modulators of Inflammation in Multiple Sclerosis, *Cells*, Volume 9, Issue 2.

Lamb CA, et al., 2021, British Society of Gastroenterology consensus, *BMJ Publishing Group Ltd & British Society of Gastroenterology*, Volume 68, Issue Supply 3.

Lee M and Eugene B., 2021, Chang, Inflammatory Bowel Diseases and the Microbiome: Searching the Crime Scene for Clues, *Gastroenterology*, Volume 160, Issue 2.

Lega S, et al., 2020, Diagnostic Approach to Monogenic Inflammatory Bowel Disease in Clinical Practice: A Ten-Year Multicentric Experience, *Inflammatory Bowel Diseases*, Volume 26, Number 5.

Li Yim, et al., 2020, Whole-Genome DNA Methylation Profiling of CD14+ Monocytes Reveals Disease Status and Activity Differences in Crohn's Disease Patients, *Journal of Clinical Medicine*, Volume 9.

Lopez- Siles, M., Duncan, S. H., Garcia- Gil, L. J. & Martinez- Medina, M, 2017, Faecalibacterium prausnitzii: from microbiology to diagnostics and prognostics. ISME J. 11, 841–852.

Loy L, et al., 2019, Detection and management of early stage inflammatory bowel disease: an update for clinicians, *Expert Review of Gastroenterology & Hepatology*, Volume 13.

Ma, Y, et al, 2021, Metagenome Analysis of Intestinal Bacteria in Healthy People, Patients With Inflammatory Bowel Disease and Colorectal Cancer, Journal Frontieres in Cellular and Infectious Microbiology, Volume 11.

Marcin W and Michael S, 2018, Genetics and epigenetics of inflammatory bowel disease, *Biomedical intelligence, Swiss Medical Weekly.*

Markelova M, et al, 2023, Association between Taxonomic Composition of Gut Microbiota and Host Single Nucleotide Polymorphisms in Crohn's Disease Patients from Russia, *international journal of molecular sciences*, Volume 29, Issue 9, https://www.mdpi.com/1422-0067/24/9/7998.

Matsuoka K, et al., 2018, Evidence-based Clinical Practice Guidelines For Inflammatory Bowel Disease, *Journal of Gastroenterology*, Volume 53.

Max Roser., et al., 2019, World Population Growth, *Published online at Our World in Data.org. Retrieved from: 'https://ourworldindata.org/world-population-growth*.

Mazziotto M, et al., 2018, Comparison of two bacterial DNA extraction methods from non-polluted and polluted soils, *Folia Microbiol (Praha*), Volume 63.

Moret-Tatay I, et al., 2019, Identification of Epigenetic Methylation Signatures with Clinical Value in Crohn's Disease, *Clinical and Translational Gastroenterology*, Volume 10, Issue 10.

Nascimento, R.P, et al, 2020, Ulcerative Colitis: Gut Microbiota, Immunopathogenesis and Application of Natural Products, *Journal of Life Sciences*, Volume 258.

Nelson, K.E, 2011, Metagenom of the Human Body, A call to replace the original Shannon's formula with unbiased estimator in the population genetics studies. https://pubmed.ncbi.nlm.nih.gov.Date accessed 23/102022, Book.

Nguyen, G C, et al., 2019, The impact of inflammatory bowel disease in Canada 2018: IBD in seniors, *Journal of the Canadian Association of Gastroenterology*, Volume 2.

Nielsen O H, et al., 2020, Efficacy and Safety of Methotrexate in The Management Of Inflammatory Bowel Disease: A Systematic Review and Meta-Analysis of Randomized, Controlled Trials, *EClinicalMedicine (Part of the Lancet Discovery Science),* Volume 20.

Osunmakinde C O, et al., 2018, Overview of Trends in the Application of Metagenomic Techniques in the Analysis of Human Enteric Viral Diversity in Africa's Environmental Regimes, *Viruses (Animal Viruses)*, Volume 10.

Park S H, et al., 2019, A 30-year trend analysis in the epidemiology of inflammatory bowel disease in the Songpa- Kangdong district of Seoul, Korea in 1986–2015, *Journal of Crohns Colitis*, Volume 13, Issue 11.

Piovani D, et al., 2019, Environmental risk factors for inflammatory bowel diseases: an umbrella review of meta- analyses, *Gastroenterology*, Volume 157, Issue 3.

Porter R J, et al., 2020, Ulcerative colitis: Recent advances in the Understanding of Disease Pathogenesis [version 1, peer review 2 approved], *F1000Research*, Volume 9.

Qingdong Guan, 2019, A Comprehensive Review and Update on the Pathogenesis of Inflammatory Bowel Disease, *Hindawi Journal of Immunology Research*.

R. Balfour Sartor , MD 1 and Sarkis K. Mazmanian, 2012, Intestinal Microbes in Infl ammatory Bowel Diseases, The American Journal of GASTROENTEROLOGY Supplements, 1:15– 21.

Ram N. Bharagava, et al., 2019, Applications of Metagenomics in Microbial Bioremediation of Pollutants: From Genomics to Environmental Cleanup, Chapter 26, Microbial Diversity in the Genomic Era.

Roda G, et al., 2020, Crohn's disease, *Nature Reviews Disease Primers*, Volume 6, Article number 22.

Schirmer, M. et al. 2018, Compositional and temporal changes in the gut microbiome of pediatric ulcerative colitis patients are linked to disease course. Cell Host Microbe 24, 600–610.

Segal A W, 2019, Studies on patients establish Crohn's disease as a manifestation of impaired innate immunity, *Journal of Internal Medicine*, Volume 286, Issue 4.

Seyed Tabib NS., et al., 2019, Big data in IBD: big progress for clinical practice, *Recent advances in clinical practice*, Volume 69, Issue 8.

Shah R, et al, 2016. Composition and Function of the Pediatric Colonic Mucosal Microbiome in Untreated Patients with Ulcerative Colitis, *Gut Microbes*, Volume 7, Issue 5.

Siew C Ng, et al., 2018, Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population- based studies, *Lancet*, Volume 390, Issue 10114.

Tasha M. Santiago-Rodriguez and Emily B. Hollistera, 2020, Potential Applications of Human Viral Metagenomics and Reference Materials: Considerations for Current and Future Viruses, *Applied and Environmental Microbiology*, Volume 86 Issue 22.

Vrakas S, et al, 2017, Intestinal Bacteria Composition and Translocation of Bacteria in Inflammatory Bowel Disease, Journal pone, Volume 12, Issue 1.

Wael N. Hozzein, 2020, Introductory Chapter: Metagenomics and Metagenomic Approaches, *IntechOpen*, Book. ISBN978-1-83880-056-7.

Wainberg M, 2019, Sinnott-Armstrong N, Mancuso N, et al. Opportunities and challenges for transcriptome-wide association studies, *Nature Genetics*, Volume 51.

Whitney D. Lynch and Ronald Hsu, 2021, Ulcerative colities, A service of the National Library of Medicine, *National Institutes of Health*, Bookshelf ID: NBK459282 .

William A Rowe, et al., 2020, Inflammatory Bowel Diseases, *American Gastroenterological Association.*

Xu, et al, 2022, The gut metagenomics and metabolomics signature in patients with inflammatory bowel disease, Gut Pathogens, Volume, 14, Article number, 26.

Younis N, et al., 2020, Inflammatory bowel disease: between genetics and microbiota, *Molecular Biology Reports*, Volume 47.

Zahraa Al Bander, et al., 2020, The Gut Microbiota and Inflammation: An Overview, *International Journal of Environmental Research and Public Health*, Volume 17, Issue 20.

Zhang, et al., 2021, Advances in Metagenomics and Its Application in Environmental Microorganisms, *Frontiers in Microbiology*, Volume 12.

Zhao M and Burisch J, 2019, Impact of Genes and the Environment on the Pathogenesis and Disease Course of Inflammatory Bowel Disease, *Digestive Diseases and Sciences*, Volume 64.

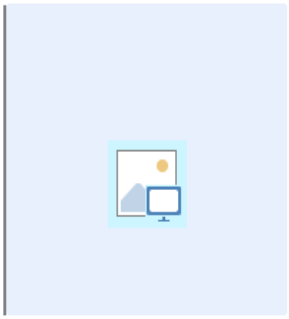
Zhu F, et al., 2019, The risk of Parkinson's disease in inflammatory bowel disease: A systematic review and meta-analysis, *Digestive and Liver Disease*, Volume 51, Issue 1.

https://acikders.ankara.edu.tr, 2023,

APPENDICES

CURRICULUM VITAE

Personal Information	
Name Surname	Omar Sajer Naser
Place of Birth	
Date of Birth	Tarih girmek için tıklayın veya dokunun.
Nationality	☐ T.C. ☒ Other: IRAQ
E-Mail Adress	
Web Page	



Education Information	
Undergraduate	
University	Tikrit University
Faculty	Education for Pure Sciences
Department	Biology
Year of Graduation	21.08.2011

Masters	
University	Fatih University
Name of Institute	Graduate Institute of Sciences and Engineering
Department	Biology
Program	Biology

Ph.D.	
University	Istanbul University
Name of Institute	Graduate School of Natural and Applied Sciences
Department	Department of Molecular Biology and Genetics
Program	Programme of Biotechnology

Articles and Publications
Omar N, et al, 2023, Application of Metagenomics approaches for Crohn's Disease, *Journal of Population Therapeutics & Clinical Pharmacology*, Volume 30, Issue 3. Effect of Environmental Conditions on the Growth of Toxic Cyanobacteria, by S. Aktan1, Ayşe Büşra Şengül, Omar Sajer, 2016. The Removal of Cyanobacterial Toxins from Drinking Water by Alum, Ferric Chloride

and Ferrous Sulfate Coagulation, By Ayşe Büşra Şengül, Sevil Aktan, Omar Sajer, 1st International Conference on Environmental Science and Technology, ICOEST -2015.