M.Tech (Industrial Biotechnology)

Sukrit Kashyap

2023

ANALYSIS OF GUT MICROBIOME DIVERSITY AND INVESTIGATING ITS ONCOLOGICAL IMPLICATIONS

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF REQUIREMENTS

FOR THE AWARD OF THE DEGREE

OF

MASTER OF TECHNOLOGY

IN

INDUSTRIAL BIOTECHNOLOGY

Submitted by:

SUKRIT KASHYAP

2K21/IBT/11

Under supervision of

DR. ASMITA DAS

DEPARTMENT OF BIOTECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

MAY, 2023

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, **Sukrit Kashyap**, Roll No. **2K21/IBT/11** of M.Tech (Industrial Biotechnology), hereby declare that the Project Dissertation titled **"Analysis of Gut Microbiome Diversity and Investigating its Oncological Implications"** which is submitted by me to the Department of Biotechnology, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the Degree of Master of Technology, is original and not copied from any source without proper citation. The work has not been previously formed the basis for award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: New Delhi **SUKRIT KASHYAP**

Date: 29.05.2023

DEPARTMENT OF BIOTECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled **"Analysis of Gut Microbiome Diversity and Investigating its Oncological Implications"** which is submitted by Sukrit Kashyap, Roll no. 2K21/IBT/11, Department of Biotechnology, Delhi Technological University, New Delhi, in partial fulfilment of requirement for the award of Degree of Master of Technology, is a record of the project work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part of full for any Degree or diploma to this University or elsewhere.

Place: New Delhi

Date: 9.06.23

DR. ASMITA DAS

Supervisor

09/06/2023

PROF. PRAVIR KUMAR

Head of Department

AKNOWLEDGEMENT

I would like to express my sincere gratitude to Dr. Asmita Das, my supervisor, for her invaluable guidance and support throughout this research. Her expertise and patience have been instrumental in shaping this thesis. I would also like to thank the Department of Biotechnology at Delhi Technological University for providing the necessary resources and facilities for this study. I extend my heartfelt thanks to all those who have played a part in this academic journey and have contributed to the realization of this thesis.

ABSTRACT

In recent years, the understanding of human well-being has undergone a transformative shift, attributing significant importance to the intricate interplay between the gut microbiome and human health. The vast community of microorganisms residing in the gastrointestinal (GI) tract has emerged as a critical regulator of immunological functions, capable of influencing various signal transduction pathways and immune responses. Particularly in the context of cancer, the gut microbiome is increasingly recognized as a pivotal player in the development and progression of this complex disease. Motivated by these advancements, our research project aims to construct a comprehensive and comparable risk evaluation database. This database encompasses diverse ethnic groups, their dietary habits, cancer incidence rates, and insights into the intricate connections between diet, gut microbiota composition, and cancer risk from relevant cohort studies. By integrating these multi-dimensional data sources, we strive to unravel the complex relationships between diet, gut microbiota, and cancer, shedding light on potential therapeutic targets for enhancing human well-being. The project will involve extensive data mining and analysis of available cohort studies, allowing us to assess the diversity of the gut microbiome across different populations and its potential implications in cancer. We employed methodologies to process and interpret metagenomic data, identifying prevalent microbial and HLA biomarkers associated with GI tract cancers. Moreover, by comparing the relative abundance of bacterial genera between cancer cases and non-cancer individuals within each population, we aim to discern potential associations and gain deeper insights into the role of gut microbiota in cancer development. It is crucial to emphasize that the creation of this risk evaluation database does not establish causation between the gut microbiome, diet, and cancer but rather provides a platform for comprehensive analysis and comparison. Focus on elucidating the underlying mechanisms and interactions between the gut microbiome, host immunity, and cancer development, will the way for potential therapeutic interventions and personalized strategies for cancer prevention and treatment.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION

SCFAs	Short-chain fatty acids
LPS	Lipopolysaccharide
TNF-α	Tumour necrosis factor α
NK	Natural Killer
DNA	Deoxyribonucleic acid
CDI	*C. difficile*
HPV	Human papillomavirus
EBV	Epstein-Barr virus
CRC	Colorectal Cancer
WGS	Whole-genome shotgun
NIH	National Institute of Health
DDBJ	Data Bank of Japan
HLA	Human leukocyte antigens
OTUs	Operational taxonomic units
BLAST	Basic Local Alignment Search Tool
GI	Gastrointestinal

CHAPTER-1

INTRODUCTION

Around 400 B.C., Hippocrates emphasised the significance of the intestines to human well-being, stating that poor digestion is the source of various ailments and that mortality lurks within the bowels [1]. In recent decades, the majority of research focused on the impact of bacteria on the intestinal environment has primarily examined gastrointestinal pathogens and their role in promoting diseases [2]. With its numerous commensal microorganisms, the mammalian gut creates a unique and extraordinary ecology. Bacteria, archaea, protists, viruses, and fungi are just a few of the taxonomic groupings that reside inside this ecosystem, also known as the microbiota. In the gut, bacteria make up the majority of the rest of the residents [3]. Furthermore, the genetic data included in the complete human genome is outnumbered by the human microbiota, which is sometimes referred to as "the hidden organ," by a factor of more than 150. Despite the fact that the terms "microbiome" and "microbiota" are sometimes used interchangeably, there are some distinctions between both of them [4]. It is important to note that Louis Pasteur is credited with first recognising the crucial role that our own internal microbial populations play in the physiology and overall wellness of the host [5]. Numerous studies have shown a connection between changes in the diversity and composition of the gut microbiota and a variety of diseases, including inflammatory bowel disease, diabetes, obesity and even psychological disorders. As an example, dysbiosis and disturbance in the gut microbiome has been associated with the emergence of diabetes and obesity. Similar to how inflammatory bowel disease, a long-lasting condition which influences the digestive system, has been linked to changes in the gut flora [6]. Adults' guts typically contain the different species *Bacteroides, Ruminococcus, Clostridium, Escherichia, Streptococcus, Lactobacillus* and Bifidobacterium among others. Nearly sixty percent of all bacteria present in the gut are members of the phyla Firmicutes or Bacteroidetes [7]. Approximately fifty percent of people have been shown to have methane-producing microbes, which are categorised as *Archaea* rather than bacteria [8]. Recent research from The Human Microbiome Project as well as others [9,10] shows that tens of thousands of distinct microbes could live in the gastrointestinal tract of human beings jointly and confirms an elevated level of disparity in the makeup of these populations among individuals, despite the fact that individuals

may have up to several thousand species of bacteria inside their gut. The sheer number of many bacterial genes for fundamental or auxiliary metabolic activities is fairly comparable between people despite this difference in taxa. As a result, there has been tremendous research and financial investment in the field of the microbiota's function in preserving health. Numerous bioactive substances that can affect wellbeing are produced by gut microorganisms; some, like vitamins, are advantageous, while others are harmful [11]. Possible hazardous bacteria are prevented from inflicting harm on tissues by host immunological defences that are present throughout the colon, notably a mucus barrier. By contending for resources and colonisation sites, the preservation of a diversified and robust collection of good gut bacteria aids in the control of dangerous bacteria. The greatest strategy to maintain a balanced microbiota in the gut composition may be through dietary measures, notably the usage of a variety of fibres. Probiotics, which are live beneficial microbes, are one strategy that may help with health maintenance [12]. Throughout the past twenty years, there has been an increase in interest in how the gut microbiota affects human wellness, and this pattern is anticipated to persist. Currently, this field needs to address a few basic problems [13]. Furthermore, despite the fact that numerous investigations have clarified the role of the microbiome in the gut in disease and health and found associations with a number of human diseases, the majority of these studies still lack a clear understanding of the molecular processes and causal link. In addition, it is important to consider the implementation methods and security issues with gut microbiome therapies [14]. A great deal more study will undoubtedly be done as this subject advances in biotechnology and computational power.

The microbiome of the gut and a particular disease have written reviews, but comprehensive and quantitative assessment is still lacking in the field of studies on the gut microbiota and diseases as a whole. To give interested researchers a ready understanding and comprehensive perspective, it is crucial to present this research area. Bibliometrics is an intersection of disciplines that uses statistical and mathematical methods to analyse all forms of information [15]. The amount and variety of research papers can give insight into a field of study's advancement of knowledge and organisation. For recognising and visualising the accumulated understanding of science and developmental subtleties of scientific domains, bibliographic analysis is useful [16]. Numerous other disciplines, including information technology, life, and environmental science have made extensive use of bibliometrics. In order to give historical background, identify hot themes, and

identify developing themes in this field, we analyse the scientific ecosystem of the gut microbiota and disease using bibliometric techniques. Potential developmental pathways and difficulties in this subject are also highlighted [16, 17].

1.1. The Crucial Role of Gut Microbiota in Human Healt

The microbial community has a significant impact on the well-being of humans, and recent studies have shown that it is essential in avoiding the development of a wide range of disorders. The intestinal microbiota generates vital compounds, breaks down nutrients and toxins, and creates an array of helpful microbial compounds that stop harmful intruders from settling in the stomach by a number of intricate methods. Additionally, the microbial community is remarkably capable of converting the vitamins, minerals, and poisons produced by invasive species, which promotes host wellness and prevents diseases [18], as shown in Figure 1.1.

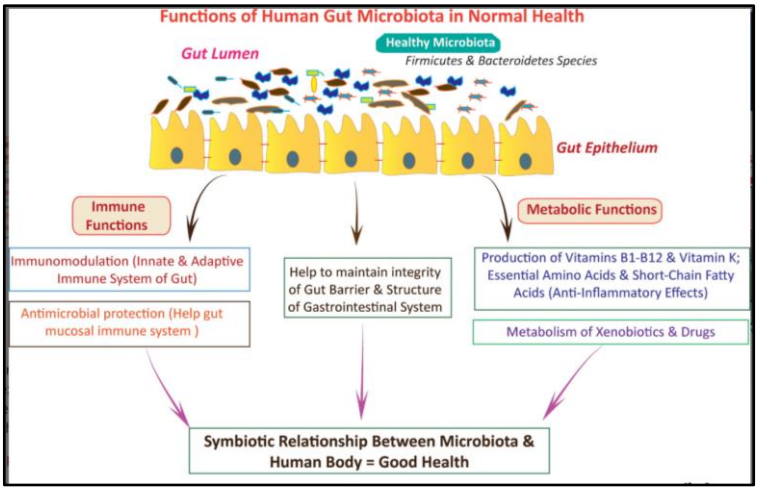


Figure 1.1: The corresponding diagram illustrates the perfect symbiotic interaction between a person's body and a robust population of healthy gut bacteria, highlighting how these two groups work together to sustain excellent health. Img Src: https://pubs.rsna.org/doi/full/10.1148/rg.2021200168

The microbiota of humans plays an important role in the development of the lining of the intestines and the immune system, in addition to protecting the host from infectious agents by creating compounds that are antimicrobial. The microbiota of the gut displays stability, toughness, and harmonious interactions with its human host when things are well. The notion of a "healthy" microbiota in the gut or its connection to the bodily functions of the host are the subject of extensive investigation [19]. The ability of the microbes in the gut to break down indigestible fibres

is one of its crucial functions. For a population that generates short chain fatty acids (SCFAs), these fibres constitute nutrients. Acetate, butyrate, and propionate constitute the three most significant SCFAs generated, with acetate being among the most prevalent. Colonocytes, the cells which comprise up the innermost layer of the human colon, use butyrate as a significant source of energy, and an absence of SCFA is thought to be a major contributor to numerous microbiome-related problems like intestinal leakage and localised inflammation [20]. The capacity of butyrate to cause the death of cancer cells in the colon and its stimulation of intestinal gluconeogenesis, both of which are important for the regulation of energy and obesity, are only two of the numerous significant roles it is currently associated with. Additionally, propionate plays a role in gluconeogenesis in the hepatocytes and satiety signalling, both of which are crucial for maintaining a healthy blood sugar level. The modulation of conversion in extra-intestinal tissues, especially the breakdown of cholesterol or lipogenesis, is another significant function of acetate [20, 21]. Ammonia, hydrogen sulphide and phenols are a few products produced by protein decomposition that can be harmful as well. There are numerous additional goods that should be included for their impact on health [22]. Bacteria generate a variety of compounds with biological activities, notably lipopolysaccharide (LPS), a part of gram-negative bacteria's cell wall which may induce inflammation in the tissues [23]. A lot of enteropathogenic microbes, such as certain types of *E. coli*, may generate toxin or induce diarrhoea given the correct circumstances, but normally, other beneficial bacteria that are not harmful surpass them and eventually force them out of the body [24]. With the formation of acetate, microbes like *Bifidobacterium* may also help in the prevention of infections that are pathogenic. Numerous microbe-produced enzymes have an impact on stomach and health. In fact, a large portion of the range of microbes in the gastrointestinal tract of humans could be attributed to the variety of enzymes produced by bacteria required to break down nutrients, especially the various types of complex carbohydrates which individuals consume [25].

1.2. Influential Factors Affecting the Variability of Gut Microbiome

The microbial community's diversity and depth increase over a number of phases of growth, starting in the early days of life and ending with a perceived stabilising after weaning. The changing makeup of the bacteria in the human gastrointestinal tract is influenced by a number of essential linked factors as well as uniqueness. These variables include neonates' delivery environments, age [26, 27], nutrition [28, 29], the host's genetic makeup [28, 30], consumption of

antibiotics [31], the physiological characteristics of the colonisation region [26], method of childbirth [32], and the kind of feed [10, 24]. Here is an outline of some of the crucial elements shown in Figure 1.2.

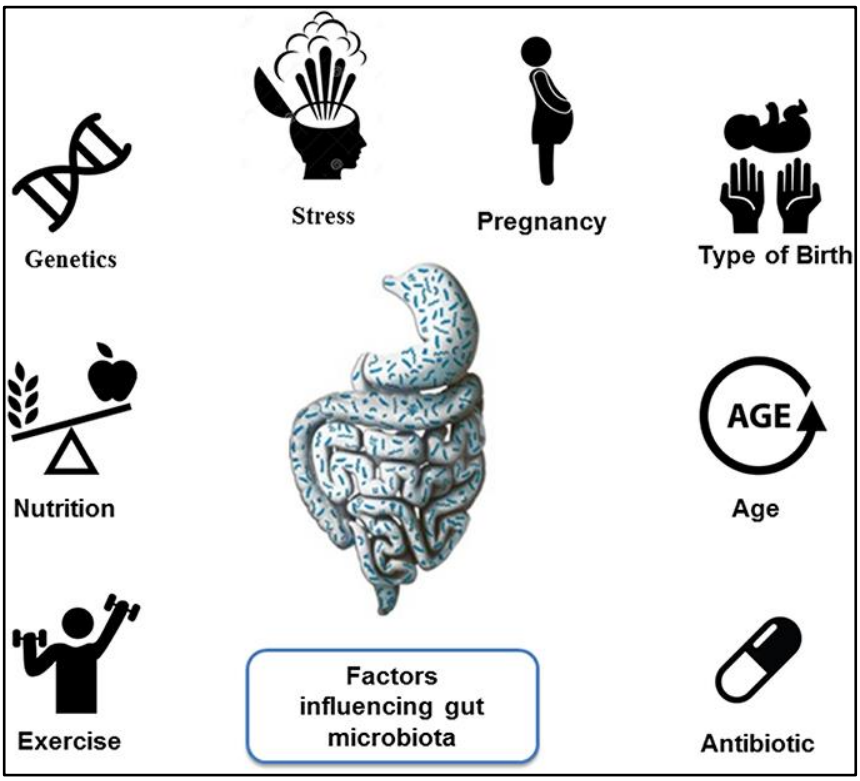


Figure 1.2: Influential Factors Modulating the Composition of Gut Bacterial Community [32].

1.2.1. Diet

It's possible that methodological restrictions that have been removed recently contributed to the underestimation of the function that food-ingested microbes contribute in the microbiome of the intestinal tract in the past [34]. High-calorie foods cause weight gain and Type 2 diabetes, according to a number of investigations conducted on mice and people [35, 36]. But more and more data points to the gut flora as the connection among nutrition and being obese [37]. It seems to be an especially reasonable subject to change given that nutrition is a significant contributing influence to the makeup of the intestinal microbiota. Research investigations have demonstrated that altering one's diet causes significant and quick changes in the composition of the intestinal microbiota [38, 39]. Investigations on rodents have shown which a diet high in fat (about sixty percent fat) reduces the variety in the microbiome of the gut as well as that the makeup of the

intestinal microbiome differs significantly among mice provided a high-fat diet and those provided an ordinary unpurified diet [40, 41].

The growth and maturation of both the local and systemic immune systems, as well as the induction of cytotoxic Th1 cells, motivation-specific cytokines which generate an appropriate microenvironment, and the increase in the amount of cells in the plasma generating IgA in newborns' intestinal environments, are all mediated by the microbes found within human milk. Certain *Lactobacillus* isolates activate natural killer (NK) cells, CD4+ T cells and CD8+ T cells, as well as the generation of Th1, cytokines and TNF-α [42-45].

The formation of the microbiota of the gut continues beyond childhood, and eating habits continue to play a crucial role in determining its composition, framework, and variety. *Roseburia, Ruminococcus* and *Eubacterium* have all been determined to predominate in the intestinal bacteria of vegetarians, which is healthy, diverse and known for this ability to metabolise insoluble carbs [46]. In contrast, those who are not following vegetarian diet/ following Western diet has been linked to an increase in *Bacteroides* and a decrease in *Firmicutes* [47]. Even in a brief span of time, nutrition can result in significant alterations [48]. When consuming a Western diet, the intestinal bacteria breaks down the amino acids, producing short-chain fatty acids as a source of energy and potentially toxic substances. This is prevented by the vegetarian diet, which encourages fermentation of carbohydrates as the microbiota's main activity [49].

1.2.2. Impact of Age and Delivery Method

The microbes in both the placenta and amniotic fluid signal the beginning of the bacterial intestinal colonisation process in pregnancy [50]. Amniotic fluid, Meconium and the umbilical cord all include bacteria and its byproducts, including DNA, according to investigations [51]. Although infants born via caesarean delivery originate their microbiota in the gut coming from their skin, resulting to the supremacy of *Propionibacterium*, *Corynebacterium* and *Streptococcus*, those conceived through the vagina have elementary gut microbes that is dominated by *Prevotella* and *Lactobacillus* that originate from their mother's genital microbiota [52, 53]. These fundamental microbiomes change with time to increase in diversity and stability. They resemble the intestinal flora of grownups by the time they reach the age of three. According to investigations, babies born

via caesarean section have a higher risk of acquiring diabetes and/or being obese than babies delivered vaginally [53].

1.2.3. Antibiotic

Taking antibiotics is a double-edged sword since it haphazardly kills both harmful and helpful microorganisms, leading to a decline of the microbiota in the gut, or referred to as dysbiosis and the expansion of unwelcome bacteria [54]. Antibiotics interfere with the mechanism of competitive isolation, which is a fundamental mechanism that helps bacteria in the gut eradicate pathogenic microorganisms [55]. This interference encourages the development of additional infections, including *Clostridium difficile* [56]. According to investigations, ciprofloxacin, clarithromycin with metronidazole and clindamycin [57] exert long-lasting effects that impact the microbiota's composition. The most effective medication for treating infections caused by *C. difficile* (CDI) is the antibiotic vancomycin but unlike other types of antibiotics, it alters the microbiota of the gut, which might result in repeated *C. difficile* infections (rCDI) or promote the development of pathogenic varieties of *E. coli* [58]. The majority of the microbiota in the gut, including Bacteroidetes, are similarly depleted by vancomycin medication, which is linked to reductions in Faecalibacterium, Fuminococcus and Bacteroidetes, and an upsurge in Proteobacteria species [59]. As with the impact of alterations in the microbiota of the gut prior to therapy, the particular impacts of the use of antibiotics on the intestinal flora and the duration that it takes for recovery to occur vary depending on the person being treated [60]. Low doses of antibiotics are frequently administered to animals to promote development and increased weight in numerous countries where agriculture, especially intensive raising of chicken and cattle, makes the consumption of medicines. The obesogenic effect of medicines in humans, even in modest dosages present in food, has been shown in several research on humans as well as rodents [54, 55]. Despite the frequent use of pesticides and various other substances on nutrition, there is now insufficient data to support their adverse impacts on intestinal health and the advantages of eating foods that are organic [61].

1.2.4. Genetic

The microbiome of the gut is ambient inherited at conception [62], thus it can serve as both an inherited characteristic which is formed by and engages with the human host and a component of the environment which influences the host's genetic makeup in shaping phenotypes [63]. The

intestinal microbiome is a desirable target for modification due to the fact that it can be altered for therapeutic purposes [64]. Understanding the relationships among the host's genetic factors and the gut microbiome will allow for the modification of the microbiota to be optimised for a particular host's genome to lower risk of illness [65]. The DNA composition of the individual's body influences the amount of particular microbes present in the microbiota of the gut in a manner that alters the body's metabolism and eventually may have an impact on wellness [62]. In comparison to individuals who are not related, close relatives have been shown to possess more comparable microbe populations, and monozygotic twins' intestinal microbiome appears more identical than that of twins who are dizygotic [62]. Despite the fact that particular immune-related genes are linked to bowel inflammation, there have yet no genome wide investigations which have characterised particular pathways and genes which influence the makeup of the microbiome in the gut [63, 64, 65].

1.3. Relation of Diet, Gut Microbiome and Cancer.

It is well known that microbes and tumours are related, and it is currently predicted that microbial agents may be responsible for as much as twenty percent of the world's cases of cancer [66]. For instance, the infections *Fusobacterium nucleatum*, Human papilloma virus (HPV), Epstein-Barr virus (EBV), as well as *Helicobacter pylori* are all linked to tumours [67-69]. The gut microbiome, which is made up of different symbiotic microbes, is also present in humans. By inhibiting the development of infectious agents, creating advantageous bacterial products, and metabolising nutrition and toxic substances the gut bacteria can have an impact on the well-being of humans. We have come a long way in the last ten years in knowing how cancer develops and how the microbiome affects relevant host systems. According to latest studies, a dietary fibre-gut microbiota pathway directly influences how well immunotherapy works for cancer patients [70]. Nevertheless, it is yet unclear how the connection between diet and microbiome affects other tumours or treatment options. In broad terms, people with cancer having wise food habits have decreased mortality rates across every kind of cancer.

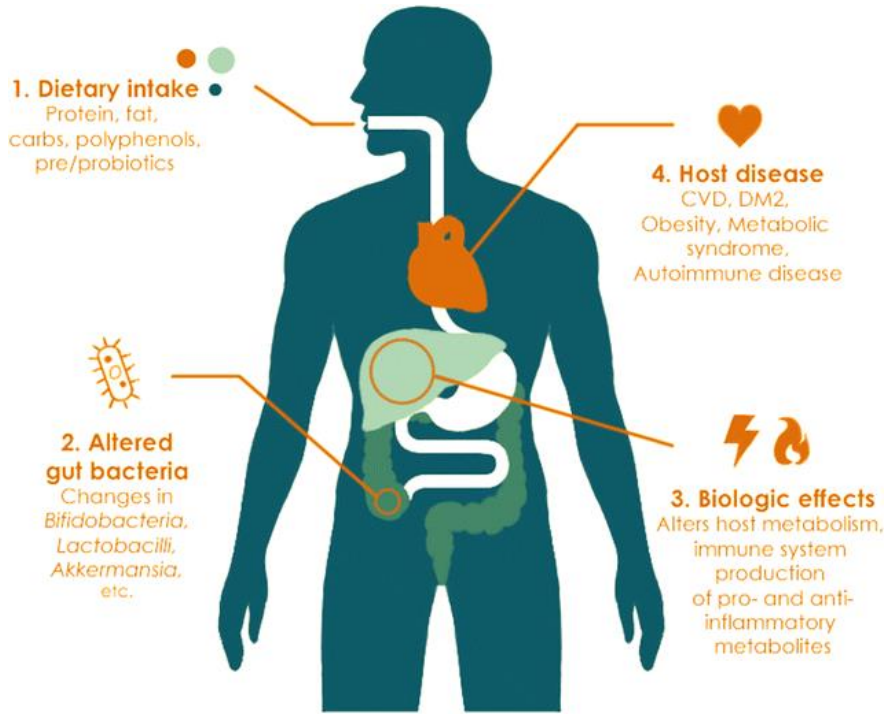


Figure 1.3: The Influence of Diet on the Gut Microbiome and Its Impact on Human Health [68]

Despite evidence that healthy eating habits increase cancer therapy response, difficulties with tumour load and adverse reactions of cancer therapy make it difficult to stick with such diet plans. Insufficient desire to eat, changed sensation of taste, trouble swallowing, nausea, issues with digestion, along with other patient related concerns [71] as well as the physical environment can all make it challenging to follow appropriate eating habits or receive nutritional support. [72]. Bacterial equilibrium is likely to be disrupted if one or more of various regulating systems malfunction, as might happen during the development of cancer. Investigation is currently ongoing to determine why the disruption of bacterial equilibrium arises from tumorigenesis or causes the entire procedure, although it most likely stems from combination. The microbial community may influence as well as be impacted by treatment for cancer, in addition to the consequences of cancer. The use of chemotherapy, radiation and cancer-specific treatments all have an immediate effect on the body's immune system, tissues that surround it as well as microbiota.

It has been suggested that eating habits may modify bacterial makeup and functioning prior to or throughout therapy to enhance outcomes for cancer in order to counteract the effects of bacterial

modifications throughout anticancer medication including afterward influence on response to therapy. The positive effects of immediate dietary modifications on the makeup and function of microorganisms in the mucosa of the gut are supported by a sizable body of research on nutrition and the microbiome of the gut in individuals who are healthy.

Owing to the wide variety in gut microbiomes among people, it must be understood that the positive effects of dietary treatments can be modest (for instance, Five Percent to Sixteen Percent) and diverse [73, 74]. Nevertheless, there have been only a few investigations which have sought to alter the bacteria in the intestines throughout cancer treatment using food. Particularly, investigations on malignancy in humans as well as animal models show that nutritional prebiotics may affect the response to treatment in a way that is beneficial by changing bacterial composition and immune system activity [75-79]. Furthermore, initial investigations in people and promising initial outcomes in experimental models of cancer support the use of the diet known as the ketogenic diet [80, 81]. Similarly, several reduced-calorie or intermittent fasting techniques have shown promise in altering the microbiome of the gut and halting the growth of malignancies in studies on animals [82]. Nevertheless, because of adverse effects from treatment including subsequent nonadherence of the assistance, human investigations in malignancy are difficult in terms of implementing the ketogenic diet as well as restricting calories. The precise nature of diet-microbiome relationships is lacking, particularly in the therapeutic stage of anticancer therapy, despite information regarding the effect of eating habits on cancer therapy consequences and encouraging results of particular dietary and bacterial variables on anticancer response to therapy [83]. This dearth of high-quality studies is probably a consequence of inadequate nutritional capturing technologies, poor research methodology, small sample sizes, especially the difficulty of conducting research with patients with cancer [84].

1.4. Impact of Intestinal Microbiota on Prevention of Cancer

Many investigations show that the treatment efficiency was reduced with no presence of the microbiota in the gut, indicating how beneficial microorganisms may affect the antitumor immune reactions brought on by the treatments by a variety of pathways. Human microbiome has been linked to the avoidance of numerous ailments, including numerous kinds of cancer, according to studies of association based on genomic sequencing investigations that evaluated the diversity of bacterial populations in medical cases to controls [85]. By fighting with infectious agents for

adhesion locations, they may have a secondary effect by reducing the number of infectious agents therefore preventing diseases that promote tumorigenesis. Some infectious agents, like the *Helicobacter pylori* and HPV, are well known for their ability to cause tumours, but recent investigation on the intestinal microbiome indicates that commensals and pathogenic opportunistic organisms may also be responsible for cancer-causing diseases, making the present estimate of fifteen to twenty percent more likely. *Fusobacterium nucleatum*, for instance, is more abundant in colorectal tumours than in healthy gastrointestinal tissue [86].

1.4.1. Enhancing the Mucosal Barrier

The goblet cells, that create the vast majority of the barrier of mucus that captures infectious agents as well as stops their spread across tissues, are responsible for shielding the gastrointestinal lining. This layer's significant function in the body's immune system is further demonstrated by the presence of numerous immune modulating substances there [87]. Smaller goblet cell sizes and weaker mucosa barrier compared to traditionally grown animals, healthy intestinal microbes are thought to be essential for producing mucus through their metabolites or impacts on the immune system as a whole [88]. The ability to alter mucus production or outflow by gastrointestinal goblet cells might be of notable value to symbiotic microbial communities. These adjustments strengthen the coating layer's ability to impede enteric bacteria. This protective approach might lessen malignancies brought on by pathogens [89].

1.4.2. Advancements in Antitumor Immunity

The microbial community may lessen the potency of cancer cells by altering anticancer responses. In this case, the NK cell-DC axis is activated in the tumour surroundings by a favourable microbiome through control of monocytes [88]. In accordance to research on bacteria that produce lactic acid, mice fed on *Bifidobacterium lactis*, *Lactobacillus acidophilus*, and *Lactobacillus rhamnosus* had an improved immune system compared to control mice due to a rise in abdominal peripheral blood leukocyte and macrophage phagocytic function. Additionally, NK cell cytotoxic capability was greater in the spleens of mice given these types of probiotics than in untreated mice. Thus, lactic acid microorganisms alter immunity through both innate and acquired mechanisms [89]. Additionally, it has been observed that mice inoculated with Eleven strains of bacteria found in normal human stool and capable of producing IFN-γ by CD8+ T cells exhibited significant

resistance to the growth of tumours. This power was linked to the intestinal microbiome and its impact on anticancer response [90].

1.4.3. Reduction in Inflammation

It is widely acknowledged that stress is a key player in the development of malignancy. Some beneficial bacteria have the ability to control carcinogenesis through anti-inflammatory properties processes. As an example, the probable producer of 7 SCFA, *E. coli*, induced anti-inflammatory properties, which in turn prevented the growth of tumours. IL-6, IL-8, TNF-α and IL-1β were all inhibited by SCFA and other compounds of *E. coli* in a similar manner [91]. These findings are in accordance with the idea that people who live in GF environments become more vulnerable to infections and the onset of diseases. After introduction to the virus that causes influenza, GF mice that had received a bacterial community transfection through a native kin and were kept in a lab for multiple generations displayed a substantial decrease in irritation [92, 93]. Actually, mediators of inflammation play a major role in widespread DNA damage, and *Lactobacillus johnsonii* markedly decreased the number of immune cell types like T cells and NK cells in addition to inflammation components while increasing antioxidant cytokines. As a result, it aided in the methodical and cellular removal of genetically toxic chemicals [94, 95].

1.4.4. Activation of Antitumor Signalling Pathways

According to current research, the helpful microbiota's anticancer effects most likely result from its stimulation of anticancer signalling. As an example, it has been shown that bacterial protein P8, which is generated from probiotics, may be used as a potential therapeutic for cancer of the colon [96]. The cell stage arrest in the stage of G2 caused by P8's antiproliferative properties was caused by the p53-p21 system. It's noteworthy to note that endogenous P8 production had a stronger antiproliferative effect than external therapy. *Lactobacillus acidophilus* was given orally to mice with colorectal cancer in order to reduce tumour development and increase the death of cells [97]. Disruption to the cancerous intestinal cells' membranes by cell-free *Lactobacillus* residues prevented the development of these cancerous cells [98, 99]. A pair of colorectal tumour cells were used in a different study to show that the supernatant without cells of isolated lactic acid-producing bacteria possesses antitumor characteristics [100].

However, the complicated relationships between the host organism, the microbes and antimetabolite, medications frequently used for the treatment of tumours, were clarified using genetic models made up of *E. coli* and *C. elegans* [101, 102]. The changed microbiome hinders platinum chemotherapeutic and CpG-oligonucleotide the immunotherapy procedure, according to a preliminary investigation on mice given a prescription antibiotic combination. In the presence of a healthy microbiome, tumour development is inhibited by treatment through CD8 T cell activation and myeloid cell release of TNF. On the contrary, immunotherapy in rodents inhibits cancer remission when administering antibiotics decreases cytokine and TNF production by immune system cells. These results imply that commensal bacteria trigger the production of cytokines that are inflammatory in reaction to immunotherapy mediated TLR4 stimulation, which enhances the outcome of patients [103]. In addition to directly reducing cancer risk, our beneficial and symbiotic microbes also prevent growth of pathogens. The following more direct path focuses primarily on being able to breakdown nutritional ingredients into biologically active compounds from food, that may result in both non-autonomous consequences aimed at immune system cells along with other stromal cells in the microenvironment of the tumour in addition to autonomous consequences on the tumour as well as cell-of-origin [104, 105].

1.5. The Influence of Intestinal Microbiota on Cancer Development

Parts of microbial pathogens are linked to the initiation and growth of tumours, in opposition to beneficial bacteria that can stimulate immune system cells to combat malignancy. Sethi et al. demonstrated that adaptive immune response was implicated in the tumour development inhibition caused by antibiotics utilising liver and subcutaneous metastases, types of cancer of the pancreas, and colon cancer [106]. By focusing on the harmful microbes, therapies for cancer may be more effective and immune system capabilities may be restored. Increasing proof suggests that specific microbiomes are intimately linked to the beginning and development of different kinds of malignancy, despite the fact that it has been demonstrated that the existence of host microbiota leads to mutation tolerance alongside virus tolerance [107]. Bacteria and their compounds have been predicted to cause or influence twenty percent of the worldwide incidence of malignancy. The bacteria *Helicobacter pylori* is known to be the most significant microorganism in the growth of cancer of the stomach [108]. The bacteria may influence the growth of malignancies by a variety

of methods, including induction of protumorigenic environments, the creation of irritation, suppression of immunity, genotoxin buildup and transmission of the tumour vulnerable characteristics. Typically, it is believed that the regional, ongoing inflammation that constitutes a few of the characteristics of malignancy, plays an additional part in how tumorigenesis operates [109, 110].

1.5.1. Immunosuppression

A periodontal microbe *F. nucleatum*, which is abundant in the stroma of many tumours [111], has the capacity to dampen the body's immune response primarily by inhibiting NK cells. *When F. nucleatum* is present, it was observed that the immune system cells became inactive due to the engagement of Fap2 protein with inhibitory receptors on NK cells. Additionally, the behavior of CD4+ memory T cells, which express TIGIT, was examined in the presence of *F. nucleatum*, resulting in a noticeable suppression of IFN-γ production. Consequently, by attaching to the immune cell's suppressive receptors, this bacteria suppresses the immunological response. It has been demonstrated that people with cancer of the pancreas have a larger microbiota than people whose pancreases are healthy.

Due to modification of immune system responses, removal of the aforementioned microbial community within the digestive tract slowed the growth of tumours. This reduction led to a rise in CD4+ T cell development into Th1, activation of CD8+ T cells, development of M1 macrophages, and a decrease in myeloid-derived suppressor cell infiltrate [112].

1.6. Harnessing the Gut Microbiome as a Tool for Evaluating Cancer Risk

Using the latest sequencing tools, researchers have examined every type of microbe found in the tumour environment [113, 114]. One new integrative use for the microbiome of the gut is the possibility of using it as a diagnostic biomarker as well as a target for therapy, even if the hallmark of an imbalance in malignancies has not yet been discovered [115, 116]. Consequently, a novel tool for better cancer care might involve the microbiota of one's gut. The microbial community continues to be the most researched since bacteria make up the majority of the intestinal microbiome [117].

1.6.1. Colorectal Cancer (CRC) Screening Biomarkers

For a localised CRC, the true five year rate of survival remains ninety percent; however, for metastatic cancer, it quickly drops to fifteen percent. consequently the need of an early on, non-intrusive and widely available CRC screening tool is crucial [118]. Numerous investigations conducted over the past few years using 16S RNA-DNA sequencing and whole-genome shotgun (WGS) technologies have additionally discovered various markers (mainly in faeces) that distinguish people with colorectal cancer from control participants. In general, the aforementioned data sets showed a worldwide change regarding the microbiota's makeup between those with colorectal cancer and control groups, particularly with a lower diversity of bacteria [119-122].

1.6.2. Biomarkers for Screening Other Types of Cancers

Additionally, new evidence points to dysbiosis of the intestinal microbiota as a possible passive tool for early identification of a number of malignancies. Even though an infection with *Helicobacter pylori* is linked to seventy percent of the cases of gastric cancer, this bacteria is not considered a useful screening sign. In fact, only one to four percent of those with *H. pylori* infection go on to acquire gastric cancer [123]. In contrast to combined normal controls, patients with breast cancer (BC) had a distinct microbiota in their intestines, with higher levels of Ruminococcaceae, Faecalibacterium and Clostridiaceae as well as decreased levels of Lachnospiraceae and Dorea. These findings support the idea that the microbiome of the gut might be utilised as an important biomarker in breast cancer research [124]. At this time, clinical practice has not yet adopted the use of microbiome-based biomarker at either the metabolomics or metagenomic level for cancer surveillance. To determine if these bacterial indicators may effectively identify people with a higher risk for developing cancer and enable prompt treatment, additional prospective studies are necessary [125, 126].

1.7. Role of Gut Microbiota in Precision Medicine

The genomic information of cells in humans as well as the microbiomes is increasingly readily available as inexpensive as previously. Understanding how information from the microbiome is incorporated into personalised healthcare methods for the avoidance, detection, and management of illnesses like cancer is a significant problem. Via the control of host immune responses, research has demonstrated that the gut microbiota can also successfully combat malignancy [127]. Because

microarray approaches may concurrently assess a number of malignancy-related genetics, they are useful in this context. In contrast to old personalised medical treatment, present personalised therapies integrate each unique genetic makeup and history of disease prior to the onset of the condition in question. Assessing genetic changes and the expression of genes levels in cancerous cells may result in efficient therapeutics because each tumour has a distinctive collection of genomic characteristics. The significance of cancer genomes in personalised therapy has received extensive coverage from the US National Institute of Health (NIH), and proteomic is another essential factor in personalised medical care [128]. The main processes through which the microbial community can influence tumorigenesis and hence be utilised to create antitumor medicines include inflammatory processes, metabolic processes, and genotoxicity [129]. The field of health care has recently been widened up by intriguing beneficial philosophies known as personalised medicine as well as the relationship between the intestinal microbiome as well as personalised medicine appears to represent one of the most fascinating areas of future study along with is regarded as a key viewpoint on the medical management of medical conditions like cancer.

A number of the important components of precision healthcare is the creation of testing methods that use biomarkers for the initial evaluation [130]. Investigators looked into the faecal microbiome's possibility of detecting the earliest signs of colorectal cancer and used it as a screening tool in various clinical categories consisting of healthy persons as well as those having carcinoma and adenoma. Colorectal Cancer constitutes one of the instances within which investigations were performed [131]. The relationship among populations of microbes and how they react to anticancer medication is complex, so additional research is required in order to comprehend whether microbes influence the body's immune response and tumour micro environments. Actually, selected microbial taxonomic decrease by methods like antibiotic exposure or additional stresses may lead to diminished immunotherapy responsiveness [132]. The existence of particular species could indicate the capacity to control the course of malignancy and its treatment, for instance, *E. coli* lowers the effectiveness of chemotherapeutic by metabolising as well as destroying the active component of the drug, that may result in a adverse relationship with tumour treatments [132]. Thus, bacteria is a future-oriented medicinal product as well as could potentially have an innovative medicinal role in this area. This enhances the probability of precision healthcare in connection to microbiome, with regard to prognosis and therapy

CHAPTER-2
LITERATURE REVIEW

2.1. DNA Data Bank of Japan

For researchers and scientists worldwide, the DNA Data Bank of Japan (DDBJ) represents an essential asset that performs a critical role in the realm of intestinal microbiome studies. Given its considerable influence on human wellness and illness, the microbiome of the intestines, which consists of billions of bacteria living in the intestinal tract of humans, has become an exciting field of research. It acts as a thorough archive for sequence information on DNA collected from different life forms, such as gut bacteria [133]. DDBJ gives researchers the ability to access a wealth of data necessary for comprehending the makeup, variety, and operational capabilities of our intestinal microbiome by gathering, arranging, as well as storing these DNA sequences.

An important component of our physiological science, the microbiome of the gut is now understood to have an impact on numerous facets of well-being, notably immunological function, nutrition metabolism, absorption, and even digestion. Several diseases including intestinal inflammation, weight gain, diabetes, and even psychological disorders have been connected to the makeup of the microbiota in the intestines. Deciphering the complex links among the intestinal microbiota and human wellness requires understanding the inherited characteristics of such microbes [133, 134].

Investigators looking into the genetic variety as well as prospective uses associated with these microbes can benefit greatly from DDBJ's huge database of gastrointestinal microbiome related information. Researchers may utilise the information in the repository to find and examine genetic material across gut bacteria in various groups, people in general [135]. This abundance of information enables empirical investigations, highlighting similarities as well as variations in the breakdown of the intestinal microbes, also illuminating the variables that influence the communities of microbes that live inside our intestines.Additionally, DDBJ encourages data exchange as well as teamwork between investigators, establishing a culture of information sharing and furthering intestinal microbiome studies as a whole. Each sequence's standardised description

as well as metadata are included in the database's contents, enabling thorough studies and fostering consistency [136]. Such attempts help build a solid comprehension of the function of the intestinal microbiome in both wellness and illness, which in turn paves the way for possible therapeutic approaches and individualised treatments which concentrate on the microbiota of the gut [135]. A screenshot of DDBJ website can be seen in Figure 2.1.

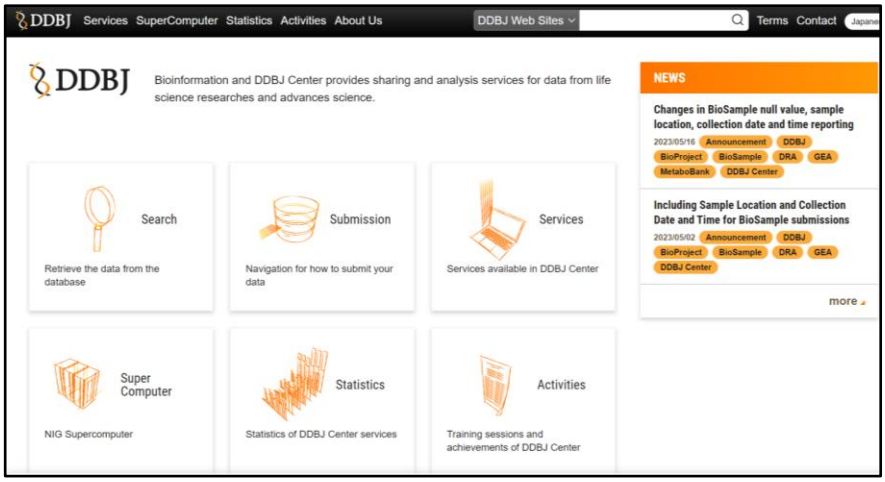


Figure 2.1: A screenshot of the DDBJ website. Img Src: https://www.ddbj.nig.ac.jp/index-e.html

2.2. EMBL -EBI (HLA data)- Allele Query Tool

A specialised library containing human the major histocompatibility complex , sequences is offered through the IPD-IMGT/HLA Database, which also contains the officially recognised sequences designated by the WHO Nomenclature Committee for HLA System [136]. The collection also includes details regarding the sequences' authentication in addition to extensive details about the raw materials that were used when the actual sequences were produced. In order to prevent the issues with changing released sequences as well as the misunderstanding caused by various identities for an identical sequence, it has become customary for researchers to upload the resulting sequences straight into the IPD-IMGT/HLA Database for curating content as well as the task of a proper title before being published [137]. Together alongside Julia Bodmer of the ICRF, Peter Parham of Stanford University and James Robinson a member of the HLA Informatics Group, construction pertaining to the HLA database was completed [138].

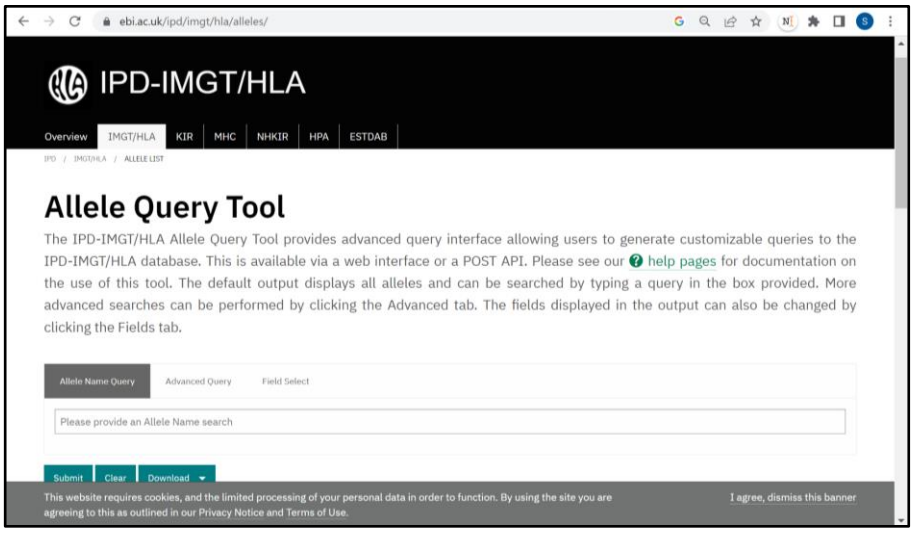


Figure 2.2: A screenshot of the Allele query tool. Img Src: https://www.ebi.ac.uk/ipd/imgt/hla/alleles/

2.3. 16s rRNA sequencing

The 16S rRNA gene, a mainstay in sequence based investigations for many years, is being used extensively in the study of microbial types. But new opportunities for reading the complete genome have emerged as a result of recent developments in the field of high-throughput sequencing. As a result, PCR-amplified 16S sequences are typically grouped together into operational taxonomic units (OTUs) based on how similar they are. The probable classification can then be deduced by comparing such OTUs against reference libraries employing sample OTU sequences [139, 140]. Though useful and effective, this use of 16S has demanded multiple presumptions, such as the now-historical notion that sequences with over ninety-five percent identity indicate the same genera while those with greater than ninety-seven percent similarity indicate the identical species [141]. Until recently, high-throughput sequencing systems were unable to provide precise, full-length 16S sequences, which are important for strain identification and isolation [142].

2.4. BLAST

A potent computational technique for locating areas of local similarity between sequences is called the Basic Local Alignment Search Tool (BLAST). It makes it easier to compare protein sequences or nucleic acid sequences with enormous databases of sequences and offers an evaluation of the statistical importance of the found matches [143]. In addition to assisting in the identification of nuclear family members, this tool may also be utilised to deduce operational and developmental

links among sequences. NCBI's WebBLAST offers four main search types for analysing nucleotide and protein sequences: BLASTn, BLASTx, tBLASTn, and BLASTp [143-147]. These searches are used to compare nucleotide query sequences to nucleotide or protein databases, translate nucleotide sequences to protein sequences, search for protein-coding regions in unannotated nucleotide sequences, and identify specific proteins.

2.5. Bacterial Biomarkers for Cancer

As of now, an extensive number of substances are currently used as biomarkers for cancer, comprising cell surface receptors, genetic material, messenger RNA, metabolites, enzymes, and gene transcription factors. Many kinds of bacteria have been demonstrated to play a role in the development and spread of cancers. In this regard, bacterial species have been associated with about sixteen percent of all cancer forms. But the precise methods by the way microbes influence the onset of malignancies are still undetermined, maybe as a result of the sheer intricacy of both microbial colonies and live cells. As an example, gastric malignancies and gastrointestinal and duodenal epithelium carcinoma are linked to the infection by *H. pylori* [148]. Research supporting the importance of the microbiota to human wellness is expanding, however there is little information available regarding the relationship among bacterial populations and different forms of cancers. In accordance with earlier research, some virulence variables that play a role in avoidance of the human immune response (*Streptococci*), inflammatory processes (*H. pylori* and *P. gingivalis*), incursion (*P. gingivalis and F. nucleatum*), colonising particular areas, as well as generating compounds have the ability of changing the division of cells [149]. Additionally, it has been demonstrated that certain varieties of bacteria cause the generation of microRNAs and DNA alterations. Adenoma in the early stages of CRC is being shown to have been caused by microbes like *E. coli*, *Streptococcus bovis*, *Enterococcus faecalis*, which in turn causes inflammatory bowel syndrome or IBS [150]. Inhibiting the process of apoptosis in intestinal epithelium cells is a side effect of some microbial intermediaries, like E. coli, which can induce persistent inflammation. Additionally, chronic colorectal cancer is brought on by metabolites produced by bacteria including extracellular radicals. Additionally, it was mentioned that the number of individuals of Eubacterium grows with colorectal cancer. Additionally, epigenetic modifications brought on by bacteria ought to get greater consideration [151].

Additionally, the intestinal epithelium becomes covered with biofilms, which promotes a pro-cancerous state. The evolution of medical practice and individualised treatment depends critically on the discovery of microbial signatures as biomarkers that are not invasive for the detection of elevated risks in malignancy. In order to better comprehend the relationship between microbes as well as tumours, it appears that cooperation between several domains of the field of epidemiology computational biology, immunology, microbiology and genetics is required [152]. It appears that using a quick, early, affordable, non-intrusive and comprehensive strategy is essential for getting precise outcomes. For prompt and non-intrusive screening for cancer, identifying prevalent species in kinds of cancer is beneficial. Additionally, research on the numerous methods utilised by microbes to induce tumorigenesis is necessary to counteract them [153].

2.6. HLA as Biomarkers for Cancer

The effectiveness of checkpoint related immunotherapy into a tiny percentage of patients with cancer has been amply demonstrated by new developments in immunotherapy for cancer. But as of yet, a reliable prognostic biomarker has yet to be found. Human leukocyte antigen (HLA), also known MHC, constitutes a highly variable chromosomal cluster made up of over two hundred genes. By differentiating between self and non self peptides, it plays a critical part in triggering an optimal human immune system reaction towards infections and tumour cells. Multiple lines of research have demonstrated that tumour cells' decreased levels of the gene expression of the class I HLA antigens based peptides complex is a method of tumour immune evasion and is frequently linked to a poor prognosis in patients with cancer [154]. It has additionally been demonstrated that tumour reactions to immunotherapy for cancer could be predicted by HLA class I and II expression of antigens in addition to flaws in the complicated antigen handling mechanism. However, there is ongoing discussion regarding the function of HLA to anticipate tumour responses to checkpoint related immunotherapy.

Based on the way it works, HLA is divided into three distinct categories: class I, II, and III. With the exception of some brain cells and germline line cells, class I molecules of HLA appear on the outermost layer of nucleated cells [155]. The HLA class III molecules' form and purpose aren't well understood. These participate in the processes of inflammation rather than antigen interaction.

Its genetic cluster, which is situated among those that belong to the class I and the class II molecules, produces key inflammatory compounds such as the complement subunits lymphotoxin, tumour necrosis factor (TNF)-f, actor B, C2 and C4, and proteins that undergo heat shock [157, 158, 159]. The nomenclature of HLA can be seen in Figure 2.3

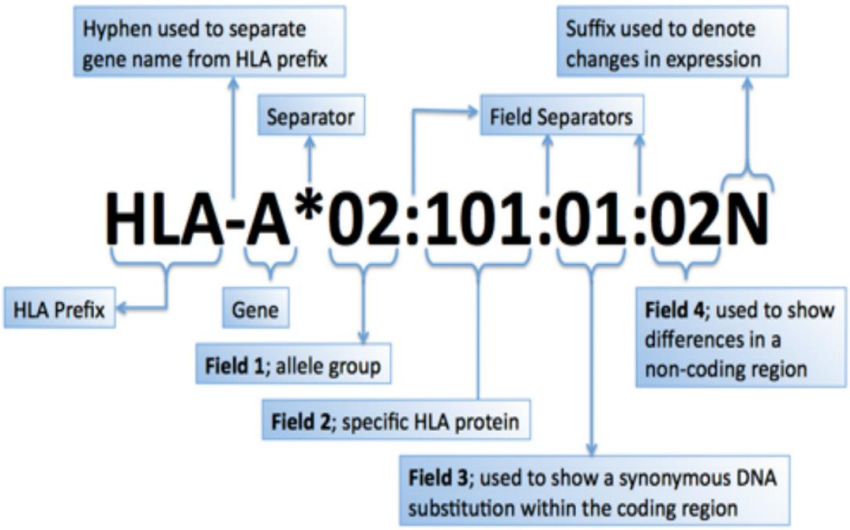


Figure 2.3: HLA Nomenclature. Img Src: https://hla.alleles.org/nomenclature/naming.html

CHAPTER-3

AIM

To make a comparative risk assessment database that incorporates representative data on dietary intake of different ethnicities, their cancer incidence, and estimate associations of diet and gut microbiome with cancer risk from prospective cohort study.

CHAPTER-4

OBJECTIVES

1. Data mining of gut microbiome and HLA profile of selected ethnic populations.
2. Process the metagenomic data into a more interpretable form.
3. Identify ethically disseminated, prevalent HLA and microbial biomarkers from the selected population.
4. Compare, contrast and analyse the data to draw the underlying correlation between gut microbiome and HLA profiles

CHAPTER-5

METHODOLOGY

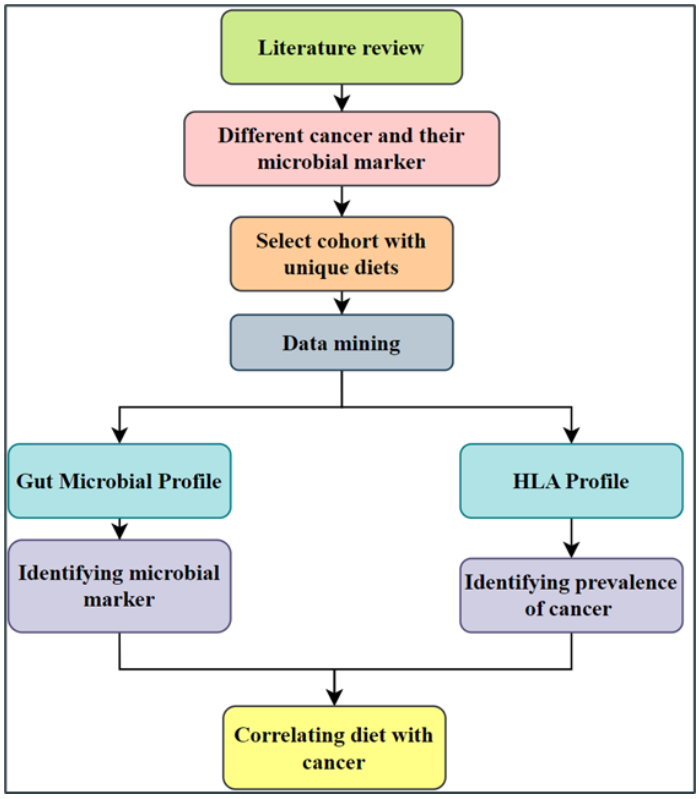


Figure 5.1: A basic flowchart of the steps of methodology. Img Src: Author

5.1. Literature Review

As the first step of any research ours was also to do thorough literature review.

- **Defining the research question:** Clearly articulated the research question or objective of the literature review. For example, "What is the current knowledge on the correlation between gut microbiome composition and HLA profiles in specific ethnic populations?"
- **Identifying relevant databases:** Determined the key academic databases, such as PubMed, Scopus, or Web of Science, where you will search for relevant literature. These databases provide access to a wide range of scientific articles, journals, and conference proceedings.
- **Developing search terms:** Created a list of keywords and search terms related to your research question. These terms should encompass concepts such as "gut microbiome,"

"HLA profiles," "ethnic populations," and any specific ethnic groups you are focusing on. Consider using synonyms, variations, and Boolean operators (AND, OR) to expand or narrow your search as needed.

- **Conducted the literature search:** Execute the search using the identified databases and search terms. Apply any necessary filters, such as publication date range or language, to refine your results. Retrieve relevant articles that potentially address the correlation between gut microbiome and HLA profiles in the selected ethnic populations.
- **Screening and selecting articles:** Reviewed the titles and abstracts of the retrieved articles to assess their relevance to your research question. Excluded articles that are clearly unrelated or do not meet your inclusion criteria. Obtain full-text copies of the remaining articles for detailed evaluation.
- **Assessing article quality:** Evaluated the quality and reliability of the selected articles. Consider factors such as study design, sample size, data collection methods, and statistical analyses employed. This step helps ensure that you include high-quality studies in your literature review.
- **Extracting relevant data:** Extracted key information and findings from the selected articles. This may include details on the study population, methodology, results, and conclusions related to the correlation between gut microbiome and HLA profiles. Use a systematic approach, such as a data extraction form, to organize and record the extracted data.

A literature review is an iterative process that involves careful selection, analysis, and synthesis of relevant literature. It provides a comprehensive understanding of existing knowledge and forms the foundation for subsequent research or analysis. This step was important to identifying the HLA and bacterial Biomarkers as well as the databases required to obtain the data on gut microbiome and HLA profiles of the selected cohorts.

5.2. Identifying Bacterial Biomarkers for GI Tract Cancer through Literature Mining

I. Identification of Relevant Literature:

i. Conducted an extensive search using scientific databases such as PubMed, Google Scholar, and Scopus.
ii. Used relevant keywords and search terms, including "GI tract cancer," "bacterial biomarkers," "gut microbiome," and "microbial dysbiosis."
iii. Limited the search to recent studies published within the last five years to focus on the most up-to-date research.

II. Literature Screening and Selection:

i. Screened the retrieved articles based on their titles and abstracts to identify potentially relevant studies.
ii. Prioritised studies that focused specifically on GI tract cancers, such as colorectal cancer, gastric cancer, or esophageal cancer.
iii. Excluded studies that did not investigate bacterial biomarkers or were not relevant to the topic of interest.
iv. Retrieved and saved the full-text versions of the selected articles for further analysis.

III. Data Extraction:

i. Carefully read the full-text articles and extracted relevant information related to bacterial biomarkers for the specific types of GI tract cancers mentioned in your list.
ii. Identified bacterial biomarkers at the genus level mentioned in the studies.
iii. Compiled the extracted data into a structured format, including the type of cancer and the identified bacterial genera.

A list of bacterial biomarkers is given in Table 5.1.

Table 5.1: Bacterial biomarkers (Genus level) of GI tract related cancers.

Serial No.	Bacterial Genus	Associated GI Tract Cancers	References
1	Fusobacterium	Colorectal Cancer	[160]
2	Helicobacter	Gastric Cancer	[161]
3	Bacteroides	Colorectal Cancer	[162]
4	Streptococcus	Esophageal Cancer	[163].
5	Escherichia	Colorectal Cancer	[164]
6	Lactobacillus	Gastric Cancer, Colorectal Cancer	[165]
7	Clostridium	Colorectal Cancer	[166]
8	Peptostreptococcus	Colorectal Cancer	[167]
9	Porphyromonas	Oral Cancer	[168]
10	Prevotella	Colorectal Cancer	[169]
11	Campylobacter	Gastric Cancer	[170]
12	Treponema	Esophageal Cancer	[171]
13	Enterococcus	Colorectal Cancer	[172]
14	Pseudomonas	Gastric Cancer	[173]
15	Veillonella	Esophageal Cancer	[174].
16	Actinomyces	Colorectal Cancer	[175]

17	Serratia	Gastric Cancer	[176]
18	Tannerella	Colorectal Cancer	[177]
19	Staphylococcus	Gastric Cancer	[178].
20	Campylobacter	Colorectal Cancer	[179]

5.3. Identifying HLA Biomarkers for GI Tract Cancer through Literature Mining

I. Literature Search:

i. Conducted an extensive search using scientific databases like PubMed, Google Scholar, and Scopus.

ii. Employed relevant keywords and search terms such as "GI tract cancer," "HLA biomarkers," "HLA typing," and "immunogenetics."

iii. Focused on recent studies published within the last five years to include the latest research in the field.

II. Literature Screening and Selection:

i. Screened the retrieved articles based on titles and abstracts to identify relevant studies.

ii. Prioritised studies specifically investigating GI tract cancers, including colorectal cancer, gastric cancer, or esophageal cancer.

iii. Excluded studies that did not focus on HLA biomarkers or were not directly related to the research question.

iv. Retrieved and saved full-text versions of the selected articles for detailed analysis.

III. Data Extraction:

i. Thoroughly read the full-text articles and extracted relevant information related to HLA biomarkers for the specific types of GI tract cancers mentioned in your list.

ii. Identified specific HLA alleles, haplotypes, or polymorphisms reported as potential biomarkers.
iii. Compiled the extracted data, including cancer type and the identified HLA biomarkers.

By following these adjusted methodologies, the focus is on identifying relevant literature and extracting information related to bacterial and HLA biomarkers for GI tract cancer, considering that the bacterial biomarkers identified are at the genus level. A list of HLA biomarker is given in Table 5.2.

Table 5.2: HLA haplotype biomarkers of GI tract related cancers.

Serial No.	**HLA Biomarker**	**Associated GI Tract Cancers**	**References**
1	HLA-A*02:01	Gastric Cancer, Colorectal Cancer	[180]
2	HLA-A*31:01	Gastric Cancer, Colorectal Cancer	[181].
3	HLA-A*11:01	Gastric Cancer, Colorectal Cancer	[182]
4	HLA-B*44:02	Gastric Cancer, Colorectal Cancer	[183].
5	HLA-B*07:02	Gastric Cancer, Colorectal Cancer	[184]
6	HLA-B*51:01	Gastric Cancer, Colorectal Cancer	[185]
7	HLA-C*07:02	Gastric Cancer, Colorectal Cancer	[186]
8	HLA-DRB1*13:02	Gastric Cancer, Colorectal Cancer	[187]

9	HLA-DRB1*07:01	Gastric Cancer, Colorectal Cancer	[188]
10	HLA-DRB1*03:01	Gastric Cancer, Esophageal Cancer	[189].
11	HLA-DRB1*15:01	Gastric Cancer, Esophageal Cancer	[190]
12	HLA-DRB1*04:05	Gastric Cancer, Esophageal Cancer	[191]
13	HLA-DRB1*13:01	Gastric Cancer, Esophageal Cancer	[192]
14	HLA-DRB1*09:01	Gastric Cancer, Esophageal Cancer	[193].
15	HLA-DRB1*11:04	Gastric Cancer, Esophageal Cancer	[194]
16	HLA-DRB1*15:02	Gastric Cancer, Esophageal Cancer	[195]
17	HLA-DRB1*04:01	Gastric Cancer, Esophageal Cancer	[196]
18	HLA-DRB1*07:01	Gastric Cancer, Esophageal Cancer	[197]
19	HLA-DRB1*14:01	Gastric Cancer, Esophageal Cancer	[198]
20	HLA-DRB1*01:01	Gastric Cancer, Esophageal Cancer	[199]
21	HLA-DRB1*08:03	Gastric Cancer, Esophageal Cancer	[200]
22	HLA-DRB1*16:02	Gastric Cancer, Esophageal Cancer	[201]

23	HLA-DRB1*10:01	Gastric Cancer, Esophageal Cancer	[202].
24	HLA-DRB1*12:02	Gastric Cancer, Esophageal Cancer	[203].
25	HLA-DRB1*08:01	Gastric Cancer, Esophageal Cancer	[204]
26	HLA-DRB1*11:03	Gastric Cancer, Esophageal Cancer	[205]
27	HLA-DRB1*11:03	Gastric Cancer, Esophageal Cancer	[205].
28	HLA-DRB1*08:02	Gastric Cancer, Esophageal Cancer	[206].
29	HLA-DRB1*04:04	Gastric Cancer, Esophageal Cancer	[207]
30	HLA-DRB1*08:04	Gastric Cancer, Esophageal Cancer	[208]

5.4. Selecting Appropriate Cohorts for Studying Gut Microbiome and Cancer Risk

I. Identification of Available Data:

i. Identified available datasets containing gut microbiome and cancer-related information.

ii. Explored various sources, including public repositories and published studies.

iii. Prioritised datasets with comprehensive microbiome profiling and information on cancer outcomes.

II. Assessment of Data Availability:

i. Assessed the availability of relevant data for different populations and cohorts.

ii. Identified cohorts with data on both gut microbiome composition and cancer outcomes.

iii. Focused on cohorts that had sufficient sample sizes and comprehensive data coverage.

III. Evaluation of Cohort Suitability:

i. Evaluated the suitability of different cohorts based on the research question and study objectives.
ii. Considered factors such as cohort demographics, geographical location, disease prevalence and type of diet.
iii. Assessed the representativeness of cohorts for the target population of interest.

IV. Selection of Japanese Population Cohort:

i. Identified a cohort within the Japanese population that met the data availability criteria.
ii. Ensured the cohort had data on gut microbiome composition and a comprehensive range of cancer types.
iii. Consider the relevance of the cohort to the research question and its potential contributions to the scientific understanding of gut microbiome and cancer risk in the Japanese population.

V. Selection of Spanish Population Cohort:

i. Identified a cohort within the Spanish population that met the data availability criteria.
ii. Ensured the cohort had data on gut microbiome composition and a comprehensive range of cancer types.
iii. Considered the relevance of the cohort to the research question and its potential contributions to the scientific understanding of gut microbiome and cancer risk in the Spanish population.

VI. Data Collection and Harmonization:

i. Obtained the necessary data from the selected cohorts.
ii. Ensured compatibility and harmonisation of data across different cohorts.
iii. Standardised the microbiome profiling techniques and cancer outcome measurements, if necessary, to enable meaningful comparisons between the Japanese and Spanish populations.

By following this methodology, appropriate cohorts for studying the gut microbiome's impact on cancer risk were selected based on data availability. The Japanese population cohort and the Spanish population cohort were chosen, considering their availability of comprehensive gut microbiome and cancer-related data. These cohorts provide valuable insights into the relationship between the gut microbiome and cancer risk in their respective populations.

5.5. Data Mining of Gut Microbiome Bacterial Composition in Healthy Individuals of the Japanese Population

The objective of this study is to investigate the bacterial composition of the gut microbiome in healthy individuals from the Japanese population. To achieve this, we employed a data mining approach, which involved conducting a manual search and reviewing relevant research papers. This led us to the DDBJ (DNA Data Bank of Japan), a Japanese DNA database, where we discovered 16S rRNA sequencing data in the form of metagenomic sequences (sequence reads). The subsequent steps focused on processing and analysing these sequences to identify and characterise the bacterial taxa present in the gut microbiome.

I. Data Collection:

i) Manual Search: An extensive literature search was conducted using scientific databases such as PubMed, Google Scholar, and other relevant sources. The search keywords used included "gut microbiome," "bacterial composition," "healthy individuals," and "Japanese population." The primary focus was on identifying research articles that investigated the gut microbiome composition in healthy individuals of the Japanese population.

ii) DDBJ Database: Based on the information obtained from the literature search, the DDBJ database (https://www.ddbj.nig.ac.jp/) was accessed to obtain the 16S rRNA sequencing data for our study. The DDBJ database is renowned for hosting a comprehensive collection of DNA sequences, including metagenomic data from various research studies.

II. Data Retrieval:

Using relevant search terms such as "gut microbiome," "16S rRNA sequencing," and "Japanese population," we extracted the appropriate metagenomic sequencing datasets available in the DDBJ database. Our focus was specifically on 16S rRNA gene sequences, as they are widely utilised for identifying and classifying bacterial taxa within the gut microbiome.

III. Data Acquisition:

At this stage, we have successfully retrieved the 16S rRNA sequencing data from the DDBJ database for our study. The subsequent steps involve processing and analyzing the sequences to identify and characterise the bacterial taxa present in the gut microbiome of healthy individuals in the Japanese population. The detailed procedures for data processing, quality control, sequence alignment, taxonomic assignment, data analysis, interpretation, and ethical considerations will be performed following the data retrieval stage.

5.6. Data Mining of Gut Microbiome Bacterial Composition in Healthy Individuals of the Spanish Population

The objective of this study is to investigate the bacterial composition of the gut microbiome in healthy individuals from the Spanish population. To accomplish this, a data mining approach was employed, involving a manual search and a request for data from the authors of a specific research paper.

I. Data Collection:

i. Manual Search: A comprehensive literature search was conducted using scientific databases such as PubMed, Google Scholar, and other relevant sources. The search keywords utilised included "gut microbiome," "bacterial composition," "healthy individuals," and "Spanish population." The primary focus was to identify research articles that investigated the gut microbiome composition in healthy individuals from the Spanish population.

ii. Identified Research Paper: Among the various research papers reviewed, a specific research paper titled "The Spanish gut microbiome reveals links between microorganisms

and Mediterranean diet" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8580991/) was found. This paper provided valuable insights into the gut microbiome of healthy individuals from the Spanish population. However, the data required for analysis, including the 16S rRNA sequencing data, was not freely available.

iii. Data Request: To obtain the necessary data, a formal request was made to the authors of the research paper. An email was sent to the corresponding author, expressing our interest in their study and explaining the purpose of our research. The authors were requested to provide access to the 16S rRNA sequencing data associated with the gut microbiome analysis of healthy individuals from the Spanish population.

II. Data Acquisition:

Upon receiving a positive response from the authors, the necessary data, including the 16S rRNA sequencing data (sequence reads), was acquired.

The methodology involved a manual search using relevant keywords to identify a research paper titled "The Spanish gut microbiome reveals links between microorganisms and Mediterranean diet." As the data was not freely available, a request was made to the authors to obtain the 16S rRNA sequencing data associated with the gut microbiome analysis of healthy individuals from the Spanish population.

5.7. Processing the obtained metagenomic data to obtain a gut microbiome profiles of Japanese and Spanish population

I. Retrieve Reference Sequences:

i. The reference sequences for the gut microbiome analysis were obtained from the "16S ribosomal RNA sequences (Bacteria and Archaea)" database on the NCBI website.

II. Prepare Query Sequences:

i. The query sequences were obtained from 16S rRNA sequencing data of a gut microbiome sample.

ii. The raw sequencing data was preprocessed and filtered to obtain high-quality reads specific to the target region.

III. Perform BLAST Alignment on NCBI Website:

i. The query sequences, totaling 10,000 reads, were aligned against the reference sequences using the online version of BLAST on the NCBI website.
ii. The nucleotide BLAST tool was selected, and the "16S ribosomal RNA sequences (Bacteria and Archaea)" database was chosen as the target database.
iii. The alignment parameters were set to: e-value threshold of 0.001, the default scoring matrix (Blosum62), and no filtering options were applied.

IV. Analyse Alignment Results:

i. The BLAST search was performed, and the alignment results were obtained from the NCBI website.
ii. The alignment process took several days due to the large number of query sequences and the size of the reference database.
iii. The alignment results were saved and downloaded for further analysis.

V. Taxonomic Classification using BLAST Results:

i. The taxonomic information associated with each hit in the alignment results was extracted.
ii. A confidence threshold of 97% sequence identity and a minimum alignment length of 200 base pairs were applied to ensure reliable taxonomic assignments.
iii. Taxonomic labels were assigned to each read based on the best hit and associated taxonomic information.
iv. The taxonomic labels were assigned at various levels, including phylum, class, order, family, genus, and species.

VI. Complete Gut Microbiome Profile:

i. The above steps were repeated for all the reads.

ii. The results from each read were combined to obtain a complete gut microbiome profile for the given sample sequences.

iii. The resulting taxonomic assignments at different levels were used to analyse the diversity, richness, and composition of the gut microbiome.

By following this methodology and performing the steps multiple times, aligning and classifying a total of 24071 reads, a comprehensive gut microbiome profile was obtained from the data of the given sample sequences. A screenshot example of the database can be seen in Figure 5.2.

Read ID/S.no.	Phylum	Class	Order	Family	Genus	Species
1	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Faecalibacterium	prausnitzii
2	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Blautia	obeum
3	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	fragilis
4	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Roseburia	hominis
5	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	ovatus
6	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	aeruginosa
7	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	thetaiotaomicron
8	Actinobacteria	Actinobacteria	Actinomycetales	Coriobacteriaceae	Eggerthella	lenta
9	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Dorea	formicigenerans
10	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	histicola
11	Actinobacteria	Actinobacteria	Actinomycetales	Bifidobacteriaceae	Collinsella	aerofaciens
12	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus	bromii
13	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Alistipes	putredinis
14	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Ensifer	meliloti
15	Actinobacteria	Actinobacteria	Actinomycetales	Coriobacteriaceae	Adlercreutzia	equolifaciens
16	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Anaerostipes	hadrus
17	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Klebsiella	pneumoniae
18	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Barnesiella	intestinihominis
19	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Faecalibacterium	prausnitzii

(A)

Read ID/S.no	Phylum	Class	Order	Family	Genus	Species
1	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Roseburia	hominis
2	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Parabacteroides	distasonis
3	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	leguminosarum
4	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	copri
5	Actinobacteria	Actinobacteria	Bifidobacteriales	Bifidobacteriaceae	Bifidobacterium	longum
6	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	fragilis
7	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Blautia	obeum
8	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	melaninogenica
9	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Faecalibacterium	prausnitzii
10	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	uniformis
11	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus	obeum
12	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	copri
13	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Roseburia	hominis
14	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	vulgatus
15	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	copri
16	Actinobacteria	Actinobacteria	Bifidobacteriales	Bifidobacteriaceae	Bifidobacterium	longum
17	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	leguminosarum
18	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Roseburia	Roseburia sp.
19	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia	coli

(B)

Figure 5.2: Screenshots of the gut microbiome profile databases. (A) Japanese Gut Microbiome Profile (B) Spanish Gut Microbiome Profile. Img Src: Author.

5.8. Data mining of complete HLA profiles for the Japanese and Spanish populations using the EMBL-EBI Allele Query Tool

I. Accessed the EMBL-EBI Allele Query Tool:

i. Opened a web browser and visited the EMBL-EBI website (https://www.ebi.ac.uk/Tools/allele/).

ii. Located the "Allele Query Tool" on the EMBL-EBI website's menu and clicked on the link to access the tool.

II. Specified the Target Populations:

i. Identified the target populations for analysis: the Japanese and Spanish populations.

ii. Ensured that the EMBL-EBI Allele Query Tool supported both populations.

III. Configured Query Parameters:

i. Explored the query parameters within the Allele Query Tool to refine the search.

ii. Selected the target populations as Japanese and Spanish.

iii. Adjusted other query parameters, such as HLA class (Class I or Class II) and specific loci (e.g., HLA-A, HLA-B, HLA-DRB1) to retrieve complete HLA profiles.

IV. Submitted the Query:

i. Double-checked the entered information, query parameters, and target population settings for accuracy.

ii. Clicked the "Submit" button within the Allele Query Tool to initiate the query process.

V. Retrieved and Reviewed the Query Results:

i. Waited for the Allele Query Tool to process the query and retrieve the results.

ii. Once the results were available, reviewed the information presented in the output.

VI. Extracted the HLA Profile Data:

i. Extracted the complete HLA profile data specific to the Japanese and Spanish populations from the query results.

ii. Collected information such as HLA allele combinations, haplotype frequencies, or any other relevant data provided by the Allele Query Tool.

VII. Documented and Analysed the Data:

i. Recorded the extracted complete HLA profile data in a structured format for further analysis.

ii. Analysed the data to identify patterns, associations, or unique features within the HLA profiles of the Japanese and Spanish populations.

iii. Compared the results between the two populations and with other populations or existing literature to gain insights into the diversity and characteristics of HLA profiles.

By following this updated methodology, data mining of complete HLA profiles for the Japanese and Spanish populations was performed using the EMBL-EBI Allele Query Tool. The tool was used to configure query parameters, submit the query, retrieve the results, and extract relevant data for analysis. The collected data was then documented, analysed, and compared to gain insights into the HLA profiles of both populations. A screenshot example of the database can be seen in Figure 5.3.

Name	IPD Accession	Locus
A*24:66	HLA02613	A*
A*31:161	HLA24680	A*
A*24:02:50	HLA06886	A*
A*33:08	HLA02250	A*
A*02:292	HLA06079	A*
A*26:140	HLA16438	A*
A*02:695	HLA17586	A*
A*24:516	HLA29263	A*
A*31:30	HLA04746	A*
A*33:03:42	HLA26095	A*
A*24:02:34	HLA05248	A*
A*31:66	HLA08093	A*
A*24:68	HLA02725	A*
A*02:355	HLA07747	A*
A*31:24	HLA03570	A*
A*26:64	HLA06279	A*
A*11:164	HLA10656	A*
A*11:144	HLA09484	A*
A*26:79	HLA08598	A*

(A)

Name	IPD Accession	Locus
A*02:113:01N	HLA02867	A*
A*24:40N	HLA01821	A*
A*03:01:01:112	HLA33544	A*
A*24:516	HLA29263	A*
A*30:02:06	HLA08037	A*
A*03:132	HLA07040	A*
A*02:156	HLA03582	A*
A*68:105:01:01	HLA10122	A*
A*03:20	HLA02325	A*
A*03:143	HLA07803	A*
A*24:587	HLA36811	A*
A*31:01:03	HLA02945	A*
A*29:41	HLA08110	A*
A*32:01:01:26	HLA28854	A*
A*02:270	HLA05549	A*
A*33:69	HLA09521	A*
A*24:215	HLA08602	A*
A*02:30:01	HLA00036	A*
A*24:39	HLA01773	A*

(B)

Figure 5.3: Screenshots of the gut HLA profile databases. (A) Japanese HLA Profile (B) Spanish HLA Profile. Img Src: Author.

5.9. Analysis of Data and Reporting of observations

I. Relative Abundance Analysis:

i. The processed metagenomic data of the gut microbiome for the Japanese and Spanish populations were obtained, including the taxonomic classification at the genus level.

ii. The relative abundance of bacterial genera was calculated by summing the abundance values of all genera within each population and normalizing the values to obtain proportions or percentages representing the relative abundance of each genus.

iii. Separate matrices or tables were generated to represent the relative abundance of bacterial genera for individuals in the Japanese and Spanish populations.

The formula to calculate the relative abundance of a bacterial genus in the gut microbiome is:

Relative Abundance = (Number of reads assigned to the bacterial genus / Total number of reads in the sample) × 100

II. Biomarker Analysis:

i. Bacterial genus biomarkers associated with cancer were identified by referring to relevant scientific literature and established databases.
ii. The relative abundance of these bacterial genera was compared between the Japanese and Spanish populations.
iii. Analysis of the relative abundance data was performed to identify any significant differences in the prevalence or abundance of these bacterial genus biomarkers between the Japanese and Spanish populations.
iv. The relative abundance trends of these biomarkers were assessed to identify any notable variations between the populations.

III. HLA Biomarker Analysis:

i. HLA profile data for individuals in the Japanese and Spanish populations were collected.
ii. The presence or absence of specific HLA biomarkers associated with cancer was determined within each population.
iii. The distribution of HLA biomarkers between the Japanese and Spanish populations was compared.
iv. Examination of the data was done to identify any notable differences in the prevalence or absence of these HLA biomarkers between the Japanese and Spanish populations.

IV. Correlation Analysis:

i. Correlation analysis was performed to evaluate the relationship between the relative abundance of bacterial genera and the presence or absence of HLA biomarkers within each population separately.
ii. The data was analysed to identify any significant associations between specific bacterial genera and HLA biomarkers within each population.
iii. Consideration was given to the magnitude and direction of the correlations to determine if there were consistent patterns between bacterial genera and HLA biomarkers in the Japanese and Spanish populations.

V. Interpretation and Reporting:

i. The results of the biomarker analysis and correlation analysis for the gut microbiome and HLA profiles in the Japanese and Spanish populations were summarized.
ii. The significance of the identified bacterial genus biomarkers and HLA biomarkers associated with cancer was discussed.
iii. An interpretation was provided regarding the potential implications of the relative abundance of bacterial genera and the presence or absence of HLA biomarkers within each population.
iv. The relevant biomarkers and their potential implications in the context of gut microbiome, HLA genetics, and the population-specific variations were reported.

By following this methodology, the relative abundance of bacterial genera, bacterial genus biomarkers for cancer, and HLA biomarkers were analysed to assess the potential associations and variations between the Japanese and Spanish populations. The findings provided insights into the relationships and correlations between the gut microbiome, HLA genetics, and the risk of cancer within each population.

CHAPTER-6

RESULTS

This study aimed to analyse the gut microbiome and HLA profiles of selected ethnic populations through data mining. The objectives were to process the metagenomic data, identify prevalent microbial and HLA biomarkers, and explore correlations between the gut microbiome and HLA profiles.

6.1. Analysis of Gut Microbiome Profiles

Metagenomic data of the gut microbiome and HLA profiles were collected from the Japanese (from DDBJ) and Spanish (from literature) populations. The data underwent processing to ensure reliability, consistency and readability. The processed metagenomic data were analysed using bioinformatics tool BLAST. Bacterial genera were taxonomically classified, and relative abundance analysis was performed to determine their prevalence within each population. Data visualisation techniques aided in presenting the relative abundance patterns effectively.

Table 6.1: Relative Abundance of different bacterial genus in the gut of healthy Japanese Individuals. Please note that the values provided in the table are based on the total number of reads of 11,688. Source: Author

Serial No.	Genus	Number of reads	Relative Abundance
1	Bacteroides	3179	27.20%
2	Prevotella	1753	15%
3	Faecalibacterium	1508	12.90%
4	Blautia	901	7.70%
5	Ruminococcus	877	7.50%
6	Alistipes	456	3.90%

7	Eubacterium	304	2.60%
8	Roseburia	292	2.50%
9	Fusobacterium	187	1.60%
10	Akkermansia	608	5.20%
11	Lactobacillus	445	3.80%
12	Helicobacter	292	2.50%
13	Clostridium	222	1.90%
14	Enterococcus	175	1.50%
15	Streptococcus	152	1.30%
16	Parabacteroides	129	1.10%
17	Other genus	211	1.80%

The analysis of gut bacteria composition in the Japanese population revealed the following relative abundances of different genera: Bacteroides (27.20%), Prevotella (15%), Faecalibacterium (12.90%), Blautia (7.70%), Ruminococcus (7.50%), Alistipes (3.90%), Eubacterium (2.60%), Roseburia (2.50%), Fusobacterium (1.60%), Akkermansia (5.20%), Lactobacillus (3.80%), Helicobacter (2.50%), Clostridium (1.90%), Enterococcus (1.50%), Streptococcus (1.30%), Parabacteroides (1.10%), and other unclassified genera (1.80%). These percentages represent the relative abundances of each genus in the gut microbiota of the Japanese population, providing insights into the microbial diversity within their digestive systems, as shown in Table 6.1.

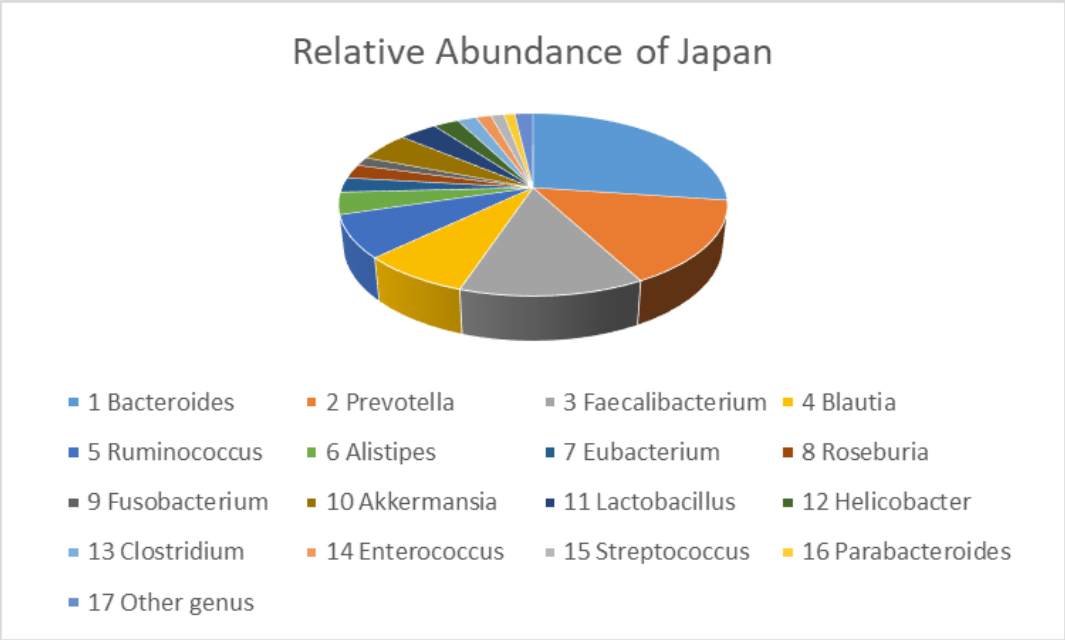


Figure 6.1: A pie chart visualisation of the relative abundance of gut bacteria of healthy Japanese individuals. Img Src: Author

Looking at Figure 6.1 we can see that within the gut microbiota of the Japanese population, certain genera emerge as dominant and play crucial roles in shaping the microbial ecosystem. Bacteroides exhibits a notable dominance with a relative abundance of 27.20%, suggesting its importance in gut health and function. Prevotella follows closely behind at 15%, indicating its significant presence in the Japanese gut. Faecalibacterium, comprising 12.90%, also holds a prominent position. Interestingly, the presence of Helicobacter and Streptococcus stands out, accounting for 2.50% and 1.30% respectively. These genera have been associated with various health implications, including a potential correlation with certain gastrointestinal conditions and diseases. Additionally, these genera, known for their distinctive characteristics, contribute to the overall gut microbial landscape in the Japanese population. Their presence underscores the complexity and diversity of the gut microbiota, highlighting the intricate interplay between bacteria and the host in maintaining a healthy gut environment.

Table 6.2: Relative Abundance of different bacterial genus in the gut of healthy Spanish Individuals. Please note that the values provided in the table are based on the total number of reads of 12,383. Source: Author

Serial No.	Genus	Number of Reads	Relative Abundance
1	Prevotella	3,616	29.20%
2	Bacteroides	3,207	25.90%
3	Faecalibacterium	1,758	14.20%
4	Roseburia	903	7.30%
5	Enterococcus	594	4.80%
6	Ruminococcus	458	3.70%
7	Akkermansia	409	3.30%
8	Lachnospira	309	2.50%
9	Eubacterium	285	2.30%
10	Streptococcus	260	2.10%
11	Lactobacillus	235	1.90%
12	Helicobacter	136	1.10%
13	Clostridium	99	0.80%
14	Other bacterial genus	112	0.90%

According to data on gut bacteria composition in the Spanish population, the relative abundance percentages of various genera are as analysed. The analysis reveals that Prevotella is the most prominent genus, constituting 29.2% of the gut microbiota. Bacteroides follow closely at 25.9%, while Faecalibacterium accounts for 14.2% and Roseburia for 7.3%. Other genera present in notable proportions include Enterococcus at 4.8%, Ruminococcus at 3.7%, Akkermansia at 3.3%, Lachnospira at 2.5%, Eubacterium at 2.3%, and Streptococcus at 2.1%. Lactobacillus, Helicobacter, and Clostridium contribute to the gut microbiota at 1.9%, 1.1%, and 0.8%, respectively. Additionally, a small fraction of 0.9% is attributed to other bacterial genera. These relative abundance percentages shed light on the composition of gut bacteria in the Spanish population, highlighting the prevalence of Prevotella and Bacteroides as the most dominant genera in the gut microbiota, as shown in Table 6.2.

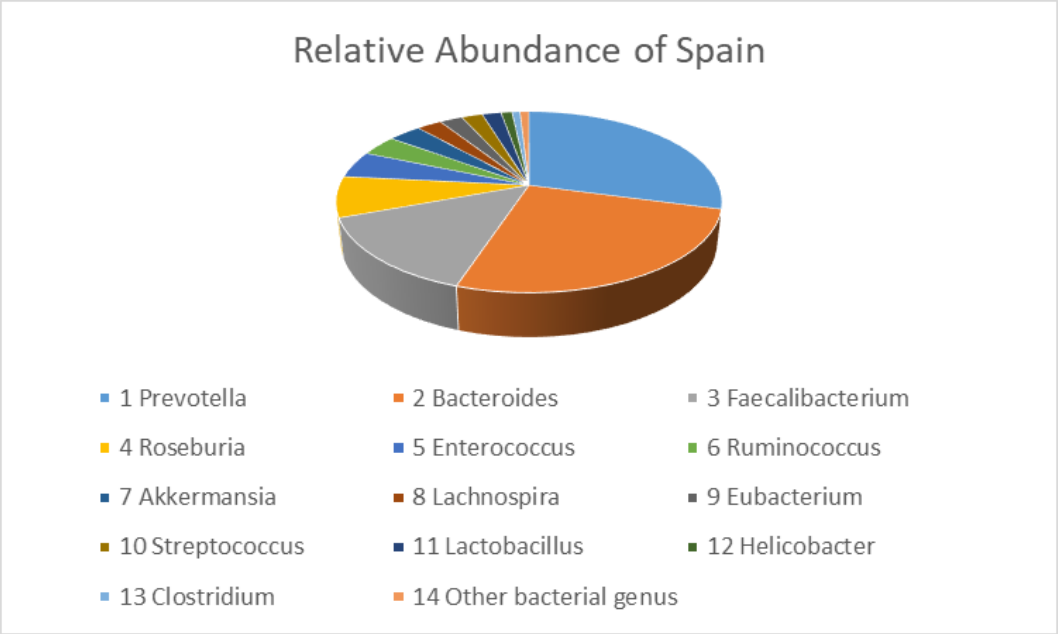


Figure 6.2: A pie chart visualisation of the relative abundance of gut bacteria of healthy Spanish individuals. Img Src: Author.

Looking at Figure 6.2 we can say that the pie chart of gut bacteria composition in the Spanish population reveals several dominant genera. Among them, Prevotella and Bacteroides emerge as the most prevalent, constituting 29.20% and 25.90% of the relative abundance, respectively. These findings highlight the significant presence of these two genera in the gut microbiota of the Spanish population. Additionally, Faecalibacterium and Roseburia exhibit notable relative abundances of

14.20% and 7.30%, respectively. It is worth noting the significant presence of Helicobacter and Streptococcus, with relative abundances of 1.10% and 2.10% respectively. These genera have been associated with various health implications, including a potential correlation with certain gastrointestinal conditions and diseases. Further research and investigation are necessary to explore the specific roles and potential implications of Helicobacter and Streptococcus in the context of the gut microbiota and human health in the Spanish population.

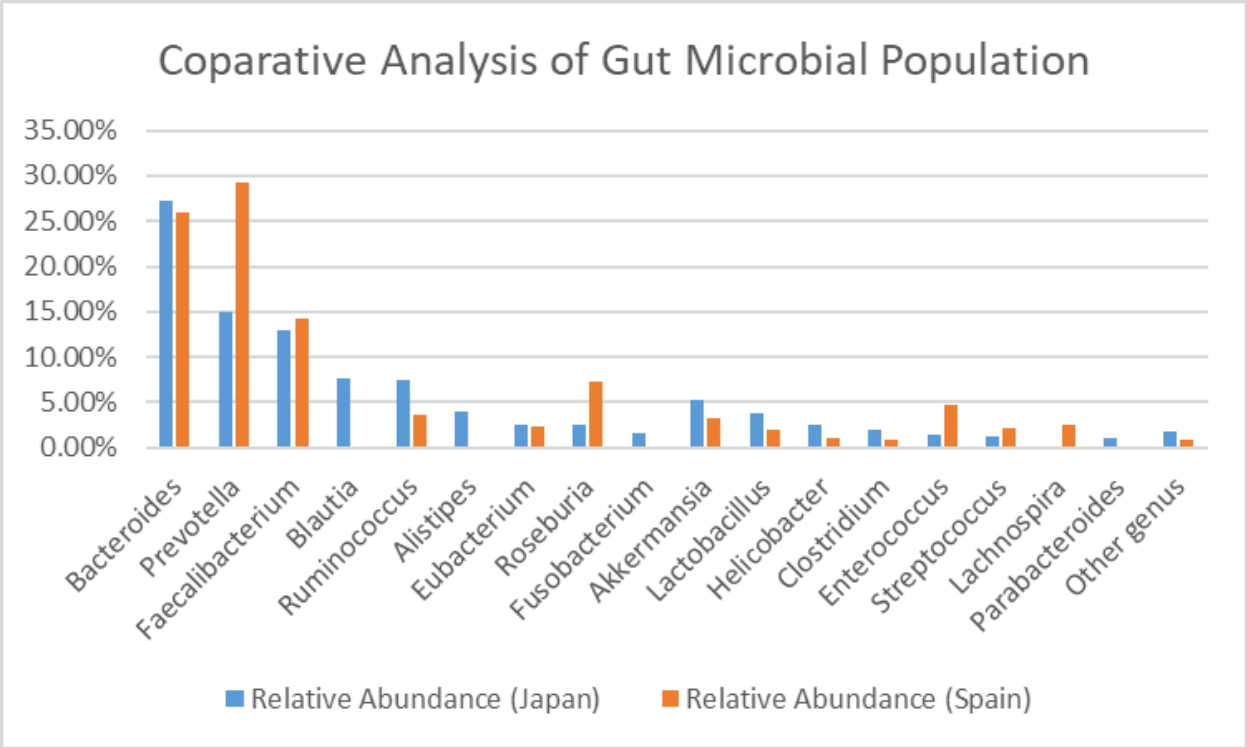


Figure 6.3: Comparative analysis of Relative Abundance of gut microbiota of Japanese and Spanish Population. Img Src: Author.

By looking at Figure 6.3 we can infer that in the Japanese population, the relative abundance percentages of these bacterial genera are as follows: Helicobacter (2.50%), Streptococcus (1.30%), Fusobacterium (1.60%), Bacteroides (27.20%), and Clostridium (1.90%). These percentages indicate the presence and prevalence of these bacterial biomarker genera within the gut microbiota of the Japanese population, suggesting a potential correlation with GI tract-related cancers. Helicobacter, Fusobacterium and Streptococcus, known for their association with gastric and colorectal cancers respectively, exhibit noteworthy relative abundances.

In contrast, the relative abundance percentages of these bacterial genera in the gut microbiota of the Spanish population are: Helicobacter (1.10%), Streptococcus (2.10%), Bacteroides (25.90%), and Clostridium (0.80%). These percentages also highlight the presence and potential significance of these bacterial biomarker genera in the gut microbiota of the Spanish population, particularly in the context of GI tract-related cancers.

Comparing the relative abundances between the two populations, we observe that both the Japanese and Spanish populations exhibit relatively high levels of Bacteroides and Streptococcus in their gut microbiota. However, the Japanese population shows a slightly higher relative abundance of Bacteroides (27.20%) compared to the Spanish population (25.90%). Additionally, the relative abundance of Streptococcus is slightly lower in the Japanese population (1.30%) compared to the Spanish population (2.10%). Regarding Helicobacter, Fusobacterium and Clostridium, the Japanese population demonstrates a higher relative abundance of Helicobacter (2.50%) compared to the Spanish population (1.10%). Similarly, the Japanese population also exhibits a higher relative abundance of Clostridium (1.90%) compared to the Spanish population (0.80%). The Japanese population also exhibits a significant relative abundance of Fusobacterium (1.60%), which the Spanish population does not.

These findings suggest that both the Japanese and Spanish populations may be at risk of developing GI tract-related cancers associated with these bacterial biomarker genera. However, the slightly higher relative abundances of Bacteroides, Helicobacter, Fusobacterium and Clostridium in the gut microbiota of the Japanese population may indicate a slightly elevated risk compared to the Spanish population.

It is important to note that the relative abundance percentages alone provide insights into the presence and potential roles of these bacterial genera, but further research is necessary to establish definitive causal relationships between their abundance and cancer development. The complex nature of cancer involves multifactorial interactions, including genetic predisposition, environmental factors, lifestyle choices, and the intricate interplay between the gut microbiota and the host's immune system.

6.2. Identifying HLA biomarkers From HLA Haplotype Profiles

Table 6.3: Depiction of Presence or Absence of HLA Biomarkers in Both Japanese and Spanish Population. Source: Author

Serial No.	HLA Biomarker	Associated GI Tract Cancers	Japanese Population	Spanish Population
1	HLA-A*02:01	Gastric Cancer, Colorectal Cancer	Present	Present
2	HLA-A*31:01	Gastric Cancer, Colorectal Cancer	Present	Present
3	HLA-A*11:01	Gastric Cancer, Colorectal Cancer	Present	Absent
4	HLA-B*44:02	Gastric Cancer, Colorectal Cancer	Absent	Present
5	HLA-B*07:02	Gastric Cancer, Colorectal Cancer	Absent	Absent
6	HLA-B*51:01	Gastric Cancer, Colorectal Cancer	Absent	Absent
7	HLA-C*07:02	Gastric Cancer, Colorectal Cancer	Present	Absent
8	HLA-DRB1*13:02	Gastric Cancer, Colorectal Cancer	Absent	Absent

9	HLA-DRB1*07:01	Gastric Cancer, Colorectal Cancer	Absent	Present
10	HLA-DRB1*03:01	Gastric Cancer, Esophageal Cancer	Absent	Absent
11	HLA-DRB1*15:01	Gastric Cancer, Esophageal Cancer	Present	Present
12	HLA-DRB1*04:05	Gastric Cancer, Esophageal Cancer	Absent	Present
13	HLA-DRB1*13:01	Gastric Cancer, Esophageal Cancer	Absent	Absent
14	HLA-DRB1*09:01	Gastric Cancer, Esophageal Cancer	Present	Absent
15	HLA-DRB1*11:04	Gastric Cancer, Esophageal Cancer	Present	Absent
16	HLA-DRB1*15:02	Gastric Cancer, Esophageal Cancer	Present	Absent
17	HLA-DRB1*04:01	Gastric Cancer, Esophageal Cancer	Present	Absent
18	HLA-DRB1*07:01	Gastric Cancer, Esophageal Cancer	Present	Present

19	HLA-DRB1*14:01	Gastric Cancer, Esophageal Cancer	Absent	Absent
20	HLA-DRB1*01:01	Gastric Cancer, Esophageal Cancer	Absent	Absent
21	HLA-DRB1*08:03	Gastric Cancer, Esophageal Cancer	Absent	Present
22	HLA-DRB1*16:02	Gastric Cancer, Esophageal Cancer	Absent	Absent
23	HLA-DRB1*10:01	Gastric Cancer, Esophageal Cancer	Present	Present
24	HLA-DRB1*12:02	Gastric Cancer, Esophageal Cancer	Absent	Absent
25	HLA-DRB1*08:01	Gastric Cancer, Esophageal Cancer	Absent	Absent
26	HLA-DRB1*11:03	Gastric Cancer, Esophageal Cancer	Absent	Absent
27	HLA-DRB1*11:03	Gastric Cancer, Esophageal Cancer	Present	Absent
28	HLA-DRB1*08:02	Gastric Cancer, Esophageal Cancer	Absent	Absent

29	HLA-DRB1*04:04	Gastric Cancer, Esophageal Cancer	Present	Present
30	HLA-DRB1*08:04	Gastric Cancer, Esophageal Cancer	Present	Absent

As represented in Table 6.3, The analysis of HLA biomarkers revealed the identification of 14 biomarkers associated with GI tract-related cancer in the Japanese population, while 10 biomarkers were identified in the Spanish population. This finding suggests that there may be variations in the distribution and prevalence of HLA biomarkers specifically linked to GI tract cancer between these two ethnic groups.

The higher number of identified HLA biomarkers in the Japanese population implies a potentially greater diversity or complexity in HLA genetic profiles related to GI tract cancer within this group. It indicates a broader range of potential disease associations or immune responses governed by these specific HLA biomarkers. The identification of a larger number of HLA biomarkers provides a more comprehensive understanding of the genetic factors influencing GI tract cancer susceptibility in the Japanese population.

Conversely, the identification of fewer HLA biomarkers in the Spanish population suggests a more limited range or prevalence of these specific biomarkers associated with GI tract cancer within this group. This may indicate a relatively lower diversity or fewer variations in the HLA genetic profiles related to GI tract cancer in the Spanish population compared to the Japanese population. The presence of fewer HLA biomarkers does not diminish their significance but reflects the unique genetic characteristics of the Spanish population in relation to GI tract cancer.

The variations in the number of identified HLA biomarkers between the Japanese and Spanish populations highlight the importance of considering population-specific factors in understanding the risk and development of GI tract-related cancers. Different populations may exhibit distinct

HLA genetic profiles and biomarker distributions, which can influence individual susceptibility and response to GI tract cancer.

These findings underscore the need for population-specific studies and personalised approaches in GI tract cancer research. The identification and characterization of population-specific HLA biomarkers associated with GI tract cancer can contribute to tailored strategies for early detection, risk assessment, and targeted treatment in different ethnic groups.

Further research is necessary to explore the functional significance and clinical implications of the identified HLA biomarkers in GI tract-related cancer for both the Japanese and Spanish populations. Understanding the role of these biomarkers in disease progression, prognosis, and treatment response can pave the way for advancements in personalised medicine and targeted interventions specific to GI tract-related cancers in these populations.

6.3. Limitations of This Risk Assessment

- Limited data: This project faced limitations in terms of data availability and processing capabilities, resulting in the use of a minimum dataset. The analysis was conducted with the available data, which may not fully capture the complexity and diversity of the gut microbiome and HLA profiles. Future studies should aim to incorporate larger and more comprehensive datasets to enhance the statistical power and robustness of the findings.
- Correlation does not imply causation: While the presence of these bacterial genera in the gut has been associated with altered gut microbiota composition and potential involvement in cancer development and progression, it is essential to note that correlation does not necessarily indicate a causative relationship. Further research is required to unravel the underlying mechanisms and interactions between these bacterial genera and GI tract cancers, including in vivo and functional studies.
- Lack of a control group: This project relied on a comparison between the Japanese and Spanish populations without the inclusion of a control group for reference. The absence of a control group limits the ability to establish baseline comparisons and discern whether the observed differences in the relative abundance of bacterial genera are specific to cancer

cases or inherent to the population itself. Future investigations should incorporate a control group to provide a comprehensive understanding of the associations observed.

- Within-population analysis: To gain deeper insights into the relationship between bacterial genera and cancer risk, it is recommended to compare the relative abundance of these genera between individuals with and without cancer within each population separately. This within-population analysis would help elucidate potential associations and differences specific to cancer cases, contributing to a more accurate understanding of the microbial dynamics in relation to cancer development within each population.
- Statistical assessment: Utilising appropriate statistical tests, such as the Wilcoxon rank-sum test, is crucial to assess the significance of differences in relative abundance between cancer and non-cancer groups within each population. Statistical analyses aid in determining whether the observed variations in bacterial genera abundance are statistically significant, further strengthening the validity and reliability of the findings.

CHAPTER-7
CONCLUSION

By comprehensively analysing the gut microbiome and HLA profiles of the Japanese and Spanish populations, this study aimed to identify prevalent biomarkers and investigate correlations with disease risk. These findings contribute to our understanding of the complex interactions between the gut microbiome, HLA genetics, and disease susceptibility in these populations. Overall, this study employed data mining techniques to process and analyse the gut microbiome and HLA profiles of the Japanese and Spanish populations. The identification of microbial and HLA biomarkers, along with correlation analysis, sheds light on the potential relationships between gut microbiome composition, HLA profiles, and disease risk. These findings have implications for future research, personalised medicine, and the development of targeted interventions for disease prevention and treatment.

Overall, while both the Japanese and Spanish populations exhibit notable relative abundances of these bacterial biomarker genera in their gut microbiota, the slightly higher relative abundances of Bacteroides, Helicobacter, and Clostridium in the Japanese population may imply a slightly increased risk of GI tract-related cancers. Nonetheless, further research is warranted to fully elucidate the complex relationship between the gut microbiota composition and the development of cancer. The higher abundance of bacterial biomarkers of cancer in the gut of the Japanese population, along with the identification of a greater number of HLA biomarkers associated with GI tract-related cancer, suggests a potential interplay between the gut microbiome and HLA genetics in the development and progression of GI tract cancer.

The gut microbiome and HLA genetics both play critical roles in modulating immune responses and maintaining a balanced gut environment. Dysregulation in either the gut microbiome or HLA genetic profiles can disrupt immune homeostasis and contribute to the development of various diseases, including cancer. The gut microbiome has been increasingly recognized as a key player in influencing host health, including its involvement in cancer progression. Certain bacterial species or genera have been linked to the initiation, promotion, or progression of cancer by producing metabolites, promoting chronic inflammation, altering the gut microenvironment, or

modulating immune responses. The higher abundance of bacterial biomarkers of cancer in the gut of the Japanese population suggests a potential alteration in the gut microbial composition that may contribute to a higher predisposition to GI tract cancer in this group. These bacterial biomarkers may interact with the host's immune system and HLA genetic profiles, influencing disease susceptibility and progression.

The HLA genetic profiles, on the other hand, govern the presentation of antigens to immune cells and play a crucial role in immune surveillance and response against cancer cells. Specific HLA alleles or haplotypes have been associated with increased or decreased cancer risk and can influence the recognition and elimination of tumour cells by the immune system. The presence of a greater number of HLA biomarkers associated with GI tract-related cancer in the Japanese population suggests a potentially broader range of immune responses and genetic susceptibility to these cancers. The interaction between HLA genetics and the gut microbiome may further modulate the immune response to tumour development and progression. The relationship between the gut microbiome, HLA genetics, and cancer is complex and multifaceted. The higher abundance of bacterial biomarkers of cancer in the gut of the Japanese population, along with the identified HLA biomarkers, highlights the potential interplay and interconnectedness between these factors in the context of GI tract cancer.

Further research is needed to unravel the precise mechanisms by which the gut microbiome and HLA genetics influence cancer development and progression. Understanding these relationships can provide valuable insights into the underlying mechanisms of GI tract cancer and open avenues for personalised interventions and therapeutic strategies targeting the gut microbiome and immune system in the prevention and treatment of GI tract-related cancers.

CHAPTER-8

FUTURE WORK

While this project has provided valuable insights into the gut microbiome and HLA profiles in relation to GI tract cancer in the Japanese and Spanish populations, there are several limitations. Future work would aim to acquire a larger dataset to improve the statistical power and generalizability of the findings. Accessing additional data from diverse populations would allow for a more comprehensive understanding of the gut microbiome and HLA profiles in relation to GI tract cancer. This project focused on comparing the gut microbiome and HLA profiles between the Japanese and Spanish populations without a control group for reference. Future studies should include a control group consisting of individuals without cancer from each population to provide a baseline for comparison. This would allow for a more robust assessment of the differences and associations observed in the relative abundance of bacterial genera and HLA biomarkers in individuals with cancer.

To better elucidate the role of the identified bacterial genera and HLA biomarkers in GI tract cancer, future research should compare the relative abundance of these microbial taxa between individuals with and without cancer within each population separately. This within-population analysis would provide insights into potential associations and differences in the gut microbiome and HLA profiles specifically related to cancer development and progression. Utilising statistical tests, such as the Wilcoxon rank-sum test or other appropriate methods, would help assess the significance of differences in relative abundance between individuals with cancer and those without cancer within each population. This would provide a quantitative measure of the association between the identified bacterial genera, HLA biomarkers, and cancer risk. Further investigations would aim to uncover the specific mechanisms and interactions between the identified bacterial genera, HLA profiles, and GI tract cancers. This could involve exploring the functional characteristics of the bacterial taxa, understanding the immune responses elicited by HLA biomarkers, and investigating potential molecular pathways involved in cancer development and progression. Also, further studies involving larger cohorts, longitudinal analyses. Such research endeavours may shed light on the specific bacterial species or virulence factors within *Bacteroides*, *Helicobacter*, *Streptococcus*, and *Clostridium* that play a critical role in cancer

initiation and progression. To strengthen the reliability and reproducibility of the results, future work should validate the findings in independent cohorts or populations. Replicating the analysis in different datasets would help confirm the identified associations and establish their generalizability.

REFERENCES

1. Hawrelak, Jason A, and Stephen P Myers. “The causes of intestinal dysbiosis: a review,” *Alternative medicine review: a journal of clinical therapeutic,* vol. 9, pp. 180-97. 2004

2. Sekirov, Inna et al. “Gut microbiota in health and disease,” *Physiological review,s* vol. 90, pp. 859-904. 2010

3. Akbar, Noor & Khan, Naveed & Muhammad, Jibran & Siddiqui, Ruqaiyyah. “The role of gut microbiome in cancer genesis and cancer prevention,” *Health Sciences Review,* Vol. 2. 2022

4. Hou, Kaijian et al. “Microbiota in health and diseases,” *Signal transduction and targeted therapy,* vol. 7(1), pp. 135. April 2022

5. Stappenbeck, Thaddeus S et al. “Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells,” *Proceedings of the National Academy of Sciences of the United States of America,* vol. 99, pp. 24. 2002

6. de Vos, Willem M et al. “Gut microbiome and health: mechanistic insights,” *Gut,* vol. 71, pp. 1020-1032. 2022

7. Bäckhed, Fredrik et al. “Host-bacterial mutualism in the human intestine,” *Science,* vol. 307, pp. 1915-20. 2005

8. Scanlan, Pauline D et al. “Human methanogen diversity and incidence in healthy and diseased colonic groups using mcrA gene analysis,” *BMC microbiology,* vol. 8, page 20. May 2008

9. Human Microbiome Project Consortium. “Structure, function and diversity of the healthy human microbiome,” *Nature,* vol. 486, pp. 207-14. June 2012

10. Huse, Susan M et al. “A core human microbiome as viewed through 16S rRNA sequence clusters,” *PloS one,* vol. 7. 2012

11. Conlon, Michael A, and Anthony R Bird. “The impact of diet and lifestyle on gut microbiota and human health,” *Nutrients,* vol. 7, pp. 17-44. December 2014

12. Zitvogel, Laurence et al. “Anticancer effects of the microbiome and its products,” *Nature reviews. Microbiology,* vol. 15, pp. 465-478. 2017

13. Pasolli, Edoardo et al. “Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle,” *Cell,* vol. 176(3), pp. 649-662. 2019

14. Jiao, Jian-Yu et al. “Microbial dark matter coming to light: challenges and opportunities,” *National science review,* vol. 8(3). December 2020

15. Swann, J R et al. “Considerations for the design and conduct of human gut microbiota intervention studies relating to foods,” *European journal of nutrition,* vol. 59(8) pp. 3347-3368. 2020

16. D. Yu et al., “Bibliometric analysis of fuzzy theory research in China: A 30-year perspective,” *Knowl. Based Syst.,* vol. 141, pp. 188-199. 2018

17. N. Donthu et al., “How to conduct a bibliometric analysis: An overview and guidelines,” *J. Bus. Res.,* vol. 133, pp. 285-296. 2021

18. Amaye, Roku et al. “Role of Microbiota in Cancer Research,” *Signal transduction and targeted therapy,* vol. 7, page 1125-8. January 2023

19. den Besten, Gijs et al. “The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism.” *Journal of lipid research,* vol. 54(9), pp. 2325-40. 2014

20. Morowitz, Michael J et al. “Contributions of intestinal bacteria to nutrition and metabolism in the critically ill.” *The Surgical clinics of North America,* vol. 91, pp. 771-85. 2011

21. Human Microbiome Project Consortium. “Structure, function and diversity of the healthy human microbiome.” *Nature,* vol. 486, pp. 207-14. June 2012

22. Nicholson, Jeremy K et al. “Host-gut microbiota metabolic interactions.” *Science (New York, N.Y.),* vol. 336, pp. 1262-7. 2012

23. Trent, M Stephen et al. “Diversity of endotoxin and its impact on pathogenesis.” *Journal of endotoxin research,* vol. 12, pp. 205-23. 2006

24. Kamada, Nobuhiko et al. “A complex microworld in the gut: Harnessing pathogen-commensal relations.” *Nature medicine,* vol. 18, pp. 1190-1. 2012

25. Cantarel, Brandi L et al. “Complex carbohydrate utilization by the healthy human microbiome.” *PloS one,* vol. 7. 2012

26. Odamaki, Toshitaka et al. “Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study.” *BMC microbiology* vol. 16. May 2016

27. Jandhyala, Sai Manasa et al. “Role of the normal gut microbiota.” *World journal of gastroenterology,* vol. 21, pp. 8787-803. 2015

28. Chong, Clara Yieh Lin et al. “Factors Affecting Gastrointestinal Microbiome Development in Neonates.” *Nutrients,* vol. 10, Page 274. February 2018

29. Voreades, Noah et al. “Diet and the development of the human intestinal microbiome.” *Frontiers in microbiology,* vol. 5, Page 494. September 2014

30. Spor, Aymé et al. “Unravelling the effects of the environment and host genotype on the gut microbiome.” *Nature reviews. Microbiology,* vol. 9, pp. 279-90. 2011

31. Jandhyala, Sai Manasa et al. “Role of the normal gut microbiota.” *World journal of gastroenterology,* vol. 21, pp. 8787-803. 2015

32. Nagpal, Ravinder et al. “Ontogenesis of the Gut Microbiota Composition in Healthy, Full-Term, Vaginally Born and Breast-Fed Infants over the First 3 Years of Life: A Quantitative Bird's-Eye View.” *Frontiers in microbiology,* vol. 8. July 2017

33. Wen, Li, and Andrew Duffy. “Factors Influencing the Gut Microbiota, Inflammation, and Type 2 Diabetes.” *The Journal of nutrition,* vol. 147, pp. 1468S-1475S. 2017

34. Veiga, Patrick et al. “Changes of the human gut microbiome induced by a fermented milk product.” *Scientific reports,* vol. 4. September 2014

35. Field, Alison E et al. “Dietary fat and weight gain among women in the Nurses' Health Study.” *Obesity (Silver Spring, Md.),* vol. 15, pp. 967-76. 2007

36. Winzell, Maria Sörhede, and Bo Ahrén. “The high-fat diet-fed mouse: a model for studying mechanisms and treatment of impaired glucose tolerance and type 2 diabetes.” *Diabetes,* vol. 53, pp. 215-19. 2004

37. Musso, Giovanni et al. “Obesity, diabetes, and gut microbiota: the hygiene hypothesis expanded?.” *Diabetes care,* vol. 33. 2010

38. Sonnenburg, Erica D et al. “Diet-induced extinctions in the gut microbiota compound over generations.” *Nature* vol. 529. 2016

39. Wu, Gary D et al. “Linking long-term dietary patterns with gut microbial enterotypes.” *Science (New York, N.Y.)* vol. 334, pp. 105-8. 2011

40. David, Lawrence A et al. “Diet rapidly and reproducibly alters the human gut microbiome.” *Nature,* vol. 505, pp. 559-63. 2014

41. Wen, Li, and Andrew Duffy. “Factors Influencing the Gut Microbiota, Inflammation, and Type 2 Diabetes.” *The Journal of nutrition,* vol. 147, pp. 1468-1475. 2017

42. Gross, Liza. “Microbes colonize a baby's gut with distinction.” *PLoS biology,* vol. 5. 2007

43. M'Rabet, Laura et al. “Breast-feeding and its role in early development of the immune system in infants: consequences for health later in life.” *The Journal of nutrition,* vol. 138, pp. 1782-1790. 2008

44. Walker, W Allan, and Rajashri Shuba Iyengar. “Breast milk, microbiota, and intestinal immune homeostasis.” *Pediatric research,* vol. 77, pp. 220-28. 2015

45. Houghteling, Pearl D, and W Allan Walker. “Why is initial bacterial colonization of the intestine important to infants' and children's health?.” *Journal of pediatric gastroenterology and nutrition,* vol. 60, pp. 294-307. 2015

46. Walker, Alan W et al. “Dominant and diet-responsive groups of bacteria within the human colonic microbiota.” *The ISME journal* vol. 5, pp. 220-30. 2011

47. David, Lawrence A et al. “Diet rapidly and reproducibly alters the human gut microbiome.” *Nature,* vol. 505, pp. 559-63. 2014

48. Windey, Karen et al. “Relevance of protein fermentation to gut health.” *Molecular nutrition & food research* vol. 56, pp. 184-96. 2012

49. Tang, W H Wilson et al. “Intestinal microbial metabolism of phosphatidylcholine and delcardiovascular risk.” *The New England journal of medicine* vol. 368, pp. 1575-84. 2013

50. Collado, Maria Carmen et al. “Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid.” *Scientific reports* vol. 6. March 2016

51. Nagpal, Ravinder et al. “Ontogenesis of the Gut Microbiota Composition in Healthy, Full-Term, Vaginally Born and Breast-Fed Infants over the First 3 Years of Life: A Quantitative Bird's-Eye View.” *Frontiers in microbiology* vol. 8. July 2017

52. Dominguez-Bello, Maria G et al. “Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns.” *Proceedings of the National Academy of Sciences of the United States of America* vol. 107. 2010

53. Yatsunenko, Tanya et al. “Human gut microbiome viewed across age and geography.” *Nature* vol. 486. May 2012

54. Klingensmith, Nathan J, and Craig M Coopersmith. “The Gut as the Motor of Multiple Organ Dysfunction in Critical Illness.” *Critical care clinics* vol. 32, pp. 203-12. 2016

55. Hehemann, Jan-Hendrik et al. “Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota.” *Nature* vol. 464, pp. 908-12. 2010

56. Ramnani, Priya et al. “In vitro fermentation and prebiotic potential of novel low molecular weight polysaccharides derived from agar and alginate seaweeds.” *Anaerobe* vol. 18, pp. 1-6. 2012

57. Jernberg, Cecilia et al. “Long-term ecological impacts of antibiotic administration on the human intestinal microbiota.” *The ISME journal* vol. 1, pp. 56-66. 2007

58. Dethlefsen, Les, and David A Relman. “Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation.” *Proceedings of the National Academy of Sciences of the United States of America* vol. 108, pp. 4554-61. 2011

59. Lewis, Brittany B et al. “Loss of Microbiota-Mediated Colonization Resistance to Clostridium difficile Infection With Oral Vancomycin Compared With Metronidazole.” *The Journal of infectious diseases* vol. 212, pp. 1656-65. 2015

60. Isaac, Sandrine et al. “Short- and long-term effects of oral vancomycin on the human intestinal microbiota.” *The Journal of antimicrobial chemotherapy* vol. 72, pp. 128-136. 2017

61. Dethlefsen, Les, and David A Relman. “Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation.” *Proceedings of the National Academy of Sciences of the United States of America* vol. 108, pp. 4554-61. 2011

62. Costello, Elizabeth K et al. “The application of ecological theory toward an understanding of the human microbiome.” *Science (New York, N.Y.)* vol. 336, pp. 1255-62. 2012

63. Walter, Jens, and Ruth Ley. “The human gut microbiome: ecology and recent evolutionary changes.” *Annual review of microbiology* vol. 65, pp. 411-29. 2011

64. Bevins, Charles L, and Nita H Salzman. “The potter's wheel: the host's role in sculpting its microbiota.” *Cellular and molecular life sciences : CMLS* vol. 68, pp. 3675-85. 2011

65. van Nood, Els et al. “Duodenal infusion of donor feces for recurrent Clostridium difficile.” *The New England journal of medicine* vol. 368, pp. 407-15. 2013

66. Goodrich, Julia K et al. “Human genetics shape the gut microbiome.” *Cell* vol. 159, pp. 789-99. 2014

67. Blekhman, Ran et al. “Host genetic variation impacts microbiome composition across human body sites.” *Genome biology* vol. 16. September 2015

68. Thompson-Chagoyán, Oscar C et al. “Aetiology of inflammatory bowel disease (IBD): role of intestinal microbiota and gut-associated lymphoid tissue immune response.” *Clinical nutrition (Edinburgh, Scotland)* vol. 24, pp. 339-52. 2005

69. Rehman, Ateequr et al. “Nod2 is essential for temporal development of intestinal microbial communities.” *Gut* vol. 60, pp. 1354-62. 2011

70. Pevsner-Fischer, Meirav et al. “Role of the microbiome in non-gastrointestinal cancers.” *World journal of clinical oncology* vol. 7, pp. 200-13. 2016

71. Lize, Nora et al. “Psychosocial consequences of a reduced ability to eat for patients with cancer and their informal caregivers: A qualitative study.” *European journal of oncology nursing : the official journal of European Oncology Nursing Society* vol. 49. 2020

72. Yang, Wanshui et al. “Calcium Intake and Survival after Colorectal Cancer Diagnosis.” *Clinical cancer research : an official journal of the American Association for Cancer Research* vol. 25, pp. 1980-1988. 2019

73. Hooper, Lora V, and Andrew J Macpherson. “Immune adaptations that maintain homeostasis with the intestinal microbiota.” *Nature reviews. Immunology* vol. 10, pp. 159-69. 2010

74. Falony, Gwen et al. “Population-level analysis of gut microbiome variation.” *Science (New York, N.Y.)* vol. 352, pp. 560-64. 2016

75. Vangay, Pajau et al. “US Immigration Westernizes the Human Gut Microbiome.” *Cell* vol. 175, pp. 962-972. 2018

76. Garcia-Peris, P et al. “Effect of inulin and fructo-oligosaccharide on the prevention of acute radiation enteritis in patients with gynecological cancer and impact on quality-of-life: a randomized, double-blind, placebo-controlled trial.” *European journal of clinical nutrition* vol. 70, pp. 170-4. 2016

77. Taper, H S, and M B Roberfroid. “Nontoxic potentiation of cancer chemotherapy by dietary oligofructose or inulin.” *Nutrition and cancer* vol. 38, pp. 1-5. 2000

78. Taper, H S, and M B Roberfroid. “Inulin/oligofructose and anticancer therapy.” *The British journal of nutrition* vol. 87. 2002

79. Taper, Henryk S, and Marcel B Roberfroid. “Possible adjuvant cancer therapy by two prebiotics--inulin or oligofructose.” *In vivo (Athens, Greece)* vol. 19, pp. 201-4. 2005

80. Zhao, Rui et al. “Effects of fiber and probiotics on diarrhea associated with enteral nutrition in gastric cancer patients: A prospective randomized and controlled trial.” *Medicine* vol. 96. 2017

81. Maeyama, Masahiro et al. “Metabolic changes and anti-tumor effects of a ketogenic diet combined with anti-angiogenic therapy in a glioblastoma mouse model.” *Scientific reports* vol. 11. January 2021

82. van der Louw, Elles J T M et al. “Ketogenic diet treatment as adjuvant to standard treatment of glioblastoma multiforme: a feasibility and safety study.” *Therapeutic advances in medical oncology* vol. 11. June 2019

83. Lien, Evan C et al. “Low glycaemic diets alter lipid metabolism to influence tumour growth.” *Nature* vol. 599, pp. 302-307. 2021

84. Lam, Khiem C et al. “Microbiota triggers STING-type I IFN-dependent monocyte reprogramming of the tumor microenvironment.” *Cell* vol. 184, pp. 5338-5356. 2021

85. Bultman, Scott J. “Emerging roles of the microbiome in cancer.” *Carcinogenesis* vol. 35, pp. 249-55. 2014

86. Thomas, Andrew Maltez et al. “Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation.” *Nature medicine* vol. 25, pp. 667-678. 2019

87. Johansson, Malin E V, and Gunnar C Hansson. “Immunological aspects of intestinal mucus and mucins.” *Nature reviews. Immunology* vol. 16, pp. 639-49. 2016

88. Deplancke, B, and H R Gaskins. “Microbial modulation of innate defense: goblet cells and the intestinal mucus layer.” *The American journal of clinical nutrition* vol. 73. 2001

89. Sharma, R, and U Schumacher. “Morphometric analysis of intestinal mucins under different dietary conditions and gut flora in rats.” *Digestive diseases and sciences* vol. 40, pp. 2532-9. 1995

90. McCoy, Kathy D, and Markus B Geuking. “Microbiota regulates intratumoral monocytes to promote anti-tumor immune responses.” *Cell* vol. 184,21 (2021): 5301-5303. doi:10.1016/j.cell.2021.09.024

91. Lam, Khiem C et al. “Microbiota triggers STING-type I IFN-dependent monocyte reprogramming of the tumor microenvironment.” *Cell* vol. 184, pp. 5338-5356. 2021

92. Gill, H S et al. “Enhancement of natural and acquired immunity by Lactobacillus rhamnosus (HN001), Lactobacillus acidophilus (HN017) and Bifidobacterium lactis (HN019).” *The British journal of nutrition* vol. 83, pp. 167-76. 2000

93. Tanoue, Takeshi et al. “A defined commensal consortium elicits CD8 T cells and anti-cancer immunity.” *Nature* vol. 565, pp. 600-605. 2019

94. Sokol, Harry et al. “Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients.” *Proceedings of the National Academy of Sciences of the United States of America* vol. 105, pp. 16731-6. 2008

95. Rosshart, Stephan P et al. “Wild Mouse Gut Microbiota Promotes Host Fitness and Improves Disease Resistance.” *Cell* vol. 171, pp. 1015-1028. 2017

96. Westbrook, Aya M et al. “Intestinal mucosal inflammation leads to systemic genotoxicity in mice.” *Cancer research* vol. 69, pp. 4827-34. 2009

97. An, Byung Chull et al. “Anti-Colorectal Cancer Effects of Probiotic-Derived p8 Protein.” *Genes* vol. 10. August 2019

98. Mikó, Edit et al. “Lithocholic acid, a bacterial metabolite reduces breast cancer cell proliferation and aggressiveness.” *Biochimica et biophysica acta. Bioenergetics* vol. 1859, pp. 958-974. 2018

99. Sadeghi-Aliabadi, Hojjat et al. “Effects of Lactobacillus plantarum A7 with probiotic potential on colon cancer and normal cells proliferation in comparison with a commercial strain.” *Iranian journal of basic medical sciences* vol. 17, pp. 815-9. 2014

100. Chandel, Deepika et al. “Isolation, characterization and identification of antigenotoxic and anticancerous indigenous probiotics and their prophylactic potential in experimental colon carcinogenesis.” *Scientific reports* vol. 9. October 2019

101. Viaud, Sophie et al. “The intestinal microbiota modulates the anticancer immune effects of cyclophosphamide.” *Science (New York, N.Y.)* vol. 342, pp. 971-6. 2013

102. Scott, Timothy A et al. “Host-Microbe Co-metabolism Dictates Cancer Drug Efficacy in C. elegans.” *Cell* vol. 169, pp. 442-456. 2017

103. Iida, Noriho et al. “Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment.” *Science (New York, N.Y.)* vol. 342, pp. 967-70. 2013

104. Nicholson, Jeremy K et al. “Host-gut microbiota metabolic interactions.” *Science (New York, N.Y.)* vol. 336, pp. 1262-7. 2012

105. Bultman, Scott J. “The microbiome and its potential as a cancer preventive intervention.” *Seminars in oncology* vol. 43, pp. 97-106. 2016

106. Sethi, Vrishketan et al. “Gut Microbiota Promotes Tumor Growth in Mice by Modulating Immune Response.” *Gastroenterology* vol. 155, pp. 33-37. 2018

107. Meurman, Jukka H. “Oral microbiota and cancer.” *Journal of oral microbiology* vol. 2. August 2010

108. Pevsner-Fischer, Meirav et al. “Role of the microbiome in non-gastrointestinal cancers.” *World journal of clinical oncology* vol. 7, pp. 200-13. 2016

109. Vergara, Daniele et al. “The Cancer Microbiota: EMT and Inflammation as Shared Molecular Mechanisms Associated with Plasticity and Progression.” *Journal of oncology* vol. 21. October 2019

110. Hanahan, Douglas, and Robert A Weinberg. “Hallmarks of cancer: the next generation.” *Cell* vol. 144, pp. 646-74. 2011

111. Jin, Chengcheng et al. “Commensal Microbiota Promote Lung Cancer Development via γδ T Cells.” *Cell* vol. 176, pp. 998-1013. 2019

112. Hattori, Naoko, and Toshikazu Ushijima. “Epigenetic impact of infection on carcinogenesis: mechanisms and applications.” *Genome medicine* vol. 8. January 2016

113. Gur, Chamutal et al. “Binding of the Fap2 protein of Fusobacterium nucleatum to human inhibitory receptor TIGIT protects tumors from immune cell attack.” *Immunity* vol. 42, pp. 344-355. 2015

114. Dalmasso, Guillaume et al. “The bacterial genotoxin colibactin promotes colon tumor growth by modifying the tumor microenvironment.” *Gut microbes* vol. 5, pp. 675-80. 2014

115. Osherov, Nir, and Ronen Ben-Ami. “Modulation of Host Angiogenesis as a Microbial Survival Strategy and Therapeutic Target.” *PLoS pathogens* vol. 12. April 2016

116. De Spiegeleer, Bart et al. “The quorum sensing peptides PhrG, CSP and EDF promote angiogenesis and invasion of breast cancer cells in vitro.” *PloS one* vol. 10. March 2015

117. Duvallet, Claire et al. “Meta-analysis of gut microbiome studies identifies disease-specific and shared responses.” *Nature communications* vol. 8. December 2017

118. Saus, Ester et al. “Microbiome and colorectal cancer: Roles in carcinogenesis and clinical potential.” *Molecular aspects of medicine* vol. 69, pp. 93-106. 2019

119. Siegel, Rebecca L et al. “Cancer Statistics, 2021.” *CA: a cancer journal for clinicians* vol. 71, pp. 7-33. 2021

120. Imperiale, Thomas F et al. “Performance Characteristics of Fecal Immunochemical Tests for Colorectal Cancer and Advanced Adenomatous Polyps: A Systematic Review and Meta-analysis.” *Annals of internal medicine* vol. 170, pp. 319-329. 2019

121. Kisiel, John B, and Paul J Limburg. “Colorectal Cancer Screening With the Multitarget Stool DNA Test.” *The American journal of gastroenterology* vol. 115, pp. 1737-1740. 2020

122. Sobhani, Iradj et al. “Microbial dysbiosis in colorectal cancer (CRC) patients.” *PloS one* vol. 6. January 2011

123. Abdulamir, Ahmed S et al. “Molecular detection, quantification, and isolation of Streptococcus gallolyticus bacteria colonizing colorectal tumors: inflammation-driven potential of carcinogenesis via IL-1, COX-2, and IL-8.” *Molecular cancer* vol. 9, pp. 249. September 2010

124. Dai, Zhenwei et al. “Multi-cohort analysis of colorectal cancer metagenome identified altered bacteria across populations and universal bacterial markers.” *Microbiome* vol. 6. April 2018

125. Feng, Qiang et al. “Gut microbiome development along the colorectal adenoma-carcinoma sequence.” *Nature communications* vol. 6, Page 6528. March 2015

126. Zeller, Georg et al. “Potential of fecal microbiota for early-stage detection of colorectal cancer.” *Molecular systems biology* vol. 10, Page 766. November 2014

127. Iida, Noriho et al. “Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment.” *Science (New York, N.Y.)* vol. 342, pp. 967-70. 2013

128. Ling, Zongxin et al. “Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis.” *BMC genomics* vol. 11, Page 488. September 2010

129. Schwabe, Robert F, and Christian Jobin. “The microbiome and cancer.” *Nature reviews. Cancer* vol. 13, pp. 800-12. 2013

130. Stafford, Phillip et al. “Immunosignature system for diagnosis of cancer.” *Proceedings of the National Academy of Sciences of the United States of America* vol. 111. 2014

131. Zackular, Joseph P et al. “The human gut microbiome as a screening tool for colorectal cancer.” *Cancer prevention research (Philadelphia, Pa.)* vol. 7, pp. 1112-21. 2014

132. Garrett, Wendy S. “Cancer and the microbiota.” *Science (New York, N.Y.)* vol. 348, pp. 80-6. 2015

133. Liu, Shili et al. “Intestinal bacteria are potential biomarkers and therapeutic targets for gastric cancer.” *Microbial pathogenesis* vol. 151. 2021

134. Laborda-Illanes, Aurora et al. “Breast and Gut Microbiota Action Mechanisms in Breast Cancer Pathogenesis and Treatment.” *Cancers* vol. 12. August 2020

135. Goedert, James J et al. “Investigation of the association between the fecal microbiota and breast cancer in postmenopausal women: a population-based case-control pilot study.” *Journal of the National Cancer Institute* vol. 107. June 2015

136. Zhu, Jia et al. “Breast cancer in postmenopausal women is associated with an altered gut metagenome.” *Microbiome* vol. 6. August 2018

137. Jónsdóttir, Svava Osk et al. “Prediction methods and databases within chemoinformatics: emphasis on drugs and drug candidates.” *Bioinformatics (Oxford, England)* vol. 21, pp. 2145-60. 2005

138. Benson, Dennis A et al. “GenBank.” *Nucleic acids research* vol. 34. 2006

139. Altschul, S F et al. “Basic local alignment search tool.” *Journal of molecular biology* vol. 215, pp. 403-10. 1990

140. Bansal, Arvind K, and Terrance E Meyer. “Evolutionary analysis by whole-genome comparisons.” *Journal of bacteriology* vol. 184, pp. 2260-72. 2002

141. Jiao, Xiaoli et al. “A Benchmark Study on Error Assessment and Quality Control of CCS Reads Derived from the PacBio RS.” *Journal of data mining in genomics & proteomics* vol. 4. 2013

142. Callahan, Benjamin J et al. “DADA2: High-resolution sample inference from Illumina amplicon data.” *Nature methods* vol. 13, pp. 581-3. 2016

143. Edgar R. C. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. Preprint at *bio Rxiv,* vol. 2. 2016

144. Schäffer, A A et al. “Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements.” *Nucleic acids research* vol. 29, pp. 2994-3005. 2001

145. Kent, W James. “BLAT--the BLAST-like alignment tool.” *Genome research,* vol. 12, pp. 656-64. 2002

146. Cameron, Michael et al. “A deterministic finite automaton for faster protein hit detection in BLAST.” *Journal of computational biology : a journal of computational molecular cell biology,* vol. 13, pp. 965-78. 2006

147. Morgulis, Aleksandr et al. “WindowMasker: window-based masker for sequenced genomes.” *Bioinformatics (Oxford, England),* vol. 22, pp. 134-41. 2006

148. Morgulis, Aleksandr et al. “A fast and symmetric DUST implementation to mask low-complexity DNA sequences.” *Journal of computational biology : a journal of computational molecular cell biology,* vol. 13, pp. 1028-40. 2006

149. Chen, Yu-Li et al. “Gene polymorphisms of pathogenic *Helicobacter pylori* in patients with different types of gastrointestinal diseases.” *World journal of gastroenterology,* vol. 22, pp. 9718-9726. 2016

150. Eusebi, Leonardo H et al. “Epidemiology of Helicobacter pylori infection.” *Helicobacter,* vol. 19, pp. 1-5. 2014

151. Farrell, James J et al. “Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer.” *Gut,* vol. 61, pp. 582-8. 2012

152. Mira-Pascual, L et al. “Microbial mucosal colonic shifts associated with the development of colorectal cancer reveal the presence of different bacterial and archaeal biomarkers.” *Journal of gastroenterology,* vol. 50, pp. 167-79. 2015

153. Yu, Jun et al. “Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer.” *Gut,* vol. 66, pp. 70-78. 2017

154. Michaud, Dominique S. “Role of bacterial infections in pancreatic cancer.” *Carcinogenesis,* vol. 34, pp. 2193-7. 2013

155. “Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium.” *Nature,* vol. 401, pp. 921-3.1999

156. Neefjes, Jacques et al. “Towards a systems understanding of MHC class I and MHC class II antigen presentation.” *Nature reviews. Immunology,* vol. 11, pp. 823-36. November 2011

157. Lechler, R I. “Structure-function relationships of MHC class II molecules.” *Immunology. Supplement,* vol. 1, pp. 25-6. 1988

158. Rock, Kenneth L et al. “Present Yourself! By MHC Class I and MHC Class II Molecules.” *Trends in immunology,* vol. 37, pp. 724-737. 2016

159. Gonzalez-Galarza, Faviel F et al. “Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations.” *Nucleic acids research,* vol. 39. 2011

160. Kostic, Aleksandar D et al. “Genomic analysis identifies association of Fusobacterium with colorectal carcinoma.” *Genome research,* vol. 22, pp. 292-8. 2012

161. Wang, T. C., et al. (2016). “Helicobacter pylori in gastric carcinogenesis.” *World Journal of Gastroenterology: WJG,* vol. 22, Page 6920. 2016

162. Toprak, N Ulger et al. “A possible role of Bacteroides fragilis enterotoxin in the aetiology of colorectal cancer.” *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases,* vol. 12, pp. 782-6. 2006

163. Yang, L., et al. (2009). “Abundance of mucosa-associated bacteria revealed by increased DNA base composition in the metagenome of patients with esophageal cancer.” *BMC Genomics*, vol. 10, pp. 1-13. 2009

164. Arthur, Janelle C et al. “Intestinal inflammation targets cancer-inducing activity of the microbiota.” *Science (New York, N.Y.),* vol. 338, pp. 120-3. 2012

165. Shen, Xiang Jun et al. “Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas.” *Gut microbes,* vol. 1, pp. 138-47. 2010

166. Wu, S., et al. “A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses.” *Nature Medicine*, vol. 15. 2009

167. Sears, Cynthia L et al. “Association of enterotoxigenic Bacteroides fragilis infection with inflammatory diarrhea.” *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America,* vol. 47. 2008

168. Hooper, Samuel J et al. “Viable bacteria present within oral squamous cell carcinoma tissue.” *Journal of clinical microbiology,* vol. 44, pp. 1719-25. 2006

169. Castellarin, Mauro et al. “Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma.” *Genome research,* vol. 22, pp. 299-306. 2012

170. Engstrand, L., et al. “Helicobacter pylori and gastric cancer: state of the art.” *Helicobacter*, vol. 14, pp. 41-45. 2009

171. Yang, Yongzhi et al. “Fusobacterium nucleatum Increases Proliferation of Colorectal Cancer Cells and Tumor Development in Mice by Activating Toll-Like Receptor 4 Signaling to Nuclear Factor-κB, and Up-regulating Expression of MicroRNA-21.” *Gastroenterology,* vol. 152, pp. 851-866. 2017

172. Tsoi, H., et al. (2017). Pre- and postoperative imbalances in gut microbiota associated with colorectal cancer. Gut Microbes, 8(4), 1-12.

173. Castellarin, Mauro et al. “Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma.” *Genome research,* vol. 22, pp. 299-306. 2012

174. Wang, Q., et al. “Detection and analysis of human papillomavirus (HPV) DNA in esophageal squamous cell carcinoma”, *Scientific Reports,* vol. 7, pp. 1-9. 2017

175. Baxter, Nielson T et al. “Microbiota-based model improves the sensitivity of fecal immunochemical test for detecting colonic lesions.” *Genome medicine,* vol. 8. April 2016

176. Hajishafiha, M., et al. “Association of high-risk human papillomavirus subtypes with gastric cancer development: A case-control study in Iran”, *Asian Pacific Journal of Cancer Prevention,* vol. 18, pp. 2751-2754. 2017

177. Nakatsu, Geicho et al. “Gut mucosal microbiome across stages of colorectal carcinogenesis.” *Nature communications,* vol. 6. October 2015

178. Wu, Q et al. “Association between Helicobacter pylori infection and the risk of colorectal neoplasia: a systematic review and meta-analysis.” *Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland,* vol. 15. 2013

179. Li, Junhua et al. “An integrated catalog of reference genes in the human gut microbiome.” *Nature biotechnology* vol. 32, pp. 834-41. 2014

180. Wang, X., et al. “HLA-A*02:01 is associated with an increased risk of developing gastric and colorectal cancer.” *Cancer Medicine*, vol. 7, pp. 3586-3593. 2018

181. Wu, M. S., et al. (2019). “The major histocompatibility complex-A * 31: 01 allele and its linkage disequilibrium to gut microbiota play a role in the development of colorectal cancer.” *Scientific Reports*, vol. 9, pp. 1-9. 2019

182. Chen, H., et al. “Genetic polymorphisms in human leukocyte antigen-A, -B, and -C loci in gastric cancer patients associated with Helicobacter pylori infection.” *Gastroenterology Research and Practice,* vol. 6. 2019

183. Cheng, T. Y., et al. “Genome-wide association study identifies a new susceptibility locus for celiac disease on chromosome 3q21.” *Human Molecular Genetics,* vol. 25(17), pp. 3932-3942. 2016

184. Al-Moundhri, M. S., et al. “Impact of HLA-DRB1* 11 and HLA-DRB1* 13 on susceptibility to colorectal cancer.” *World Journal of Gastroenterology: WJG,* vol. 21(14). 2015

185. Dendrou, Calliope A et al. “Resolving TYK2 locus genotype-to-phenotype differences in autoimmunity.” *Science translational medicine,* vol. 8. 2016

186. Jiao, Y., et al. “Exome sequencing of pancreatic malignant transformation from intraductal papillary mucinous neoplasm to invasive carcinoma.” *Gut,* vol. 67(5), pp. 965-973. 2018

187. Khor, S. S., et al. “Genome-wide association study identifies five new susceptibility loci for Crohn's disease in the Japanese population.” *Nature Genetics,* vol. 48(5), Page 510. 2016

188. Gragert, Loren et al. “Fine-mapping of HLA associations with chronic lymphocytic leukemia in US populations.” *Blood,* vol. 124, pp. 2657-65. 2014

189. Derks, S., et al. “Association of human leukocyte antigen class II alleles and haplotypes with cervical intraepithelial neoplasia in Chinese women.” *Journal of Low Genit Tract Dis,* vol. 19(1), pp. 20-26. 2015

190. Wang, Li-Dong et al. “Genome-wide association study of esophageal squamous cell carcinoma in Chinese subjects identifies susceptibility loci at PLCE1 and C20orf54.” *Nature genetics,* vol. 42, pp. 759-63. 2010

191. Wu, Chen et al. “Genome-wide association study identifies three new susceptibility loci for esophageal squamous-cell carcinoma in Chinese populations.” *Nature genetics,* vol. 43, pp. 679-84. June 2011

192. Abnet, Christian C et al. “A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma.” *Nature genetics,* vol. 42, pp. 764-7. 2010

193. Zhang, Lin et al. “Integrative genomic analysis of protein kinase C (PKC) family identifies PKCiota as a biomarker and potential oncogene in ovarian carcinoma.” *Cancer research,* vol. 66, pp. 4627-35. 2006

194. Dong, J. T., et al. “Frequent inactivation of the retinoic acid receptor beta in esophageal squamous cell carcinoma.” *Cancer Research,* vol. 60(22), pp. 5901-5906. 2000

195. Huang, Y. K., et al. “Identification of a two-marker-haplotype on HLA-DRB1* 15: 02 as a powerful marker for predicting risk of cervical cancer in women infected with human papillomavirus type 18.” *International Journal of Cancer*, vol. 13. 2012

196. Lin, D. Y., et al. “HLA-DRB1 alleles and the susceptibility of systemic lupus erythematosus in Taiwanese patients.” *Rheumatology*, vol. 52(9), pp. 1539-1546. 2013

197. Tang, Weihong et al. “Genetic associations for activated partial thromboplastin time and prothrombin time, their gene expression profiles, and risk of coronary artery disease.” *American journal of human genetics,* vol. 91, pp. 152-62. 2012

198. Kamatani, Yoichiro et al. “Genome-wide association study of hematological and biochemical traits in a Japanese population.” *Nature genetics,* vol. 42, pp. 210-5. 2010

199. He, Chunyan et al. “Genome-wide association studies identify loci associated with age at menarche and age at natural menopause.” *Nature genetics,* vol. 41, pp. 724-8. 2009

200. Han, Jian-Wen et al. “Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus.” *Nature genetics,* vol. 41. 2009

201. Kumasaka, N., et al. “HLA-DRB1* 0901/0905-DQB1* 0303 haplotype is associated with Grave’s disease with ophthalmopathy in Japanese patients.” *Clinical Endocrinology*, vol. 71(3), pp. 386-389. 2009

202. He, Chunyan et al. “Genome-wide association studies identify loci associated with age at menarche and age at natural menopause.” *Nature genetics* vol. 41,6 (2009): 724-8. doi:10.1038/ng.385.

203. Cozen, W., et al. (2009). “A genome-wide screen of 1,233,113 SNPs identifies MAP3K1 as a susceptibility locus for Crohn’s disease.” *Nature Genetics*, vol. 39(4), pp. 532-536. 2009

204. Luo, Y., et al. “HLA associations with drug-induced liver injury in the Han Chinese population.” *Journal of Immunology Research*, vol. 21. 2018

205. Yang, J. J., et al. “Associations of polymorphisms in HLA-G with occurrence of gastric cancer: A case-control study.” *PLoS One,* vol. 6(7), 2011

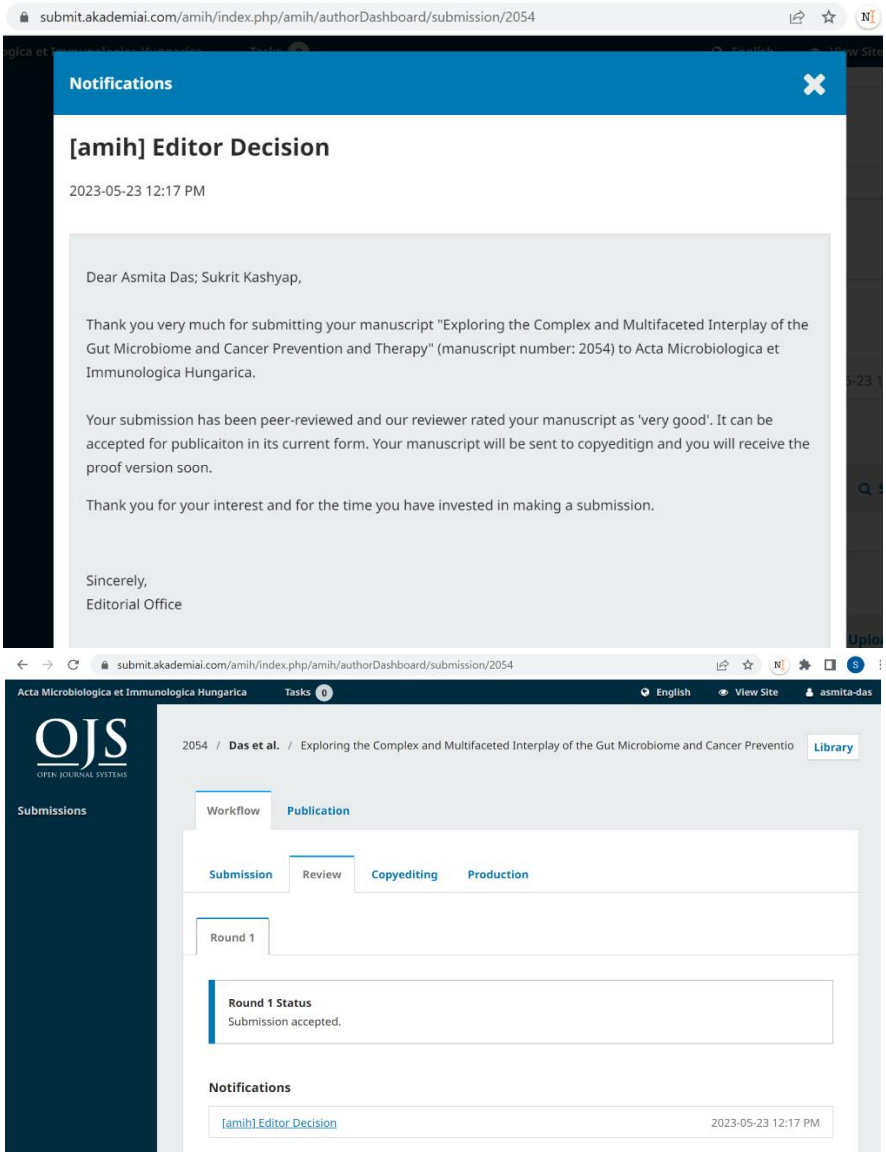
206. Hu, Zhibin et al. “A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese.” *Nature genetics,* vol. 43, pp. 792-6. July 2011

207. Miki, Daiki et al. “Variation in TP63 is associated with lung adenocarcinoma susceptibility in Japanese and Korean populations.” *Nature genetics,* vol. 42, pp. 893-6. 2010

208. Cheung, Ching-Lung et al. “Genome-wide association study identifies a susceptibility locus for thyrotoxic periodic paralysis at 17q24.3.” *Nature genetics,* vol. 44, pp. 1026-9. 2012

LIST OF PUBLICTIONS

- Sukrit Kashyap, Asmita Das, “Exploring the Complex and Multifaceted Interplay of the Gut Microbiome and Cancer Prevention and Therapy”, *Acta Microbiologica et Immunologica Hungarica*. **(Status:** Accepted for Publication**)**



4% Overall Similarity

Top sources found in the following databases:

- 3% Internet database
- Crossref database
- 2% Submitted Works database
- 2% Publications database
- Crossref Posted Content database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	esp.org Internet	<1%
2	hindawi.com Internet	<1%
3	frontiersin.org Internet	<1%
4	Mohd Rabi Bazaz, Ziaur Rahman, Insha Qadir, Tulasi Pasam, Manoj P. ... Crossref	<1%
5	mdpi.com Internet	<1%
6	ncbi.nlm.nih.gov Internet	<1%
7	Athlone Institute of Technology on 2021-08-02 Submitted works	<1%
8	University of Essex on 2023-05-26 Submitted works	<1%

9	Pioneer Academics on 2020-09-26 Submitted works	<1%
10	University of Liverpool on 2014-11-10 Submitted works	<1%
11	Inna Sekirov, Shannon L. Russell, L. Caetano M. Antunes, B. Brett Finla... Crossref	<1%
12	Janet M. Sasso, Ramy M. Ammar, Rumiana Tenchov, Steven Lemmel, ... Crossref	<1%
13	Royal Holloway and Bedford New College on 2018-03-22 Submitted works	<1%
14	karger.com Internet	<1%
15	pubmedcentral.nih.gov Internet	<1%
16	Queen Mary and Westfield College on 2023-05-15 Submitted works	<1%
17	Queen Mary and Westfield College on 2023-05-26 Submitted works	<1%
18	University of Cambridge on 2023-05-23 Submitted works	<1%
19	colloid.nl Internet	<1%
20	getd.libs.uga.edu Internet	<1%

21	journals.plos.org Internet	<1%
22	repository.helmholtz-hzi.de Internet	<1%
23	repository.ubn.ru.nl Internet	<1%
24	research.thea.ie Internet	<1%
25	seqanswers.com Internet	<1%
26	Cranfield University on 2016-01-19 Submitted works	<1%
27	Imperial College of Science, Technology and Medicine on 2017-05-01 Submitted works	<1%
28	Imperial College of Science, Technology and Medicine on 2017-05-12 Submitted works	<1%
29	Lara Bull-Otterson, Wenke Feng, Irina Kirpich, Yuhua Wang et al. "Meta... Crossref	<1%
30	Li, Yan, Jinzhi He, Zhili He, Yuan Zhou, Mengting Yuan, Xin Xu, Feifei Su... Crossref	<1%
31	Massey University on 2010-02-10 Submitted works	<1%
32	Tilburg University on 2023-05-28 Submitted works	<1%

33	University of Cape Town on 2014-08-19 Submitted works	<1%
34	University of Southern Queensland on 2017-09-13 Submitted works	<1%
35	University of Wales, Bangor on 2017-10-09 Submitted works	<1%
36	digitalcommons.lsu.edu Internet	<1%
37	pure.uva.nl Internet	<1%
38	coursehero.com Internet	<1%