CHARACTERIZATION OF GUT MICROBIOTA OF THE INDIAN TYPE 2 DIABETIC PATIENTS: SUSTAINABILITY OF HUMAN HEALTH

Thesis submitted for the award of the Degree of

Doctor of Philosophy (Science)

Jadavpur University

By

Debjit De

Registration Index No.: 200/15/Life Sc./24

Department of Life Science and Biotechnology

Jadavpur University

Kolkata – 700032, India

2022

যাদবপুর বিশ্ববিদ্যালয়
কলকাতা-৭০০০৩২, ভারত

*JADAVPUR UNIVERSITY
KOLKATA-700 032, INDIA

DEPARTMENT OF LIFE SCIENCE AND BIOTECHNOLOGY

CERTIFICATE FROM THE SUPERVISOR

This is to certify that the thesis entitled "***Characterization Of Gut Microbiota Of The Indian Type 2 Diabetic Patients: Sustainability Of Human Health***" submitted by **Sri Debjit De** who got his name registered on **15.10.2015** with **Registration Index No.- 200/15/Life Sc./24** for the award of Ph.D. (Science) degree of Jadavpur University, is absolutely based upon his own work under the supervision of **Dr. Paltu Kumar Dhal, Assistant Professor, Department of Life Science and Biotechnology, Jadavpur University** and that neither this thesis nor any part of it has been submitted for either any degree / diploma or any other academic award anywhere before.

Paltu Kr. Dhal 29/7/22

Signature of supervisor with date and official seal

Dr. Paltu Kumar Dhal
Assistant Professor, Ph. D
Life Science & Biotechnology Dept.
Jadavpur University, Kolkata - 32

Dedicated To . . .

My Family

Acknowledgement

*Undertaking doctoral research has been a lifetime experience for me, and it would never have been possible without continuous and spontaneous support, supervision, and guidance from many people. Hence, it gives me immense pleasure to acknowledge all those who have guided, contributed and supported me in carrying out my doctoral research entitled "**Characterization of Gut Microbiota of the Indian Type 2 Diabetic Patients: Sustainability of Human Health**" at the Jadavpur University, Kolkata, India.*

*First of all, I would like to express my sincere gratitude and indebtedness to my supervisor **Dr. Paltu Kumar Dhal**, Assistant Professor, Department of Life Science and Biotechnology, Jadavpur University, Kolkata, India, for his keen interest, perennial encouragement and scholastic guidance all through the tenure of this investigation. I have been amazingly fortunate to have the advisors who gave me the freedom to explore on my own and, at the same time, the guidance to recover when my steps faltered. I would also like to express my gratitude to **Dr. (Prof). Subhankar Chowdhury**, Endocrinologist, MBBS, DTM, DM – Endocrinology and MRCP of Endocrine Department of IPGMER and SSKM Hospital, Kolkata, for his encouragement and guidance all through the tenure of this investigation.*

*I owe my deepest gratitude to **Prof. Ratan Gachhui**, Professor and Head of the Department, Department of Life Science and Biotechnology, Jadavpur University, for extending his close cooperation and pursuing my research work. My special thanks to **Prof. Parimal Karmakar**, **Prof. Biswadip Das** and **Dr. Arunima Sengupta**, faculty members of our Department, for their wholehearted and unfailing assistance throughout my research work.*

I am immensely grateful to all of my lab mates ***Mr. Tilak Nayak****,* ***Mrs. Pranamita Kunda****,* ***Mr. Indraneel Sengupta****,* ***Miss Gargi Das and Miss Annesha Guray*** *and all the Departmental staff of Jadavpur University for their goodwill and cooperative interaction. I would like to extend a special thanks to my other Departmental research scholars and dear friends* ***Mr. Avishek Mukherjee****,* ***Mr. Uttariya Roy****,* ***Mr. Soumyadev Sarkar****,* ***Mr. Subhadeep Das****,* ***Miss Jhilam Majumdar****,* ***Mrs. Rubia Parvin and Miss Arunima Mondal*** *Department of Life Science & Biotechnology, Jadavpur University for their undue cooperation and support at various phases during my research work.*

*I will be failing my duty if I do not acknowledge the kind cooperation of my parents, who have always been a source of inspiration for me to work hard. I would also like to thank my very close friends (****Mr. Arnab Mukherjee****,* ***Mr. Agni Mitra*** *and* ***Mr. Rabisankar Singh Sardar****), all my relatives, and well-wishers for their blessings and support wishes and everything they did for me so that I could carry out this research work successfully. Above all, I would like to thank the Almighty for his enormous blessings.*

Thanking you,

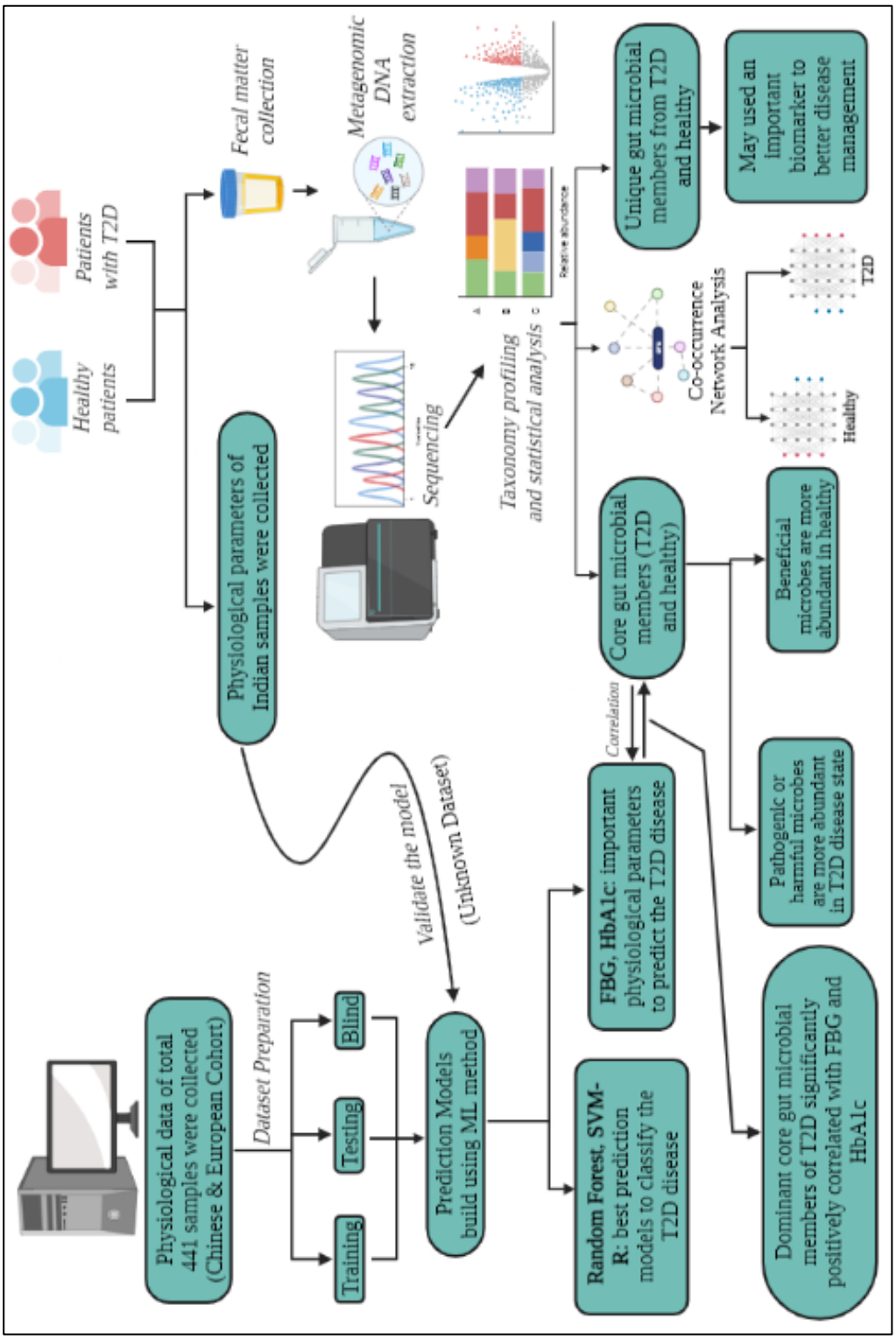
DEBJIT DE

CONTENTS

LIST OF TABLES

LIST OF FIGURES

GRAPHICAL ABSTRACT



ABSTRACT

Index No.: 200/15/Life Sc./24

Title: Characterization of gut microbiota of the Indian Type 2 Diabetic patients: Sustainability of Human Health

Type 2 Diabetes (T2D) is a severe global public health issue due to the rising incidence of overweight/obesity and poor lifestyles in the twenty-second century. The relationship between changed gut flora and a higher incidence of T2D is well addressed by several epidemiological studies. Understanding of the causative gut microbiota and their interaction with host as well as important host physiological parameters for early detection of disease are emerging research topic for the better management of T2D. With this view this study aimed to develop efficient models for identifying essential physiological markers for improved T2D classification using machine learning algorithms. In addition to that using amplicon metagenomic approaches, an effort has also been made to understand the alterations in core gut microbial members in Indian T2D patients with respect to their control (NGT). Our data indicate the level of fasting blood glucose (FBG) and glycated hemoglobin (HbA1c) were the most useful physiological indicators while Random Forest (SEN: 1.00; ACC: 0.76) and Support Vector Machine with RBF kernel (SEN: 1.00; ACC: 0.67)were effective predictions models for identifications of T2D.*Prevotella_9*, *Alloprevotella*, *Bacteroides*, *PrevotellaIncertaeSedis*, *Rikenellaceae RC-9 gut group*, *Eubacterium*, *UCG-002*, *Phascolarctobacterium*, and *Asteroleplasma* genera were the most dominant gut microbial members in T2D disease. In contrast *Prevotella*, *Roseburia*, *LachnospiraceaeIncertaeSedis*, *Butyrivibrio*, *Faecalibacterium*, *Klebsiella*, *Succinivibrio*,*Megasphaera*, *SelenomonadaceaeIncertaeSedis*, *Treponema*, and*Akkermansia* genera were the most dominant in healthy individuals. The

dominating gut microbialmembers *Allopreotella*, *Rikenellaceae RC9 gut group*, *Haemophilus*, *Ruminococcustorques group*, etc. in Indian T2D patients showed a strong association with both FBG and HbA1c.These members have been reported to have a crucial role in gut barrier breakdown, blood glucose, and lipopolysaccharide level escalation, or as biomarkers. While the dominant NGT microbiota (*Akkermansia*, *Ligilactobacillus*, *Enterobacter*, etc.) in the colon has been shown to influence inflammatory immune responses by acting as an anti-inflammatory agent and maintaining the gut barrier. The co-occurrence network analysis indicates that changes in network complexity in T2D lead to variations in the different gut microbial members compared to NGT. *Firmicutes*, *Bacteroidota*, *Proteobacteria*, *Actinobacteriota* and *Spirochaetota* gut microbial phyla network were identified as keystone taxa in T2D co-occurrence network. In contrast, *Bacteroidota*, *Firmicutes*, *Proteobacteria*, *Patescibacteria* and *Desulfobacterota* as keystone microbial phyla for healthy. The metabolic pathway prediction revealed that abundant microbial metabolic pathways in T2D diseases condition are mostly associated with insulin resistance and inflammation. This study may provide a better understanding of the gut-microbial diversity in Indian T2D patients and show the way for the development of valuable diagnostics strategies to improve the prediction and modulation of the T2D.

Paltu Kr. Dhal
29/7/22

Dr. Paltu Kumar Dhal
Assistant Professor, Ph. D
Life Science & Biotechnology Dept.
Jadavpur University, Kolkata - 32

Debjit De
29/07/2022

Signature

General Introduction

Type 2 diabetes (T2D) is a category of metabolic disorders characterized by persistent hyperglycemia caused by decreased insulin secretion, impaired insulin action, or both processes at the same time, resulting in long-term consequences (Kerner et al., 2014). Chronic micro and macrovascular problems are related with persistent hyperglycemia. Diabetes puts people at a higher risk of having a variety of potentially fatal health problems, such as vascular damage to the heart, eyes, kidneys, and nerves (Ballan et al., 2020). T2D is characterized by insulin resistance and pancreatic-cell dysfunction, resulting in unstable hyperglycemia (Hameed et al., 2015; Akbari et al., 2020). Insulin secretion is reduced in the case of -cell dysfunction, restricting the body's ability to maintain physiological plasma glucose levels, whereas insulin resistance contributes to increased glucose synthesis in the liver and decreased glucose absorption in the muscle, adipose tissue, and liver (Galicia-Garcia et al., 2020). It has a complicated and multifaceted etiology that includes genetic and environmental factors, and it mainly affects adults in their forties, while there has been an increase in the incidence of diabetes in children and young people. Genetic predisposition, age, obesity, physical inactivity, a previous diagnosis of pre-diabetes or gestational diabetes (DMG), an inadequate diet, and stress are all risk factors for T2D (Kolb et al., 2017; Ballan et al., 2020).

According to the International Diabetes Federation (IDF) 2019, around 463 million individuals (20–79 years) in the world, representing to 9.3 percent of the world population, are living with diabetes (Williams et al., 2020); it is anticipated that this figure will increase to 700 million in 2045. In 2019, 374 million people were at risk of having T2D, and this proportion has increased in several countries. The biggest number of patients with diabetes are between 40 and 59 years old. For every two persons with diabetes, one does not know that they have the disease, or 263 million people. The most

probable factors for the increased incidence of diabetes include social and economic changes, including changes towards a sedentary lifestyle, an imbalanced diet leading to the degradation of nutritional status, an increased prevalence of being overweight, and growing urbanization. On the other hand, greater health care has enhanced the life expectancy of patients with diabetes (Cho et al., 2018).

In general, T2D is associated with higher levels of pro-inflammatory cytokines, chemokines, and inflammatory proteins. Patients with T2D frequently have a high-fat diet that is related with increased lipopolysaccharide synthesis by Gram-negative bacteria in the gut, and its transit to the blood circulation promotes inflammatory responses that lead to insulin resistance (Galicia-Garcia et al., 2020). Both genetic and epigenetic variables have been involved in the development of inflammation associated with T2D. The disruption of the epigenetic regulatory mechanisms that control the expression of a significant number of genes has been linked to the pathogenesis of numerous disorders connected to the immune system, including T2D. It is well recognized that the existence of a pro-inflammatory phenotype is highly related with the development of insulin resistance, β cells, and vascular problems in a patient with T2D disease (Bicsak et al., 2017). Hyperglycemia and dyslipidemia generate aberrant epigenetic modifications that increase the activation of the main inflammatory pathways and that lead to the formation of a state of low-grade chronic inflammation in T2D (Ballan et al., 2020).

In recent few years, the medical data were rising tremendously due to digitalisation which is known as big data. It contains massive number of data that includes hospital records, patient medical records, and results of medical examinations and that cannot be processed neither by human or by conventional computers (Chahal

and Gulia 2016). The manual decisions can be erroneous and susceptible. Although the concealed pattern in such data can be overlooked, this can be major adverse impact on decision making for the right treatment of a patient. Hence automated computational prediction approaches are vital for the early identification of diabetes (Huang et al., 2007; Contreras et al., 2018; Chaki et al., 2020). Machine learning (ML) methodology is a one of the most useful automated computational prediction methods which can be implemented on all accessible data and can helps in prediction of diabetes disease states. ML algorithms are less time intensive and swiftly process the symptoms using current data and knowledge to help in early identification of disease. Early identification of the diabetes condition is very much significant since the disease might create more and severe complication with time (Heydari et al., 2010).

There are recent several studies endorsed the discrimination between T2D and normal person (normal glucose tolerance, NGT) using different ML models based on patients' physiological conditions (Zhang et al., 2021). However, most of those studied models made their observations based on the limited number of samples from a single geographical location. Additionally, none of them attempted to identify important physiological parameters out of their prediction model that significantly differentiates T2D disease from NGT. While best prediction model with high accuracy essentially needed a large sample size with variant coverage (Wei et al., 2013, Arbabshirani et al., 2017). However, the deep study on predicting the most important influencing physiological parameters is incompletely explained while none from India. Nevertheless, this study attempted to make the contribution in introduce the best ML methods for better prediction of T2D and NGT, and identify the most important physiological parameters to detect the disease condition irrespective of their geographical location.

The gut microbiota plays an essential metabolic function, whether it be through its ability to decompose non-digestible carbohydrates and to synthesis micronutrients or through its involvement with the immune system (Rowland et al., 2018). The term microbiota refers to the assemblage of living microorganisms, including bacteria, archaea, protozoa, fungus, and algae, that is present in a given habitat (Berg et al., 202). Recently, changes in the human intestinal microbiota have been connected with pathological states such as obesity and other metabolic illnesses such as T2D, metabolic syndrome, and insulin resistance (Munoz-Garach et al., 2016; Marchesi et al., 2016). Among the mechanisms that relate the intestinal microbiota with T2D and insulin resistance, there is an increase in the permeability of the intestinal barrier, resulting in metabolic endotoxemia. In addition, an increased generation of branched-chain amino acids (BCAA), imidazole propionate, and of trimethylamine N-oxide (TMAO) as well as interaction with bile acids, changes in fatty acid metabolism, and intestinal hormones also occur. These alterations may contribute to higher levels of obesity and poor insulin signaling (Munoz-Garach et al., 2016; Heianza et al., 2019; Gurung et al., 2020).

Studies on the Indian population's gut microbiota reveal that it differs dysbiosis of microbial members from western population in terms of composition (Patil et al., 2012; Bhute et al., 2016, Gaike et al., 2020). Therefore, we were the first to provide the preliminary information on the functional role of gut microbiome of Indian T2D patients from the eastern region of the Indian Subcontinent, especially, in and around the Kolkata, West Bengal, with almost similar dietary status. This is because the Indian population has a particular gut microbial characteristics (Bhute et al., 2016). With 69.1 million projected diabetic patients in 2015, India is one of the world's diabetes capitals (International Diabetes Federation, 2015). Although several possible causes are put up, the rapidly spreading diabetes epidemic in India is not fully understood. The

characteristics of Indian diabetes patients are distinct and paradoxical when compared to those of diabetic patients in the west. They include the possibility of an increased genetic propensity (Ramachandran et al., 2012), intrauterine undernutrition (thrifty phenotype) leading to an epigenetic propensity (Yajnik, 2001), the onset of diabetes at an earlier age and at a lower body mass index (BMI) than in white Caucasians, and more (Yajnik, 2004). Rapid changes in the economy, nutrition, and rural-urban migration appear to be contributing factors in the development of diabetes in this population (Anjana et al., 2011).

Since T2D is the variation of the disease that affects the most diabetes people and that is usually related with obesity and cardiovascular problems, efforts have been undertaken to find new medicines to control and prevent the condition (Yaribeygi et al., 2019). With the introduction of the metagenomics techniques, the whole-genome sequencing of all the DNA included in a sample as well as a taxonomic inquiry at the species and strain level became possible, providing a functional profile of the metabolic pathways present in a community. With these traits, a greater knowledge of the association between the gut microbiota and T2D should lead to breakthroughs in therapeutic techniques and the development of new medicines, such as the use of probiotics.

Chapter 1

Literature Review on Association of Gut Microbial Members with The Type 2 Diabetes (T2D) Disease

1.1. Introduction

Type 2 Diabetes (T2D) is a diverse metabolic disorder characterized by elevated blood glucose levels and insulin resistance that leads to sedentary living and excess body weight. The primary cause of this disease is a reduction in insulin-producing beta cells, which promotes insulin resistance and hepatic glucose production. The aetiology of T2D disease is strongly linked to environmental, hereditary, and to a lesser extent, geographical factors (due to food habits, lifestyle, and so on) (Guo et al., 2016; Petersen et al., 2018). Chronic and low-grade inflammation, as well as cytokine production via lipotoxicity and promoting macrophage infiltration into adipose tissue by changing lymphocyte cells (B-cell and T-cell), are the hallmarks of this metabolic disease (Yoon et al., 2006; Boulangé et al., 2016).

In recent studies, intestinal flora was identified as the largest and most complex organ composed of more than 1000 microbial species and many studies have reported that there is a relationship between various metabolic immune disorders and intestinal microbial dysbiosis (Shreiner et al., 2015). The human body contains trillions of microorganisms (3.8×10^{13}) with a complex ecosystem that resides in our bodies during and after birth (Baothman et al., 2016; Meijnikman et al., 2018). They are colonized on all surfaces of the human body that are exposed to the environment, with most residing in the gastrointestinal (GI) tract (10^{14}). Microbial communities are more similar at particular body sites among different subjects than in the same subject at different body sites; for example, oral microbiota of different individuals are more similar than microbial communities of skin and mouth in a single individual (Costello et al., 2015; Indias et al., 2016). This fact gained the attention of many scientists and they proceeded to characterize the gut microbial communities profile in disease states with compare to

healthy people by using Next Generation Sequencing (NGS) technology. Specifically, Metagenomics 16S rRNA variable region sequencing and Whole Genome Shotgun (WGS) sequencing substantially help to identify their complexity in the gastrointestinal tract. The Human Microbiome Project (USA) (website), and metaHIT Consortium (Europe) (website) were developed to help in the characterization and deep understanding of the shifts of gut-microbiota profile in disease states with compare to healthy and also help to identification of key gut-microbial biomarker to predict the disease condition to improve the disease diagnosis with already established methods (Morgan et al., 2012).

So, in this review, we retrieved the reported taxonomy information about the gut microbial diversity in T2D disease states in various geographical locations (Africa, Asia, Australia, Europe, and the US) because they have their different natural habitat like food consumption, lifestyle, socioeconomics, environment, etc. By using that information, we tried to create a clear global picture of the gut-microbial diversity pattern in T2D conditions in different geographical locations for the improvement of diseases prediction and diagnosis with already established methods. Along with that, here we summarized the probable future therapeutic strategies to mitigate the condition.

1.2. Gut microbial association with Type 2 Diabetes (T2D)

The intestine harbouring trillions of microorganisms is important to the metabolic health of the host because the commensal microbiota of a healthy gut is linked with vital activities, such as the production of water-soluble vitamins, digestion, harvesting energy from food components, xenobiotic degradation, and production of metabolites (Patterson et al. 2016). It is also proved that these vital activities can support and promote the functional capacity of the gut epithelium and intestinal barrier integrity

respectively as well as provide protection from other harmful organisms (Xu et al., 2016). Gut inflammation, use of antibiotics, menopause, toxin, stress, and others triggers disruption of the host microbiota equilibrium called dysbiosis cause disorders such as cardiovascular, autoimmune disorders, autism, obesity, and T2D (Hegde et al. 2018; Battson et al. 2018; Opazo et al. 2018; Sgritta et al. 2019; Bianchi et al. 2018; Karlsson et al. 2013). T2D may be linked to the composition of the intestinal microbiota is becoming clearer with more studies showing the involvement of microbiota in obesity and their role in insulin resistance and is directly responsible for the induction of low-grade inflammation (Roager et al. 2017). The main dysbiosis condition observed in T2D patients have reduced butyrate-producing bacteria (like *Faecalibacterium prausnitzii* and *Roseburia intestinalis*), a pro-inflammatory environment with increased expression of microbial genes involved in oxidative stress, serum lipopolysaccharide (LPS) concentration, and increased intestinal permeability, on the other hand, reduced expression of genes involved in vitamin synthesis (Sabatino et al., 2017). The LPS of Gram-negative bacteria can stimulate the inactive immune system by activating toll-like receptors with inflammatory cytokines production. Along with that LPS further promotes the activation of the c-Jun N-terminal kinase pathways and nuclear factor kappa-B both of these pathways are associated with insulin resistance and the deficiency of insulin signalling in the muscle, liver, adipose tissue, and hypothalamus (Figure 1.1) (Caricilli and Saad 2013; Newsholme et al. 2016). The previous report from Chinese T2DM patients demonstrated a decrease in short-chain fat acids (SCFA) producing bacteria, mainly butyrate-producing bacteria *Eubacterium rectale*, *F. prausnitzii*, *Clostridiales sp.*, and *R. intestinalis*. SCFA involves in the anaerobic breakdown of dietary fibre, protein, and peptides when the gut microbiota is in dysbiosis is directly related to the alteration of SCFA production (Alexander et al.

2020). They also help to produce acetate, propionate, and butyrate where acetate and propionate are mostly produced by the Bacteroidetes phylum, and butyrate is produced from phylum Firmicutes (Baxter et al. 2019).

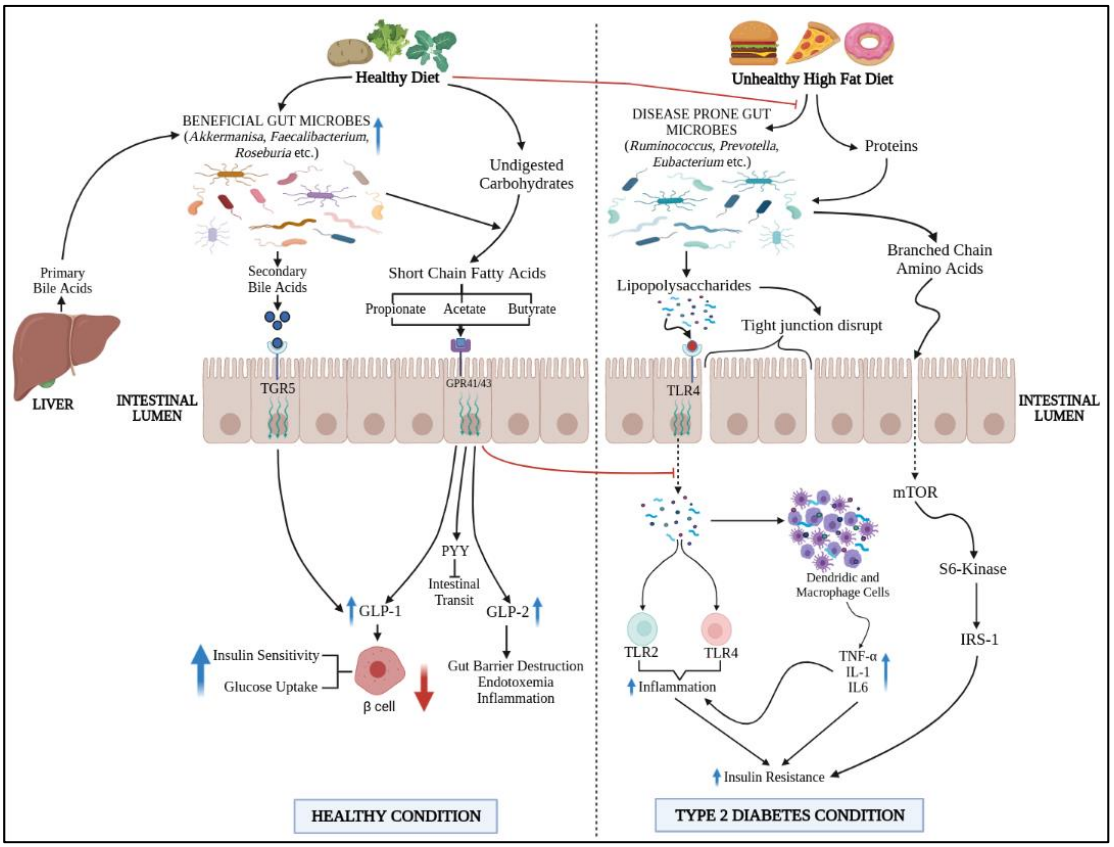


Figure 1.1 – Schematic diagram of gut microbial association under healthy and diseased conditions. Due to alteration of dietary pattern and intestinal microbial members, the junction protein expression down regulated which increased gut permeability at epithelial layers. This activity raises the level of lipopolysaccharides (LPS) in the bloodstream, resulting in metabolic endotoxemia and insulin resistance.

SCFA and butyrate improve insulin sensitivity and secretion by stimulating the secretion of peptide 1 like glucagon (GLP-1) and reducing the inflammation of adipocytes (Ríos-Covián et al., 2016; Tolhurst et al., 2012; Wang et al., 2015; Gao et al., 2018). So, we can conclude from all these studies that factors that can increase levels of SCFA, especially butyrate, are important for relieving T2D symptoms.

1.3. Composition of the gut microbial community of Type 2 Diabetic patients from varying geographical regions

Different ethnicity, feeding habits, and socioeconomic status in different geographical locations seem to increase the diabetes burden. For example, Asiatic populations reported a lower prevalence of disease compared with European populations (Connolly et al., 2019). While in urban Western societies, the rate of T2D is increased due to food selection, obesity, physical inactivity, and lifestyle (Connolly et al., 2000). However, a recent cross-sectional study indicates that variations of T2D disease between India's different states showed ranging from 4.3% to 11.8% and the main reason behind this is low socioeconomic status groups living in unfavourable urban areas showed a higher rate of diabetes (Anjana et al., 2021). According to WHO [21], Germany had the highest rates of T2D disease in Europe in the year 2019 (15.3%) and followed by Portugal (14.2%), Malta (2.2%), while on the other hand, Ireland showed 4.4%. It seems to be that the urban development and not the urbanization by itself determine disease prevalence (Anjana et al., 2021).

In recent several studies indicated that along with these different factors, differential gut microbial abundance in the host gastrointestinal tract also plays an important role in the progression/severity of the disease by modulating the normal pathways in the host and elevating the level of insulin resistance. So, we intensively study the gut microbial community-related articles reports from five different geographical locations Africa, Asia, Australia, Europe, and the US zone due to their different natural habitat like food consumption, lifestyle, socioeconomics, environment, etc. for better and deeply understand the microbial diversity (Table 1.1). During the literature survey, we included only those studies that had case-control studies of fecal

gut-microbiota that had included only T2D (case) and healthy (control). No studies were included that only presented abstracts with no full report; also, ongoing clinical, in vivo, or in vitro studies, review papers and other than T2D disease (Gestational Diabetes Mellitus (GDM), Type 1 Diabetes Mellitus (T1D) or metabolic diseases) studies were also excluded in this review. To make a global picture of gut-microbial patterns in T2D disease conditions we compared these five geographical locations with each other using the Venn diagram (Figure 1.2). This review will help to get a clear picture of the variation of gut microbial members in disease states globally.

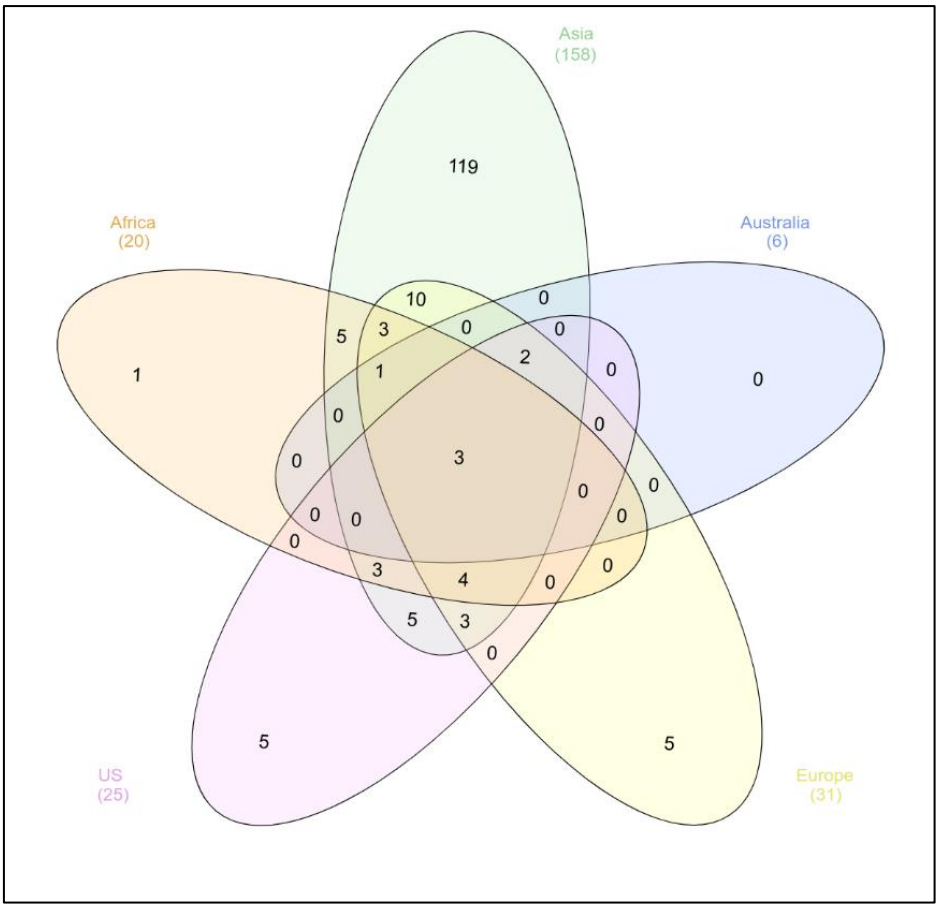


Figure 1.2 – Venn diagram for zone specific and common gut microbial members from five different geographical location. According to the Venn diagram, 27 gut microbes were common for both T2D and NGT groups, and 7 and 7 gut microbes were unique for T2D and NGT groups respectively.

Table 1.1 – List of the studies included in this review

Literature	Tittle	Year	Zone
Qin et al.	A metagenome-wide association study of gut microbiota in type 2 diabetes	2012	ASIA
Zhang et al.	Human Gut Microbiota Changes Reveal the Progression of Glucose Intolerance	2013	ASIA
Qian et al.	Association analysis of dietary habits with gut microbiota of a native Chinese community	2018	ASIA
Zhao et al.	The Alteration in Composition and Function of Gut Microbiome in Patients with Type 2 Diabetes	2020	ASIA
Zhang et al.	Characteristics of the gut microbiome in patients with prediabetes and type 2 diabetes	2021	ASIA
Bhute et al.	Gut microbial diversity assessment of Indian type-2-diabetics reveals alterations in eubacteria, archaea, and eukaryotes	2017	ASIA
Gaike et al.	The Gut Microbial Diversity of Newly Diagnosed Diabetics but Not of Prediabetics Is Significantly Different from That of Healthy Nondiabetics	2020	ASIA
Das et al.	Alterations in the gut bacterial microbiome in people with type 2 diabetes mellitus and diabetic retinopathy	2021	ASIA
Navab-Moghadam et al.	The association of type II diabetes with gut microbiota composition	2017	ASIA
Ahmad et al.	Analysis of gut microbiota of obese individuals with type 2 diabetes and healthy individuals	2019	ASIA
Wang et al.	A comparative study of microbial community and functions of type 2 diabetes mellitus patients with obesity and healthy people	2020	ASIA
Zhang et al.	Decreased Abundance of *Akkermansia muciniphila* Leads to the Impairment of Insulin Secretion and Glucose Homeostasis in Lean Type 2 Diabetes	2021	ASIA
Hoang et al.	Metagenomic 16S rDNA amplicon data of microbial diversity of guts in Vietnamese humans with type 2 diabetes and nondiabetic adults	2020	ASIA
Naderpoor et al.	Faecal microbiota are related to insulin sensitivity and secretion in overweight or obese adults	2019	Australia
Lecamwasam et al.	Gut Microbiome Composition Remains Stable in Individuals with Diabetes-Related Early to Late Stage Chronic Kidney Disease	2020	Australia
Karlsson et al.	Gut metagenome in European women with normal, impaired and diabetic glucose control	2013	EUROPE
Larsen et al.	Gut Microbiota in Human Adults with Type 2 Diabetes Differs from Non-Diabetic Adults	2010	EUROPE
Sroka-Oleksiak et al.	Metagenomic Analysis of Duodenal Microbiota Reveals a Potential Biomarker of Dysbiosis in the Course of Obesity and Type 2 Diabetes: A Pilot Study	2020	EUROPE
Lambeth et al.	Composition, Diversity and Abundance of Gut Microbiome in Prediabetes and Type 2 Diabetes	2015	US
Almugadam et al.	Alterations of Gut Microbiota in Type 2 Diabetes Individuals and the Confounding Effect of Antidiabetic Agents	2020	US
Doumatey et al.	Gut Microbiome Profiles Are Associated with Type 2 Diabetes in Urban Africans	2020	Africa
Afolayan et al.	Insights into the gut microbiota of Nigerian elderly with type 2 diabetes and non-diabetic elderly persons	2020	Africa

From the comparison study, we found three gut-microbial families as common in all geographical locations *Lachnospiraceae*, *Ruminococcaceae*, and *Veillonellaceae*. Gut-microbial family *Streptococcaceae* is found as common in Africa, Asia, Australia, and Europe; on the other hand, *Bifidobacteriaceae* and *Fusobacteriaceae* are T2D associated gut-microbial families were found mainly in Asia, Australia, Europe, and the US. *Bacteroidaceae*, *Enterobacteriaceae*, *Lactobacillaceae*, and *Prevotellaceae* gut-microbial families were found as common members in Africa, Asia, Europe, and the US. *Clostridiaceae*, *Desulfovibrionaceae*, and *Eubacteriaceae* T2D gut-microbial families were found as common members in Africa, Asia, and Europe; whereas *Coriobacteriaceae*, *Peptostreptococcaceae* and *Succinivibrionaceae* gut-microbial families were found as common T2D gut-microbial members in Africa, Asia, and the US. *Coprobacillaceae*, *Erysipelotrichaceae*, and *Tannerellaceae* T2D gut-microbial families were also found as common members in Asia, Europe, and US respective geographical locations. When we zoom in on this T2D associated gut-microbial

diversity picture we found *Acidaminococcaceae, Christensenellaceae, Eggerthellaceae, Eubacteriales* Incertae Sedis, and *Oscillospiraceae* share as common T2D associated gut-microbial members in both Africa and Asia geographical locations. In the same manner *Akkermansiaceae, Bacillales* Family XI. Incertae Sedis, *Carnobacteriaceae, Moraxellaceae, Pseudomonadaceae, Rikenellaceae, Sphingomonadaceae, Staphylococcaceae, Sutterellaceae,* and *Synergistaceae* families were found as common T2D gut-microbial members in Asia and Europe. *Acidobacteriaceae, Elusimicrobiaceae, Enterococcaceae, Methanobacteriaceae*, and *Peptoniphilaceae* gut-microbial families were found as common T2D gut-microbial members in both Asia and US.

From the literature survey and Venn diagram, we observed that there are few T2D associated gut-microbial members present in each zone. For example, total of 119 T2D associated gut microbial members were found from the pool of reported data in Asia Zone specific locations, which includes *Actinomycetaceae*, *Aerococcaceae*, *Halobacteroidaceae, Corynebacteriaceae*, *Enterococcaceae*, *Methanomicrobiaceae*, *Odoribacteraceae, Spirochaetaceae,* etc. (Qin et al., 2012; Zhang et al., 2013; Qian et al., 2017; Bhute et al., 2017; Navab-Moghadam et al., 2017; Ahmad et al., 2019; Gaike et al., 2020; Wang et al., 2020; Zhao et al., 2020; Zhang et al., 2021; Das et al., 2021; Zhang et al., 2021). In the same manner, *Bradyrhizobiaceae, Chrysiogenaceae, Coriobacteriales, Fabaceae*, and *Promicromonosporaceae* T2D associated gut-microbial families were found in Europe zone specific (Karlsson et al., 2013; Larsen et al., 2010; Sroka-Oleksiak et al., 2020); whereas as US zone-specific *Eubacteriales, Muribaculaceae, Pseudonocardiaceae, Saprospiraceae,* and *Streptococcus* these gut-microbial members are found associated with T2D (Lambeth et al., 2015; Almugadam et al., 2020). In Africa, *Mogibacteriaceae* family is found in that particular zone-

specific T2D-associated gut microbial member (Doumatey et al., 2020; Afolayan et al., 2020). In summary from the literature survey, we acquired the knowledge about the global pattern of gut-microbial diversity in T2D disease states from different geographical locations, which may play an important role in the development of T2D disease, and also this review suggests that they could be a strong target to improve the diagnosis of disease along with already established methods.

1.4. Future direction of possible microbiome based strategies for T2D prevention

1.4.1. Microbes based

From our literature survey data *Anaerostipes*, *Blautia*, *Coprococcus*, *Epulopiscium*, *Lachnospira*, *Marvinbryantia*, *Oribacterium*, *Pseudobutyrivibrio*, and *Roseburia* under the family *Lachnospiraceae*; *Faecalibacterium* and *Ruminococcus* under the family *Ruminococcaceae* gut microbial genera were reported as potential beneficial gut microbial members associated with healthy individuals. Several reports stated that the reduction of these microbes in the T2D gut-microbiome with compares to healthy disrupt the host-microbiota homeostasis which provides too many human diseases beyond the digestive system (Gurung et al., 2020). So, if we somehow increase the abundance of those important beneficial gut-microbial members in diseases states, the disease can be diagnosed and cured very well. For example, Bacteroides under the family *Bacteroidaceae* and *Ruminococcus* under the family *Ruminococcaceae* are the animal-based gut-microbial members while *Prevotella* under the family *Prevotellaceae* is a plant-based gut-microbial member. Those gut microbial members are important for human health as well as protection from gut barrier destruction (Gurung et al., 2020). Gut-microbial members Clostridia, *Faecalibacterium*, *Roseburia*, *Butyricoccus*, Lactobacillus, and Bifidobacterium are butyrate-producing microbes, play important

beneficial roles in the host by anti-inflammatory, anti-tumorigenic, and pathogen exclusion activity. These microbes are represented as the most beneficial microbial genera which are frequently reported in several recent T2D studies (Cummings et al., 2002; Ha et al., 2022). Bifidobacterium, according to our literature survey we gained knowledge about its negative association and potential protection against the T2D disease. From the survey, we also understood that this potential gut-microbial member has an important role in the improvement of glucose tolerance and till now not been used as a probiotic against the T2D disease (Moya-Perez et al., 2015; Kikuchi et al., 2018). There are several reports where researchers suggested that introducing the Bifidobacterium as a probiotic into the human gut in the T2D disease state, will improve human health by protecting from disease and several animal studies support this thought (Wang et al., 2015; Aoki et al., 2017; Kikuchi et al., 2018). Along with the *Bifidobacterium, Bacteroides* and *Roseburia* gut-microbial genera are also reported as beneficial members of the human host and have the protective ability against the T2D disease and also play an important role in the improvement of glucose tolerance (Cano et al., 2012; Yang et al., 2017). Recently in some research papers, *Faecalibacterium* gut-microbe was reported as a popular probiotic for colitis through improving the haptic function and decreasing the liver fat inflammation in mice models and is highly abundant in healthy individuals with compare to T2D disease (Zhang et al., 2013; Karlsson et al., 2013; Graessler et al., 2013; Remely et al., 2014; Rossi et al., 2015). Another beneficial gut-microbe *Akkermansia*, a mucin-degrading agent that is present in the mucus layer of the gastrointestinal tract, is reported to play an important role in the prevention of high-fat diet-induced metabolic disorders (Couzin-Frankel, 2010). The beneficial effect of *Akkermansia* on host glucose metabolism was first reported in animal models and then in human studies. Several reports stated that these beneficial

activities enhanced the intestinal levels of the gut barrier, and gut peptide secretion (Tanca et al. 2017; Li et al. 2020). It is also reported that this human mucus colonizer can be used for the prevention or treatment of obesity and its associated metabolic disorders (Sun and Chang, 2014). Although some recent reports indicate that a decrease in this genus in diabetes is associated with inflammation and metabolic disorders in mice models it can be used as a biomarker for impaired glucose tolerance (Anhê et al. 2015; Schneeberger et al. 2015; Sonnenburg and Bäckhed 2016; Plovier et al. 2017). The diversity of Lactobacillus microbe in the human gut is high among the other potential probiotic microbes. There is one report that stated that restricting the lifelong food intake increase the Lactobacillus abundance (Sun and Chang, 2014; Kesika et al., 2019). Although some other studies reported that this microbe has a positive impact on the human host but this beneficial effect of probiotics are species specific and also related to hosting physiology. For example, *L. plantarum, L. reuteri, L. casei, L. curvatus, L. gasseri, L. paracasei, L. rhamnosus* and *L. sakei* species are showed a beneficial effect in mice model of T2D disease (Naito et al., 2011; Fak et al., 2012; Park et al., 2013; Okubo et al., 2013; Park et al., 2015; Lim et al., 2016; Martinic et al., 2018; Dang et al., 2018). Also, few species of Lactobacillus have been tested as probiotics. For example, *L. plantarum* present in fermented food products improves the glucose metabolism in diet-induced and genetic animal models of T2D (Martinic et al., 2018; Lee et al., 2018; Balakumar et al., 2018). These gut-microbial members play a wide variety of important beneficial functions that includes synthesis of some nutritional factors (eg. Vitamins); detoxification of the human host from some harmful xenobiotics and improve the intestinal immune system; providing signals for maintaining gut integrity by secretion of anti-microbial products, which prevent the pathogenic bacteria in the development of their colonization (Makishima et al., 2002; Cipriani et al., 2010).

1.4.2. Diet Based

The composition of the microbial community ecosystem is dynamic and its composition is dependent upon many factors including genes, medication, and diet (Cox and Blaser, 2015). Within them, dietary changes can induce temporary shifts in a large number of microorganisms as rapidly as within 24 h because it is the main source of energy for individuals and a crucial method for humans to maintain health and growth (Singha et al., 2017; Makki et al., 2018). Studies have shown these age-related gut microflora changes could occur due to changes in the diet at different ages and changes in inflammation due to some age-related diseases and changes leading to decreased immune system function (Vaiserman et al., 2017). The varying composition of gut microorganisms has been identified according to geographical regions and this may also be due to different regional eating habits (Herath et al., 2020). The gut microbiota plays a key role in the body's metabolism and immunity responses provide a potential impact on the beginning of metabolic diseases like diabetes (Sonnenburg and Bäckhed, 2016).

The main energy source of the gut microflora is dietary carbohydrates and it is inversely associated with the incidence of T2D. The impact of dietary fibre is established on intestinal microflora populations, and research indicates that fibre intake is associated with an increase in microbial diversity and the ratio of Firmicutes to Bacteroidetes (Martínez et al., 2013). Previous studies confirmed that an increase in dietary fibre intake also increases the abundance of the human intestinal microflora and leads to higher microflora richness with higher microflora stability (Tap et al., 2015). Dietary fibre intake promotes the fermentation of intestinal microbes and this appears to cause an increase in short-chain fatty acids (SCFAs) that have the regulation

mechanism of glucose homeostasis (Gholizadeh et al., 2019). Studies have reported that soluble fibre has a direct blood-glucose-lowering effect and it can increase the viscosity of gastric juices, more viscous fibre leads to longer gastric emptying times and improve starch digestion (Fuller et al., 2016). Additionally, it is associated with a reduced rate of glucose absorption, leading to changes in blood glucose as well as cholesterol (Fuller et al., 2016). That's why consuming more dietary fibre appears to reduce the risk of T2D and it is also associated with maintaining proper body weight. Nevertheless, some SCFAs appear to be involved in some of the mechanisms associated with diabetes, which also establishes the link between microbiota and diabetes (Kasubuchi et al., 2015; Neis et al., 2015). In addition to SCFAs, intestinal microflora appears to regulate lipopolysaccharide (LPS) levels and these levels are also thought to be involved in the development of diabetes (Canni et al., 2009). Patients with T2D have fewer butyrate-producing bacteria and the ratio of Firmicutes to Bacteroidetes is also significantly lower than non-diabetic patients (Yoo et al., 2016).

Healthy adults and children can increase their intake of plant foods rich in fibre while reducing total energy intake that is more often associated with high-sugar, high-fat, and low-fibre foods (Dahl et al., 2015). A recent study combined measurements of intestinal microbiome diversity with diet history, and blood test parameters from volunteers. This report indicated that a personalized diet can successfully improve the blood glucose level of T2D patients (Zeevi et al., 2015). Combining big data analysis and the use of more specific medicinal nutrition recommendations shows the possible prevention and management of T2D more efficiently.

1.5. Conclusion

Epidemiological studies provide a clear indication of the association between gut microbiota disturbance and increased incidence of T2D. Impaired energy metabolism has been proposed as a driving force for this metabolic disease may be due to a change in gut microbiota which causes obesity that in turn induces T2D. On the other hand, supplementation of prebiotics and/or the use of probiotics is an important mechanism for the rehabilitation of gut microbiota, and for the harmony of the body, homeostasis may be beneficial for T2D treatment. In general, early detection, understanding of the mechanisms of relationship, and screening of the causative gut microbiota are recommended for the future management of T2D patients.

1.6. Novelty and objective of this study

1.6.1. Novelty of the present study

Several mathematical and statistical models were built utilizing human physiological characteristics to predict the risk of the disease; machine learning (ML) is one of them. Although multiple studies approved the discriminating between T2D and normal person (NGT) using different ML models depending on patients' physiological conditions (Stolfi et al. 2020; Tigga and Garg 2020; Zhang et al. 2021b). However, most of those analyzed models made observations based on the small number of samples from a specific geographical site. Additionally, none attempted to uncover critical physiological factors out of their prediction model that significantly separates T2D disease from NGT. While best prediction model with high accuracy generally needs a large sample size with variant coverage (Wei et al. 2013; Arbabshirani et al. 2017).

The recent development of studies has indicated that along with the host's genetics, gut microbiota plays an essential role in establishing obesity and T2D (Karlsson et al., 2013; Bhute et al. 2017; Sroka-Oleksiak et al. 2020). Over the past decade around the world, many groups have devoted significant efforts to identify the gut microbiota's structural and functional features individuals with compared to understand the disease progression (Bhute et al. 2017; Gaike et al. 2020). Most of this research aimed to analyses the differences in gut microbial members either between T2D and pre-T2D with NGT or between gut microbiome following treatment of disease. However, the thorough investigation on predicting the essential influencing physiological aspects and their association with gut flora in disease states is incompletely explained, while none from India.

1.6.2. Objective of this present study

Objective 1 - Introducing best Machine Learning methods to predict T2D more accurately and identify the most important physiological parameters for early detection of the disease condition irrespective of their geographical location.

Objective 2 - Identification of the core and unique gut microbial members of Indian T2D with compare to healthy individuals and identify the differentially abundant core gut microbial genera as well as their association with the important physiological parameters for the improvement of diseases prediction and diagnosis along with established methods.

Objective 3 - Study of structural and functional alteration of gut microbial members associated with Indian T2D with compare to healthy individuals for

Chapter 2

Classification of Type 2 Diabetes (T2D) Disease & Prediction of Important Physiological Parameters Via Machine Learning Approaches

2.1. Introduction

Machine learning (ML) is subfield of Artificial Intelligence that solves the real world problems by "providing learning ability to computer without additional programming" (Choi et al., 2019). The machine learning has developed from the efforts of researching whether computers could gather knowledge to mimic the human brain. The first attempts of ML were in 1952 when Arthur Samuel developed the first game-playing program for checkers, to accomplish enough skills to win against a world checker champion. Later in 1957, Frank Rosenblatt created an electronic device which has the ability to learn how to solve complex problems by imitating the process in human brain (Choi et al., 2019). Development of ML contributed to the greater use of computers in medicine (Tigga and Garg, 2020).

According to artificial intelligence market research firm 'Tech Emergence' and the researcher from the paper, the major machine learning applications in medicine are: smart electronic health records, drug discovery, biomedical signal processing and disease identification and diagnosis (Chaudhury et al., 2017; Deberneh and Kim, 2021). In most cases of disease identification and diagnosis, the development of ML systems is considered as an attempt to imitate the medical experts' knowledge in the identification of disease. Since ML allows computer programs to learn from data developing a model to recognize common patterns and being able to make decisions based on gathered knowledge, it does not have difficulties with the incompleteness of used medical database (Deberneh and Kim, 2021). In medical application, the most famous machine learning technique is classification because it corresponds to problems appearing in everyday life, among which the most usually applied techniques are Artificial Neural Networks (ANNs) and Bayesian Network (BNs).

The usage of machine learning in disease classification is very frequent and scientists are even more interested in the development of such systems for easier tracking and diagnosis of diabetes and cardiovascular diseases (Kumari et al., 2013; Sharma et al., 2013; Fatima et al., 2017). According to World Health Organization (WHO), both diabetes and cardiovascular disease (CYD) are among top ten causes of death worldwide [14]. The research from the January 2017 showed that the number one cause of death worldwide are CYDs. The world's biggest killer is taking the leading position in the list of top ten causes of deaths in the last 15 years and in 2015 was counting for 15 million deaths (Tan et al., 2009; Iyer et al., 2015; Otoom et al., 2015; Vembandasamy et al., 2015). On the other hand, the first WHO Global report on diabetes demonstrated that in the period from 1980 to 2014, the number of adults with diabetes has risen from 108 million to 422 million, and the number of victims of diabetes in period from 2000 to 2015 increases from less than 1 million to 1.6 million people (Kononenko et al., 2001). The morbidity and mortality from diabetes and CYD indicate the need for early classification of patients which can be achieved developing machine learning models. These models enable analysis of bigger and more complex data in order to achieve more accurate results and guide better decisions in real time without human intervention.

This study was designed to perform a review of Artificial Neural Network and Bayesian Network and their application in classification of diabetes and CVD diseases. The purpose is to show the comparison of these machine learning techniques and to discover the best option for achieving the highest output accuracy of the classification.

2.2. Materials and methods

2.2.1. Data Collection

The relevant physiological information of a total of 441 samples (T2D: 224 and NGT: 217) of patients was selected for this study. Among them, 345 data were obtained from Chinese cohorts (Qin et al. 2012) and 96 data from European cohorts (Karlsson et al. 2013). The physiological parameters that include in our study were age, gender, body mass index (BMI), fasting blood glucose (FBG), fasting insulin (FI), glycated hemoglobin (HbA1c), total cholesterol (CHL), high-density lipoproteins (HDL), low-density lipoproteins (LDL), triglycerides (TG) and c – peptide (CP).

2.2.2. Preparation of Training, Testing and blind / identification dataset

After getting total 441 physiological parameters data, we were made a training dataset to build up a prediction model by training it (150 samples data) and a testing dataset (150 samples data), to evaluate the performance and ability of discrimination between two different classes (T2D and NGT) of that trained prediction model. To avoid prediction biasness, training and testing datasets were made by random sampling with keeps in mind that there were no common samples data present in both datasets and each of them consists with 1:1 ratio of T2D and NGT samples physiological data (Barman et al., 2014). From the remaining 141 samples, a known blind/identification dataset was build but they were treated as an unknown dataset to evaluate better efficiency of our prediction models. Lastly, we used this prediction model on our own dataset, collected from in and around Kolkata, West Bengal (see in the Chapter 3) to verify their performance for real unknown datasets.

2.2.3. Feature selection and ML methods

Feature selection improves the discrimination ability of the prediction model to relieve the over-fitting problem and help to better understand data by examining the importance of the features (Guyon and Elisseeff 2003; Saeys, Inza and Larranaga 2007). Here we used the Recursive Feature Elimination (RFE) algorithm (Granitto et al. 2006; Chen and Jeong 2007) as a feature selection method to find out what are the best physiological parameters that showed higher discrimination ability between two classes using the "caret" R package (Kuhn 2008). Random Forest (RF) (Svetnik et al. 2003) and Support Vector Machine (SVM) (Statnikov et al. 2013) were used for the prediction of T2D and NGT based on physiological data. The prediction models were built up using 10 fold cross-validation methods.

2.2.4. Performance checking of the prediction model

The performance of the prediction model was evaluated using the testing and blind datasets. To evaluate the performance of the prediction models, they were assessed via sensitivity (SEN), specificity (SPF), accuracy (ACC), precision (PRC) and F1 – Score values by using following formulas,

Where, TP - True Positive (NGT samples correctly identified as NGT), FP - False Positive (NGT samples incorrectly identified as T2D), TN - True Negative (T2D samples correctly identified as T2D), FN - False Negative (T2D samples incorrectly identified as NGT). All those statistical analyses were performed in R (R version 3.6.3) with the packages "randomForest" (Liaw, Wiener and others 2002), "rfUtilities" (Evans and Murphy 2019), "caret" (Kuhn 2008), "caTools" (Tuszynski and Tuszynski 2007),

"e1071" (Meyer et al. 2012), "verification" (Gilleland 2015) and "pROC" (Robin et al. 2011).

$$SEN = \frac{TP}{TP + FN}$$

$$SPF = \frac{TN}{TN + FP}$$

$$ACC = \frac{TP + TN}{TP + FP + TN + FN}$$

$$PRC = \frac{TP}{TP + FP}$$

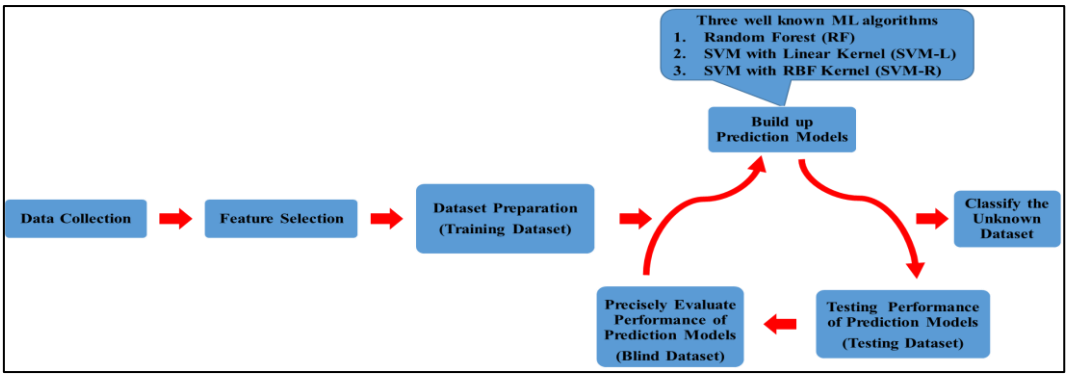


Figure 2.1 – Schematic flowchart of different phases for building a prediction models by using Machine Learning Algorithms.

2.3. Results

2.3.1. Selection of optimal features, construction and performance evaluation of MLT models to classify between T2D and NGT

Feature Selection (FS) is a pattern recognition application to remove the irrelevant or noise from the original features data. RFE FS is a multivariate approach that incorporates all variables in the algorithm and gradually excludes those variables which are not able to discriminate between the different classes. In this study, nine (9)

physiological parameters (BMI, FBG, HbA1c, FI, CP, CHL, HDL, LDL and TGL) of the total of 441 samples were considered to identify the best physiological parameters having the discriminatory ability between T2D and NGT and we have found five best physiological parameters (through RFE FS) that includes FBG, HbA1c, CP, FI and CHL with high accuracy (ACC = 95%).

For this investigation, those five important physiological parameters were further used to build as well as to evaluate the performance of the prediction models using three different MLT methods, i.e. RF, SVM – L, SVM – R. The prediction models were built with 150 training datasets (75 T2D and 75 NGT) and performance of these prediction models were tested using the same number of the testing dataset (75 T2D and 75 NGT) by measuring their SEN, SPF, ACC and PRC with 10 – fold cross-validation. However, the best prediction models were measured by their performance checking of precision (PRC) and recall (also known as SEN), since they were directly proportional to the true positive (Barman, Saha and Das 2014). All the prediction models worked very well and their values of SEN, SPF and ACC of the three prediction models were nearly the same. But the PRC score in SVM – L (100%) was higher than RF (94%) and SVM – R (94%), while the recall score of RF (100%) was higher than the SVM – L and SVM – R (Table 2). However, they were further evaluated to confirm their discriminatory abilities between T2D and NGT using a blind dataset.

2.3.2. Evaluation of prediction methods with blind dataset and classification of unknown samples

We used the same approach to avoid any bias in the performance of our proposed models and observed how well they could distinguish between two classes. Our analysis reported that all three prediction models worked very well to classify the

T2D and NGT in blind. Both RF and SVM – R models were able to identify the total 74 T2D samples correctly, (100% SEN values) while SVM – L showed the best prediction efficiency (97% SPF value) compared to the other two (Table 2.1).

Table 2.1: Comparative performance measurements among three different MLT methods using three different datasets with 10 – fold cross validation. Here RF: Random Forest, SVM – L: Support Vector Machine with Linear Kernel, SVM – R: Support Vector Machine with RBF Kernel.

Datasets	MLT	Sensitivity	Specificity	Accuracy	Precision
Test Dataset	RF	1.00	0.98	0.97	0.94
	SVM – L	0.97	1.00	0.98	1.00
	SVM – R	0.98	0.94	0.96	0.94
Blind Dataset	RF	1.00	0.88	0.94	0.90
	SVM – L	0.81	0.97	0.88	0.96
	SVM – R	1.00	0.88	0.94	0.90
Unknown Dataset	RF	1.00	0.52	0.76	0.68
	SVM – R	1.00	0.35	0.67	0.60

Overall, this investigation reported that the best two effective prediction models are Random Forest (RF) and SVM – R (SVM with RBF Kernel) as indicated on precision (PRC) and recall (SEN) values. The collected physiological parameters of 34 samples (17 T2D and 17 NGT) as unknown datasets were used to further evaluate the efficiency of RF and SVM – R prediction models using the top five physiological data that were identified in RFE – FS. Both the prediction models were successful in classifying all T2D samples as a true positive with 100% SEN or recall (Table 2.1).Interestingly from the above study, it is observed that FBG and HbA1c were demonstrated as the most important discriminative parameters with the highest mean decrease scores (95.2 and 75.2 respectively) among the two study groups.

2.4. Discussion

Many reports endorsed the usefulness of different Machine learning techniques to discriminate between T2D and NGT using patient's physiological conditions, but none has been attempted to identify the important parameters that can alone predict and diagnose the T2D (Patil, Joshi and Toshniwal 2010; Soni et al. 2011; Meng et al. 2013; Choubey and Paul 2017; Sisodia and Sisodia 2018; Choi et al. 2019; Tigga and Garg 2020). In this study, we are first to attempt to develop an MLT based prediction model using the conventional classification algorithms as well as identification of most important physiological parameters (using feature selection method: Recursive Features Elimination) to classify diabetes status. Our prediction models are developed and verified using two different regions of datasets (Chinese and European) and we applied these models to the studied Indian samples, to avoid any geographic biases.

Our proposed prediction models Random Forest (RF) and Support Vector Machine with RBF Kernel (SVM-R) have outperformed other already established models with high accuracy (94%) (Patil, Joshi and Toshniwal 2010; Soni et al. 2011; Meng et al. 2013; Choubey and Paul 2017; Sisodia and Sisodia 2018; Choi et al. 2019; Tigga and Garg 2020). Those models also identify the two most important physiological parameters FBG and HbA1c which have a greater role in the classification of T2D and diagnosis of the disease which is in line with the American Diabetes Association (ADA) and the World Health Organization (WHO) recommendation as well as previous investigations; stating that both fasting blood glucose (FBG) and glycated hemoglobin (HbA1c) are critical to classify the T2D patients (Inzucchi 2012; Olokoba, Obateru and Olokoba 2012; Deberneh and Kim 2021). The HbA1c is a convenient physiological parameter along with the FBG and represent the level of blood glucose in T2D patients

for 3 to 4 months which reflect the cumulative measure of chronic hyperglycemia and correlate well with the risk of long-term diabetes complications (Toque 2011; Sherwani et al. 2016; Chaudhury et al. 2017).

So, from this investigation we were hypothesized that during the development of diabetes significant changes in the level of both FBG and HbA1c, can be used as critical physiological measurements to identify the T2D patients or risk of disease in an impaired state of the patients around the world.

2.5. Conclusion

In this chapter we finally hypothesised that significant changes in the level of both fasting blood glucose (FBG) and glycated haemoglobin (HbA1c) during the development of diabetes can be used as critical physiological measurements to identify T2D disease or risk of illness in an impaired state around the world and that RF and SVM – R ML methods can also be used for better disease prediction.

Chapter 3

Study of Dominant Gut-Microbial Members Association with Indian Type 2 Diabetes (T2D) Via Metagenomic Approach

3.1. Introduction

When it comes to the human body, the microbiota is a collection of microbes that can be found on and inside the body, where microbial communities can form in a variety of niches. These are some of the most extensively studied microbiotas on the planet, with the gut and oral cavity receiving the majority of the research attention. If we look at it from the perspective of cell count, we are more microbe than humans, as microbe cells outnumber human cells by a factor of ten (Lepage et al., 2013). Besides that, a large number of viruses and a variety of micro-eukaryotes can be found all over the human body, as well as a variety of other organisms (Lepage et al., 2013). When it comes to functional potential, the coding genes of the microbiome outnumber those of the human host by a factor of 100 (Kim et al., 2013; Sweeney et al., 2013).

The majority of metagenomic studies involving the human microbiota have relied on targeted approaches that focused on the 16S marker gene (Kim et al., 2013; Robles2013). These results have provided a reasonably clear picture of the organisms that can be found in the various parts of the body. Although whole genome sequencing is still in its early stages, the field has begun to shift in recent years, adding a functional layer to the information already available. It has been discovered that the microbial members are generally more stable concerning the functional composition of the microbiome than they are with concerning species composition of the microbiome when they are compared across a population using both taxonomy annotation and functional annotation.

Type 2 diabetes (T2D) is a metabolic disease whose primary cause is insulin resistance. T2D is the most common type of diabetes. Other factors, such as mental stress, infection, and genetic predisposition, may also play a role in the development of

T2D disease (Wellen et al., 2005; Cani et al., 2009; Tilg et al., 2009; Tsukumo et al., 2009). As a result of chronic low-grade inflammation, T2D is characterised by increased levels of a variety of inflammatory mediators such as tumour necrosis factor and interleukins in the bloodstream (Dandona et al., 2004). The importance of the gut microbiome has received a great deal of attention in the last decade all over the world. Understanding the interaction between the gut microbiome and diabetes would provide new insights into the development of diabetes therapeutics. The composition of the intestinal microbiota has been linked to the development of metabolic diseases such as diabetes in recent studies based on large-scale 16S rRNA gene sequencing as well as more limited techniques such as quantitative real-time PCR (qPCR) and fluorescent in situ hybridization (FISH). The complete picture of the gut microbiota is slowly being revealed.

The gut microbiota is made up of more than 1000 microbial species that are primarily distributed across nine phyla, with the majority of them belonging to the *Firmicutes*, *Bacteroidetes*, *Proteobacteria*, and *Actinobacteria* families (Schloissnig et al., 2013). Several studies in mice models and humans found that increasing body weight was associated with a higher proportion of *Firmicutes* and a lower proportion of *Bacteroidetes* in the gut (Backhed et al., 2004; Ley et al., 2005; Turnbaugh et al., 2009). Zhang and colleagues demonstrated that *Firmicutes* were significantly reduced in post-gastric-bypass individuals and *Prevotellaceae* were highly enriched in obese individuals in their findings (Zhang et al., 2009). The increased ability of the obesity-associated microbiome to harvest energy from the diet was hypothesised to explain the differences in microbial composition between the two groups (Turnbaugh et al., 2006). Schwiertz and colleagues have published data that has sparked controversy (Schwiertz et al., 2009). The researchers discovered lower *Firmicutes* to *Bacterodetes* ratios in

overweight human adults when compared to lean controls. They concluded that using weight-loss diets as a test, another study found no evidence of a link between the proportion of *Bacteroidetes* and *Firmicutes* in the gut and human obesity (Duncan e al., 2008). As a result, the composition of the obese microbiome is still up for debate, and more scientific evidence is required to fully understand the relationship between gut microbial members and metabolic diseases in obese individuals. For example, in prebiotic-treated mice, *Bifidobacterium* levels were found to be significantly and positively correlated with improved glucose tolerance as well as reduced low-grade inflammation (Cani et al., 2008; Cani et al., 2009). It has also been reported that higher levels of *Bacteroides sp.* in rats were associated with the development of diabetes type 1 in these animals (Brugman et al., 2006). Some researchers believe that the gut microbiota directed increased monosaccharide uptake from the gut and instructed the host to increase hepatic production of triglycerides, both of which are associated with the development of insulin resistance in humans (Membrez et al., 2008).

In addition to food digestion and absorption, the gut microbiota has been shown to have a variety of other physiological functions, including enhanced host immune responses, biological antagonisms, strengthened antitumor responses, and the production of beneficial compounds (A M O'Hara and F. Shanahan, 2006; S. M. Jandhyala et al., 2015). Once the gut microbiota is out of balance, a cascade of diseases, including metabolic diseases, cardiovascular and cerebrovascular diseases, autoimmune diseases, inflammatory bowel disease, psychotic disorders, and cancer, would be induced as a result (A. M. Valdes et al., 2018). Several studies observed the alteration of the gut microbiota along with the development of T2D. The microbiota in the gut is involved in the regulation of glucose and insulin sensitivity. T2D patients' symptoms can be improved by altering their gut microbiota, which also aids in the

reversal of impaired glucose tolerance and fasting glucose levels in those with prediabetes. However, the deep study on the association of important influencing physiological factors with gut microbes in disease states is incompletely explained while none from India. So, in this chapter, we are attempting to give an accurate depiction of the human gut microbial diversity that influences T2D disease, particularly those in West Bengal, India. Also, we have proposed a few unique gut microbes that act as a key biomarker to improve the disease diagnosis along with already established methods.

3.2. Material and Methods

3.2.1. Sample selection and collection

The samples were selected as per suggestion from the doctors of the endocrine department of IPGMER and SSKM Hospital, Kolkata, India based on World Health Organization (WHO) criteria, and anthropometric measurements were done from 34 samples (17 NGT and 17 T2D) from West Bengal at IPGMER and SSKM Hospital. Only newly diagnosed cases of T2D in males of age group above 25 years and up to 55 years, willing to take participation, were included in our study. The patients, in the age group below 25 years and above 55 years, already diagnosed or treated with insulin, were excluded from this study.

The physiological parameters of all these samples were measured in the Endocrinology Lab of IPGMER and SSKM Hospital. The FI and CP were measured using Siemens Immulite Insulin and C-Peptide Kit and other remaining physiological data such as BMI, FBG, CHL, HDL, LDL, and TGL were measured by normal testing procedure (Zhang et al., 2013). The protocol and the project were approved by the ethics committee at SSKM Hospital.

3.2.2. DNA extraction and amplicon metagenomic sequencing

The metagenomic DNA was extracted from the patients' fecal samples by using Power Fecal DNA Isolation Kit (Mo Bio, Catalog No. 12830-50) following the manufacturer's instructions. The extracted metagenomic DNA was pooled for the amplification of hypervariable V3–V4 regions of the bacterial 16S rRNA gene and sequenced them using the Illumina MiSeq platform (2 × 300 bp paired-end). The raw paired-end primer trimmed sequences were provided by Eurofins, India. All raw metagenomic DNA sequences were submitted to SRA–NCBI database (Accession No. PRJNA486712).

3.2.3. Sequence processing and taxonomy classification

All the raw fastq datasets were processed by the following sequence processing protocol (Dhal et al., 2020; Nayak et al., 2021). For all 16S rRNA amplicon gene sequences from each sample, the quality screening was done by using Trimmomatic, version 0.33 (parameters: SLIDINGWINDOW: 4:15) (Bolger et al., 2014). High-quality sequence reads were then merged with PEAR, version 0.9.5 (Zhang et al., 2014), using default parameters. For OTU clustering, SWARM, version 2.0, was used with default parameters (Mahé et al., 2014). Moreover, SINA tool was used for alignment and taxonomic classification using the SILVA ribosomal RNA gene database, version 138, as a reference sequence using the representative sequence per OTU (Pruesse et al., 2012). Absolute singletons OTUs, as well as unclassified sequences on phylum level, were removed from our dataset using our standardized R script.

3.2.6. Statistical analysis

Principal component analysis (PCA) was done to understand the pattern among the two groups (T2D and NGT) of samples by utilizing their respective physiological data. To compare the physiological data of T2DandNGT groups, we used the Kruskal–Wallis rank–sum test.

Alpha (α) diversity analysis was done based on the rarefied data (minimum number of sequences among the samples) by sub-sampling the dataset. To assess the microbial communities' richness and evenness, OTU number (nOTU), inverse Simpson (invS), and Shannon diversity (shannon) were measured. The differences in diversity between T2D and NGT were assessed by Wilcoxon rank–sum test. The unique and core bacterial members among the two groups (T2D and NGT) were identified by using Venny, version 2.1 (Oliveros, 2007), with genera that had >0.5% abundance. Spearman rank correlation was calculated to assess if there were any relationship between alpha-diversity and the physiological parameters and to identify the association between the physiological parameters and microbial genera.

For beta (β) diversity, OTUs data were pruned to exclude the rare biosphere by retaining OTUs that were present in one or more than one sequence in three or more than three samples. This reduction of the datasets did not change diversity patterns (Mantel test; $r > 0.9$, $p = 0.001$). To test the differences in community-level (diversity) among T2D and NGT groups permutational multivariate analysis of variance (PERMANOVA) was calculated. The contribution of physiological parameters for explaining the variation in community structure redundancy analysis (RDA) was calculated based on their centered log transformed of pruned data using *aldex.clr* function with a median of 128 Monte Carlo Dirichlet of ALDEx2 R package. Forward

model selection was carried out to assess which are the best physiological parameters to explain this variation in the community based on maximum adjusted R^2 and minimum Akaike Information Criterion (AIC).

The differentially abundant OTUs among the T2D and NGT groups were identified by using Dotplot. All statistical analyses, as well as figure visualizations, were performed in R, version 3.6.3, with the packages "vegan" (Oksanen et al., 2019) and "ALDEx2" (Fernandes et al., 2014), and the PCA plot was made using OriginPro 2021 software, version 9.8.0.200.

3.3. Results

3.3.1. Physiological parameters of Indian T2D and NGT samples

The pathophysiological conditions of diabetes patients were assessed via nine different parameters (BMI, FBG, FI, HbA1c, CP, CHL, HDL, LDL, and TGL) of T2D with respect to NGT (Table 3.1). Among them, the average level of FBG and HbA1c in the T2D group (168 mg/dl and 8.1% respectively) were found significantly higher (p-value ≤ 0.05) than NGT (Table 3.2). The PCA analysis indicates first three principal components accounted for 72.8% variation among the two groups of samples based on their measured physiological parameters (Figure 3.1). The PC1 alone explained 33.1% variation, majorly contributed by BMI, CP, CHL, and LDL; PC2 explained 23.7% of the total variation that was mainly driven by FBG, HbA1c, and TGL; and PC3 was responsible for the remaining 16% variation explained by FI and HDL. It was also evident that the T2D group was separated as a single cluster from the NGT group along the FBG and HbA1c parameters.

Table 3.1 – Physiological characteristics of type 2 diabetes and control used in this study: Age, Body Mass Index (BMI), Fasting Blood Glucose (FBG), Fasting Insulin (FI), HbA1c, C – Peptide (CP), Cholesterol (CHL), High Density Lipoprotein (HDL), Low Density Lipoprotein (LDL), Triglycerides (TGL).

Sample	Type	Age	Sex	BMI (kg/m2)	FBG (mg/dl)	FI (uIU/mL)	HbA1c (%)	CP (ng/mL)	CHL (mg/dl)	HDL (mg/dl)	LDL (mg/dl)	TGL (mg/dl)
A2	NGT	32	M	22.2	95	8.26	6.03	2.53	199	48	133	208
B2	NGT	40	M	28.9	87	3.02	6	1.51	127	29	71	170
C2	NGT	23	M	22.7	70	18.8	4.6	2.68	141	184	50	51
D2	NGT	43	M	22	72	2.89	5.7	1.03	197	33	153	146
E2	NGT	40	M	20.5	71	6.98	5.1	1.71	139	35	95	98
F2	NGT	45	M	38.2	82	18	5.8	5.15	231	44	172	179
G2	NGT	35	M	23.6	89	2	5.8	1.25	197	38	137	150
H2	NGT	43	M	24.3	93	13.4	5.7	3.11	276	44	185	164
I2	NGT	30	M	27.7	90	6.27	5.45	2.52	241	31	123	651
J2	NGT	54	M	23.8	98	6.06	5.8	2.35	351	32	239	252
K2	NGT	45	M	38.2	82	18	5.8	5.15	231	44	172	179
L2	NGT	44	M	25.8	69	8.77	7.6	3.04	239	37	158	347
M2	NGT	45	M	21.88	93	9.41	8.8	3.16	274	40	192	305
N2	NGT	52	M	26.7	70	6.57	6.8	3.17	128	41	110	250
O2	NGT	45	M	19.1	100	3.27	5.2	1.45	171	49	111	173
P2	NGT	42	M	20.9	110	3.09	5.5	1.8	133	43	85	101
Q2	NGT	28	M	23.5	110	5.7	5.5	1.37	120	42	72	117
A1	T2D	48	M	21.4	184	4.18	6.5	1.06	188	38	129	187
B1	T2D	42	M	24.1	115	2.26	7.1	1.33	180	31	99	240
C1	T2D	57	M	23.8	114	13	5.3	3.21	92	38	34	117
D1	T2D	48	M	19.5	156	6.33	8.8	2.21	259	33	193	264
E1	T2D	48	M	21.2	125	4.11	7.4	1.65	105	41	50	65
F1	T2D	36	M	29.7	119	9.96	6.9	3	200	31	169	136
G1	T2D	42	M	24	279	2.36	8.3	0.853	115	34	60	84
H1	T2D	46	M	24.2	167	14.7	7.3	3.35	248	40	214	156
I1	T2D	45	M	23.4	273	4.28	8.2	1.25	109	36	70	110
J1	T2D	51	M	27.9	150	11.9	7.7	2.68	326	40	289	201
K1	T2D	40	M	29	135	4.48	10.2	1.99	208	46	149	172
L1	T2D	47	M	22.6	145	6	8.7	2.73	193	37	118	297
M1	T2D	42	M	24.19	185	7.29	11.7	2.09	240	40	198	190
N1	T2D	42	M	21.5	110	13.5	6.9	3.32	96	28	122	127
O1	T2D	42	M	22.3	228.8	10.4	9.6	3.53	264	40	166	382
P1	T2D	48	M	24.4	128	6.29	6.1	2.21	166	32	101	258
Q1	T2D	53	M	24	248.9	19.1	10.3	2.92	203	49	139	142

Table 3.2 – Differences in physiological parameters between Q11 diabetes subjects and controls assess by Kruskal–Wallis rank–sum test. Here '*' indicates highly significant.

Parameters	χ^2	DF	*p* – Value
Body Mass Index (BMI)	0.001	1	0.9725
Fasting Blood Glucose (FBG)	11.640	1	0.0006 *
Fasting Insulin (FI)	0.050	1	0.8228
Glycated hemoglobin (HbA1c)	13.233	1	0.0003 *
C – Peptide (CP)	0.015	1	0.9040
Cholesterol (CHL)	0.323	1	0.5698
High Density Lipoprotein (HDL)	1.909	1	0.1671
Low Density Lipoprotein (LDL)	0.001	1	0.9725
Triglycerides (TGL)	0.058	1	0.8094

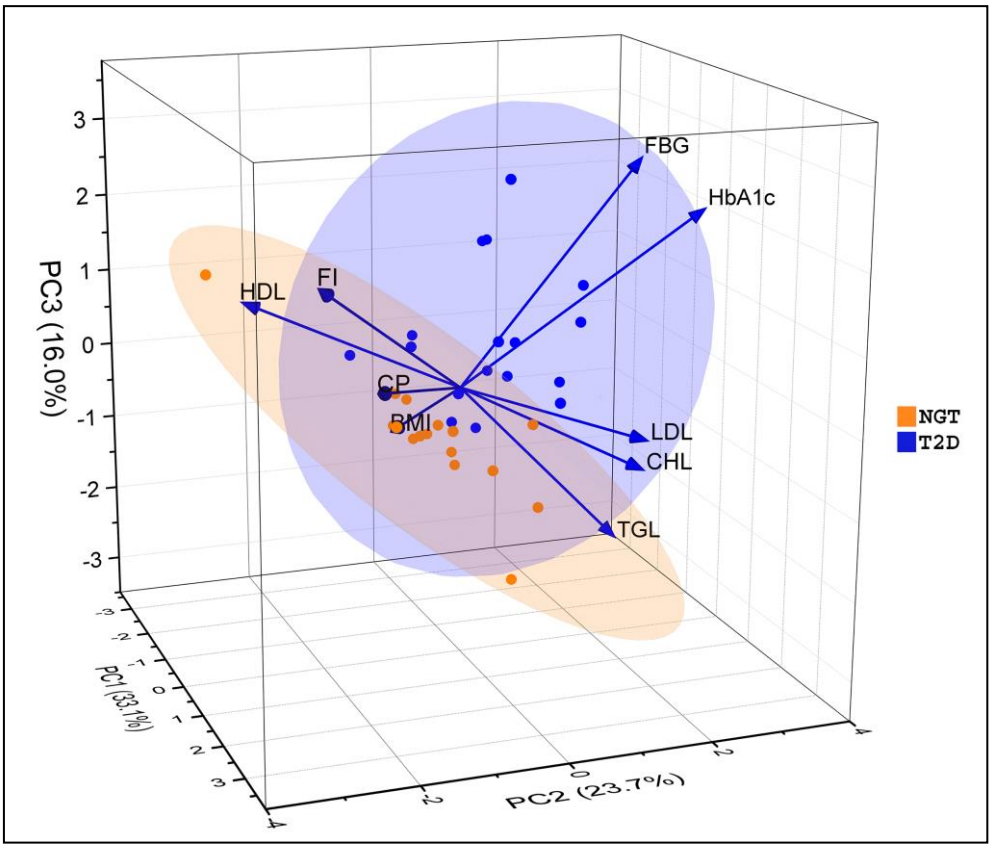


Figure 3.1: Principal Component Analysis (PCA) based on physiological parameters of the Indian diabetes subjects and controls. The samples were divided into two groups along with three principal components (PCs). PC1, PC2 and PC3 explained 33.1, 23.7 and 16 percent of the total variation respectively. Here BMI: Body Mass Index, FBG: Fasting Blood Glucose, FI: Fasting Insulin, HbA1c: Glycated Hemoglobin, CP: C – Peptide, CHL: Cholesterol, HDL: High-Density Lipoprotein, LDL: Low-Density Lipoprotein, TGL: Triglyceride.

3.3.2. Diversity analysis and taxonomy composition of the Indian T2D and NGT

By removing primer sequences of microbial hypervariable V3–V4 region of 16S rRNA gene amplicon sequences, a total of 71,30,226 clipped pair-end reads were generated. After trimming and merging the paired-end reads, a total of 44,00,731 merged sequences were obtained (Table 3.3). The high-quality reads were then clustered using > 97% sequence identity which generated 7,71,043 OTUs. A total of 43,467 swarm OTUs was obtained by removing the absolute singletons and unclassified sequence at the phylum level to avoid the rare biosphere, potential chimera effects, and PCR artifact (Dhal et al., 2020; Nayak et al., 2021).

Table 3.3 – Step by step sequence processing information

Sample name	Raw reads		Trimmed reads		Merged reads
	R1	R2	R1	R2	
NGT_A2	439815	439815	408185	408185	282099
NGT_B2	188218	188218	177819	177819	114826
NGT_C2	203916	203916	192060	192060	126306
NGT_D2	359010	359010	337672	337672	215326
NGT_E2	259562	259562	245596	245596	159454
NGT_F2	268907	268907	255313	255313	168866

NGT_G2	311594	311594	289635	289635	209913
NGT_H2	122664	122664	112373	112373	80967
NGT_I2	433617	433617	396827	396827	293562
NGT_J2	398206	398206	370119	370119	374688
NGT_K2	130171	130171	117736	117736	73222
NGT_L2	92324	92324	79979	79979	57980
NGT_M2	107082	107082	94305	94305	58283
NGT_N2	109650	109650	97842	97842	63353
NGT_O2	159926	159926	142419	142419	102672
NGT_P2	109491	109491	98428	98428	72689
NGT_Q2	108642	108642	98155	98155	72736
T2D_A1	159569	159569	148540	148540	89266
T2D_B1	401345	401345	378999	378999	24224
T2D_C1	398441	398441	377084	377084	245136
T2D_D1	411988	411988	392645	392645	269060
T2D_E1	306343	306343	287331	287331	192160
T2D_F1	160288	160288	152237	152237	99571
T2D_G1	107250	107250	97719	97719	67772
T2D_H1	117728	117728	107507	107507	76057
T2D_I1	128132	128132	117798	117798	83202
T2D_J1	282064	282064	260293	260293	190675
T2D_K1	119656	119656	108908	108908	73283
T2D_L1	120560	120560	107822	107822	79813
T2D_M1	135500	135500	123074	123074	88987
T2D_N1	105784	105784	93019	93019	60754
T2D_O1	111605	111605	100007	100007	74793
T2D_P1	151930	151930	135809	135809	93380
T2D_Q1	109248	109248	96292	96292	65656
Total Reads	7130226		6599547		4400731

α diversity i.e., diversity within the sample, was measured through nOTUs, Shannon diversity index as well as inverse Simpson index. It was observed that the average nOTU was higher in the T2D group (1960) than in the NGT (1565). Similar results were observed for Species richness and evenness in T2D and NGT groups as indicated by the Shannon diversity and inverse Simpson index (Figure 3.2). Spearman rank correlations test indicated a strong association of FBG with alpha diversity of the T2D group ($r = 0.54$, p-value ≤ 0.05) but none in NGT.

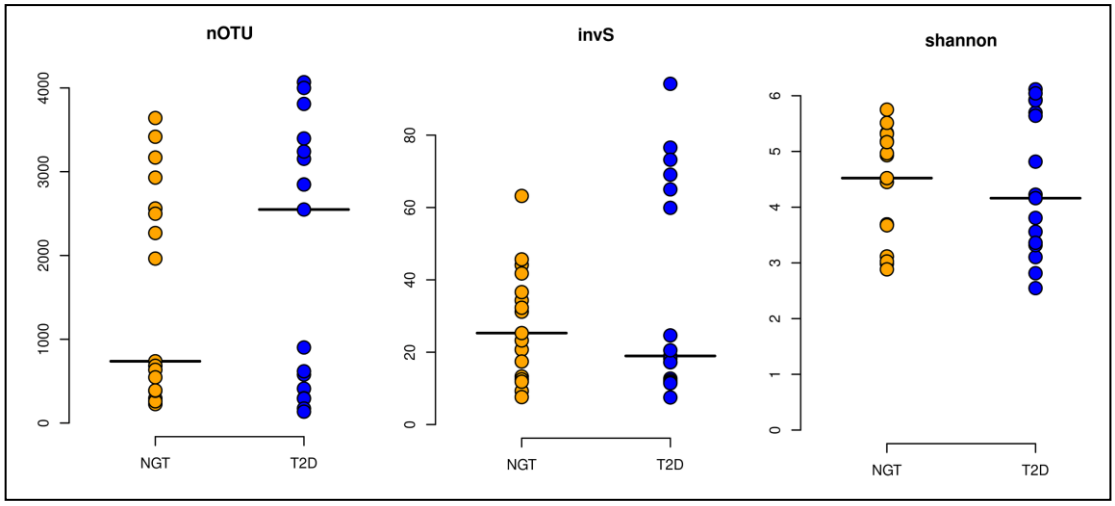


Figure 3.2 – Alpha diversity Indices. The alpha diversity of the studied groups was measured based on their richness (nOTU and inverse Simpson index (invS)) and evenness (Shannon Index (shannon)). Here horizontal lines in the plot represent their respective mean value.

The bacterial communities of gut microbiota were dominated by the members of *Bacteroidota*, *Firmicutes*, *Proteobacteria*, and *Actinobacteria* which represented almost 97% of sequences (Figure 3.3 (A)). In this study, we also observed 27 bacterial genera representing the core gut microbiome in the studied samples while each of 7 bacterial genera was found as unique for the T2D and NGT microbiome (Figure 3.3 (B)). The core microbiome was mainly dominated by *Prevotella_9*, *Prevotella*, *Prevotellaceae Incertae Sedis*, *Bacteroides*, and *Alloprevotella* of *Bacteroidia*; *Lachnospiraceae Incertae Sedis*, *Roseburia*, and *Faecalibacterium* of *Clostridia*; *Megasphaera* of *Negativicutes* and *Succinivibrio* of *Gammaproteobacteria* (Figure 3.4). The unique bacterial member for the T2Dmicrobiome was composed of *Eubacterium eligens group*, *Lachnoclostridium*, *Ruminococcus torques group*, and *Clostridia vadinBB60 group Incertae Sedis*, and *Lachnospira* under the class *Clostridia*; *Haemophilus* of *Gammaproteobacteria* and *Catenibacterium* of *Bacilli*. While *Alistipes* and *Muribaculaceae Incertae Sedis* under the class *Bacteroidia*;

Ligilactobacillus and *Holdemanella* of *Bacilli*; *Enterobacter* of *Gammaproteobacteria*; *Blautia* and *Coprococcus* of *Clostridia* were observed only in the NGT group.

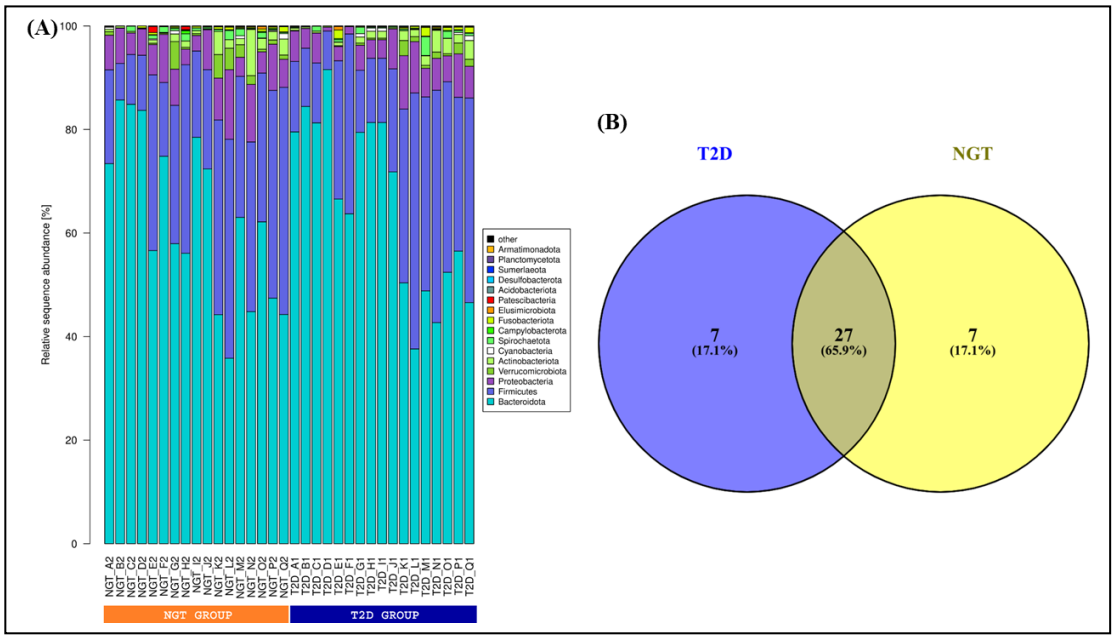


Figure 3.3 – (A) Phylum level taxonomic composition. Relative sequence abundance of most (top 10 based) dominant gut microbes in phylum level of studied samples. **(B) Venn diagram for unique and common gut microbes**. According to the Venn diagram, 27 gut microbes were common for both T2D and NGT groups, and 7 and 7 gut microbes were unique for T2D and NGT groups respectively.

β diversity was a measure to determine the intra sample variation of the gut microbial community using the pruned 6903 OTU datasets. The differential OTUs using the ALDEx2 test reported a total of 61 OTUs representing 68.1% of total communities for T2D and NGT gut-microbiome, that include class *Bacteroidia* (34 OTUs), *Clostridia* (13 OTUs), *Gammaproteobacteria* (5 OTUs), *Negativicutes* (4 OTUs), *Spirochaetia* (2 OTUs), *Bacilli* (2 OTUs) and *Verrucomicrobiae* (1 OTU) which were deferred as differential abundant between T2D and NGT (Figure 3.5).

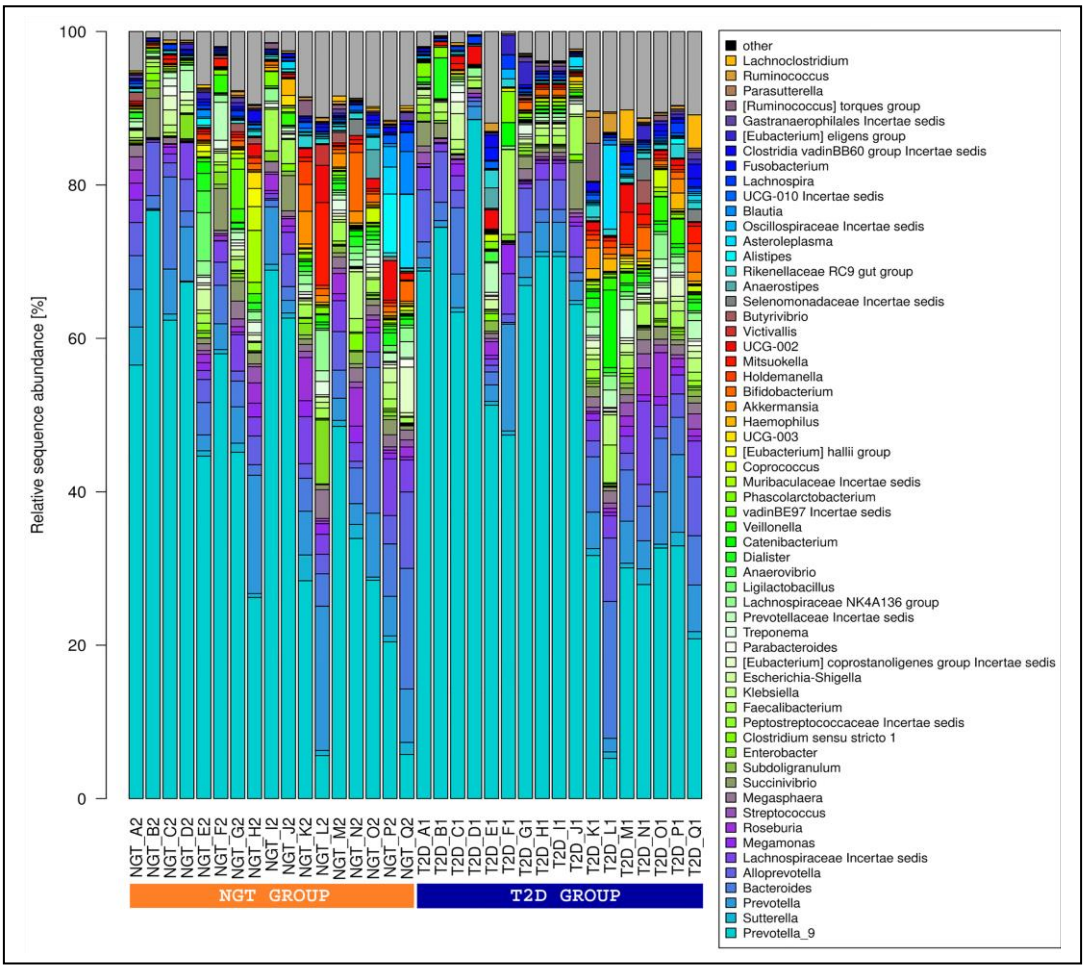


Figure 3.4 – Taxonomic composition at the genus level. Relative sequence abundance of most (top 10 based) dominant gut microbes in genus level of studied samples.

Within *Bacteroidia*, OTU affiliated with genus *Prevotella_9* (15 OTUs), *Alloprevotella* (otu18 and otu36), *Bacteroides* (otu28), *Prevotella Incertae Sedis* (otu48), and *Rikenellaceae RC-9 gut group* (otu82) significantly enriched in the T2D microbiome whereas *Prevotella* (otu22, otu24 and otu116) significant enriched in NGT microbiome. Within the *Clostridia* class, *Eubacterium* (otu49 and otu59) and *UCG-002* (otu46) genera were found dominant in the T2D microbiome, whereas *Roseburia* (otu38 and otu51), *Lachnospiraceae Incertae Sedis* (otu43 and otu112), *Butyrivibrio* (otu55) and *Faecalibacterium* (otu42) genera were found significantly enriched in NGT

microbiome. Similarly, *Gammaproteobacteria*, *Haemophilus* (otu237) showed dominance in the T2D microbiome whereas *Klebsiella* (otu83) and *Succinivibrio* (otu17) genera were found highly enriched in NGT. It was also observed that within *Negativicutes*, genera *Phascolarctobacterium* (otu33) was significantly dominant in the T2D microbiome but in the same class *Megasphaera* (otu25) and *Selenomonadaceae Incertae Sedis* (otu150) genera were significantly dominant in the NGT microbiome. Within Bacilli, the genus *Asteroleplasma* (otu64) significantly enriched in the T2D group whereas under the class Spirochaetia and *Verrucomicrobiae, Treponema* (otu81 and otu104) and *Akkermansia* (otu100) genera showed most dominance in the NGT group respectively.

Similarities or dissimilarities between two groups were projected in an ordination space as well as their associated physiological parameters on the NMDS plot (Figure 3.6). Moreover, *Envfit* result showed that FBG (R^2 = 0.2022, p – Value = 0.025) and HbA1c (R^2 = 0.1480, p – Value = 0.086) coincided with microbial community composition, but the association seems to be weak. Redundancy analysis which was performed to assess the significant contribution of the tested parameters in describing the variation in microbial communities revealed that only HbA1c had the explanatory power for bacterial communities of T2D microbiota with 2.1% (Adj R2 = 0.021, F = 1.34, AIC = 168.51, p = 0.05). Together NMDS and RDA supported each other's results and suggested that HbA1c, as well as FBG, were the responsible variable among the parameters for variation in the microbial composition in the T2D group.

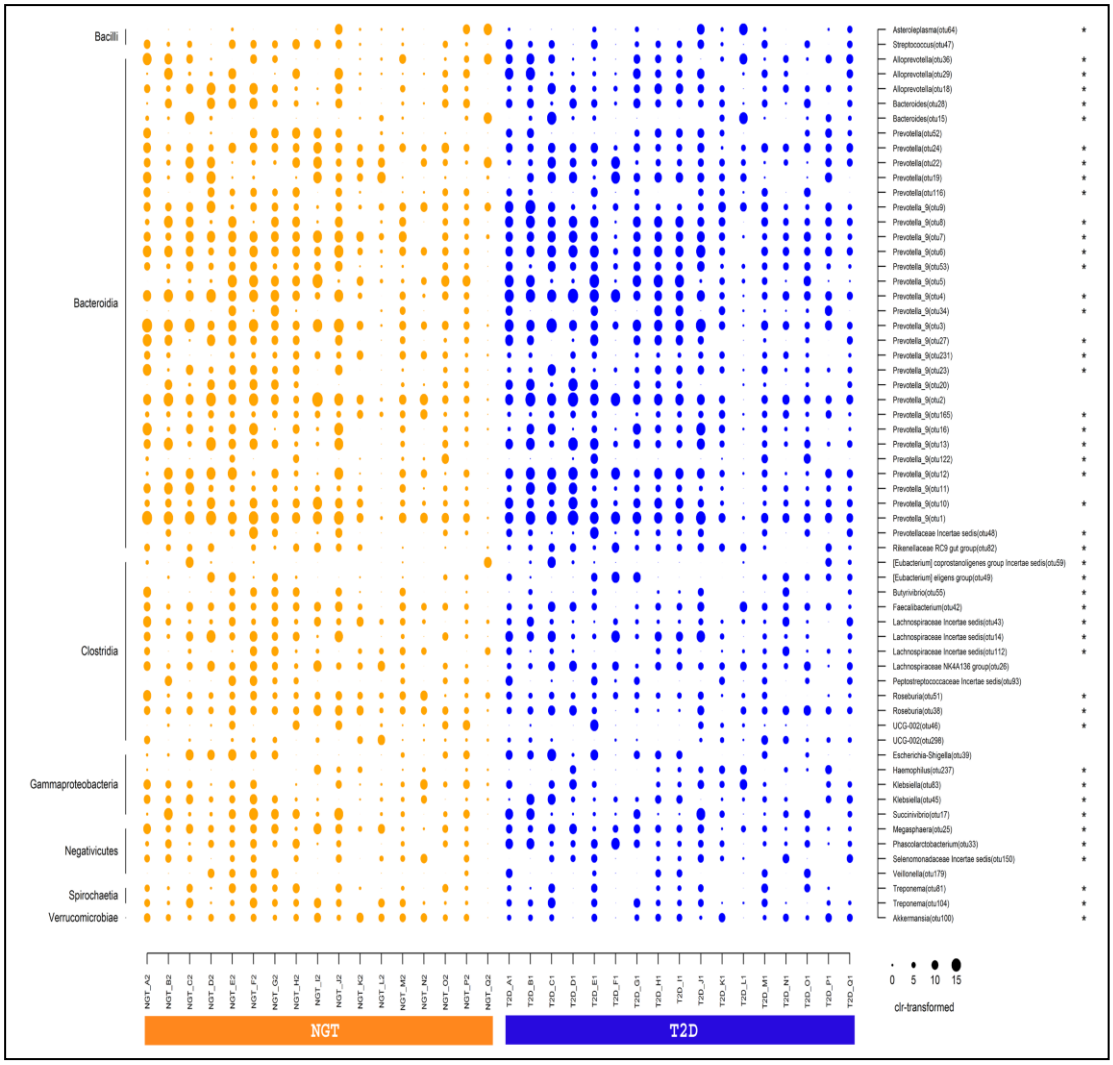


Figure 3.5 – Dominant bacterial community between two groups (T2D and NGT). Differentially abundant OTUs within two groups are represented in Dotplot using ALDEx2. Dotplot represents class level taxonomy on the left side and genus level on the right side. The size of each dot (0, 5, 10, 15) represents centered log-ratio (clr) – transformed sequence counts. P-values of ≤ 0.05 were indicated by the '*' symbols.

The significant correlation between the significant differentially abundant OTUs with the most important physiological parameters (FBG and HbA1c, as they were found as the most significant influence in our statistical analysis) was measured by calculating the Spearman correlation coefficient (SCC). As indicated in Figure 3.7, otu10, otu27, and otu231 represent Prevotella_9, otu28 represent the Bacteroidandes,

otu48 represent the *Prevotella Incertae Sedis* showed a significantly positive correlation with FBG (p - Value ≤ 0.05) while out53, otu122, and otu231 representing *Prevotella_9*, otu64 representing *Asteroleplasma* and otu28 representing Bacteroides were highly positively correlated with the HbA1c (p - Value ≤ 0.05).

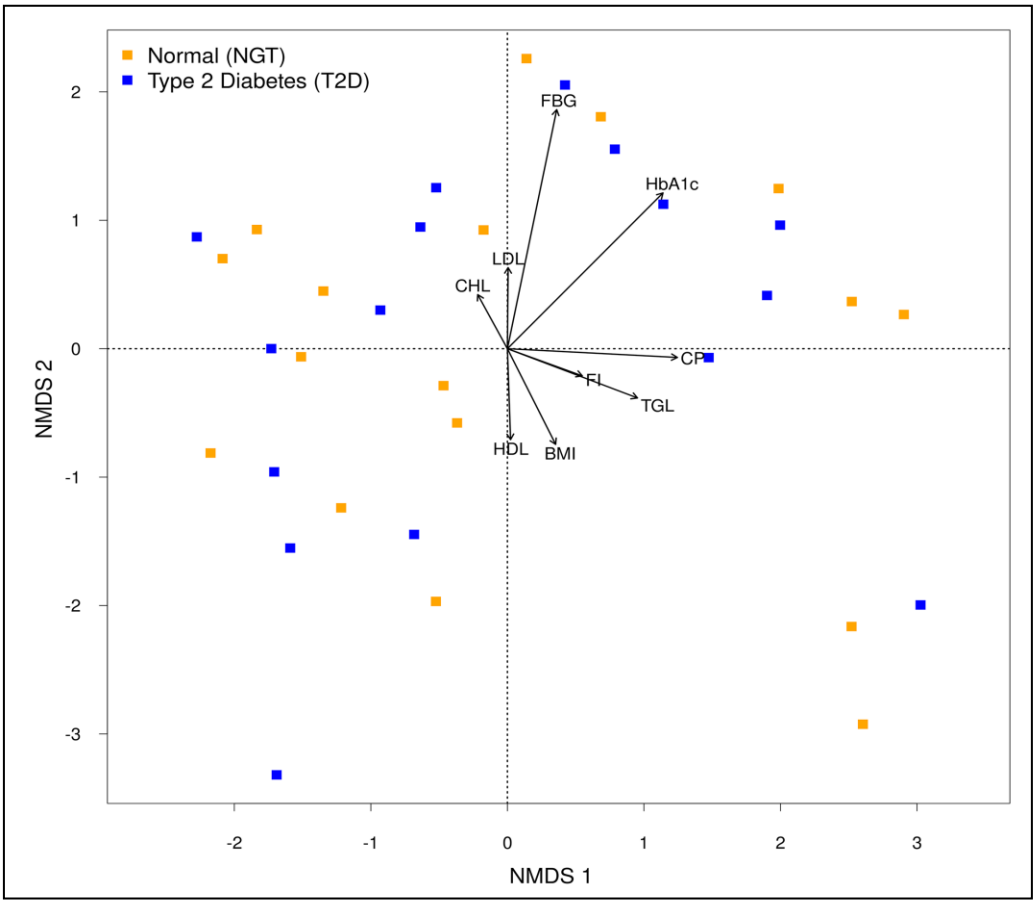


Figure 3.6 – Non-metric multidimensional scaling (NMDS) plot of the bacterial communities of each group. Arrows of the NMDS plot indicate *envfit* correlations of bacterial community composition with physiological parameters.

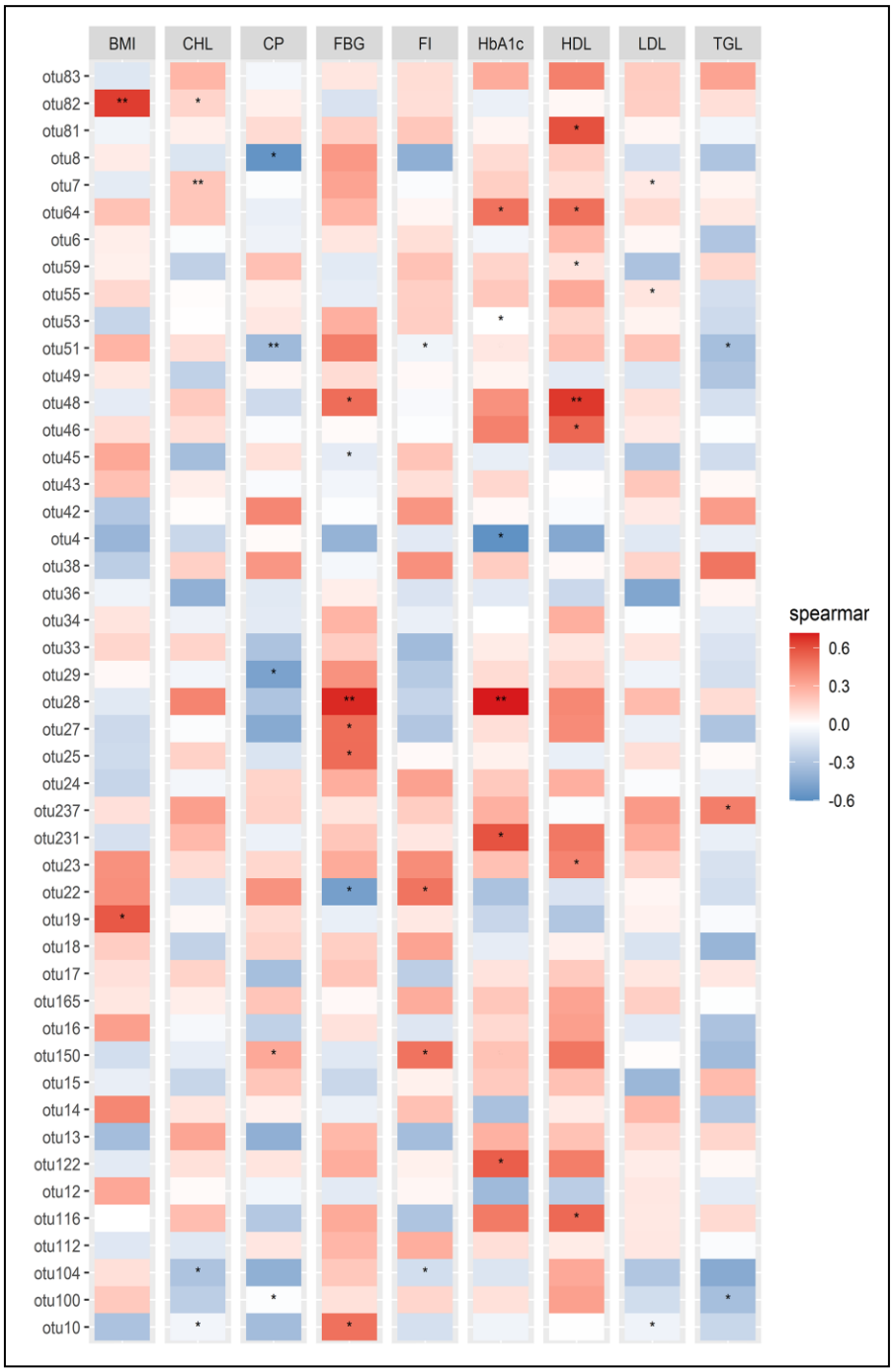


Figure 3.7 – Correlation Heatmap of physiological parameters with the significant differential abundant OTUs identified in DotPlot analysis. Spearman correlation analysis based on differentially abundant significant OTUs and the measured physiological parameters. Spearman correlation values were shown in the vertical heatmap panel to the right. P-values of ≤ 0.05 were indicated by the '*' symbols.

3.4. Discussions

Alterations of gut microbiota and their association with T2D are well-established around the world (Karlsson et al., 2013; Bhute et al., 2017; Gaike et al., 2020; Sroka-Oleksiak et al., 2020). However, the microbial dynamism of T2D patients from normal as well as their correlation with the important physiological parameters (FBG and HbA1c) is not reported, which is another novelty of our investigation. In this study, we were the first to provide the preliminary information on the gut microbiome of T2D patients from the eastern region of the Indian Subcontinent, especially in and around Kolkata, West Bengal. The T2D patients from this region have unique dietary status compared to other regions and this seems to restrict us from collecting the samples from different regions which is also reflected in our sample size. The microbial community of the studied samples was dominated by the members of the bacterial groups under phylum *Bacteroidota*, *Firmicutes*, *Proteobacteria*, and *Actinobacteria*. *Bacteroidota* and *Firmicutes* are the well known dominant bacteria phylum found in obesity, diabetes, and also in normal gut microbiome around the world (Gaike et al., 2020; Sroka-Oleksiak et al., 2020). Although there are reports on the differences in abundance among *Bacteroidota* and *Firmicutes* in T2D patients to NGT (Zhang et al., 2013; Ahmad et al., 2019). However, some other reports stated that such differences are not significant in T2D from NGT, which is in line with our results, as this investigation mostly focused on T2D irrespective of their obesity status (Turnbaugh et al., 2006; Ley et al., 2008; Zhang et al., 2013). The members of phyla *Firmicutes* play an important key role in fat digestion and their higher abundance is directly associated with obesity whereas *Bacteroidota* is associated with the production of short-chain fatty acids (SCFAs) (Ahmad et al., 2019).

Among the 27 core bacterial genera, the taxonomy of the associated genera with significantly dominated OTUs in studied T2D samples is *Prevotella_9, Alloprevotella, Bacteroides, Prevotella Incertae Sedis, Rikenellaceae RC-9 gut group, Eubacterium, UCG-002, Phascolarctobacterium,* and *Asteroleplasma*. They are also reported to be well-associated with T2D; for example, *Allopreotella* and *Bacteroides* are reported as risk factors for diabetes as these are reported to increase the level of lipopolysaccharides (LPS) and insulin resistance, which are detrimental to human health (Cheng et al., 2017; Wang et al., 2020). The *Prevotella_9* is reported to be associated with a plant-based low-fat diet and represents key bacterial members during human gut microbiota maturation in infants to young adults (Qian et al., 2018; Li et al., 2020b). However, the biological significance in the human gut enterocyte of both *Prevotella_9* and *Asteroleplasma* has not been well elucidated. While *Rikenellaceae RC9 gut group* bacterial genera showed an association with a high-fat diet and play an important role in lipid metabolism (Zhao et al., 2018). The genus *Phascolarctobacterium* is reported as an enriched bacterial genus in the T2D mice model and negatively correlated with fasting insulin (Naderpoor et al., 2019; Song et al., 2020). We found OTUs representing *Prevotella_9, Bacteroides, Prevotella Incertae Sedis* and *Asteroleplasma* bacterial genera have a significantly positive correlation with important established physiological parameters FBG and HbA1c. Interestingly, this observation supported the correlation analysis of alpha-diversity (richness and evenness) of the gut microbial community of studied T2D patients with FBG. Also, the results of NMDS *envfit* and RDA reflect that FBG and HbA1c both coincided most strongly with the microbial community composition of the T2D microbiome.

On the other hand, *Prevotella, Roseburia, Lachnospiraceae Incertae Sedis, Butyrivibrio, Faecalibacterium, Klebsiella, Succinivibrio, Megasphaera,*

Selenomonadaceae Incertae Sedis, Treponema, and *Akkermansia* genera are found as dominant bacterial genera in the NGT microbiome. A similar result was observed in the study by Almugadam et al. (2020) where they reported that short-chain fatty acid (SCFA) and butyrate producers such as *Faecalibacterium, Roseburia, Selenomonadaceae Incertae Sedis, Succinivibrio,* and *Megasphaera* genera were abundant in the healthy gut microbiome (Almugadam et al., 2020). *Prevotella, Succinivibrio, Treponema,* and *Lachnospiraceae Incertae Sedis* major contributes to inter-individual variation in gut microflora and are associated with better digestion of plant-derived complex carbohydrates and fibres diet for glucose homeostasis along with the production of butyric acid in the human colon for intestinal barrier protection (Arumugam et al., 2011; Schnorr et al., 2014; De Filippo et al., 2017; Zhao et al., 2020). Several investigators report the enrichment of butyrate-producing bacterial genera such as *Roseburia, Butyrivibrio, Faecalibacterium, Lachnospiraceae Incertae Sedis,* and *Megasphaera* are responsible for the reduction of inflammatory symptoms as well as insulin resistance. These bacterial genera play an important key role in intestinal health maintenance, immune defence, regulation of the dynamic balance of T-cells, and promote Treg cell differentiation by butyrate production (Canani et al., 2011; Karlsson et al., 2013).

Klebsiella bacteria are also found in the healthy human intestines and are not reported to be pathogenic as long the person is sick because of pneumonia, bloodstream infections, wound, or surgical site infections, etc. (Canani et al., 2011). A high abundance of mucin degrading *Akkermansia* bacterial genus in healthy human guts is well documented as they play a vital role in insulin resistance as well as intestinal barrier and LPS leakage reduction (Tanca et al., 2017; Gurung et al., 2020). Although some recent reports indicate that a decrease in this genus in diabetes is associated with

inflammation and metabolic disorders in the mice model, it can be used as a biomarker for impaired glucose tolerance (Sonnenburg and Bäckhed, 2016; Plovier et al., 2017).

Several unique bacterial genera are identified in T2D compared to the NGT microbiome and probably play some roles in the structural and functional attributes of the gut microbes in the human intestine for the development of disease. The unique genera for the T2D microbiome are *Catenibacterium*, *Eubacterium eligens group*, *Lachnoclostridium*, *Ruminococcus torques group*, *Clostridia vadinBB60 group Incertae Sedis*, *Lachnospira*, and *Haemophilus*. Several investigators reported that a few of these bacterial genera such as *Ruminococcus torques group*, *Lachnospira*, and *Haemophilus* act in mucus degradation by decreasing the gut barrier integrity, and they can be used as bacterial biomarkers to study their involvement in the human gut or their uses as diagnostic tools should be encouraged (Chen et al., 2020; Vacca et al., 2020). *Haemophilus* bacterial genus reported highly abundant in the Chinese T2D cohort is a particular biomarker for them (Chen et al., 2020). While for NGT, the unique bacterial genera are *Enterobacter*, *Ligilactobacillus*, *Alistipes*, *Muribaculaceae Incertae Sedis*, *Blautia*, *Holdemanella*, and *Coprococcus* are found in this investigation. Few of those genera including, *Alistipes*, *Blautia*, and *Holdemanella* are observed in the normal human gastrointestinal tract and they have an important key role in protection from many diseases such as liver and cardiovascular fibrotic disorders and also from various pathogens (Arumugam et al., 2011; Parker et al., 2020). *Coprococcus*, *Muribaculaceae Incertae Sedis*, and Enterobacter bacterial genera are having the ability for metabolic improvements and consorted with a higher quality of life indicators supported by previous reports (Valles-Colomer et al., 2019; Wang et al., 2020).

This investigation gives a well-resolved picture of the bacterial diversity and their correlation with important physiological parameters that influence the decrease of SCFA and butyrate producing core bacteria which are beneficial for the human gut in T2D patients, in West Bengal, India. Also, we suggest that along with the well-established physiological parameters, the unique gut microbes can be used as a key biomarker to improve the disease diagnosis.

3.5. Conclusion

From the investigation in this study, following conclusions can be drawn:

(1) Both of Fasting Blood Glucose (FBG) and Glycated Hemoglobin (HbA1c) physiological parameters coincided with the microbial community composition of the T2D microbiome by decreasing the beneficiary core gut microbial members.

(2) *Catenibacterium*, *Eubacterium eligens group*, *Lachnoclostridium*, *Ruminococcus torques group*, *Clostridia vadinBB60 group Incertae Sedis*, *Lachnospira*, and *Haemophilus* can be used as important biomarkers for Indian T2D patients.

Chapter 4

Study of Structural and Functional Alteration of The Gut Microbial Members Associated With Indian Type 2 Diabetes Patients

4.1. Introduction

Type 2 Diabetes (T2D), which is characterised by low-grade inflammation, insulin resistance, and β-cell failure, is becoming more common worldwide (Butler et al., 2003; Shoelson, 2006; Xu et al., 2013). In 2010, the estimated prevalence of adults with diabetes was 8.3%, with T2D accounting for at least 90% (Alberti and Zimmet, 1998; Whiting et al., 2011). This percentage is expected to rise to 9.9% by 2030. (Whiting et al., 2011). Obesity, which causes low-grade inflammation and insulin resistance, is primarily responsible for the development of T2D. (Hotamisligil, 2006).

Obesity, T2D, and other metabolic illnesses have been linked to the gut microbiota, according to some recent research studies (Turnbaugh et al., 2006; Vijay-Kumar et al., 2010; Zhao et al., 2013). In germfree mice, the gut microbiota from an obese adult can induce the development of obese phenotypes, demonstrating the causal significance of the gut microbiota in the development of obesity and metabolic disorders (Turnbaugh et al., 2006). According to these studies, the gut microbiota can directly cause insulin resistance, obesity, and type 2 diabetes.

The development of T2D may be greatly influenced by the alteration of gut microbiome (Backhed et al., 2004; Collins et al., 2013; Le Chatelier et al., 2013; Zhao, 2013). For instance, when a pure form of endotoxin was subcutaneously pumped into mice, it caused fat and insulin resistance since it was created by an opportunistic infection in the gut like Escherichia coli (Cani et al., 2007). In T2D patients' guts compared to healthy controls, there were more opportunistic infections, like Betaproteobacteria (Larsen et al., 2010). The most recent comparative metagenomic investigation of faecal samples from T2D patients and healthy controls revealed that the diseased samples contained higher levels of opportunistic pathogenic

microorganisms such *Clostridium* and *Desulfovibrio* but lower levels of butyrate-producing microorganisms (Qin et al., 2012). Another study discovered that an enhanced microbial translocation from the colon into tissues was indicative of the early beginning of high-fat diet-induced T2D. (Amar et al., 2011). Obesity and insulin resistance were brought on in germ-free mice by the opportunistic infection Enterobacter cloacae B29, which was found in the patient's gut who was both diabetic and morbidly obese (Fei and Zhao, 2013). Together, these results suggest that a dysbiosis of intestinal microbiota may be a primary factor in the development of obesity and diabetes, suggesting a new area of focus for the prevention and treatment of these conditions. Furthermore, nutritional changes to the gut flora have been linked to a reduction in both genetic and non-genetic childhood obesity, according to a recent study (Zhang et al., 2015).

The goal of this chapter was to help in better understand and identify the keystone taxa that are responsible for disease conditions in the intestinal microbiome as compared to the study's healthy participants. By anticipating the microbial metabolic pathways for creating and refining the microbiome-based diagnostic and treatment for T2D disease states, we also studied the functional modifications caused by differences in gut microbiota in T2D disease condition. In order to treat diabetes, hyperlipidaemia, and other metabolic illnesses, a unique treatment strategy that targets the gut microbiota using prebiotics, probiotics, diets, and medications might be effective.

4.2. Material and Methods

4.2.1. Microbial co-occurrence network analysis

The co-occurrence network analysis was performed to assess the complexity of the microbiome and identify potential keystone taxa for each group. The co-occurrence

network was constructed with the OTUs that were present in 10% of samples and had more than 10 sequences for each group. We used Spearman's rank correlation to assess the association among microbial OTUs from each group. A p-value of ≤0.05 and a Spearman's rank correlation coefficient (ρ) of ≥0.6 were selected as the thresholds between two OTUs (Jiao et al., 2016; Li et al., 2021). Two co194 occurrence network was built, the T2D Co-occurrence Network (TCN) and NGT Co-occurrence Network (NCN). The network's topology was measured by calculating the nodes, edges, average weighted degree, network diameter, graph density, modularity, average clustering coefficient, and average path length for each network. The network visualization and topology analysis were performed in the Gephi 0.9.2 (https://gephi.org/) visualization tool (Bastian et al., 2009).

The role of nodes in individual co-occurrence network topology was determined by evaluating the within-module connectivity (Zi) and among-module connectivity (Pi) using a web-based tool Molecular Ecological Network Analysis Pipeline (MENAP) (http://ieg4.rccc.ou.edu/mena) (Deng et al., 2012; Qiu et al., 2022). Based on this analysis the nodes are classified into four groups, they are – (a) peripheral nodes ($Zi < 2.5$, $Pi < 0.62$), (b) connectors ($Zi < 2.5$, $Pi > 0.62$), (c) module hubs ($Zi > 2.5$, $Pi < 0.62$), and (d) network hubs ($Zi > 2.5$, $Pi > 0.62$) (Qiu et al., 2022). Module hubs are densely connected to many nodes within r own modules, whereas network hubs serve as both connectors and module hubs. Together with network hubs, module hubs and connectors were termed keystone nodes/taxa (Olesen et al., 2007; Zhou et al., 2010; Deng et al., 2012; Qiu et al., 2022).

4.2.2. Metabolic pathway prediction based on amplicon 16S rRNA metagenome

To predict the microbial metabolic pathways based on the representative sequences of pruned OTUs from T2D and NGT groups, Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt2) tool was used (Douglas et al., 2020). Also, the relative abundance or copies per million (CPM) of the microbial pathways was calculated based on pruned OTU table data.

4.3. Results

4.3.1. Co-occurrence network analysis and keystone taxa of the Indian T2D and NGT

To understand potential interactions among gut microbial community members for each group,334 we constructed co-occurrence networks based on OTU to OTU correlations. The T2D co-occurrence network (TCN) consisted of 168 nodes and 213 edges, while the NGT co-occurrence network (NCN) consisted of 217 nodes and 233 edges (Table 3). The modularity of TCN is 0.93 decreased from NCN modularity (0.96), accompanying the increase of average weighted degree in TCN (1.268) compared to NCN (1.074).

The nodes present in both TCN and NCN networks were mostly dominated by phyla *Firmicutes*, *Bacteroidota*, *Proteobacteria*, *Actinobacteriacteriota*, *Verrucomicrobiota*, *Spirochaetota*, *Fusobacteriota*, and *Desulfobacterota* (Figure 6; Figure 7). But their percentage in each network was different, like the Firmicutes present in TCN and NCN is 57.14% and 48.39% respectively, the same trend also observed in *Bacteroidota* (TCN vs NCN: 28.57% vs 36.87%), *Proteobacteria* (TCN vs NCN: 8.33% vs 7), *Actinobacteria* (TCN vs NCN: 2.98% vs 2.3%), *Verrucomicrobiota*

(TCN vs NCN: 1.19% vs 0.46%), *Spirochaetota* (TCN vs NCN: 0.6% vs 0.46%), *Fusobacteriota* (TCN vs NCN: 0.6% vs 0.46%) and *Desulfobacterota* (TCN vs NCN: 0.6% vs 0.92%). *Cyanobacteria* (0.92%), *Campylobacterota* (0.46%), *Patescibacteria* (0.46%) and *Elusimicrobiota* (0.46%) gut microbial phyla were found only in the NCN, while none from TCN.

Table 4.1: Characteristics information of two gut microbial co-occurrence network; TCN – T2D Co-Occurrence Network, NCN – NGT Co-Occurrence Network.

Network Topology Parameters	NCN	TCN
No. of nodes	217	168
No. of edges	233	213
Average Weighted Degree	1.074	1.268
Network Diameter	3	2
Graph Density	0.005	0.008
Modularity	0.96	0.93
Average clustering co-efficient	0.226	0.208
Average path length	1.084	1.082

We also identified 14 and 8 OTUs as keystone nodes from TCN and NCN networks respectively based on within-module connectivity (Zi) and among-module connectivity (Pi) values. Among them, 6 OTUs as module hubs and 8 OTUs as connector nodes were identified in the TCN network, whereas in the NCN network 7 OTUs as module hubs and 1 OTU as connector nodes were identified. The identified keystone taxa, 5 OTUs were found under the phylum *Firmicutes*, 4 for *Bacteroidota*, 3 for *Proteobacteria*, one for *Actinobacteriota* and one for *Spirochaetota* gut microbial phyla in TCN network. In contrast, 2 OTUs were found under the phylum *Bacteroidota*, 3 for *Firmicutes*, one for *Proteobacteria*, one for *Patescibacteria* and one for *Desulfobacterota* as keystone microbial phyla for NCN. Due to the decrease in network

topology and different gut microbial compositions, the network stability also decreases in TCN compared to NCN.

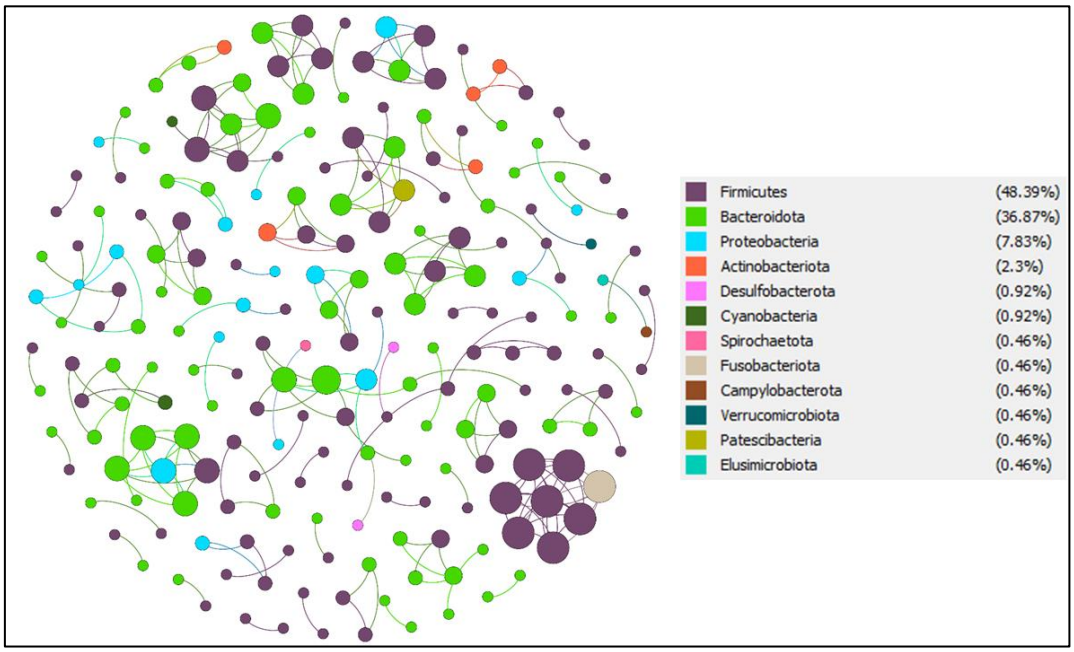


Figure 4.1 – NGT Co-occurrence Network (NCN). From total OTU abundance data, we select the NGT specific OTUs using the specified criteria, and a co-occurrence microbial network was constructed in Gephi.

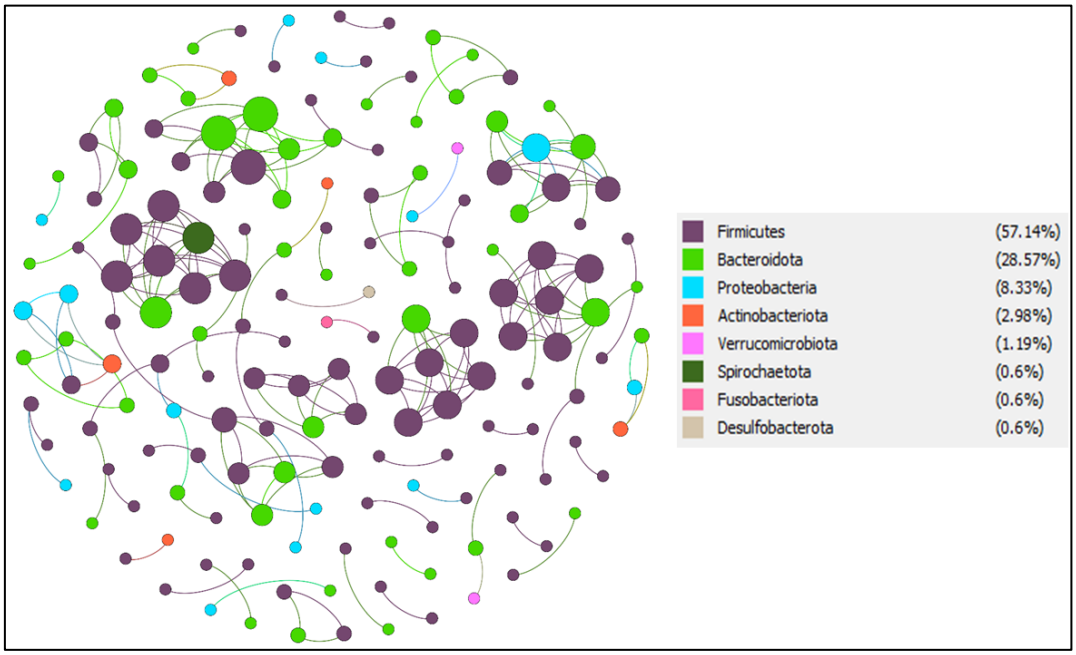


Figure 4.2 – T2D Co-occurrence Network (TCN). From total OTU abundance data, we select the T2D specific OTUs using the specified criteria, and a co-occurrence microbial network was constructed in Gephi.

4.3.2. Metabolic profile

Total of 6664 microbial pathways were predicted at community level through pathway prediction analysis using PICRUSt2 (Douglas et al., 2020) based on 16S rRNA marker gene. Among them total of 713 microbial metabolic pathways were identified as significantly (p-Value ≤ 0.05) differentially abundant pathways in between T2D and NGT group. 699 metabolic pathways were enriched in NGT group and from that top 20 highly significantly abundant microbial metabolic pathways were selected for study (Table 4.3). Remaining 14 microbial metabolic pathways were significantly enriched in T2D group (Table 4.2).

Table 4.2: Significantly enriched metabolic pathway prediction from 16S rRNA marker gene in T2D group considered in this study.

Pathway ID	Pathways	p-Value
K18302	membrane fusion protein, multidrug efflux system	0.02
K07051	uncharacterized protein	0.02
K15331	tRNA (uracil-5-)-methyltransferase [EC:2.1.1.35]	0.05
K17243	alpha-1,4-digalacturonate transport system permease protein	0.04
K17241	alpha-1,4-digalacturonate transport system substrate-binding protein	0.04
K19169	DNA sulfur modification protein DndB	0.04
K16874	2,5-furandicarboxylate decarboxylase 1	0.03
K10105	lipoyltransferase 1	0.04
K17311	trehalose transport system substrate-binding protein	0.04
K17312	trehalose transport system permease protein	0.04
K17313	trehalose transport system permease protein	0.04
K17335	nuclear factor of activated T-cells 5	0.04
K18022	glyceraldehyde dehydrogenase small subunit [EC:1.2.99.8]	0.04
K18282	cyanide dihydratase [EC:3.5.5.-]	0.05

Table 4.3: Top 20 significantly enriched metabolic pathway prediction from 16S rRNA marker gene in NGT group considered in this study.

Pathway ID	Pathways	p-Value
K06147	ATP-binding cassette, subfamily B, bacterial	0.03
K01990	ABC-2 type transport system ATP-binding protein	0.04
K02003	putative ABC transport system ATP-binding protein	0.03
K07024	sucrose-6-phosphatase [EC:3.1.3.24]	0.04
K01091	phosphoglycolate phosphatase [EC:3.1.3.18]	0.05
K03406	methyl-accepting chemotaxis protein	0.01
K02030	polar amino acid transport system substrate-binding protein	0.04
K03657	DNA helicase II / ATP-dependent DNA helicase PcrA [EC:3.6.4.12]	0.03
K00850	6-phosphofructokinase 1 [EC:2.7.1.11]	0.04
K03497	chromosome partitioning protein, ParB family	0.05
K03091	RNA polymerase sporulation-specific sigma factor	0.03
K03498	trk system potassium uptake protein	0.04
K03499	trk system potassium uptake protein	0.05
K04759	ferrous iron transport protein B	0.04
K03798	cell division protease FtsH [EC:3.4.24.-]	0.05
K02028	polar amino acid transport system ATP-binding protein [EC:3.6.3.21]	0.04
K03569	rod shape-determining protein MreB and related proteins	0.05
K03686	molecular chaperone DnaJ	0.04
K00945	CMP/dCMP kinase [EC:2.7.4.25]	0.05
K07258	serine-type D-Ala-D-Ala carboxypeptidase (penicillin-binding protein 5/6) [EC:3.4.16.4]	0.04

4.4. Discussions

In the past twenty years, diabetes mellitus has become the most common metabolic condition. Obesity and unhealthy weight gain have been linked to sedentary urban lifestyles, increased intake of processed and fried foods, and diets high in fat and protein, which disturb the normal physiological processes governing metabolic homeostasis. The importance of the gut microbiota in maintaining a healthy immune and metabolic system cannot be overstated. Gut microbiota has been shown to impact the pancreas directly. Gut microbiota has been proposed to modulate glucose

homeostasis through multiple mechanisms (Kootte et al., 2012; Harsch et al., 2018; Aydin et al., 2018; Aw et al., 018; Gérard et al., 2019). Four distinct methods by which the gut microbiome affects glucose homeostasis are supported by experimental evidence, those are –

(1) The metabolites produced by anaerobic microbial fermentation in the gut that have β cell-modulating effects (Priyadarshini et al., 2018; Brubaker et al., 2018; Gérard et al., 2019).
(2) Inflammatory cascades stimulate cytokine functions in the islets of Langerhans (Åkerfeldt et al., 2008; Ehses et al., 2008; Maslowski et al., 2009; Kamada et al., 2013; Tesi et al., 2021).
(3) Direct islet signalling influences insulin and glucagon secretion via incretin modulation (Gao et al., 2009; Tolhurst et al., 2012).
(4) Through modulation of incretins, direct islet signalling affects insulin and glucagon secretion (Myers et al., 2003).

Among these four mechanisms, 1 and 3 were mainly observed increased mainly in T2D susceptibility (Achparaki et al., 2012). Eubiosis is commonly used to describe an ideal bacterial population made up of 95% of Bacteroidetes and 5% Firmicutes that produce important natural microbial metabolites such as short-chain fatty acids (SCFAs) and branched-chain amino acids (BCAAs) and influence lipid metabolism. SCFAs like butyrate, acetate, and propionate are derived by anaerobic fermentation of undigested carbohydrates (dietary fibres) and help in maintaining intestinal integrity, prevent the epithelial layer, form tight junctions, and protect intestinal permeability (Macfarlane et al., 2012). These microbial secondary metabolites function as key components of the microorganism, activating signalling pathways. In T2D patients, the

majority of such specific microbiota associated with the production of these essential secondary metabolites are reduced which is also in line with our co-occurrence network (CN) analysis results.

Our co-occurrence network analysis showed that in T2D disease condition, significant changes in microbial network topological properties leads to a decrease in network stability and alteration in the microbial community in the human gastrointestinal tract, which is also in line with previous studies where they were reported, network complexity of the gut microbial community association was decreased in T2D (Li et al., 2020[a]). Interestingly co-occurrence network analysis also revealed that there are significant differences present in the proportion of taxonomic abundance of *Firmicutes* and *Bacteroidota* phylum in T2D compared to the NGT group which is also in line with the previously reported data (Turnbaugh et al., 2006; Ley et al., 2008; Zhang et al., 2013; Ahmad et al., 2019). The same trend was also observed in identified keystone taxa from the two co-occurrence networks and they might play an essential role in maintaining the microbial structure links, information transmission, and ecological function of the entire ecological communities in the gastrointestinal tract (Li et al., 2020[a,b], 2021).

The disruption of *Firmicutes* and *Bacteroidota* (B/F) ratio has been directly linked to insulin resistance (Irfan et al., 2022). Intestinal permeability is impacted by altered B/F ratio, and lipopolysaccharide (LPS) from proteobacteria is transmitted from the gut. The result of our metabolic prediction analysis from both T2D and NGT was in line with the previously reported studies. For example, through the activation of the immune system by LPS translocation, interleukin-1 (IL-1), tumour necrosis factor (TNF), Jun N-terminal kinases (JNK), and IkB kinase are all involved (IKK) in T2D

disease states (Irfan et al., 2022). The insulin signalling cascade is rendered inefficient as a result of JNK and IKK activation caused by LPS, which phosphorylates the insulin receptor substrate (IRS) but does not activate downstream effector molecules like PI3K and AKT (Velloso et al., 1996; Folli et al., 1997). IKK also stimulates nuclear factor kappa B translocation (NF-kB) and several other genes were involved in inflammatory and apoptotic responses are made more active by the transcription factor NF-kB. From our metabolic pathway prediction result we observed that significantly abundant microbial metabolic pathways are mostly associated with insulin resistance and inflammation in human T2D diseases condition which is also in line with the previous reports and the researchers were named this type of action as metabolic endotoxemia (Karin et al., 2000; Malle et al., 2015; Meyerovich et al., 2016; Li et al., 2020)

The results of this study demonstrate the significance of abundant harmful gut microbial members in T2D disease condition and their important role in elevation of disease condition. Also revealed the importance of a balanced gut microbiome and the use of the right amount of dietary fibre to promote fermentation and the production of advantageous SCFAs, which not only affect intestinal permeability but also directly and indirectly affect cell activity.

4.5. Conclusion

From the investigation in this study, the conclusions can be drawn that for communities and individuals suffering from obesity and diabetes, the use of pre or probiotics is necessary, as is a balanced diet that includes enough dietary fibre. More participants are anticipated to be included in future analyses, and research into the causal relationship is progressing. As a result, new strategies for treating and preventing microorganisms are being developed.

Chapter 5

Summary and Future Scopes of This Work

5.1. Summary of this work

Type 2 Diabetes (T2D) is a class of metabolic illnesses usually defined by chronic hyperglycemia resulting from diminished insulin production or impaired insulin action or both processes together, generating long-term implications. Persistent hyperglycemia is connected with regular micro and macrovascular issues. People with diabetes are at an increased risk of developing various health problems that may be life threatening, such as vascular disease that affects the heart, eyes, kidneys, and nerves. It has a complicated and multifaceted aetiology that incorporates genetic and environmental components and usually affects persons from the fourth decade of life; however, there has been a rise in the incidence of diabetes in children and young people. Recently, changes in the human gut microbiota have been connected with pathological states such as obesity and other metabolic disorders such as T2D, metabolic syndrome, and insulin resistance. Among the mechanisms that relate the intestinal microbiota with diabetes and insulin resistance, there is an increase in the permeability of the intestinal barrier, resulting in metabolic endotoxemia. These alterations may contribute to higher levels of obesity and impaired insulin signalling.

In *Chapter 1*, a systematic literature review on the microbial relationship with Type 2 Diabetes (T2D) disease status in various geographical locations. This thesis chapter also summarised existing diagnostic tools and suggested future treatments for the disease. This chapter aimed to summarise the pattern of the gut-microbial diversity in T2D conditions in different geographical regions, to improve disease diagnosis using currently available approaches.

In *Chapter 2*, the best T2D disease prediction models, as well as the most and the most relevant physiological parameters for better disease prediction regardless of geographic locational of 441 patient samples (T2D: 224 and NGT: 217) were chosen for this investigation, among them 345 data from Chinese cohorts and 96 data from European cohorts. This study used nine physiological parameters: BMI, FBG, FI, HbA1c, CHL, HDL, LDL, TGL, and CP. The RFE algorithm, a feature selection approach, was utilised to determine the best physiological characteristics that exhibited superior discrimination between T2D and healthy people. Three separate datasets, including training, testing, and blind datasets, were created to generate the prediction models. For disease prediction, three well-known machine learning algorithms were applied for SVM-L, SVM-R, and RF. The SEN, SPF, ACC, PRC, and F1 – Scores were calculated to analyse the performance of the prediction models. An unknown dataset collected from West Bengal, India was employed, to evaluate the prediction models more precisely. RF and SVM – R have outperformed other well-known models inaccuracy. These models also identify the two most critical physiological parameters, FBG and HbA1c, which play a more prominent role in T2D classification and diagnosis, as recommended by the American Diabetes Association (ADA) and the World Health Organization (WHO). So, we hypothesised in this chapter that significant changes in the level of both FBG and HbA1c during the development of diabetes can be used as critical physiological measurements to identify T2D disease or risk of illness in an impaired state around the world and that RF and SVM – R ML methods can also be used for better disease prediction. This work was published in Frontiers in Microbiology (De et al., 2022).

In *Chapter 3*, efforts were made to present a clear, resolved depiction of the human gut microbial diversity that affects T2D disease, particularly among people

living in West Bengal, Indian. In addition, this chapter attempted to identify distinct gut microbial members that can serve as essential biomarkers to improve disease diagnosis in conjunction with existing known approaches. A total of 34 samples (17 NGT and 17 T2D) from West Bengal were studied at IPGMER and SSKM Hospital in India, following the recommendations of the doctors in the Endocrine Department of the hospital, using World Health Organization (WHO) criteria. Anthropometric measurements were taken for all 34 samples (17 NGT and 17 T2D) from West Bengal. To detect patterns, a principal component analysis (PCA) was carried out on the data to see patterns among the samples based on their physiological parameters, and Kruskal-Walis rank-sum test was used to determine the differences in physiological parameters between T2D and healthy participants in this study. The composition of the microbial community was determined by sequencing the samples in the V3-V4 region of the 16S rRNA amplicon metagenome(Illumina MiSeq platform). Following this, several bioinformatics (such as quality trimming, merging, OTU clustering, taxonomy classification, and microbial diversity determination) and statistical analysis were carried out using the R programming language. The OTU number, the inverse Simpson index, and the Shannon diversity index were calculated within the α-diversity indices to examine the species richness and evenness and, along with this, to determine whether or not there was a statistically significant difference in α-diversity indices between the two groups Kruskal-Wallis test was used. We eliminated the uncommon species from the amplicon datasets for the β-diversity analysis and examined for differences at the community level using the PERMANOVA procedure. With the help of the Bray–Curtis dissimilarity matrix, we were able to identify differences in community structure between the sites, and the nonmetric multidimensional scaling (NMDS) plot was used to display patterns in the

bacterial community composition for each group. The redundancy analysis (RDA) was carried out to determine the contribution of physiological parameters to variance in the population. To investigate the differentially abundant OTUs in T2D and NGT, Dotplot analysis was performed, followed by a Spearman rank correlation test to examine the relationship between physiological parameters and those differentially abundant OTUs. The 16S rRNA marker gene was utilised for metagenome valuable content by conducting a phylogenetic examination of communities and reconstructing unobserved states.

The measured physiological parameters reveal that the T2D group was separated as a single cluster from the NGT group and with the FBG and HbA1c parameters. The bacterial communities of gut microbiota were dominated by Bacteroidota, Firmicutes, and Proteobacteria members into both gr. Among them, 27 microbial genera were identified as core gut microbial members in the studied samples, including*Prevotella_9*, *Prevotella*, *Prevotellaceae*IncertaeSedis, *Bacteroides*, *Alloprevotella*, *Lachnospiraceae*IncertaeSedis, *Roseburia*, *Faecalibacterium*, *Megasphaera*, and *Succinivibrio*. *Eubacterium eligens* group, *Lachnoclostridium*, *Ruminococcus torques* group, *Clostridia* vadinBB60 group Incertaesedis, *Lachnospira*, *Haemophilus*, and *Catenibacterium* genera were identified as unique bacterial members for the T2D microbiome. These gut microbial genera were reported to act in mucus degradation by decreasing the gut barrier integrity and are found abundantly in the T2D disease state. They can be used as biomarkers for disease diagnosis.On the other hand,*Alistipes*, *Muribaculaceae*Incertaesedis, *Ligilactobacillus*, *Holdemanella*, *Enterobacter*, *Blautia*, and *Coprococcus* genera were observed only in the NGT group as unique gut microbial members and reported that they are dominant in the normal human

gastrointestinal tract and have an important key role in protection from many diseases like liver and cardiovascular fibrotic disorders and also from various pathogens. It is also reported that they can make metabolic improvements and consorted with a higher quality of life indicators. FBG and HbA1c were the most determinant variable among the parameters and influenced the microbial community composition. The study presented in this chapter gives a clear picture of bacterial diversity and its relationship to critical physiological parameters that determine the T2D disease in West Bengal, India, by decreasing the SCFA and butyrate-producing core bacteria, which are favourable to the human gut. Also, we proposed that, in addition to well-established physiological measures, unique gut microorganisms can be employed as an essential biomarker to aid disease diagnosis. This work was published in Frontiers in Microbiology (De et al., 2022).

In *Chapter 4*, efforts were made to present a clear, resolved depiction of the structural and functional role of most abundant human gut microbial members that may affects T2D disease, particularly among people living in West Bengal, Indian. In addition, this chapter attempted to identify distinct gut microbial members interaction and metabolic pathways during disease condition that can serve as essential biomarkers to improve disease diagnosis in conjunction with existing known approaches. We eliminated the rare biosphere to give a clear resolve picture of interaction between gut microbial members in a particular T2D disease condition with compared to healthy individuals based on OTU to OTU Spearman's Rank Correlation analysis and try to identify the keystone taxa for that particular habitat. Then the metabolic pathways had been predicted from the representative sequences of those OTUs which are selected interaction study.

The co-occurrence network analysis revealed that due to the decrease in network topology and different gut microbial compositions, the network stability also decreases in T2D disease state compared to healthy. Also identified keystone taxa for example 5 OTUs were found under the phylum *Firmicutes*, 4 for *Bacteroidota*, 3 for *Proteobacteria*, one for *Actinobacteriota* and one for *Spirochaetota* gut microbial phyla in T2D network. In contrast, 2 OTUs were found under the phylum *Bacteroidota*, 3 for *Firmicutes*, one for *Proteobacteria*, one for *Patescibacteria* and one for *Desulfobacterota* as keystone microbial phyla for healthy network. The metabolic prediction result indicates in T2D disease condition the abundant microbial metabolic pathways were mainly associated with insulin resistance and inflammation. This work was published in Frontiers in Microbiology (De et al., 2022).

From the summary of all work in this thesis, following conclusions can be drawn:

(1) Chapter 1 – Epidemiological studies provide a clear indication of the association between gut microbiota disturbance and increased incidence of T2D. Impaired energy metabolism has been proposed as a driving force for this metabolic disease associated with the perturbation in gut microbiota which causes obesity that in turn induces T2D disease.

(2) **Chapter 2 –** Random Forest (RF) and support vector machine with RBF Kernel (SVM–R) are the best prediction models to predict the T2D and normal state based on a patient's physiological condition. Also, Fasting blood glucose (FBG) and HbA1c individually or together can be used for the T2D diagnosis as well as defining the disease in an impaired state.

(3) **Chapter 3 –** Both of FBG and HbA1c physiological parameters coincided with the core microbial community composition of the T2D microbiome by

decreasing the beneficiary core gut microbial members. Also, unique gut microbial members (e.g., *Catenibacterium*, *Eubacterium eligens group*, *Lachnoclostridium* etc.) can be used as important biomarkers for Indian T2D patients.

(4) **Chapter 4 –** The topology study of co-occurrence network analysis indicates that changes in network complexity in T2D lead to variations in the different gut microbial members compared to NGT. The metabolic pathway prediction revealed that abundant microbial metabolic pathways in T2D diseases condition are mostly associated with insulin resistance and inflammation.

5.2. Future Scopes of this work

The Indian population size is large and has diverse dietary compositions or food habits with large metabolic differences. Recently, one report on the gut microbiota of T2D from the western part of India (Maharashtra, especially, in and around the city, Pune); however, none are from other regions/parts of this country (Gaike et al., 2020). In this study, we were the first to provide the preliminary information on the functional role of gut microbiome of Indian T2D patients from the eastern region of the Indian Subcontinent, especially, in and around the Kolkata, West Bengal, with almost similar dietary status and this seems to restrict us from increasing the sample size. This is a preliminary dataset that will help us formulate strategies to collect more samples from a diverse population for a deep understanding of the gut microbiome in Indian T2D patients. With the increase in the sample size, we will be able to perform more in-depth microbial diversity analysis and learn more about what governs the distribution of gut microbial taxa and how these distributions, as well as

their ecosystem contributions in Indian T2D patients, will help to improve more accurate diagnosis of T2D disease in the future.

Further and more extensive research can be carried out to understand the microbial communities present better, and their structural patterns can be expanded upon by doing detailed studies of gastrointestinal tract microbiota to look for any changes in patterns by sampling various sites across India overall timeline with a high number of samples. This will strengthen previous community pattern findings while confirming the relationship between physiological parameter change and microbial community structure. Microbial communities can be analyzed and compared across a vertical profile and a horizontal gradient.

The human gut microbiota plays a critical role in human health. Treatment of T2D metabolic disorder will be improved with a better knowledge of the mechanisms involved in this connection. Faecal transplants have already been conducted and have proven to be quite successful in treating metabolic syndrome. In the future, it will be feasible to manipulate the microbiota better to treat a variety of ailments, most likely in a more regulated manner than transplanting entire communities from one person to another. It would also be interesting to examine the differences in core gut microbial populations and identify the unique gut microbiota in disease states. A more profound and more extensive analysis of the sequences obtained from metagenomes can be carried out to identify the taxonomic identity of previously undocumented microbial diversity. The study of the human gut microbiome in health and disease has gotten off to a tremendous start. A lot of information has been gathered, particularly in the era of next-generation sequencing. However, there are many unanswered questions, and I believe we may expect more exciting results in the coming years. I am confident that

these will contribute to a better knowledge and diagnosis of disease in various ways, thereby enhancing the quality of life for many individuals.

References

- Afolayan, A. O., Adebusoye, L. A., Cadmus, E. O., and Ayeni, F. A. (2020). Insights into the gut microbiota of Nigerian elderly with type 2 diabetes and non-diabetic elderly persons. *Heliyon* 6, e03971. doi: 10.1016/J.HELIYON.2020.E03971.
- Ahmad, A., Yang, W., Chen, G., Shafiq, M., Javed, S., Zaidi, S. S. A., et al. (2019). Analysis of gut microbiota of obese individuals with type 2 diabetes and healthy individuals. *PLoS One* 14, e0226372. doi: 10.1371/JOURNAL.PONE.0226372.
- Akbari, M., and Hassan-Zadeh, V. (2020). The inflammatory effect of epigenetic factors and modifications in type 2 diabetes. *Inflammopharmacology* 28, 345–362. doi: 10.1007/S10787-019-00663-9/TABLES/3.
- Alexander, M., and Turnbaugh, P. J. (2020). Deconstructing Mechanisms of Diet-Microbiome-Immune Interactions. *Immunity* 53, 264–276. doi: 10.1016/J.IMMUNI.2020.07.015.
- Almugadam, B. S., Liu, Y., Chen, S. M., Wang, C. H., Shao, C. Y., Ren, B. W., et al. (2020). Alterations of Gut Microbiota in Type 2 Diabetes Individuals and the Confounding Effect of Antidiabetic Agents. *J. Diabetes Res.* 2020. doi: 10.1155/2020/7253978.
- Anhê, F. F., Varin, T. V., Le Barz, M., Desjardins, Y., Levy, E., Roy, D., et al. (2015). Gut Microbiota Dysbiosis in Obesity-Linked Metabolic Diseases and Prebiotic Potential of Polyphenol-Rich Extracts. *Curr. Obes. Rep.* 4, 389–400. doi: 10.1007/S13679-015-0172-9/FIGURES/1.
- Anjana, R. M., Deepa, M., Pradeepa, R., Mahanta, J., Narain, K., Das, H. K., et al. (2017). Prevalence of diabetes and prediabetes in 15 states of India: results from the ICMR-INDIAB population-based cross-sectional study. *lancet.*

Diabetes Endocrinol. 5, 585–596. doi: 10.1016/S2213-8587(17)30174-2.

- Anjana, R. M., Pradeepa, R., Deepa, M., Datta, M., Sudha, V., Unnikrishnan, R., et al. (2011). Prevalence of diabetes and prediabetes (impaired fasting glucose and/or impaired glucose tolerance) in urban and rural India: phase I results of the Indian Council of Medical Research-INdia DIABetes (ICMR-INDIAB) study. *Diabetologia* 54, 3022–3027. doi: 10.1007/S00125-011-2291-5.
- Anuja Kumari, V., and Chitra, R. Classification Of Diabetes Disease Using Support Vector Machine. *Int. J. Eng. Res. Appl. www.ijera.com* 3, 1797–1801. Available at: www.ijera.com [Accessed July 24, 2022].
- Aoki, R., Kamikado, K., Suda, W., Takii, H., Mikami, Y., Suganuma, N., et al. (2017). A proliferative probiotic Bifidobacterium strain in the gut ameliorates progression of metabolic disorders via microbiota modulation and acetate elevation. *Sci. Reports 2017 71* 7, 1–10. doi: 10.1038/srep43522.
- Arbabshirani, M. R., Fornwalt, B. K., Mongelluzzo, G. J., Suever, J. D., Geise, B. D., Patel, A. A., et al. (2018). Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. *npj Digit. Med. 2018 11* 1, 1–7. doi: 10.1038/s41746-017-0015-z.
- Balakumar, M., Prabhu, D., Sathishkumar, C., Prabu, P., Rokana, N., Kumar, R., et al. (2018). Improvement in glucose tolerance and insulin sensitivity by probiotic strains of Indian gut origin in high-fat diet-fed C57BL/6J mice. *Eur. J. Nutr.* 57, 279–295. doi: 10.1007/S00394-016-1317-7/FIGURES/9.
- Ballan, R., and Saad, S. M. I. (2021). Characteristics of the Gut Microbiota and Potential Effects of Probiotic Supplements in Individuals with Type 2 Diabetes

mellitus. *Foods 2021, Vol. 10, Page 2528* 10, 2528. doi: 10.3390/FOODS10112528.

- Baothman, O. A., Zamzami, M. A., Taher, I., Abubaker, J., and Abu-Farha, M. (2016). The role of Gut Microbiota in the development of obesity and Diabetes. 15, 1–8. doi: 10.1186/S12944-016-0278-4.
- Battson, M. L., Lee, D. M., Weir, T. L., and Gentile, C. L. (2018). The gut microbiota as a novel regulator of cardiovascular function and disease. *J. Nutr. Biochem.* 56, 1–15. doi: 10.1016/J.JNUTBIO.2017.12.010.
- Baxter, N. T., Schmidt, A. W., Venkataraman, A., Kim, K. S., Waldron, C., and Schmidt, T. M. (2019). Dynamics of human gut microbiota and short-chain fatty acids in response to dietary interventions with three fermentable fibers. *MBio* 10. doi: 10.1128/MBIO.02566-18/ASSET/88C0E415-6A89-47FC-B94A-BD9D9EE21EB4/ASSETS/GRAPHIC/MBIO.02566-18-F0005.JPEG.
- Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M. C. C., Charles, T., et al. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome* 8. doi: 10.1186/S40168-020-00875-0.
- Bezirtzoglou, E., Stavropoulou, E., Kantartzi, K., Tsigalou, C., Voidarou, C., Mitropoulou, G., et al. (2021). Maintaining Digestive Health in Diabetes: The Role of the Gut Microbiome and the Challenge of Functional Foods. *Microorg. 2021, Vol. 9, Page 516* 9, 516. doi: 10.3390/MICROORGANISMS9030516.
- Bhute, S. S., Ghaskadbi, S. S., and Shouche, Y. S. (2017a). Rare biosphere in human gut: A less explored component of human gut microbiota and its association with human health. *Min. Microb. Wealth MetaGenomics*, 133–142. doi: 10.1007/978-981-10-5708-3_8/COVER.
- Bhute, S. S., Suryavanshi, M. V., Joshi, S. M., Yajnik, C. S., Shouche, Y. S.,

and Ghaskadbi, S. S. (2017b). Gut microbial diversity assessment of Indian type-2-diabetics reveals alterations in eubacteria, archaea, and eukaryotes. *Front. Microbiol.* 8, 214. doi: 10.3389/FMICB.2017.00214/BIBTEX.

- Bianchi, F., Larsen, N., de Mello Tieghi, T., Adorno, M. A. T., Kot, W., Saad, S. M. I., et al. (2018). Modulation of gut microbiota from obese individuals by in vitro fermentation of citrus pectin in combination with Bifidobacterium longum BB-46. *Appl. Microbiol. Biotechnol.* 102, 8827–8840. doi: 10.1007/S00253-018-9234-8/FIGURES/8.
- Bianchi, F., Larsen, N., Tieghi, T. de M., Adorno, M. A. T., Saad, S. M. I., Jespersen, L., et al. (2019). In vitro modulation of human gut microbiota composition and metabolites by Bifidobacterium longum BB-46 and a citric pectin. *Food Res. Int.* 120, 595–602. doi: 10.1016/J.FOODRES.2018.11.010.
- Bicsak, T. A., Walsh, B., and Fineman, M. (2017). Metformin-associated lactic acidosis: Moving towards a new paradigm? *Diabetes. Obes. Metab.* 19, 1499. doi: 10.1111/DOM.12994.
- Bonfili, L., Cecarini, V., Gogoi, O., Gong, C., Cuccioloni, M., Angeletti, M., et al. (2021). Microbiota modulation as preventative and therapeutic approach in Alzheimer's disease. *FEBS J.* 288, 2836–2855. doi: 10.1111/FEBS.15571.
- Boulangé, C. L., Neves, A. L., Chilloux, J., Nicholson, J. K., and Dumas, M. E. (2016). Impact of the gut microbiota on inflammation, obesity, and metabolic disease. *Genome Med.* 8. Available at: /pmc/articles/PMC4839080/ [Accessed April 12, 2022].
- Cani, P. D., and Delzenne, N. M. (2009). Interplay between obesity and associated metabolic disorders: new insights into the gut microbiota. *Curr. Opin. Pharmacol.* 9, 737–743. doi: 10.1016/J.COPH.2009.06.016.

- Caricilli, A. M., and Saad, M. J. A. (2013). The Role of Gut Microbiota on Insulin Resistance. *Nutr. 2013, Vol. 5, Pages 829-851* 5, 829–851. doi: 10.3390/NU5030829.
- Chahal, A., Gulia, P., and Gill, N. S. (2022). Different analytical frameworks and bigdata model for internet of things Analysis,Design and Optimization of Classification Techniques for Performance Improvement in DataMining View project Ph.D research View project Different analytical frameworks and bigdata model for Internet of Things. *Artic. Indones. J. Electr. Eng. Comput. Sci.* 25, 1159–1166. doi: 10.11591/ijeecs.v25.i2.pp1159-1166.
- Chaki, J., Thillai Ganesh, S., Cidham, S. K., and Ananda Theertan, S. (2022). Machine learning and artificial intelligence based Diabetes Mellitus detection and self-management: A systematic review. *J. King Saud Univ. - Comput. Inf. Sci.* 34, 3204–3225. doi: 10.1016/J.JKSUCI.2020.06.013.
- Chang, H.-W., Yan, D., Singh, R., Bui, A., Lee, K., Truong, A., et al. (2022). Multi-omic Analysis of the Gut Microbiome in Psoriasis Reveals Distinct Host-Microbe Associations. *JID Innov.*, 100115. doi: 10.1016/j.xjidi.2022.100115.
- Chen, E. Y., and Seeburg, P. H. (2009). Supercoil Sequencing: A Fast and Simple Method for Sequencing Plasmid DNA. *http://www.liebertpub.com/dna* 4, 165–170. doi: 10.1089/DNA.1985.4.165.
- Chen, J. I., Yang, Y., Yu, N., Sun, W., Yang, Y., and Zhao, M. I. (2022). Relationship between gut microbiome characteristics and the effect of nutritional therapy on glycemic control in pregnant women with gestational diabetes mellitus. *PLoS One* 17, e0267045. doi: 10.1371/JOURNAL.PONE.0267045.
- Chen, J., Ryu, E., Hathcock, M., Ballman, K., Chia, N., Olson, J. E., et al.

(2016). Impact of demographics on human gut microbial diversity in a US Midwest population. *PeerJ* 2016, 1–20. doi: 10.7717/peerj.1514.

- Chen, P. C., Chien, Y. W., and Yang, S. C. (2019). The alteration of gut microbiota in newly diagnosed type 2 diabetic patients. *Nutrition* 63–64, 51–56. doi: 10.1016/J.NUT.2018.11.019.
- Chen, Z., Radjabzadeh, D., Chen, L., Kurilshikov, A., Kavousi, M., Ahmadizar, F., et al. (2021). Association of Insulin Resistance and Type 2 Diabetes With Gut Microbial Diversity. *JAMA Netw. Open* 4, e2118811. doi: 10.1001/jamanetworkopen.2021.18811.
- Cho, N. H., Shaw, J. E., Karuranga, S., Huang, Y., da Rocha Fernandes, J. D., Ohlrogge, A. W., et al. (2018). IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes Res. Clin. Pract.* 138, 271–281. doi: 10.1016/J.DIABRES.2018.02.023.
- Cipriani, S., Mencarelli, A., Palladino, G., and Fiorucci, S. (2010). FXR activation reverses insulin resistance and lipid abnormalities and protects against liver steatosis in Zucker (fa/fa) obese rats. *J. Lipid Res.* 51, 771–784. doi: 10.1194/JLR.M001602/ATTACHMENT/F0C1B013-DB1C-4AF9-94A3-33A4C985D2DB/MMC1.ZIP.
- Connolly, M. L., Lovegrove, J. A., and Tuohy, K. M. (2010). Konjac glucomannan hydrolysate beneficially modulates bacterial composition and activity within the faecal microbiota. *J. Funct. Foods* 2, 219–224. doi: 10.1016/J.JFF.2010.05.001.
- Connolly, V., Unwin, N., Sherriff, P., Bilous, R., and Kelly, W. (2000). Diabetes prevalence and socioeconomic status: a population based study showing increased prevalence of type 2 diabetes mellitus in deprived areas. *J. Epidemiol.*

Community Heal. 54, 173–177. doi: 10.1136/JECH.54.3.173.

- Costello, E. K., Lauber, C. L., Hamady, M., Fierer, N., Gordon, J. I., and Knight, R. (2009). Bacterial community variation in human body habitats across space and time. *Science* 326, 1694–1697. doi: 10.1126/SCIENCE.1177486.
- Costello, M. E., Robinson, P. C., Benham, H., and Brown, M. A. (2015). The intestinal microbiome in human disease and how it relates to arthritis and spondyloarthritis. *Best Pract. Res. Clin. Rheumatol.* 29, 202–212. doi: 10.1016/J.BERH.2015.08.001.
- Couzin-Frankel, J. (2010). Bacteria and asthma: Untangling the links. *Science (80-.).* 330, 1168–1169. doi: 10.1126/SCIENCE.330.6008.1168/ASSET/ED93A2B4-4BC8-42A5-9795-B3D95E7FA0A8/ASSETS/GRAPHIC/330_1168_F3.GIF.
- Cox, L. M., and Blaser, M. J. (2014). Antibiotics in early life and obesity. *Nat. Rev. Endocrinol. 2014 113* 11, 182–190. doi: 10.1038/nrendo.2014.210.
- Cummings, J. H., and Macfarlane, G. T. (2002). Gastrointestinal effects of prebiotics. *Br. J. Nutr.* 87, S145–S151. doi: 10.1079/BJN/2002530.
- Dahl, C. J., Lutz, A., and Davidson, R. J. (2015). Reconstructing and deconstructing the self: cognitive mechanisms in meditation practice. *Trends Cogn. Sci.* 19, 515–523. doi: 10.1016/J.TICS.2015.07.001.
- Dang, F., Jiang, Y., Pan, R., Zhou, Y., Wu, S., Wang, R., et al. (2018). Administration of Lactobacillus paracasei ameliorates type 2 diabetes in mice. *Food Funct.* 9, 3630–3639. doi: 10.1039/C8FO00081F.
- Das, T., Jayasudha, R., Chakravarthy, S. K., Prashanthi, G. S., Bhargava, A., Tyagi, M., et al. (2021). Alterations in the gut bacterial microbiome in people with type 2 diabetes mellitus and diabetic retinopathy. *Sci. Reports 2021 111*

11, 1–15. doi: 10.1038/s41598-021-82538-0.

- Diabetes Mellitus: Type 1, Type 2, and Gestational Diabetes Available at: https://www.webmd.com/diabetes/guide/types-of-diabetes-mellitus [Accessed August 20, 2021].
- Doumatey, A. P., Adeyemo, A., Zhou, J., Lei, L., Adebamowo, S. N., Adebamowo, C., et al. (2020). Gut Microbiome Profiles Are Associated With Type 2 Diabetes in Urban Africans. *Front. Cell. Infect. Microbiol.* 10, 63. doi: 10.3389/FCIMB.2020.00063/BIBTEX.
- Duparc, T., Plovier, H., Marrachelli, V. G., Van Hul, M., Essaghir, A., Ståhlman, M., et al. (2017). Hepatocyte MyD88 affects bile acids, gut microbiota and metabolome contributing to regulate glucose and lipid metabolism. *Gut* 66, 620–632. doi: 10.1136/GUTJNL-2015-310904.
- Fåk, F., and Bäckhed, F. (2012). Lactobacillus reuteri Prevents Diet-Induced Obesity, but not Atherosclerosis, in a Strain Dependent Fashion in Apoe−/− Mice. *PLoS One* 7, e46837. doi: 10.1371/JOURNAL.PONE.0046837.
- Fatima, M., Pasha, M., Fatima, M., and Pasha, M. (2017). Survey of Machine Learning Algorithms for Disease Diagnostic. *J. Intell. Learn. Syst. Appl.* 9, 1–16. doi: 10.4236/JILSA.2017.91001.
- Fuller, S., Beck, E., Salman, H., and Tapsell, L. (2016). New Horizons for the Study of Dietary Fiber and Health: A Review. *Plant Foods Hum. Nutr.* 71, 1–12. doi: 10.1007/S11130-016-0529-6/TABLES/2.
- Gaike, A. H., Paul, D., Bhute, S., Dhotre, D. P., Pande, P., Upadhyaya, S., et al. No Title. 5. Available at: https://journals.asm.org/doi/10.1128/mSystems.00578-19 [Accessed July 23, 2022].

- Galicia-Garcia, U., Benito-Vicente, A., Jebari, S., Larrea-Sebal, A., Siddiqi, H., Uribe, K. B., et al. (2020). Pathophysiology of Type 2 Diabetes Mellitus. *Int. J. Mol. Sci.* 21, 1–34. doi: 10.3390/IJMS21176275.
- Gao, X., Zhang, M., Xue, J., Huang, J., Zhuang, R., Zhou, X., et al. (2018). Body mass index differences in the gut microbiota are gender specific. *Front. Microbiol.* 9, 1250. doi: 10.3389/FMICB.2018.01250/BIBTEX.
- Gholizadeh, P., Mahallei, M., Pormohammad, A., Varshochi, M., Ganbarov, K., Zeinalzadeh, E., et al. (2019). Microbial balance in the intestinal microbiota and its association with diabetes, obesity and allergic disease. *Microb. Pathog.* 127, 48–55. doi: 10.1016/J.MICPATH.2018.11.031.
- Global report on diabetes Available at: https://www.who.int/publications/i/item/9789241565257 [Accessed August 20, 2021].
- Graessler, J., Qin, Y., Zhong, H., Zhang, J., Licinio, J., Wong, M. L., et al. (2012). Metagenomic sequencing of the human gut microbiome before and after bariatric surgery in obese patients with type 2 diabetes: correlation with inflammatory and metabolic parameters. *Pharmacogenomics J. 2013 136* 13, 514–522. doi: 10.1038/tpj.2012.43.
- Guo, X., Yang, B., Tan, J., Jiang, J., and Li, D. (2016a). Associations of dietary intakes of anthocyanins and berry fruits with risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective cohort studies. *Eur. J. Clin. Nutr. 2016 7012* 70, 1360–1367. doi: 10.1038/ejcn.2016.142.
- Guo, Y., Qi, Y., Yang, X., Zhao, L., Wen, S., Liu, Y., et al. (2016b). Association between Polycystic Ovary Syndrome and Gut Microbiota. *PLoS One* 11, e0153196. doi: 10.1371/JOURNAL.PONE.0153196.

- Gurung, M., Li, Z., You, H., Rodrigues, R., Jump, D. B., Morgun, A., et al. (2020). Role of gut microbiota in type 2 diabetes pathophysiology. *EBioMedicine* 51, 102590. doi: 10.1016/J.EBIOM.2019.11.051.
- Ha, N., Cipriani, L. A., de Oliveira, N. S., Uczay, J., Pessatti, M. L., and Fabregat, T. E. H. P. (2022). Peptide profile of the sardine protein hydrolysate affects food utilization and intestinal microbiota of Nile tilapia. *Aquac. Int.* 30, 365–382. doi: 10.1007/S10499-021-00804-4/TABLES/9.
- Hameed, I., Masoodi, S. R., Mir, S. A., Nabi, M., Ghazanfar, K., and Ganai, B. A. (2015). Type 2 diabetes mellitus: From a metabolic disorder to an inflammatory condition. *World J. Diabetes* 6, 598. doi: 10.4239/WJD.V6.I4.598.
- Hegde, S., Lin, Y. M., Golovko, G., Khanipov, K., Cong, Y., Savidge, T., et al. (2018). Microbiota dysbiosis and its pathophysiological significance in bowel obstruction. *Sci. Reports 2018 81* 8, 1–12. doi: 10.1038/s41598-018-31033-0.
- Heianza, Y., Sun, D., Li, X., Didonato, J. A., Bray, G. A., Sacks, F. M., et al. (2019). Gut microbiota metabolites, amino acid metabolites and improvements in insulin sensitivity and glucose metabolism: the POUNDS Lost trial. *Gut* 68, 263–270. doi: 10.1136/GUTJNL-2018-316155.
- Herath, M., Hosie, S., Bornstein, J. C., Franks, A. E., and Hill-Yardin, E. L. (2020). The Role of the Gastrointestinal Mucus System in Intestinal Homeostasis: Implications for Neurological Disorders. *Front. Cell. Infect. Microbiol.* 10, 248. doi: 10.3389/FCIMB.2020.00248/BIBTEX.
- Heydari, I., Radi, V., Razmjou, S., and Amiri, A. (2010). Chronic complications of diabetes mellitus in newly diagnosed patients. *Int. J. Diabetes Mellit.* 2, 61–63. doi: 10.1016/J.IJDM.2009.08.001.

- http://ljournal.ru/wp-content/uploads/2016/08/d-2016-154.pdf (2016). *http://ljournal.ru/wp-content/uploads/2016/08/d-2016-154.pdf*. doi: 10.18411/D-2016-154.
- Huang, E. S., Brown, S. E. S., Ewigman, B. G., Foley, E. C., and Meltzer, D. O. (2007). Patient perceptions of quality of life with diabetes-related complications and treatments. *Diabetes Care* 30, 2478–2483. doi: 10.2337/DC07-0499.
- Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., et al. (2012). Structure, function and diversity of the healthy human microbiome. *Nat. 2012 4867402* 486, 207–214. doi: 10.1038/nature11234.
- Iyer, A., Jeyalatha, S., and Sumbaly, R. (2015). DIAGNOSIS OF DIABETES USING CLASSIFICATION MINING TECHNIQUES. *Int. J. Data Min. Knowl. Manag. Process* 5. doi: 10.5121/ijdkp.2015.5101.
- Jiang, Z., Sun, T.-Y., He, Y., Gou, W., Zuo, L.-S.-Y., Fu, Y., et al. Dietary fruit and vegetable intake, gut microbiota, and type 2 diabetes: results from two large human cohort studies. doi: 10.1186/s12916-020-01842-0.
- Kamil, R. Z., Murdiati, A., Juffrie, M., Nakayama, J., and Rahayu, E. S. (2021). Gut microbiota and short-chain fatty acid profile between normal and moderate malnutrition children in Yogyakarta, Indonesia. *Microorganisms* 9, 1–15. doi: 10.3390/microorganisms9010127.
- Karlsson, F. H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C. J., Fagerberg, B., et al. (2013a). Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nat. 2013 4987452* 498, 99–103. doi: 10.1038/nature12198.

- Karlsson, F., Tremaroli, V., Nielsen, J., and Bäckhed, F. (2013b). Assessing the Human Gut Microbiota in Metabolic Diseases. *Diabetes* 62, 3341–3349. doi: 10.2337/DB13-0844.
- Karlsson, J., Persson, L. O., Sjöström, L., and Sullivan, M. (2000). Psychometric properties and factor structure of the Three-Factor Eating Questionnaire (TFEQ) in obese men and women. Results from the Swedish Obese Subjects (SOS) study. *Int. J. Obes. Relat. Metab. Disord.* 24, 1715–1725. doi: 10.1038/SJ.IJO.0801442.
- Kasubuchi, M., Hasegawa, S., Hiramatsu, T., Ichimura, A., and Kimura, I. (2015). Dietary Gut Microbial Metabolites, Short-chain Fatty Acids, and Host Metabolic Regulation. *Nutr. 2015, Vol. 7, Pages 2839-2849* 7, 2839–2849. doi: 10.3390/NU7042839.
- Kerner, W., and Brückel, J. (2014). Definition, classification and diagnosis of diabetes mellitus. *Exp. Clin. Endocrinol. Diabetes* 122, 384–386. doi: 10.1055/S-0034-1366278.
- Kesika, P., Sivamaruthi, B. S., and Chaiyasut, C. (2019). Do Probiotics Improve the Health Status of Individuals with Diabetes Mellitus? A Review on Outcomes of Clinical Trials. *Biomed Res. Int.* 2019. doi: 10.1155/2019/1531567.
- Kikuchi, R., Irie, J., Yamada-Goto, N., Kikkawa, E., Seki, Y., Kasama, K., et al. (2018). The Impact of Laparoscopic Sleeve Gastrectomy with Duodenojejunal Bypass on Intestinal Microbiota Differs from that of Laparoscopic Sleeve Gastrectomy in Japanese Patients with Obesity. *Clin. Drug Investig. 2018 386* 38, 545–552. doi: 10.1007/S40261-018-0638-0.
- Kitten, A. K., Ryan, L., Lee, G. C., Flores, B. E., and Reveles, K. R. (2021).

Gut microbiome differences among Mexican Americans with and without type 2 diabetes mellitus. *PLoS One* 16, 1–10. doi: 10.1371/journal.pone.0251245.

- Kolb, H., and Martin, S. (2017). Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes. *BMC Med.* 15. doi: 10.1186/S12916-017-0901-X.
- Kononenko, I. (2001). Machine learning for medical diagnosis: history, state of the art and perspective. *Artif. Intell. Med.* 23, 89–109. doi: 10.1016/S0933-3657(01)00077-X.
- Larsen, N., Vogensen, F. K., Van Den Berg, F. W. J. J., Nielsen, D. S., Andreasen, A. S., Pedersen, B. K., et al. (2010). Gut Microbiota in Human Adults with Type 2 Diabetes Differs from Non-Diabetic Adults. *PLoS One* 5, e9085. doi: 10.1371/journal.pone.0009085.
- Lee, H., Lee, Y., Kim, J., An, J., Lee, S., Kong, H., et al. (2018). Modulation of the gut microbiota by metformin improves metabolic profiles in aged obese mice. *https://doi.org/10.1080/19490976.2017.1405209* 9, 155–165. doi: 10.1080/19490976.2017.1405209.
- Li, Q., Chang, Y., Zhang, K., Chen, H., Tao, S., and Zhang, Z. (2020a). Implication of the gut microbiome composition of type 2 diabetic patients from northern China. *Sci. Reports 2020 101* 10, 1–8. doi: 10.1038/s41598-020-62224-3.
- Li, Q., Jin, M., Liu, Y., and Jin, L. (2020b). Gut Microbiota: Its Potential Roles in Pancreatic Cancer. *Front. Cell. Infect. Microbiol.* 10, 564. doi: 10.3389/FCIMB.2020.572492/BIBTEX.
- Lim, S. M., Jeong, J. J., Woo, K. H., Han, M. J., and Kim, D. H. (2016). Lactobacillus sakei OK67 ameliorates high-fat diet–induced blood glucose

intolerance and obesity in mice by inhibiting gut microbiota lipopolysaccharide production and inducing colon tight junction protein expression. *Nutr. Res.* 36, 337–348. doi: 10.1016/J.NUTRES.2015.12.001.

- Makishima, M., Lu, T. T., Xie, W., Whitfield, G. K., Domoto, H., Evans, R. M., et al. (2002). Vitamin D receptor as an intestinal bile acid sensor. *Science (80-.).* 296, 1313–1316. doi: 10.1126/SCIENCE.1070477/SUPPL_FILE/1070477S1_THUMB.GIF.
- Makki, K., Deehan, E. C., Walter, J., and Bäckhed, F. (2018). The Impact of Dietary Fiber on Gut Microbiota in Host Health and Disease. *Cell Host Microbe* 23, 705–715. doi: 10.1016/J.CHOM.2018.05.012.
- Marchesi, J. R., Adams, D. H., Fava, F., Hermes, G. D. A., Hirschfield, G. M., Hold, G., et al. (2016). The gut microbiota and host health: a new clinical frontier. *Gut* 65, 330. doi: 10.1136/GUTJNL-2015-309990.
- Martínez, I., Lattimer, J. M., Hubach, K. L., Case, J. A., Yang, J., Weber, C. G., et al. (2012). Gut microbiome composition is linked to whole grain-induced immunological improvements. *ISME J. 2013 72* 7, 269–280. doi: 10.1038/ismej.2012.104.
- Martinic, A., Barouei, J., Bendiks, Z., Mishchuk, D., Heeney, D. D., Martin, R., et al. (2018). Supplementation of Lactobacillus plantarum Improves Markers of Metabolic Dysfunction Induced by a High Fat Diet. *J. Proteome Res.* 17, 2790–2802. doi: 10.1021/ACS.JPROTEOME.8B00282/ASSET/IMAGES/LARGE/PR-2018-002829_0007.JPEG.
- Maskarinec, G., Raquinio, P., Kristal, B. S., Setiawan, V. W., Wilkens, L. R., Franke, A. A., et al. (2021). The gut microbiome and type 2 diabetes status in

the Multiethnic Cohort. *PLoS One* 16, e0250855. doi: 10.1371/JOURNAL.PONE.0250855.

- Mateo-Gallego, R., Moreno-Indias, I., Bea, A. M., Sánchez-Alcoholado, L., Fumanal, A. J., Quesada-Molina, M., et al. (2021). An alcohol-free beer enriched with isomaltulose and a resistant dextrin modulates gut microbiome in subjects with type 2 diabetes mellitus and overweight or obesity: A pilot study. *Food Funct.* 12, 3635–3646. doi: 10.1039/d0fo03160g.
- Meijnikman, A. S., Gerdes, V. E., Nieuwdorp, M., and Herrema, H. (2018). Evaluating Causality of Gut Microbiota in Obesity and Diabetes in Humans. 39, 133–153. doi: 10.1210/ER.2017-00192.
- Menni, C., Zhu, J., Le Roy, C. I., Mompeo, O., Young, K., Rebholz, C. M., et al. (2020). Serum metabolites reflecting gut microbiome alpha diversity predict type 2 diabetes. *Gut Microbes* 11, 1632–1642. doi: 10.1080/19490976.2020.1778261/SUPPL_FILE/KGMI_A_1778261_SM0210.ZIP.
- MetaHIT Consortium (Metagenomics of the Human Intestinal Tract consortium) - Wellcome Sanger Institute Available at: https://www.sanger.ac.uk/resources/downloads/bacteria/metahit/ [Accessed August 20, 2021].
- Michels, N., Zouiouich, S., Vanderbauwhede, B., Vanacker, J., Indave Ruiz, B. I., and Huybrechts, I. (2022). Human microbiome and metabolic health: An overview of systematic reviews. *Obes. Rev.* 23, e13409. doi: 10.1111/OBR.13409.
- Mokkala, K., Paulin, N., Houttu, N., Koivuniemi, E., Pellonperä, O., Khan, S., et al. (2021). Metagenomics analysis of gut microbiota in response to diet

intervention and gestational diabetes in overweight and obese women: a randomised, double-blind, placebo-controlled clinical trial. *Gut* 70, 309–318. doi: 10.1136/GUTJNL-2020-321643.

- Moreno-Indias, I., Sánchez-Alcoholado, L., García-Fuentes, E., Cardona, F., Queipo-Ortuño, M. I., and Tinahones, F. J. (2016). Insulin resistance is associated with specific gut microbiota in appendix samples from morbidly obese patients. *Am. J. Transl. Res.* 8, 5672. Available at: /pmc/articles/PMC5209518/ [Accessed April 12, 2022].
- Morgan, X. C., Tickle, T. L., Sokol, H., Gevers, D., Devaney, K. L., Ward, D. V., et al. (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol.* 13, 1–18. doi: 10.1186/GB-2012-13-9-R79/FIGURES/6.
- Moya-Pérez, A., Neef, A., and Sanz, Y. (2015). Bifidobacterium pseudocatenulatum CECT 7765 Reduces Obesity-Associated Inflammation by Restoring the Lymphocyte-Macrophage Balance and Gut Microbiota Structure in High-Fat Diet-Fed Mice. *PLoS One* 10, e0126976. doi: 10.1371/JOURNAL.PONE.0126976.
- Muñoz-Garach, A., García-Fontana, B., and Muñoz-Torres, M. (2019). Vitamin D Status, Calcium Intake and Risk of Developing Type 2 Diabetes: An Unresolved Issue. *Nutrients* 11. doi: 10.3390/NU11030642.
- Naito, E., Yoshida, Y., Makino, K., Kounoshi, Y., Kunihiro, S., Takahashi, R., et al. (2011). Beneficial effect of oral administration of Lactobacillus casei strain Shirota on insulin resistance in diet-induced obesity mice. *J. Appl. Microbiol.* 110, 650–657. doi: 10.1111/J.1365-2672.2010.04922.X.
- Navab-Moghadam, F., Sedighi, M., Khamseh, M. E., Alaei-Shahmiri, F.,

Talebi, M., Razavi, S., et al. (2017). The association of type II diabetes with gut microbiota composition. *Microb. Pathog.* 110, 630–636. doi: 10.1016/J.MICPATH.2017.07.034.

- Navina, S., Acharya, C., DeLany, J. P., Orlichenko, L. S., Baty, C. J., Shiva, S. S., et al. (2011). Lipotoxicity causes multisystem organ failure and exacerbates acute pancreatitis in obesity. *Sci. Transl. Med.* 3. doi: 10.1126/SCITRANSLMED.3002573/SUPPL_FILE/3-107RA110_SM.PDF.
- Neis, E. P. J. G., Dejong, C. H. C., and Rensen, S. S. (2015). The Role of Microbial Amino Acid Metabolism in Host Metabolism. *Nutr. 2015, Vol. 7, Pages 2930-2946* 7, 2930–2946. doi: 10.3390/NU7042930.
- Newsholme, P., and De Bittencourt, P. I. H. (2016). Gut associated bacteria are critical to metabolism, inflammation and health. *Curr. Opin. Clin. Nutr. Metab. Care* 19, 245–249. doi: 10.1097/MCO.0000000000000293.
- NIH Human Microbiome Project - Home Available at: https://www.hmpdacc.org/hmp/ [Accessed August 20, 2021].
- Nuli, R., Azhati, J., Cai, J., Kadeer, A., Zhang, B., and Mohemaiti, P. (2019). Metagenomics and Faecal Metabolomics Integrative Analysis towards the Impaired Glucose Regulation and Type 2 Diabetes in Uyghur-Related Omics. *J. Diabetes Res.* 2019. doi: 10.1155/2019/2893041.
- Okubo, H., Sakoda, H., Kushiyama, A., Fujishiro, M., Nakatsu, Y., Fukushima, T., et al. (2013). Lactobacillus casei strain shirota protects against nonalcoholic steatohepatitis development in a rodent model. *Am. J. Physiol. - Gastrointest. Liver Physiol.* 305, 911–918. doi: 10.1152/AJPGI.00225.2013/ASSET/IMAGES/LARGE/ZH30241365620006.JPEG.

- Opazo, M. C., Ortega-Rocha, E. M., Coronado-Arrázola, I., Bonifaz, L. C., Boudin, H., Neunlist, M., et al. (2018). Intestinal microbiota influences non-intestinal related autoimmune diseases. *Front. Microbiol.* 9, 432. doi: 10.3389/FMICB.2018.00432/BIBTEX.
- Otoom, A. F., Abdallah, E. E., Kilani, Y., Kefaye, A., and Ashour, M. (2015). Effective diagnosis and monitoring of heart disease. *Int. J. Softw. Eng. its Appl.* 9, 143–156. doi: 10.14257/ijseia.2015.9.1.12.
- Pai, C., Wang, C., Hung, W., Hung, W., Tsai, H., and Chang, C. (2022). Interrelationship of Gut Microbiota , Obesity , Body Composition and Insulin Resistance in Asians with Type 2 Diabetes Mellitus.
- Park, S. H., Kim, K. A., Ahn, Y. T., Jeong, J. J., Huh, C. S., and Kim, D. H. (2015). Comparative analysis of gut microbiota in elderly people of urbanized towns and longevity villages. *BMC Microbiol.* 15, 1–9. doi: 10.1186/S12866-015-0386-8/FIGURES/5.
- Parks, B. W., Nam, E., Org, E., Kostem, E., Norheim, F., Hui, S. T., et al. (2013). Genetic Control of Obesity and Gut Microbiota Composition in Response to High-Fat, High-Sucrose Diet in Mice. *Cell Metab.* 17, 141–152. doi: 10.1016/J.CMET.2012.12.007.
- Patil, H. R., Al Badarin, F. J., Al Shami, H. A., Bhatti, S. K., Lavie, C. J., Bell, D. S. H., et al. (2012). Meta-Analysis of Effect of Dipeptidyl Peptidase-4 Inhibitors on Cardiovascular Risk in Type 2 Diabetes Mellitus. *Am. J. Cardiol.* 110, 826–833. doi: 10.1016/J.AMJCARD.2012.04.061.
- Patterson, E. E., Ryan, P. M., Cryan, J. F., Dinan, T. G., Paul Ross, R., Fitzgerald, G. F., et al. (2016). Gut microbiota, obesity and diabetes. *Postgrad. Med. J.* 92, 286–300. doi: 10.1136/POSTGRADMEDJ-2015-133285.

- Petersen, K. S., Keogh, J. B., Lister, N. B., and Clifton, P. M. (2018). Dietary quality and carotid intima media thickness in type 1 and type 2 diabetes: Follow-up of a randomised controlled trial. *Nutr. Metab. Cardiovasc. Dis.* 28, 830–838. doi: 10.1016/J.NUMECD.2018.05.001.
- Petersen, M. C., and Shulman, G. I. (2018). Mechanisms of insulin action and insulin resistance. *Physiol. Rev.* 98, 2133–2223. doi: 10.1152/PHYSREV.00063.2017/ASSET/IMAGES/LARGE/Z9J00418286800 19.JPEG.
- Pinna, N. K., Anjana, R. M., Saxena, S., Dutta, A., Gnanaprakash, V., Rameshkumar, G., et al. (2021). Trans-ethnic gut microbial signatures of prediabetic subjects from India and Denmark. *Genome Med.* 13, 1–20. doi: 10.1186/S13073-021-00851-9/FIGURES/3.
- Prevalence of diabetes in Europe in 2019 | Statista Available at: https://www.statista.com/statistics/1081006/prevalence-of-diabetes-in-europe/ [Accessed August 20, 2021].
- Qian, L. J., Kang, S. M., Xie, J. L., Huang, L., Wen, Q., Fan, Y. Y., et al. (2017). Early-life gut microbial colonization shapes Th1/Th2 balance in asthma model in BALB/c mice. *BMC Microbiol.* 17, 1–8. doi: 10.1186/S12866-017-1044-0/TABLES/3.
- Radwan, S., Gilfillan, D., Eklund, B., Radwan, H. M., El Menofy, N. G., Lee, J., et al. (2020). A comparative study of the gut microbiome in Egyptian patients with Type I and Type II diabetes. *PLoS One* 15, e0238764. doi: 10.1371/JOURNAL.PONE.0238764.
- Ramachandran, A., Snehalatha, C., Shetty, A. S., and Nanditha, A. (2012). Trends in prevalence of diabetes in Asian countries. *World J. Diabetes* 3, 110.

doi: 10.4239/WJD.V3.I6.110.

- Remely, M., Aumueller, E., Jahn, D., Hippe, B., Brath, H., and Haslberger, A. G. (2014). Microbiota and epigenetic regulation of inflammatory mediators in type 2 diabetes and obesity. *Benef. Microbes* 5, 33–43. doi: 10.3920/BM2013.006.
- Ríos-Covián, D., Ruas-Madiedo, P., Margolles, A., Gueimonde, M., De los Reyes-Gavilán, C. G., and Salazar, N. (2016). Intestinal short chain fatty acids and their link with diet and human health. *Front. Microbiol.* 7, 185. doi: 10.3389/FMICB.2016.00185/BIBTEX.
- Roager, H. M., and Christensen, L. H. (2022). Personal diet–microbiota interactions and weight loss. *Proc. Nutr. Soc.*, 1–12. doi: 10.1017/S0029665122000805.
- Rold, L. S., Bundgaard-Nielsen, C., Holm-Jacobsen, J. N., Ovesen, P. G., Leutscher, P., Hagstrøm, S., et al. (2022). Characteristics of the gut microbiome in women with gestational diabetes mellitus: A systematic review. *PLoS One* 17, e0262618. doi: 10.1371/JOURNAL.PONE.0262618.
- Rossi, M., Johnson, D. W., and Campbell, K. L. (2015). The Kidney–Gut Axis: Implications for Nutrition Care. *J. Ren. Nutr.* 25, 399–403. doi: 10.1053/J.JRN.2015.01.017.
- Ruuskanen, M. O., Erawijantari, P. P., Havulinna, A. S., Liu, Y., Méric, G., Tuomilehto, J., et al. (2022). Gut Microbiome Composition Is Predictive of Incident Type 2 Diabetes in a Population Cohort of 5,572 Finnish Adults. *Diabetes Care* 45, 811–818. doi: 10.2337/DC21-2358.
- Sabatino, A., Regolisti, G., Cosola, C., Gesualdo, L., and Fiaccadori, E. (2017). Intestinal Microbiota in Type 2 Diabetes and Chronic Kidney Disease. *Curr.*

Diab. Rep. 17, 1–9. doi: 10.1007/S11892-017-0841-Z/TABLES/3.

- Schneeberger, M., Everard, A., Gómez-Valadés, A. G., Matamoros, S., Ramírez, S., Delzenne, N. M., et al. (2015). Akkermansia muciniphila inversely correlates with the onset of inflammation, altered adipose tissue metabolism and metabolic disorders during obesity in mice. *Sci. Reports 2015 51* 5, 1–14. doi: 10.1038/srep16643.
- Sedighi, M., Razavi, S., Navab-Moghadam, F., Khamseh, M. E., Alaei-Shahmiri, F., Mehrtash, A., et al. (2017). Comparison of gut microbiota in adult patients with type 2 diabetes and healthy individuals. *Microb. Pathog.* 111, 362–369. doi: 10.1016/j.micpath.2017.08.038.
- Sell, H., Habich, C., and Eckel, J. (2012). Adaptive immunity in obesity and insulin resistance. *Nat. Rev. Endocrinol. 2012 812* 8, 709–716. doi: 10.1038/nrendo.2012.114.
- Sgritta, M., Dooling, S. W., Buffington, S. A., Momin, E. N., Francis, M. B., Britton, R. A., et al. (2019). Mechanisms Underlying Microbial-Mediated Changes in Social Behavior in Mouse Models of Autism Spectrum Disorder. *Neuron* 101, 246-259.e6. doi: 10.1016/J.NEURON.2018.11.018.
- Shah, V. O., Lambeth, S. M., Carson, T., Lowe, J., Ramaraj, T., Leff, J. W., et al. (2015). Composition Diversity and Abundance of Gut Microbiome in Prediabetes and Type 2 Diabetes. *J. Diabetes Obes.* 2, 108–114. doi: 10.15436/2376-0949.15.031.
- Sharma, P., and Kaur, M. (2013). Classification in Pattern Recognition: A Review. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* 3, 2277. Available at: www.ijarcsse.com [Accessed July 24, 2022].
- Sharma, S., and Tripathi, P. (2019). Gut microbiome and type 2 diabetes: where

we are and where to go? *J. Nutr. Biochem.* 63, 101–108. doi: 10.1016/J.JNUTBIO.2018.10.003.

- Shreiner, A. B., Kao, J. Y., and Young, V. B. (2015). The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.* 31, 69. doi: 10.1097/MOG.0000000000000139.
- Singh, R. K., Chang, H. W., Yan, D., Lee, K. M., Ucmak, D., Wong, K., et al. (2017). Influence of diet on the gut microbiome and implications for human health. *J. Transl. Med. 2017 151* 15, 1–17. doi: 10.1186/S12967-017-1175-Y.
- Sonnenburg, J. L., and Bäckhed, F. (2016). Diet–microbiota interactions as moderators of human metabolism. *Nat. 2016 5357610* 535, 56–64. doi: 10.1038/nature18846.
- Sroka-Oleksiak, A., Młodzińska, A., Bulanda, M., Salamon, D., Major, P., Stanek, M., et al. (2020). Metagenomic Analysis of Duodenal Microbiota Reveals a Potential Biomarker of Dysbiosis in the Course of Obesity and Type 2 Diabetes: A Pilot Study. *J. Clin. Med.* 9, 369. doi: 10.3390/jcm9020369.
- Stanford, K. I., Rasmussen, M., Baer, L. A., Lehnig, A. C., Rowland, L. A., White, J. D., et al. (2018). Paternal Exercise Improves Glucose Metabolism in Adult Offspring. *Diabetes* 67, 2530–2540. doi: 10.2337/DB18-0667.
- Stolfi, C., Troncone, E., Marafini, I., and Monteleone, G. (2020). Role of TGF-Beta and Smad7 in Gut Inflammation, Fibrosis and Cancer. *Biomol. 2021, Vol. 11, Page 17* 11, 17. doi: 10.3390/BIOM11010017.
- Sun, J., and Chang, E. B. (2014). Exploring gut microbes in human health and disease: Pushing the envelope. *Genes Dis.* 1, 132–139. doi: 10.1016/J.GENDIS.2014.08.001.
- Tan, K. C., Teoh, E. J., Yu, Q., and Goh, K. C. (2009). A hybrid evolutionary

algorithm for attribute selection in data mining. *Expert Syst. Appl.* 36, 8616–8630. doi: 10.1016/J.ESWA.2008.10.013.

- Tanca, A., Abbondio, M., Palomba, A., Fraumene, C., Manghina, V., Cucca, F., et al. (2017). Potential and active functions in the gut microbiota of a healthy human cohort. *Microbiome* 5, 1–15. doi: 10.1186/S40168-017-0293-3/FIGURES/6.
- Tap, J., Furet, J. P., Bensaada, M., Philippe, C., Roth, H., Rabot, S., et al. (2015). Gut microbiota richness promotes its stability upon increased dietary fibre intake in healthy adults. *Environ. Microbiol.* 17, 4954–4964. doi: 10.1111/1462-2920.13006/SUPPINFO.
- Therdtatha, P., Song, Y., Tanaka, M., Mariyatun, M., Almunifah, M., Manurung, N. E. P., et al. (2021). Gut microbiome of indonesian adults associated with obesity and type 2 diabetes: A cross-sectional study in an asian city, yogyakarta. *Microorganisms* 9, 897. doi: 10.3390/MICROORGANISMS9050897/S1.
- Tolhurst, G., Heffron, H., Lam, Y. S., Parker, H. E., Habib, A. M., Diakogiannaki, E., et al. (2012). Short-chain fatty acids stimulate glucagon-like peptide-1 secretion via the G-protein-coupled receptor FFAR2. *Diabetes* 61, 364–371. doi: 10.2337/DB11-1019/-/DC1.
- Tsai, H., Tsai, W., Hung, W., Hung, W., Chang, C., and Dai, C. (2021). Gut Microbiota and Subclinical Cardiovascular Disease in Patients with Type 2 Diabetes Mellitus. 1–15.
- Vaiserman, A. M., Koliada, A. K., and Marotta, F. (2017). Gut microbiota: A player in aging and a target for anti-aging intervention. *Ageing Res. Rev.* 35, 36–45. doi: 10.1016/J.ARR.2017.01.001.

- Van De Wiele, T., Van Praet, J. T., Marzorati, M., Drennan, M. B., and Elewaut, D. (2016). How the microbiota shapes rheumatic diseases. *Nat. Rev. Rheumatol.* 12, 398–411. doi: 10.1038/NRRHEUM.2016.85.
- Vembandasamyp, K., Sasipriyap, R. R., and Deepap, E. (2015). Heart Diseases Detection Using Naive Bayes Algorithm. *IJISET-International J. Innov. Sci. Eng. Technol.* 2. Available at: www.ijiset.com [Accessed July 24, 2022].
- Wang, J., Chen, W. D., and Wang, Y. D. (2020). The Relationship Between Gut Microbiota and Inflammatory Diseases: The Role of Macrophages. *Front. Microbiol.* 11, 1065. doi: 10.3389/FMICB.2020.01065/BIBTEX.
- Wang, J., Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., et al. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nat. 2012 4907418* 490, 55–60. doi: 10.1038/nature11450.
- Wang, J., Tang, H., Zhang, C., Zhao, Y., Derrien, M., Rocher, E., et al. (2014). Modulation of gut microbiota during probiotic-mediated attenuation of metabolic syndrome in high fat diet-fed mice. *ISME J. 2015 91* 9, 1–15. doi: 10.1038/ismej.2014.99.
- Wang, L., Yu, X., Xu, X., Ming, J., Wang, Z., Gao, B., et al. (2021). The Fecal Microbiota Is Already Altered in Normoglycemic Individuals Who Go on to Have Type 2 Diabetes. *Front. Cell. Infect. Microbiol.* 11, 1. doi: 10.3389/FCIMB.2021.598672/BIBTEX.
- Wang, W. L., Xu, S. Y., Ren, Z. G., Tao, L., Jiang, J. W., and Zheng, S. Sen (2015). Application of metagenomics in the human gut microbiome. *World J. Gastroenterol.* 21, 803. doi: 10.3748/WJG.V21.I3.803.
- Wei, B., Wang, Y., Xiang, S., Jiang, Y., Chen, R., and Hu, N. (2021). Alterations of gut microbiome in patients with type 2 diabetes mellitus who had

undergone cholecystectomy. *Am. J. Physiol. - Endocrinol. Metab.* 320, E113–E121. doi: 10.1152/AJPENDO.00471.2020.

- Williams, R., Karuranga, S., Malanda, B., Saeedi, P., Basit, A., Besançon, S., et al. (2020). Global and regional estimates and projections of diabetes-related health expenditure: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res. Clin. Pract.* 162. doi: 10.1016/J.DIABRES.2020.108072.
- Wu, H., Tremaroli, V., Schmidt, C., Lundqvist, A., Olsson, L. M., Krämer, M., et al. (2020). The Gut Microbiota in Prediabetes and Diabetes: A Population-Based Cross-Sectional Study. *Cell Metab.* 32, 379-390.e3. doi: 10.1016/J.CMET.2020.06.011.
- Xu, J., Verbrugghe, A., Lourenço, M., Janssens, G. P. J., Liu, D. J. X., Van de Wiele, T., et al. (2016). Does canine inflammatory bowel disease influence gut microbial profile and host metabolism? *BMC Vet. Res.* 12, 1–10. doi: 10.1186/S12917-016-0736-2/TABLES/4.
- Yajnik, C. S. (2001). The insulin resistance epidemic in India: fetal origins, later lifestyle, or both? *Nutr. Rev.* 59, 1–9. doi: 10.1111/J.1753-4887.2001.TB01898.X.
- Yaribeygi, H., Farrokhi, F. R., Butler, A. E., and Sahebkar, A. (2019). Insulin resistance: Review of the underlying molecular mechanisms. *J. Cell. Physiol.* 234, 8152–8161. doi: 10.1002/JCP.27603.
- Yoo, H. H., Kim, I. S., Yoo, D. H., and Kim, D. H. (2016). Effects of orally administered antibiotics on the bioavailability of amlodipine: Gut microbiota-mediated drug interaction. *J. Hypertens.* 34, 156–162. doi: 10.1097/HJH.0000000000000773.

- Yoon, H., Shaw, J. L., Haigis, M. C., and Greka, A. (2021). Lipid metabolism in sickness and in health: Emerging regulators of lipotoxicity. *Mol. Cell* 81, 3708–3730. doi: 10.1016/J.MOLCEL.2021.08.027.
- Yoon, K. H., Lee, J. H., Kim, J. W., Cho, J. H., Choi, Y. H., Ko, S. H., et al. (2006). Epidemic obesity and type 2 diabetes in Asia. *Lancet* 368, 1681–1688. doi: 10.1016/S0140-6736(06)69703-1.
- Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., et al. (2015). Personalized Nutrition by Prediction of Glycemic Responses. *Cell* 163, 1079–1094. doi: 10.1016/J.CELL.2015.11.001.
- Zhang, C., Li, S., Yang, L., Huang, P., Li, W., Wang, S., et al. (2013a). Structural modulation of gut microbiota in life-long calorie-restricted mice. *Nat. Commun. 2013 41* 4, 1–10. doi: 10.1038/ncomms3163.
- Zhang, F., Wang, M., Yang, J., Xu, Q., Liang, C., Chen, B., et al. (2019). Response of gut microbiota in type 2 diabetes to hypoglycemic agents. *Endocrine* 66, 485–493. doi: 10.1007/S12020-019-02041-5/FIGURES/4.
- Zhang, J., Ni, Y., Qian, L., Fang, Q., Zheng, T., Zhang, M., et al. (2021a). Decreased Abundance of Akkermansia muciniphila Leads to the Impairment of Insulin Secretion and Glucose Homeostasis in Lean Type 2 Diabetes. doi: 10.1002/advs.202100536.
- Zhang, Q., Ma, C., Duan, Y., Heinrich, B., Rosato, U., Diggs, L. P., et al. (2021b). Gut microbiome directs hepatocytes to recruit MDSC and promote cholangiocarcinoma. *Cancer Discov.* 11, 1248. doi: 10.1158/2159-8290.CD-20-0304.
- Zhang, Q., Zhang, Y., Zeng, L., Chen, G., Zhang, L., Liu, M., et al. (2021c). The Role of Gut Microbiota and Microbiota-Related Serum Metabolites in the

Progression of Diabetic Kidney Disease. *Front. Pharmacol.* 12. doi: 10.3389/FPHAR.2021.757508/FULL.

- Zhang, X., Coker, O. O., Chu, E. S. H., Fu, K., Lau, H. C. H., Wang, Y. X., et al. (2021d). Dietary cholesterol drives fatty liver-associated liver cancer by modulating gut microbiota and metabolites. *Gut* 70, 761–774. doi: 10.1136/GUTJNL-2019-319664.
- Zhang, X., Shen, D., Fang, Z., Jie, Z., Qiu, X., Zhang, C., et al. (2013b). Human Gut Microbiota Changes Reveal the Progression of Glucose Intolerance. *PLoS One* 8, e71108. doi: 10.1371/JOURNAL.PONE.0071108.
- Zhang, Y., Gu, Y., Ren, H., Wang, S., Zhong, H., Zhao, X., et al. (2020). Gut microbiome-related effects of berberine and probiotics on type 2 diabetes (the PREMOTE study). *Nat. Commun. 2020 111* 11, 1–12. doi: 10.1038/s41467-020-18414-8.
- Zhang, Y., Lu, S., Yang, Y., Wang, Z., Wang, B., Zhang, B., et al. (2021e). The diversity of gut microbiota in type 2 diabetes with or without cognitive impairment. *Aging Clin. Exp. Res.* 33, 589–601. doi: 10.1007/S40520-020-01553-9/FIGURES/5.
- Zhao, L., Lou, H., Peng, Y., Chen, • Shihong, Zhang, Y., and Li, • Xiaobo (2020a). Comprehensive relationships between gut microbiome and faecal metabolome in individuals with type 2 diabetes and its complications. *Endocrine* 66, 526–537. doi: 10.1007/s12020-019-02103-8.
- Zhao, L., Zhang, F., Ding, X., Wu, G., Lam, Y. Y., Wang, X., et al. (2018). Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. *Science (80-.).* 359, 1151–1156. doi: 10.1126/SCIENCE.AAO5774.
- Zhao, Y., Zhou, C., Wu, C., Guo, X., Hu, G., Wu, Q., et al. (2020b). Subchronic

oral mercury caused intestinal injury and changed gut microbiota in mice. *Sci. Total Environ.* 721, 137639. doi: 10.1016/J.SCITOTENV.2020.137639.

- Zhernakova, A., Kurilshikov, A., Bonder, M. J., Tigchelaar, E. F., Schirmer, M., Vatanen, T., et al. (2016). Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science (80-.).* 352, 565–569. doi: 10.1126/SCIENCE.AAD3369/SUPPL_FILE/ZHERNAKOVA_TABLES_S1_TO_S19.XLSX.

DEBJIT DE

Research Scholar

Department of Life Science & Biotechnology

Jadavpur University

West Bengal, India

ACADEMIC QUALIFICATION

2009-2012	B.Sc. Genetics	Institute of Genetic Engineering, M.A.K.A.U.T.
2012-2014	M.Sc. Biomedical Genetics	Vellore Institute of Technology (V.I.T.) University, India
2015-Present	PhD. (Perusing)	Jadavpur University, India

PUBLICATIONS FROM THESIS

- **De, D.**, Nayak, T., Chowdhury, S., & Dhal, P. K. **(2022)**. Insights of host physiological parameters and gut-microbiome of Indian Type 2 Diabetic patients visualized via metagenomics and machine learning approaches. ***Frontiers in Microbiology***, 0, 2479. doi.org/10.3389/FMICB.2022.914124. **Impact Factor –** 6.06.

OTHER PUBLICATIONS

- Nayak, T., **De, D.**, Dhal, P. K. **(2021)**. The differences in SARS-CoV and SARS-CoV-2 specific co-expression network mediated biological process in human gut enterocytes. ***Infection, Genetics and Evolution***, doi.org/10.1016/j.meegid.2021.104892. **Impact Factor** - 3.34.

- Nayak, T., **De, D.**, Karmakar, P., Deb, A., Dhal, P. K. **(2021)**. Microbial communities of the drinking water with gradient radon concentration are primarily contributed by radon and heavy metal content. ***Frontiers in Environmental Science***, 9, doi: 10.3389/fenvs.2021.576400. **Impact Factor –** 5.41.
- Nayak, T., **De, D.**, Barman, C., Karmakar, P., Deb, A., Dhal, P. K. **(2019)**. Characterization of indigenous bacteria from radon-rich groundwater and their tolerance to physicochemical stress. ***International Journal of Environmental Science and Technology***, 17, 1627–1636, https://doi.org/10.1007/s13762-019-02445-w. **Impact Factor -** 2.86.

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 18 July 2022
doi: 10.3389/fmicb.2022.914124

Insights of Host Physiological Parameters and Gut Microbiome of Indian Type 2 Diabetic Patients Visualized *via* Metagenomics and Machine Learning Approaches

Debjit De[1], *Tilak Nayak*[1], *Subhankar Chowdhury*[2] *and Paltu Kumar Dhal*[1]*

[1] *Department of Life Science and Biotechnology, Jadavpur University, Kolkata, India,* [2] *Department of Endocrinology, Institute of Post Graduate Medical Education and Research (IPGMER) and SSKM Hospital, Kolkata, India*

OPEN ACCESS

Edited by:
Punyasloke Bhadury, Indian Institute of Science Education and Research Kolkata, India

Reviewed by:
Sundus Javed, COMSATS University, Pakistan
Rebiya Nuli, Xinjiang Medical University, China

***Correspondence:**
Paltu Kumar Dhal
paltuk.dhal@jadavpuruniversity.in

Specialty section:
This article was submitted to Microbial Symbioses, a section of the journal Frontiers in Microbiology

Received: *06 April 2022*
Accepted: *13 June 2022*
Published: *18 July 2022*

Citation:
De D, Nayak T, Chowdhury S and Dhal PK (2022) Insights of Host Physiological Parameters and Gut Microbiome of Indian Type 2 Diabetic Patients Visualized via Metagenomics and Machine Learning Approaches. Front. Microbiol. 13:914124. doi: 10.3389/fmicb.2022.914124

Type 2 diabetes (T2D) is a serious public health issue and may also contribute to modification in the structure of the intestinal microbiota, implying a link between T2D and microbial inhabitants in the digestive tract. This work aimed to develop efficient models for identifying essential physiological markers for improved T2D classification using machine learning algorithms. Using amplicon metagenomic approaches, an effort has also been made to understand the alterations in core gut microbial members in Indian T2D patients with respect to their control normal glucose tolerance (NGT). Our data indicate the level of fasting blood glucose (FBG) and glycated hemoglobin (HbA1c) were the most useful physiological indicators while random forest and support vector machine with RBF Kernel were effective predictions models for identifications of T2D. The dominating gut microbial members *Allopreotella*, *Rikenellaceae RC9 gut group*, *Haemophilus*, *Ruminococcus torques group*, etc. in Indian T2D patients showed a strong association with both FBG and HbA1c. These members have been reported to have a crucial role in gut barrier breakdown, blood glucose, and lipopolysaccharide level escalation, or as biomarkers. While the dominant NGT microbiota (*Akkermansia*, *Ligilactobacillus*, *Enterobacter*, etc.) in the colon has been shown to influence inflammatory immune responses by acting as an anti-inflammatory agent and maintaining the gut barrier. The topology study of co-occurrence network analysis indicates that changes in network complexity in T2D lead to variations in the different gut microbial members compared to NGT. These studies provide a better understanding of the gut microbial diversity in Indian T2D patients and show the way for the development of valuable diagnostics strategies to improve the prediction and modulation of the T2D along with already established methods.

Keywords: type 2 diabetes, gut microbiota, machine learning, feature selection, microbial communities

Infection, Genetics and Evolution 92 (2021) 104892

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

The differences in SARS-CoV and SARS-CoV-2 specific co-expression network mediated biological process in human gut enterocytes

Tilak Nayak [1], Debjit De [1], Paltu Kumar Dhal [*]

Department of Life Science and Biotechnology, Jadavpur University, Kolkata 700032, India

ARTICLE INFO

Keywords:
SARS-CoV
SARS-CoV-2
RNA-Seq
Disease specific co-expression network
Dietary supplementary

ABSTRACT

Novel coronavirus SARS-CoV-2 was recently outbreak worldwide causes severe acute respiratory syndrome along with gastrointestinal symptoms for some infected patients. Information on detail pathogenesis, host immune responses and responsible biological pathways are limited. Therefore, infection specific host gut responses and dietary supplements to neutralize immune inflammation demand extensive research. This study aimed to find differences in global co-expression protein-protein interaction sub-network and enriched biological processes in SARS-CoV and SARS-CoV-2 infected gut enterocytes cell line. Attempts have also been made to predict some dietary supplements to boost human health. The SARS-CoV and SARS-CoV-2 infected differential express proteins were integrated with the human protein interaction network and co-expression subnetworks were constructed. Common hubs of these sub-networks reshape central cellular pathways of metabolic processes, lipid localization, hypoxia response to decrease oxygen level and transport of bio-molecules. The major biological process enriched in the unique hub of SARS-CoV-2 significantly differ from SARS-CoV, related to interferon signaling, regulation of viral process and influenza-A enzymatic pathway. Predicted dietary supplements can improve SARS-CoV-2 infected person''s health by boosting the host immunity/reducing inflammation. To the best of our knowledge this is the first report on co-expression network mediated biological process in human gut enterocytes to predict dietary supplements/compounds.

1. Introduction

Severe acute respiratory syndrome (SARS) first emerged in 2003 caused by coronavirus SARS-CoV (Drosten et al., 2003). In late December 2019, a novel coronavirus (SARS-CoV-2) epidemic happened from China and on 30th January 2020 World Health Organization (WHO) declared COVID-19 as a pandemic (Zhu et al., 2020; Li et al., 2020). Coronaviruses (CoVs) are the single stranded RNA viruses that infects animals and humans causing respiratory, gastrointestinal and hepatic disease (Leibowitz and Weiss, 2013; Lamers et al., 2020). Till date, there have been seven human coronaviruses (HCoVs) identified, including HCoVs-NL63, HCoVs-229E, HCoVs-OC43, HCoVs-HKU1, SARS-CoV, MERS-CoV and novel SARS-CoV-2 (Ye et al., 2020). Despite some common clinical symptoms, SARS-CoV-2 has the highest pathogenicity with 106,125,682 confirmed cases and 2,320,497 deaths globally as of 10th February 2021 much more than SARS-CoV (8422 people infected in 26 countries, leading to 916 deaths) according to WHO. (https://www.who.int/emergencies/diseases/novel-coronavirus-2019). Along with their common clinical symptoms subset of patients showed severe gastrointestinal problems for SARS-CoV-2 (Lamers et al., 2020).

Although there are some reports of host responses on infected lung epithelial cells, less research has been done on human gut infection which is another important site for SARS-CoV-2 causing gastrointestinal problems. Early reports revealed that in SARS patients, there is a pulmonary infection and severe lung damage associated with elevated pro-inflammatory cytokines in serum (IL-6, IL-8, IFN-γ, IL-1β, TNF-α; Azkur et al., 2020; Prasad et al., 2020; Liang et al., 2020).

There are some recent reports on RNA-Seq expression for SARS-CoV and SARS-CoV-2 infected lung epithelial and gut enterocytes cell line to characterize the differentially expressed genes and their responsible metabolic pathways, however, lacking the global co-expression profile of intestinal cells (Lamers et al., 2020; Lieberman et al., 2020). Protein-protein interaction (PPI) network from differentially expressed datasets and their co-expression profile may provide a global picture of cellular processes that can be used as a target to improve diagnostic, prognostic

* Corresponding author at: Department of Life Science and Biotechnology, Jadavpur University, 188 Raja S.C. Mallick Road, Kolkata 700 032, India.
E-mail address: paltuk.dhal@jadavpuruniversity.in (P.K. Dhal).
[1] Authors contributed equally to this work.

https://doi.org/10.1016/j.meegid.2021.104892
Received 6 March 2021; Received in revised form 17 April 2021; Accepted 30 April 2021
Available online 3 May 2021

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 17 May 2021
doi: 10.3389/fenvs.2021.576400

Microbial Communities of the Drinking Water With Gradient Radon Concentration Are Primarily Contributed by Radon and Heavy Metal Content

*Tilak Nayak[1], Debjit De[1], Parimal Karmakar[1], Argha Deb[2] and Paltu Kumar Dhal[1]**

[1]Department of Life Science and Biotechnology, Jadavpur University, Kolkata, India, [2]Department of Physics, School of Studies in Environmental Radiation and Archaeological Sciences, Jadavpur University, Kolkata, India

OPEN ACCESS

Edited by:
Andrew Hursthouse, University of the West of Scotland, United Kingdom

Reviewed by:
Zhi Wang, Innovation Academy for Precision Measurement Science and Technology (CAS), China
Avishek Dutta, University of California, San Diego, United States

***Correspondence:**
Paltu Kumar Dhal
paltuk.dhal@jadavpuruniversity.in

Specialty section:
This article was submitted to Toxicology, Pollution and the Environment, a section of the journal Frontiers in Environmental Science

Received: 26 June 2020
Accepted: 26 April 2021
Published: 17 May 2021

Citation:
Nayak T, De D, Karmakar P, Deb A and Dhal PK (2021) Microbial Communities of the Drinking Water With Gradient Radon Concentration Are Primarily Contributed by Radon and Heavy Metal Content. Front. Environ. Sci. 9:576400. doi: 10.3389/fenvs.2021.576400

Radon and heavy metal (HM) contamination in drinking water and their impact on health have been reported earlier. However, relatively little is known about the microbial community in drinking water with gradients of radon and the drivers of microbial community patterns in such water. With this view, we first examine microbial dynamics of drinking water in the permissible level of 93 ± 2 Bq/l as control, 510 ± 1.5 6 Bq/l and 576 ± 2 Bq/l as medium, and 728 ± 3 Bq/l as high radon-containing tube wells from Dumka and Godda districts, which comes under a major fault of the eastern fringes of India. Attempts have also been made to predict the impact of the radon contamination gradient and other water environmental parameters on community structure. The measured physicochemical character revealed strong clustering by the sampling site with respect to its radon and HM content. The radon-contaminated sites represent HM-rich nutrient-limited sites compared to the control. Radon (Rn), HM (Pb, Cu, and As), and total suspended solids (TSSs) were the most determinant variable among the parameters and influenced the microbial community composition of that region. The microbial diversity of those sites was lower, and this measured diversity decreased gradually on the sites with an increased gradient of radon contamination. The dominant microbial families in the contaminated sites were Moraxellaceae, Chitinophagaceae, unclassified *Candidatus* Azambacteria, unclassified *Candidatus* Moranbacteria, unclassified *Candidatus* Collierbacteria, and Gammaproteobacterial members, which are reported to abundantly inhabit radiation and chemolithotrophic environments and pose better radionuclide protective mechanisms, while the bacterial members dominant in the control site were Comamonadaceae, Rhodocyclaceae, Nitrospirales Incertae Sedis, cvE6, unclassified Woesearchaeota (DHVEG-6), and Holophagaceae, which are reported to be abundant in natural soil and drinking water, and labile in harsh environments. Relative sequence abundance of Comamonadaceae was decreasing on the sites with an increasing radon gradient, while the opposite trend was observed for Chitinophagaceae. The distribution of such microbial assemblages is linked to radon and heavy metal, highlighting that taxa with distinct environmental preferences underlie

International Journal of Environmental Science and Technology (2020) 17:1627–1636
https://doi.org/10.1007/s13762-019-02445-w

ORIGINAL PAPER

Characterization of indigenous bacteria from radon-rich groundwater and their tolerance to physicochemical stress

T. Nayak[1] · D. De[1] · C. Barman[2] · P. Karmakar[1] · A. Deb[2] · P. K. Dhal[1]

Received: 26 November 2018 / Revised: 8 April 2019 / Accepted: 17 June 2019 / Published online: 22 June 2019

Abstract

Radiation exposure and heavy metal (HM) exposure are serious health hazards causing DNA mutation, oxidative damage and may also be responsible for various life-threatening human diseases, while bacteria can withstand such toxicity easily. This study attempted to identify inhabitant bacterial isolates from radon-contaminated groundwater of Tantloi, India, and effort has also been made to characterize their response against radiation, oxidative stress as well as heavy metal tolerance and removal. Total sixteen (16) bacterial isolates were identified as *Bacillus* spp., *Stenotrophomonas* spp., *Brevibacillus* sp., *Chryseobacterium* sp., *Escherichia* sp. and *Microbacterium* sp., which showed less number of distinct carbohydrates utilization potential but high salinity tolerance properties. In addition, Gram-positive *Bacillus* spp. can tolerate 1 kGy of γ radiation, 10 mM H_2O_2, 7 days of desiccation and different heavy metals (Cu, Pb, Cr, Zn and As). Four *Bacillus* spp. and *Microbacterium* sp., which showed total maximum tolerable concentration (MTC) > 8 out of 5 heavy metals, were considered for their HM removal property analysis. Five multimetal resistance strains had strong removal capacity of Pb and Zn (89–94%) followed by Cr (49–56%) and Cu/As (0.4–22%). This investigation may provide baseline information of radon-contaminated groundwater microbiology, thus could be used to formulate an appropriate strategy for radon and radionuclides remediation.

Keywords Radon contamination · Groundwater · γ-Radiation · Heavy metal · Maximum tolerable concentration (MTC)

Introduction

Heavy metal (HM) and radionuclides contamination of groundwater via different anthropogenic activities cause serious threats to environment because of their toxic nature (Bahadir et al. 2007). The radionuclides affect the living organisms mainly through metabolic impairment, genetic mutation and malignant tumors (Garnier-Laplace et al. 2004). A very few attempt has been made on microbiology of radon-contaminated water except one on culture-independent and -dependent survey of the bacterial community of radon-contaminated hot spring, but they did not analyze further of those bacteria (Anitori et al. 2002). Most of the work done on the topic of radon was measuring the radon concentration in groundwater, hot springs, soil, thermal plant and radon level in house (Singh et al. 2008; Singh et al. 2002). Some research groups were trying to correlate the earthquake using the radon level in groundwater and soil (Negarestani et al. 2002). Lifetime fatality risk assessment study and risk of lung cancer analysis were done against indoor radon concentration by various scientists (Kansal et al. 2012; Darby et al. 2005). In these prospects, this is the first attempt to search the stress tolerance culturable microbiology from radon-contaminated groundwater so far.

The study site is located in and around Tantloi, Dumka district of Jharkhand, which is the part of Son–Narmada–Tapti (SONATA) a major fault of eastern fringes, India. Recently, it is observed that 42% of drinking water tube well of this geothermal region contains elevated level of radionuclides specifically radon, which exceeds the

Editorial responsibility: BV Thomas.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13762-019-02445-w) contains supplementary material, which is available to authorized users.

✉ P. K. Dhal
paltuk.dhal@jadavpuruniversity.in
http://www.jdvu.ac.in

[1] Department of Life Science and Biotechnology, Jadavpur University, 188 Raja S.C. Mallick Road, Kolkata 700 032, India

[2] School of Studies in Environmental Radiation and Archaeological Sciences, Department of Physics, Jadavpur University, Kolkata 700 032, India