

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter describes the available literature covering research topics related to this study. Background information on T2DM, human gut microbiota and next-generation sequencing technology is briefed in this chapter. The gut microbiota profile of different ethnic groups in various studies from Malaysia and worldwide are reviewed. Past findings on the gut microbiota profile in T2DM and the associated changes in clinical characteristics are also summarised.

2.2 Type 2 Diabetes

Diabetes is a growing threat to global health and is among the top causes of death worldwide. It is mainly categorised as type 1 and type 2, both characterised as the body's inability to produce or use insulin effectively (WHO, 2016). It was estimated that 463 million people, or 8.4% of the world population, were living with diabetes in the year 2019 (Diabetes Atlas, 2019). This is expected to increase to 700 million by 2045, approximately 9.9% of the global population (Saeedi et al., 2019). In Malaysia, diabetes is one of the leading non-communicable diseases (NCD), with 3.9 million people aged 18 years and above suffering from it (Institute for Public Health, 2019). The alarming increase in diabetes prevalence from 13.4% (2015) to 18.3% (2019) was

highlighted in the latest National Health and Morbidity Survey (NHMS) (Institute for Public Health, 2019).

Diabetes was also found to increase cardiovascular risk which is responsible for 35% of mortality among the Malaysian population (WHO, 2018). Type 2 diabetes (T2DM) accounts for up to 90% of people with diabetes around the world (Diabetes Atlas, 2019). Worldwide, T2DM is a growing health burden that impacts overall health quality and is a predisposing factor for other metabolic diseases. The complications that follow a T2DM diagnosis cause debilitating physical disabilities and affect psychosocial well-being (Clinical Practice Guidelines, 2015; WHO, 2016).

Asia has an escalating number of diabetes patients, and it is predicted that by 2030, more than 60% of T2DM cases will originate in Asia (Hu, 2011). The development of T2DM is strongly related to various factors like obesity, age, ethnicity, family history, as well as various genetic predisposition and environmental triggers (Grarup et al., 2014). More recently, owing to obesity, unhealthy diet and physical inactivity, T2DM, which used to be more common among older adults, is also affecting younger adults and children (Diabetes Atlas, 2019; Institute for Public Health, 2019). In a study in Malaysia, T2DM prevalence was found to differ among ethnic groups. Jamal et al. (2015) reported that Malaysian Chinese has a lower prevalence of T2DM in comparison to Malaysian Indians and Malays. The Malaysian National Health and Morbidity Survey in the year 2015 reported an increased number of T2DM cases among Malaysian Indians (22.1%), followed by Malaysian Malays (14.6%) and Malaysian Chinese (12.0%) (Institute for Public Health, 2015). This was agreed by Hussein et al. (2016), who found the highest prevalence of T2DM among Malaysian Indians. These disparities seen in T2DM prevalence among different ethnic groups may be associated

with specific factors, i.e. diet, lifestyle, and genetic or gene-environmental factors (Abdullah et al., 2014). Hence it is vital to identify the variation of risk factors for T2DM among different ethnic populations. This may also explain the distribution of T2DM prevalence around the globe.

2.2.1. Pathogenesis of T2DM

T2DM is characterised by increased blood glucose levels or hyperglycaemia due to altered insulin production and resistance. Insulin is a hormone secreted from the pancreas that facilitates glucose intake by cells from the bloodstream for energy harvest. A state of unresponsiveness of the body cells toward insulin triggers the mass production of insulin. In the long term, the pancreatic beta cells are damaged, leading to low insulin production in the body. This failure of insulin homeostasis leads to high glucose levels in the bloodstream, an important diagnostic factor for T2DM (Lin et al., 2010). The ineffective insulin also decreases glucose absorption for various organs, impairing the central and autonomous nervous system. Hence, there will be an increase in hormone secretions like glucagon and incretin, resulting in hepatic glucose output and renal glucose absorption, respectively (Clinical Practice Guidelines, 2015). This further increases the blood sugar level and worsening of T2DM. The progressive pathophysiology of insulin-causing hyperglycaemia will prompt various organ damage, ultimately leading to serious health complications, including cardiovascular diseases (CVD) (Kanter et al., 2016). Thus, T2DM has a significant impact on patients' morbidity and mortality.

2.2.2 Risk Factors for T2DM Development

The exact causes of T2DM are not entirely understood. Several studies linked the increasing incidence of T2DM with increasing age (Ogurtsova et al., 2017), family history, genetic heterogeneity (Aravindalochanan et al., 2014), economic development, and urbanisation which led to a sedentary lifestyle and consumption of unhealthy foods, and obesity (Cecchini et al., 2010; Hu, 2011).

T2DM is heritable via genetic variation within families (Cole et al., 2020). Individuals with family history of diabetes were significantly associated with increased risk for T2DM development. This is due to the presence of common genetic variants associated with risk of T2DM which could be heritable among family members (Lyssenko et al., 2008). As such, genetic factors could probably be used to predict the future development of T2DM.

Meanwhile, the increasing prevalence of physical inactivity and weight gain leading to abdominal obesity is also linked to an increase in T2DM prevalence (Cecchini et al., 2010; Hu, 2011). Abdominal obesity is an established determinant in identifying high-risk individuals to develop metabolic syndrome. This body fat distribution is calculated indirectly through anthropometrical measurements, namely hip circumference, waist circumference or waist-to-hip ratio (WHR), and body mass index (BMI) (Clinical Practice Guidelines, 2015).

The ethnic-specific differences in body fat distribution or abdominal obesity among T2DM patients have been reported in European and Asian studies (Caspard et al., 2018; Gulati et al., 2017; Wang et al., 2018). This suggests that an ethnic-specific observation in assessing body fat percentage may effectively identify the increased risk for T2DM.

In Malaysia, it was found that 64.8% of females, 68.3% of people of Indian ethnicity and 71.5 % of individuals in the 60-64 years old age group had abdominal obesity (Institute for Public Health, 2019). Furthermore, Abdullah et al. (2014) found that the major contributors to increasing T2DM prevalence in Malaysia were age, gender, elevated WHR ratio, diet, sleep duration, physical inactivity, family history and urbanisation. These factors have led to a rapid increase in T2DM prevalence in the urban population and may explain the increasing incidence of T2DM in Malaysia.

Recent evidence suggests that gut microbiota could potentially be a novel contributor to T2DM development (Gurung et al., 2020). Most studies found a significant difference in gut microbiota composition and function between individuals with T2DM and healthy controls (Allin et al., 2015). The interventional modulation of gut microbiota in T2DM, found in both studies on mice (Chen et al., 2018) and humans (Candela et al., 2016; Ohlsson, 2019), was shown to restore health-promoting bacteria, improve metabolic functions and glucose intolerance. Thus, a better understanding of the specific gut microbial alterations in T2DM could provide an alternative target for the prevention and management of T2DM.

2.2.3 Diagnosis and Management of T2DM

The T2DM appearance is mostly symptomless initially due to the progressive pathophysiology of insulin production and resistance. Thus, the exact onset of the disease is difficult to be determined. This results in a long pre-diagnostic period among T2DM and an increase in undiagnosed T2DM patients (Diabetes Atlas, 2019).

Moreover, the silent hyperglycaemic period of T2DM allows the appearance of micro and macrovascular complications that may fail to heal at diagnosis and could lead to life-threatening complications (Kanter & Bornfeldt, 2016). T2DM is a risk factor for cardiovascular diseases (CVD), cerebrovascular diseases and other complications, including nephropathy, retinopathy, neuropathy and dermatopathy (Diabetes Atlas, 2019). Therefore, the American Diabetic Association (ADA) guidelines of 1997 or World Health Organization (WHO) National diabetic group criteria of 2006 recommended simple tests for screening and diagnosis of T2DM (Olokoba et al., 2012).

In clinical settings, a patient's history and physical examination are still the first steps to diagnose T2DM. Patients with risk factors such as the history of hypertension, cardiovascular disease, impaired glucose tolerance or impaired fasting glucose, HDL <0.9 mmol/L or triglycerides >2.8 mmol/L, obesity and other characteristics leading to insulin resistance, such as polycystic ovarian syndrome (PCOS), physical inactivity, medications like antipsychotic drugs, history of gestational diabetes mellitus (GDM), are all screened for T2DM early (Clinical Practice Guidelines, 2015).

Patients with clinical symptoms for T2DM, i.e. tiredness, lethargy, polyuria, polydipsia, polyphagia, weight loss, pruritus vulvae and balanitis, are encouraged for diagnostic testing of T2DM and its complications (Clinical Practice Guidelines, 2015). The clinical parameters involve blood analysis [fasting plasma glucose (FPG), HbA1c, renal profile (RP), lipid profile and liver function test (LFT)], urinalysis (for albumin or microalbuminuria) and ECG. However, to specifically diagnose T2DM, there are three important diagnostic testing, i.e., oral glucose tolerance test (OGTT) for T2DM (FPG= ≥ 7.0 mmol/L; 2-hour postprandial (2HPP)= ≥ 11.1 mmol/L), HbA1c (T2DM

≥ 6.3 %) and venous plasma glucose [fasting (≥ 7.0 mmol/L) or random (≥ 11.1 mmol/L)] (Clinical Practice Guidelines, 2015).

The management of diabetes has been established on non-pharmacological and pharmacological basis. The non-pharmacological management focuses on lifestyle changes, encompassing a healthy diet, increasing physical activity, withdrawal from smoking or drinking habits, and maintaining a healthy body weight (Clinical Practice Guidelines, 2015). Pharmacological management consists of administration of oral anti-diabetic (OAD) agents for monotherapy, i.e., metformin, sulphonylureas (gliclazide, glibenclamide, glipizide, glimepiride), meglitinides (repaglinide and nateglinide), alpha-glucosidase inhibitors (acarbose), thiazolidinediones (rosiglitazone and pioglitazone), dipeptidyl peptidase-4 (sitagliptin, vildagliptin, saxagliptin, linagliptin and alogliptin). If glucose control is poor, a combination of any two or three OADs and/or injectable agents (i.e., insulin, GLP-1 receptor agonist) is indicated. Anti-hypertensive and lipid-lowering medicines may also be shown to prevent complications like CVD (Clinical Practice Guidelines, 2015; Diabetes Atlas, 2019). Despite efforts by the stakeholders like the Ministry of Health Malaysia to educate the public regarding T2DM, the incidence of diabetes rises, as evidenced in the NHMS 2019 (Institute for Public Health, 2019).

2.3 Human Gut Microbiota

The human microbiota consists of approximately 100 trillion cells, exceeding the human somatic and germ cells by a factor of ten (Peterson et al., 2009). The integration of genome and traits from both humans and microbes is thought to form the

basis of human genetic makeup and to develop human metabolic features (Gupta et al., 2017). Therefore, several physiological, metabolic and immunological functions in the body depend on the mutualistic relationship with the microbes and provide unique functional traits that are not evolved in humans (Lynch & Pedersen, 2016). The increasing understanding that the human microbiome profoundly influences human physiology in health and disease has led to various microbiome projects globally.

Initial projects, i.e. the Human Microbiome Project (Huttenhower et al., 2012; Peterson et al., 2009) and MetaHit (Qin et al., 2010), attempted to characterise the composition, diversity and functional properties of the human microbiome. Since then, with the development of culture-independent approaches, namely 16S ribosomal RNA (16S rRNA) gene sequencing, metagenomics, meta-transcriptomes, meta-proteomics and metabolomics, multi-species microbial communities in various niches in the human body have been discovered (Lynch & Pedersen, 2016). The diversity and abundance of microbiome functionality in sites, namely gut, skin, oral and vagina, proved to be specific and personalised in each healthy individual (Huttenhower et al., 2012).

The human gut microbiota (HGM) has the most significant number and most diverse microbes compared to other body sites (Gupta et al., 2017). It is also observed that 99.1% of the gut is dominated by bacteria, mostly of phyla *Firmicutes* and *Bacteroidetes*. HGM also varies in different individuals, but a small subset of core gut bacteria and genes from the core microbiome was found to be shared among individuals (Qin et al., 2010). This further supports the notion of the mutualistic symbiotic relationship between the microbes and the host. The various endogenous and exogenous factors, i.e. age (Jandhyala et al., 2015), diet, lifestyle (Hu, 2011), exposure to antimicrobial agents (Forslund et al., 2015), host immunity (Levy et al., 2016),

ethnicity, geography and environmental influences (Gupta et al., 2017), shape the gut microbiota over time.

2.3.1 Development of Gut Microbiota

The growing knowledge of the development of the gut microbiota from the period of infancy is vital for understanding the role of HGM in disease. The sterile foetal gut is colonised with microbes soon after birth. The unstable microbial community at the infancy stage is affected by mode of delivery, diet, medical factors and gestational age at birth (Jandhyala et al., 2015). As the HGM of infants matures into an adult-like gut microbiota beginning from age one, it tends to remain stable and once again declines with old age (Lynch & Pedersen, 2016). The HGM structure and composition result from natural selection by microbes and the host's mutualistic relationship in performing physiological functions in the body. The gut microbiota fluctuates with different environmental factors and usually returns to its original composition once the perturbations are resolved (Candela et al., 2012).

The predominant phyla observed in the adult gut are *Firmicutes*, *Bacteroidetes*, *Actinobacteria*, *Proteobacteria* and *Verrucomicrobia*. The Gram-positive *Firmicutes* and the Gram-negative *Bacteroidetes* represent 90% of the total composition of the human gut bacteria and produce short-chain fatty acids (SCFA) (Qin et al., 2010). The SCFA are metabolic products of microbial fibre fermentation in the gut that benefit human metabolism and the immune system (Parada Venegas et al., 2019). In a healthy gut, *Firmicutes* are mainly represented by butyrate producers, namely genera *Clostridium*, *Eubacterium*, *Faecalibacterium* and *Roseburia* (Parada Venegas et al.,

2019). Butyrate protects against inflammation by reducing pro-inflammatory cytokine production, maintaining the integrity of gut epithelium and serving an essential role in blood glucose and lipid regulation (Lin et al., 2012).

On the other hand, phylum *Bacteroidetes* is represented by acetate and propionate producers, i.e. *Bacteroides* and *Prevotella* (Parada Venegas et al., 2019; Telle-Hansen et al., 2018). The genera under subdominant phyla, namely *Verrucomicrobia* (i.e. *Akkermansia*) and *Actinobacteria* (i.e. *Bifidobacterium*) were found to be higher in abundance while *Proteobacteria* (i.e. *Escherichia coli*) was commonly found in lower abundance in a healthy gut (Eckburg et al., 2005; Jandhyala et al., 2015). Gut microbiota has various functions, including protection against epithelial cells injury (Rakoff-Nahoum et al., 2004), regulation of host fat storage (Bäckhed et al., 2004), stimulation of intestinal angiogenesis (Stappenbeck et al., 2002), metabolism of dietary components, development of the immune system, biogenesis of energy (Jandhyala et al., 2015), and biosynthesis of biological molecules (steroid hormones, vitamins and neurotransmitters) (Lynch & Pedersen, 2016).

Numerous studies have highlighted the change in gut microbiota composition in the disease state. Dysbiosis is a term describing the alterations in the composition of HGM and its physiological functions in an individual with disease, compared to modifications to a healthy host (Vallianou et al., 2018). Hence, understanding the healthy microbiome composition could explore the plausible links of the microbiome with disease.

2.3.2 Gut Microbiota and Ethnicity

Various studies have found that the gut microbial composition varies between ethnic groups. Race or ethnicity impacts the conformation of the gut microbiota among individuals residing in various geographical locations (Gupta et al., 2017). Although not understood fully, this difference could be due to food intake, genetics, or the living environment (Hu, 2011).

A large cross-sectional study conducted across nine provinces in China on 314 volunteers belonging to 7 ethnic groups (Bai, Han, Kazakh, Mongol, Tibetan, Uyghur and Zhuang) found significant segregation of gut microbiota composition specific to the host ethnic group suggesting a strong effect of ethnicity on gut microbiota composition (Zhang et al., 2015). Although varying gut microbiota composition were found among different ethnic groups, similar genus-level phylogenetic core was found in all study participants. The core comprising of nine predominant genera, namely *Phascolarctobacterium*, *Roseburia*, *Bacteroides*, *Blautia*, *Faecalibacterium*, *Clostridium*, *Subdoligranulum*, *Ruminococcus* and *Coprococcus* were also known SCFA-producers. However, the phylogenetic core varied significantly at the species-level. This indicates that a common and heritable microbial lineage found in all the ethnic groups tend to evolve with host associated factors, i.e., lifestyle, ethnicity or geography. Thus, the identification of phylo-functional core as well as their evolutionary changes among healthy individuals could serve as an important link in understanding gut microbiota changes in health and disease (Zhang et al., 2015).

In a similar study in China, a statistically significant difference in richness and diversity of gut microbiota composition was reported between three different ethnic

groups (59 Tu Yao, 47 Han and 28 Zhuang) that lived in the same rural area with one ethnic group, Tu Yao, living more isolated in the mountains (Liao et al., 2018). The main ethnicity-associated difference the study noted was difference in the abundance of genus *Megamonas*, which positively correlated with consumption of beans among Han as compared to the Zhuang and Tu Yao groups. Also, although not associated with ethnicity, the study found that differences in the abundance of dominant genera *Bacteroides* and *Prevotella* was associated with alcohol consumption. A higher relative abundance of *Bacteroides* and lower abundance of *Prevotella* was noted among those who consumed more alcohol in both Han and Yao groups. Hence, the varying composition of gut microbiota found between different ethnic groups in this study was linked with dietary intake (Liao et al., 2018).

Few studies have also investigated gut microbiota variation in the Malaysian multi-ethnic community and investigated its association with different factors. A study led by Chong et al. (2015) analysed the gut microbiota of primary school children aged 7 to 12 from rural areas of Perak. The three ethnic groups recruited were Malay (N=24), Chinese (N=1) and Orang Asli-Temiar Tribe (N=20). The highest bacterial diversity was seen among economically deprived Orang-Asli children, while low diversity was seen among Malay and Chinese children exposed to a westernised diet and sedentary lifestyle. While, as globally noted, the phyla *Firmicutes* and *Bacteroidetes* were dominant across all study groups, the Orang-Asli children had a higher abundance of order *Aeromonadales*, unclassified order of *Bacteroidetes* and *Deltaproteobacteria* as well as an unclassified genus of family *Ruminococcaceae* in comparison to Malay and Chinese children. Hence, it was found that gut bacteria is strongly affected by ethnicity-

related socio-economical variation. This disparity is linked with host genetics and the lifestyle practised by different ethnic groups (Chong et al., 2015).

In another study, Khine et al. (2019) compared the gut microbiota of healthy Malay children from Penang (N=21) and Kelantan (N=33) with Chinese children (7-12 years old) from Penang (N=21), Kelantan (N=45) and Guangzhou city in China (N=18). This study identified that geographical-associated factors, i.e., diet and lifestyle, instead of ethnicity, affected the gut microbiota composition. This was reflected in children of Chinese ethnicity living in different cities with significantly different gut microbiota. In contrast, children of Chinese and Malay ethnicity residing in the same city had no significant difference in gut microbiota, suggesting that they share similar microbiota profiles (Khine et al., 2019). Thus, this study showed that similar dietary habits practised by the community living in the same region are a major determinant in shaping the gut microbiota composition (Khine et al., 2019).

Meanwhile, Dwiyanto et al. (2021) investigated the gut microbiota of four major ethnic groups (46 Malay, 65 Chinese, 49 Indian and 54 Orang Asli-Jakun) living in a same community (Segamat, Johor). The permutational multivariate analysis of variance (PERMANOVA) analysis revealed that ethnicity exerted the most significant effect size and the ethnic-specific enterotypes were identified. The increased proportion of *Prevotella* and *Bacteroides*-dominant enterotypes was found among Jakun and Chinese ethnic groups, respectively. Meanwhile, Malay and Indian ethnic groups were equally distributed in these two dominant enterotypes (Dwiyanto et al., 2021). The ethnic distribution in enterotypes was reflected in the *Prevotella* and *Bacteroides* ratio where Jakun, Indian, Malay and Chinese ethnic groups exhibited largest to lowest ratio, consecutively. Hence, in this Malaysian locality, ethnicity strongly influenced gut

microbiota composition, and this difference was associated with ethnic-associated unique lifestyle and dietary habits practised by the different ethnic groups in this study (Dwiyanto et al., 2021). Therefore, it is clear that the impact of ethnicity should be included in studies investigating gut microbiota patterns in disease.

2.3.3 Gut Microbiota and T2DM

Evidence suggests the involvement of gut microbiota is an essential factor for T2DM development, in addition to host genetics (Aravindalochanan et al., 2014), environmental factors (Razak et al., 2005; Wellen et al., 2005) and sedentary lifestyle (Hu, 2011). T2DM is characterised by hyperglycaemia, insulin resistance leading to inadequate insulin secretion, abnormal metabolism of glucose, lipid and amino acid (Saeedi et al., 2019). The influence of gut dysbiosis in association with metabolic endotoxemia and reduced secretion of incretins as well as butyrate may induce the low-grade inflammation seen with T2DM occurrence (Vallianou et al., 2018). This may explain changes in glucose metabolism, insulin resistance and possibly, T2DM development.

Table 2.1 shows a list of studies that analysed gut microbiota composition using 16S rDNA sequencing technology among T2DM participants when compared to healthy controls.

Table 2.1 Studies on Gut Microbiota Composition between Healthy and T2DM Adults using 16S rDNA Sequencing.

Author (year), ref	Study	Ethnicity	Gut Microbiota Composition (Phylum/Class/Order/Family/Genus/Species)	
			Abundant in T2DM	Abundant in Healthy Controls
(Afolayan et al., 2020)	This is a cross-sectional study aimed to analyse the GM composition in healthy elderly individuals (N=22) and elderly people living with T2DM (N=20) in Ibadan, Nigeria	Yoruba, Ibo and Others	i. p. <i>Actinobacteria</i> [f. <i>Coriobacteriaceae</i> (g. <i>Collinsella</i>) & f. <i>Bifidobacteriaceae</i>] ii. p. <i>Firmicutes</i> (g. <i>Ruminococcus</i>)	i. p. <i>Firmicutes</i> [f. <i>Clostridiaceae</i> (g. <i>Clostridium</i>) & f. <i>Peptostreptococcae</i>]
(Ahmad et al., 2020)	This is a case-control study aimed at investigating the GM composition in healthy (N=20) and obese individuals with T2DM (40) in Islamabad, Pakistan	Punjabi	i. p. <i>Firmicutes</i> (g. <i>Lactobacillus</i> , g. <i>Subdoligranulum</i> , g. <i>Allisonella</i> g. <i>Dialister</i> g. <i>Negativicutes</i> , & g. <i>Eubacterium</i>) ii. p. <i>Actinobacteria</i> (c. <i>Coriobacteriia</i>)	i. p. <i>Firmicutes</i> (g. <i>Bacillus</i> & g. <i>Ruminococcus</i>) ii. p. <i>Bacteroidetes</i> (g. <i>Bacteroides</i> & g. <i>Prevotella</i>) iii. p. <i>Proteobacteria</i> (c. <i>Gammaproteobacteria</i> , c. <i>Deltaproteobacteria</i> & g. <i>Escherichia-Shigella</i>) iv. p. <i>Verrucomicrobia</i> (c. <i>Verrucomicrobiae</i>) v. p. <i>Elusimicrobia</i>
(Almugadam et al., 2020a)	This is a case-control study aimed at investigating the GM composition among healthy (N=24) and T2DM individuals (N=24) as well as exploring the effect of the antidiabetic agent on gut flora in Sudan.	N/A	i. p. <i>Actinobacteria</i> (g. <i>Bifidobacterium</i>) ii. p. <i>Firmicutes</i> (g. <i>Catenibacterium</i> , g. <i>Blautia</i> , g. <i>Holdemanella</i> & g. <i>Parvimonas</i>) iii. p. <i>Fusobacteria</i> (g. <i>Fusobacterium</i>)	i. p. <i>Firmicutes</i> (g. <i>Faecalibacterium</i> & g. <i>Dialister</i>) ii. p. <i>Proteobacteria</i> iii. p. <i>Verrucomicrobia</i> (g. <i>Succinivibrio</i>) iv. p. <i>Elusimicrobia</i> (g. <i>Elusimicrobium</i>)
(Alvarez-Silva et al., 2021)	This is a case-control study comparing the GM composition of healthy adults (Danish, N=138 and Indian N=137) and adults with T2DM (Danish, N=141 and Indian, N=157) in Denmark and India	White European (Danish) and South Indian (Dravidian) ancestry	i. p. <i>Firmicutes</i> (g. <i>Lachnospirillum</i>)	i. p. <i>Firmicutes</i> (g. <i>Anaerospirillum</i> , g. <i>Subdoligranulum</i> & g. <i>Butyrivibrio</i>)

GM, gut microbiota; p, Phylum; c, Class; f, Family; g, Genus; sp, Species. All findings are significant with $p < 0.05$.

Table 2.1 continued.

Author (year), ref	Study	Ethnicity	Gut Microbiota Composition (Phylum/Class/Order/Family/Genus/Species)	
			Abundant in T2DM	Abundant in Healthy Controls
(Doumatey et al., 2020)	This is a case-control study aimed at identifying the GM composition among healthy controls (N=193) and T2DM individuals (N=98) in urban West Africa	N/A	i. p. <i>Actinobacteria</i> (f. <i>Coriobacteriaceae</i> & f. <i>Bifidobacteriaceae</i>) ii. p. <i>Bacteroidetes</i> (g. <i>Prevotella</i>) iii. p. <i>Firmicutes</i> (g. <i>Peptostreptococcus</i> & g. <i>Eubacterium</i>) iii. p. <i>Euryarchaeota</i> iv. p. <i>Tenericutes</i>	i. p. <i>Firmicutes</i> (f. <i>Clostridiaceae</i> & f. <i>Peptostreptococcaceae</i> , g. <i>Anaerostipes</i> , g. <i>Ruminococcus</i> , g. <i>Clostridium</i> & g. <i>Epulopiscium</i>)
(Fassatoui et al., 2019)	This is a case-control study aimed at identifying the GM composition among type-1 (N=10) and type-2 T2DM subjects (N=10) in comparison to healthy controls (N=10) in Tunis.	N/A	No significant findings	i. p. <i>Verrucomicrobia</i> (sp. <i>Akkermansia muciniphila</i>) ii. p. <i>Firmicutes</i> (sp. <i>Faecalibacterium prausnitzii</i>)
(Hoang et al., 2021)	This case-control study aimed to identify the GM composition of healthy (N=7) and T2DM subjects (N=7) in Hanoi, Vietnam.	Vietnamese	No significant findings	i. p. <i>Firmicutes</i> (c. <i>Clostridia</i>)
(Kitten et al., 2021)	This is a cross-sectional study aimed at identifying GM composition in healthy (N=23) and individuals with T2DM (N=14) in San Antonio, Texas	Mexican Americans	i. p. <i>Firmicutes</i> (g. <i>Blautia</i> , g. <i>Ruminococcus</i> , g. <i>Clostridium</i> , & sp. <i>Ruminococcus torques</i>)	i. p. <i>Bacteroidetes</i> (g. <i>Prevotella</i>) ii. p. <i>Firmicutes</i> (f. <i>Lachnospiraceae</i> & g. <i>Streptococcus</i>)

GM, gut microbiota; p, Phylum; c, Class; f, Family; g, Genus; sp, Species. All findings are significant with $p < 0.05$.

Table 2.1 continued.

Author (year), ref	Study	Ethnicity	Gut Microbiota Composition (Phylum/Class/Order/Family/Genus/Species)	
			Abundant in T2DM	Abundant in Healthy Controls
(Larsen et al., 2010)	This is a case-control study aimed at assessing the difference in GM composition among healthy controls (N=18) and T2DM individuals (N=18) in Denmark.	N/A	i. p. <i>Proteobacteria</i> (c. <i>Betaproteobacteria</i>) ii. p. <i>Bacteroidetes</i> (g. <i>Bacteroides</i>)	i. p. <i>Firmicutes</i> (c. <i>Clostridia</i>)
(Leite et al., 2017)	This case-control study aimed to identify GM composition in healthy controls (N=22) and Brazilian T2DM participants (N=20) from Barretos, Sao Paulo, Brazil.	Caucasian, Asian & Afro-descendent	i. p. <i>Firmicutes</i> (c. <i>Clostridia</i> , o. <i>Clostridiales</i> , f. <i>Ruminococcaceae</i>) ii. p. <i>Bacteroidetes</i> (f. <i>Prevotellaceae</i> , g. <i>Prevotella</i> , sp. <i>Prevotella copri</i> & sp. <i>Bacteroides xylanisolvens</i>)	i. p. <i>Firmicutes</i> (f. <i>Lachnospiraceae</i>) ii. p. <i>Bacteroidetes</i> (c. <i>Bacteroidia</i> , o. <i>Bacteroidales</i> , f. <i>Bacteroidaceae</i> , g. <i>Bacteroides</i> , sp. <i>Bacteroides vulgatus</i> & sp. <i>Bacteroides rodentium</i>)
(Li et al., 2021)	This case-control study aimed to identify GM composition between healthy (Han, N=10 and Mongolian, N=18) and T2DM participants (Han, N=10 and Mongolian, N=18) in two different ethnic groups living in the same area in Beijing, China.	Han Mongolian	i. p. <i>Bacteroidetes</i> (g. <i>Alistipes</i>) ii. p. <i>Actinobacteria</i> (g. <i>Bifidobacterium</i>) iii. p. <i>Firmicutes</i> (g. <i>Papillibacter</i>) i. p. <i>Bacteroidetes</i> (g. <i>Barnesiella</i>) ii. p. <i>Actinobacteria</i> (g. <i>Bifidobacterium</i>) iii. p. <i>Firmicutes</i> (g. <i>Papillibacter</i>)	i. p. <i>Firmicutes</i> (g. <i>Clostridium XIVa</i> , g. <i>Clostridium XI</i> , g. <i>Flavonifractor</i> , g. <i>Oscillibacter</i> , g. <i>Fastidiosipila</i> & <i>Lachnospiraceae incertae sedis</i>) ii. p. <i>Bacteroidetes</i> (g. <i>Prevotella</i>) i. p. <i>Firmicutes</i> (g. <i>Fastidiosipila</i>)
(Li et al., 2020b)	This is a cohort study aimed at analysing GM composition among healthy (N=30) and T2DM individuals (N=50) in Northern China	N/A	i. p. <i>Firmicutes</i> (g. <i>Faecalibacterium</i> , g. <i>Dorea</i> & sp. <i>Faecalibacterium prausnitzii</i>) ii. p. <i>Fusobacteria</i> (g. <i>Fusobacterium</i>)	i. p. <i>Bacteroidetes</i> (g. <i>Parabacteroides</i>) ii. p. <i>Firmicutes</i> (g. <i>Streptococcus</i>) iii. p. <i>Actinobacteria</i> (g. <i>Bifidobacterium</i>) iv. p. <i>Verrucomicrobia</i> (g. <i>Akkermansia</i>)

GM, gut microbiota; p, Phylum; c, Class; f, Family; g, Genus; sp, Species. All findings are significant with $p < 0.05$.

Table 2.1 continued.

Author (year), ref	Study	Ethnicity	Gut Microbiota Composition (Phylum/Class/Order/Family/Genus/Species)	
			Abundant in T2DM	Abundant in Healthy Controls
(Pushpanathan et al., 2016)	This is a case-control study aimed at determining the GM composition and the profile of monocyte chemoattractant protein-1 (MCP-1) and interferon-gamma (IFN γ) among healthy (N=13) and T2DM individuals (N=17) attending a tertiary care centre in India	N/A	i. p. <i>Proteobacteria</i> (g. <i>Escherichia</i>) ii. p. <i>Firmicutes</i> (g. <i>Lactobacillus</i>) ii. p. <i>Bacteroidetes</i> (g. <i>Prevotella</i>)	i. p. <i>Firmicutes</i> (g. <i>Faecalibacterium</i> & g. <i>Eubacterium</i>) ii. p. <i>Actinobacteria</i> (g. <i>Bifidobacterium</i>)
(Salamon et al., 2018)	This is a case-control study aimed at identifying the GM composition among type-1 (N=22) and type-2 T2DM subjects (N=23) in comparison to healthy volunteers (N=23) in Poland	N/A	i. p. <i>Firmicutes</i> (g. <i>Ruminococcus</i>) ii. p. <i>Proteobacteria</i> (f. <i>Enterobacteriaceae</i>)	i. p. <i>Bacteroidetes</i> (f. <i>Flavobacteriaceae</i> & g. <i>Bacteroides</i>) ii. p. <i>Firmicutes</i> (f. <i>Clostridiaceae</i> , f. <i>Lachnospiraceae</i> , f. <i>Peptostreptococcaceae</i> g. <i>Anaerostipes</i> , & g. <i>Roseburia</i>)
(Wang et al., 2020)	This case-control study aimed at identifying the difference in GM composition between T2DM subjects (N=134) and healthy controls (N=37) in China.	N/A	i. p. <i>Proteobacteria</i> (g. <i>Escherichia-Shigella</i>) ii. p. <i>Actinobacteria</i> (g. <i>Bifidobacteria</i>) iii. p. <i>Firmicutes</i> (g. <i>Faecalibacterium</i> & g. <i>Lactobacillus</i>)	i. p. <i>Bacteroidetes</i> (g. <i>Prevotella</i>) ii. p. <i>Firmicutes</i> (g. <i>Roseburia</i>)
(Wang et al., 2017)	This is a case-control study aimed at identifying GM composition between healthy (Uygurs and Kazaks, N=10) and T2DM participants (Uygurs and Kazaks, N=10) in two ethnic minority groups in Xinjiang, China	Uygurs Kazaks	i. p. <i>Firmicutes</i> (f. <i>Erysipelotrichaceae</i>) i. p. <i>Firmicutes</i> (f. <i>Veillonellaceae</i>) ii. p. <i>Proteobacteria</i> (f. <i>Desulfovibrionaceae</i> & f. <i>Aeromonadaceae</i>)	No significant findings i. p. <i>Firmicutes</i> (f. <i>Planococcaceae</i> & f. <i>Veillonellaceae</i>) ii. p. <i>Actinobacteria</i> (f. <i>Coriobacteriaceae</i>)

GM, gut microbiota; p, Phylum; c, Class; f, Family; g, Genus; sp, Species. All findings are significant with $p < 0.05$.

The gut microbiota composition in T2DM was found to vary in different studies. Larsen and co-researchers (2010) were the first to report that T2DM participants, compared to healthy controls, undergo gut microbial dysbiosis, whereby a shift occurs in proportions of phyla *Firmicutes* and *Bacteroidetes*. Few studies that reported either a higher (Li et al., 2020b; Salamon et al., 2018; Wang et al., 2020) or lower (Doumatey et al., 2020; Fassatoui et al., 2019) abundance of Gram-positive *Firmicutes* also found a lower and higher abundance of Gram-negative *Bacteroidetes* in T2DM participants, respectively. Two of these studies also reported a higher *Firmicutes/Bacteroidetes* (F/B) ratio in T2DM participants, when compared to healthy controls (Salamon et al., 2018; Wang et al., 2020). The F/B ratio value has been associated with a wide range of diseases including metabolic disorders such as obesity, cardiovascular diseases and type 2 diabetes (Cani, 2018; Magne et al., 2020). While some studies report it as a measure of dysbiosis, others find no significant difference between the F/B ratio of cases versus control (Almugadam et al., 2020b; Larsen et al., 2010; Zhang et al., 2013). The abundance of Gram-positive *Actinobacteria* was also seen to either increase (Almugadam et al., 2020a; Doumatey et al., 2020; Wang et al., 2020) or decrease (Li et al., 2020b; Pushpanathan et al., 2016) in T2DM participants. On the other hand, a higher abundance of Gram-negative *Proteobacteria* (Larsen et al., 2010; Pushpanathan et al., 2016; Salamon et al., 2018; Wang et al., 2020) and a lower abundance of Gram-negative *Verrucomicrobia* (Almugadam et al., 2020a; Fassatoui et al., 2019; Li et al., 2020b) were also seen in T2DM participants.

The gut microbiota composition in T2DM also seem to vary among different ethnic groups residing in various geographical locations (Table 2.1). In a study investigating Caucasian, Asian and Afro-descendent ethnic groups found in Brazil,

phylum *Firmicutes* was found in lower abundance among T2DM participants compared to the healthy control groups (Leite et al., 2017). Similar findings were also seen among Vietnamese participants in a study from Vietnam (Hoang et al., 2021). In the Punjabi ethnic population in Pakistan, a lower abundance of phyla *Bacteroidetes*, *Proteobacteria* and *Verrucomicrobia* were noted among T2DM participants (Ahmad et al., 2020), while in another study, only a lower abundance of phylum *Bacteroidetes* was noted among the Mexican American participants with T2DM (Kitten et al., 2021), in comparison to their healthy control groups, respectively. On the other hand, among the Yoruba ethnic group in Nigeria, an increased abundance of phylum *Actinobacteria* was reported among the T2DM participants compared with the healthy controls of the same ethnicity (Afolayan et al., 2020).

When the gut microbiota composition of T2DM was compared between different ethnic groups, phyla *Bacteroidetes* and *Actinobacteria* increased in abundance in both Han and Mongolian ethnic groups with T2DM in China. The F/B ratio was significantly higher in the Mongolian T2DM group than in the healthy Mongolian control group (Li et al., 2021). Meanwhile, in another study done in China, the Kazak ethnic group with T2DM had a higher abundance of phylum *Proteobacteria* and a lower abundance of phylum *Actinobacteria* in comparison to the Uygur ethnic group with T2DM, that had a lower abundance of phylum *Firmicutes* when compared to their respective healthy controls (Wang et al., 2017). Hence, it is clear that varying gut microbiota composition is found in different ethnic groups with T2DM.

Thus, it is essential to identify if a population has ethnic-specific changes in the gut microbial composition in T2DM as the development of population-based medicine could be good choice for gut microbiota-based prevention and management of T2DM.

Besides that, gut microbiota alterations in T2DM are also associated with metformin intake, the most prescribed antidiabetic medication for T2DM patients. The consumption of metformin was found to either have therapeutic or adverse effects on the intestine by modification of gut microbiota composition (Forslund et al., 2015). Hence, to fully explore the gut microbiota changes in T2DM, an observation of the gut microbial changes without the effect of drugs should be done. Therefore, it is important to analyse the gut microbiota composition among pre-diabetics (preDM) and newly diagnosed diabetics (newDM) who have yet to begun pharmacotherapy. By understanding the specific gut changes occurring at earlier stages of T2DM development, a gut microbiota-based screening tool could be established to identify individuals at high risk of developing T2DM.

2.3.4 Clinical Characteristics and Gut Microbiota in T2DM

Factors correlating with the gut microbiota enable the identification of relevant associations contributing to health and disease. The gut dysbiosis that occurs along with changes in clinical characteristics could be associated with underlying insulin resistance and a pro-inflammatory state (Vallianou et al., 2018).

Table 2.2 shows various clinical characteristics that were altered in T2DM and correlated with alteration in gut bacterial composition in several studies. The increase in weight or upper body fat measured with BMI had been established to correlate strongly with an increased insulin resistance state (Wondmkun, 2020). The visceral fat accumulation may indicate altered energy metabolism, suggesting its interplay in the development of metabolic syndromes (Jiang et al., 2020). Furthermore, obese and ageing individuals are high-risk for T2DM and cardiovascular disease development (Barlow et al., 2015). The increase in BMI, diabetes profile [FPG, insulin, HOMA-IR, HbA1C], lipid profile [total cholesterol (TC), low-density lipoprotein cholesterol (LDL), high-density lipoprotein cholesterol (HDL) and triglycerides (TG)] and markers of inflammation [C-reactive protein (CRP), IL-6 and TNF- α], as well as low diversity of gut bacteria, are often found among T2DM participants (Ahmad et al., 2020; Bhute et al., 2017; Doumatey et al., 2020; Sato et al., 2014). Sato and their team (2014) observed an increase of lipopolysaccharide-binding protein (LBP) in plasma correlated positively with BMI, HbA1C, CRP, TNF- α and IL-6.

Table 2.2 The Correlation of Clinical Characteristics with Gut Microbiota Composition in T2DM Participants.

Parameters	Correlation of Parameters with Gut Microbiota (references) in T2DM	
	Positive Correlation	Negative Correlation
Age	N/A	f. <i>Caulobacteraceae</i> (Salamon et al., 2018) g. <i>Akkermansia</i> (Salamon et al., 2018)
BMI	N/A	f. <i>Ruminococcaceae</i> (Salamon et al., 2018) g. <i>Streptococcus</i> (Salamon et al., 2018) sp. <i>Faecalibacterium prausnitzii</i> (Navab-Moghadam et al., 2017)
FPG	p. <i>Firmicutes</i> (Ahmad et al., 2020) p. <i>Actinobacteria</i> (Ahmad et al., 2020) c. <i>Betaproteobacteria</i> (Larsen et al., 2010) g. <i>Clostridium cluster XI</i> (Adachi et al., 2019)	p. <i>Bacteroidetes</i> (Ahmad et al., 2020) p. <i>Proteobacteria</i> (Ahmad et al., 2020) sp. <i>Akkermansia muciniphila</i> (Fassatoui et al., 2019)
HbA1c	f. <i>Flavobacteriaceae</i> (Salamon et al., 2018)	g. <i>Faecalibacterium</i> (Salamon et al., 2018) g. <i>Collinsella</i> (Salamon et al., 2018) g. <i>Poryomonas</i> (Salamon et al., 2018) g. <i>Bulleidia</i> (Salamon et al., 2018) sp. <i>Akkermansia muciniphila</i> (Fassatoui et al., 2019)
TC	N/A	g. <i>Clostridium cluster XVIII</i> (Adachi et al., 2019)
LDL	g. <i>Enterococcus</i> (Salamon et al., 2018)	g. <i>Clostridium cluster XVIII</i> (Adachi et al., 2019)
TG	f. <i>Gemellaceae</i> (Salamon et al., 2018) g. <i>Atopobium</i> (Salamon et al., 2018)	N/A

BMI, body mass index; HDL, high-density lipoprotein; TC, Total cholesterol; LDL, low-density lipoprotein; TG, triglyceride, p, Phylum; c, Class; f, Family; g, Genus; sp, Species. All the findings are significant with $p < 0.05$.

2.3.5 Next-Generation Sequencing, 16S rDNA Sequencing

Next-generation sequencing (NGS) is a collective term used to describe the revolutionising nucleic acid sequencing technology in genomic studies (Behjati et al., 2013). These high-throughput sequencing techniques have proven to be rapid, sensitive and less costly for large-scale sequencing of nucleic acids as compared to the previously widely used Sanger sequencing (Heather et al., 2016; van Dijk et al., 2014). NGS has been used to monitor genetic evolution and drug resistance and as a tool in the epidemiological tracing of infectious diseases (Ekblom et al., 2011).

In traditional bacterial identification methods, the bacterial morphology and phenotypic characteristics were used to identify an isolate, which has many variations in different laboratories (Pence et al., 2020). This method was only limited to the identification of microbes that could be successfully cultured, which is estimated to be less than 2% of the bacteria found in the environment (Wade, 2002). Hence, in the 1980s, the identification of a new standard in classifying bacteria based on their stable genome coding for 16S rRNA was introduced (Janda et al., 2007). Ever since, the high throughput sequencing technology provided new insights into identifying unculturable microbes living in a mixed-bacterial community.

The 16S rRNA gene is a component in all bacterial species that codes for the RNA part of the 30S subunit of the bacterial ribosome (Wang et al., 2015). This 16S rRNA gene (i.e. the 16S rDNA) which is 1.5 kilobases in length, contains both highly conserved and variable regions (Clarridge, 2004). While the conserved regions have identical sequences across species and are hence suitable sites for primer design, the nine short hypervariable region sequences are helpful in identifying the bacterial taxa

up to species level (Clarridge, 2004; Weinstock, 2012). Hence, 16S rRNA is the most common gene targeted to taxonomically classify the trillions of bacteria present in their native environments (Johnson et al., 2019).

2.4 Conclusion

This chapter reviewed the available literature on the influence of ethnicity on gut microbiota composition. It also summarised available findings on gut microbiota in T2DM. This gives an insight into past studies done around similar areas of research in Malaysia and other regions of the world. This helped to identify the research gap in this study and provided some idea on the expected study findings, herein.

