CHAPTER 4

RESULTS

4.1 Introduction

This chapter describes the findings of this study. Firstly, the characteristics of study participants based on ethnicity and/or diabetes status are reported. This is followed by microbiota diversity and compositional analysis at the phylum and genus taxonomic levels. The gut microbiota data is then used for correlational analysis with the variables investigated in this study. Lastly, the findings from the systematic review are summarized.

4.2 Characteristics of Study Participants

The clinical characteristics (anthropometry, demographic, diabetic profile and biochemical parameters) of the 90 participants recruited in this study, consisting of 45 participants each in the nonDM and T2DM groups, respectively are shown in Table 4.1. Participants were matched by ethnicity whereby each group consisted of 15 Malay, 15 Chinese and 15 Indian participants respectively. The characteristics of participants are reported in either mean ± SD for normally distributed data or median (interquartile range, IQR) for non-normally distributed (skewed) data.

In comparing the anthropometric, demographic and diabetic profile characteristics between nonDM and T2DM groups, the age and BMI were found to be significantly different. The T2DMs were older (57.02 ± 9.98 years) than nonDMs (45.07 ± 12.82 years) and had higher BMI (28.35 ± 4.99 kg/m^2) than nonDMs (26.14 ± 5.01 kg/m^2). As expected, the diabetic parameter i.e., FPG level was also significantly higher in T2DM (9.25 ± 3.43 mmol/L) when compared to nonDM (5.21 ± 0.41 mmol/L).

The biochemical parameters consisted of three major tests, the liver function test (LFT), renal profile (RP) and fasting serum lipid (FSL). In LFT, only ALP was significantly higher in T2DM (84.67 ± 19.81 U/L) when compared to nonDM (76.16 ± 19.62 U/L). In T2DM, urea was significantly increased [4.60 (2.35) mmol/L] while chloride, conversely, was decreased (101.42 ± 2.75 mmol/L) when compared to nonDM (4.08 ± 1.27 mmol/L and 102.71 ± 2.62 mmol/L, respectively). In the FSL test, only TG was found to be significantly higher in T2DM [1.50 (1.15) mmol/L] than the nonDM (1.20 ± 0.54 mmol/L).

Table 4.1 The Characteristics of Study Participants in NonDM and T2DM Groups, n=90.

Parameters	NonDM Mean ± SD or *Median (IQR)* n=45	T2DM Mean ± SD or *Median (IQR)* n=45	All participants Mean ± SD or *Median (IQR)* n=90	*p*-value	OR (95% CI)
Age (years)	45.07 ± 12.82	57.02 ± 9.98	51.06 ± 12.92	**<0.001**	1.09 (1.05 - 1.14)
Gender (M/F)*	10/35	16/29	26/64	0.163	
Height (m)	1.60 ± 0.10	1.60 ± 0.10	1.60 ± 0.10	0.940	0.85 (0.01 – 58.78)
Weight (kg)	67.34 ± 14.54	72.42 ± 12.76	69.88 ± 13.84	0.086	1.03 (0.99 – 1.06)
BMI (kg/m^2)	26.14 ± 5.01	28.35 ± 4.99	27.24 ± 5.10	**0.046**	1.10 (1.00 – 1.20)
Diabetic Profile					
FPG (mmol/L)	5.21 ± 0.41	9.25 ± 3.43	*5.70 (3.38)*	**<0.001**	8.69 (3.01 – 25.07)
HbA1c (%)	N/A	*7.40 (2.80)*			
Glucose control (good/poor)	N/A	11/34			
Years diagnosed	N/A	*3.00 (8.50)*			
Liver Function Test					
TP (g/L)	72.76 ± 4.02	73.36 ± 3.37	73.06 ± 3.70	0.441	1.05 (0.93 – 1.17)
Albumin (g/L)	41.58 ± 2.90	42.13 ± 2.83	41.86 ± 2.86	0.357	1.07 (0.93 – 1.24)
ALP (U/L)	76.16 ± 19.62	84.67 ± 19.81	80.41 ± 20.07	**0.047**	1.02 (1.00 – 1.05)
ALT (U/L)	*16.00 (9.00)*	*22.00 (14.50)*	*19.00 (13.25)*	0.080	1.03 (0.99 – 1.06)
TB (µmol/L)	*10.00 (6.00)*	10.93 ± 3.84	*10.50 (5.25)*	0.549	0.98 (0.90 – 1.06)
Renal Profile					
Urea (mmol/L)	4.08 ± 1.27	*4.60 (2.35)*	*4.20 (1.95)*	**0.014**	1.46 (1.08 – 1.96)
Na (mmol/L)	137.84 ± 2.72	137.9 ± 2.43	137.67 ± 2.57	0.512	0.95 (0.80 – 1.12)
K (mmol/L)	4.03 ± 0.28	4.10 ± 0.41	4.06 ± 0.35	0.353	1.77 (0.53 – 5.93)
Cl (mmol/L)	102.71 ± 2.62	101.42 ± 2.75	102.07 ± 2.75	**0.030**	0.83 (0.71 – 0.98)
Cr (mmol/L)	64.78 ± 14.68	*65.00 (38.50)*	*64.50 (26.25)*	0.062	1.02 (0.99 – 1.04)
Fasting Serum Lipids					
TC (mmol/L)	5.64 ± 0.83	5.56 ± 1.31	5.60 ± 1.09	0.756	0.94 (0.64 – 1.38)
TG (mmol/L)	1.20 ± 0.54	*1.50 (1.15)*	*1.30 (0.73)*	**0.002**	3.75 (1.64 – 8.58)
HDL (mmol/L)	1.37 ± 0.33	1.25 ± 0.29	1.31 ± 0.31	0.076	0.28 (0.07 – 1.15)
LDL (mmol/L)	3.72 ± 0.75	3.42 ± 1.33	3.57 ± 1.08	0.201	0.77 (0.52 – 1.15)
HDL/TC (%)	24.84 ± 7.35	22.96 ± 6.46	23.90 ± 6.94	0.203	0.96 (0.90 – 1.02)

Data are expressed as mean ± SD for normal data or median [interquartile range (IQR)] for skewed data, analysed using Simple Logistic Regression for continuous variables and *Chi-square test for categorical variables in SPSS, p-value < 0.05 is significant.

M/F, male/female; BMI, body mass index; FPG, fasting plasma glucose; HbA1c, hemoglobin A1c; TP, total protein; ALP, Alkaline phosphatase; ALT, Alanine aminotransaminase; TB, total bilirubin; Na, sodium; K, potassium; Cl, chloride; Cr, creatinine; TC, total cholesterol; TG, triglyceride; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

In Table 4.2, the clinical characteristics of all study participants were compared by ethnicity. The Chinese participants were older (56.00 ± 12.12 years) than the Indian participants (48.97 ± 12.69 years) in this study. Meanwhile, the mean weight (75.30 ± 12.91 kg) and BMI (29.06 ± 6.06 kg/m^2) of the Indian participants were significantly higher when compared with the Chinese participants (63.57 ± 11.55 kg and 24.94 ± 3.99 kg/m^2, respectively).

In LFT, the levels of albumin and ALP were found to be significantly different across the ethnic groups. The albumin level was higher in Chinese (42.47 ± 2.61 g/L) when compared to Indians (40.77 ± 2.60 g/L). The Indian participants had significantly higher ALP (89.20 ± 22.97 U/L) than the Malay participants (75.00 ± 19.25 U/L). Meanwhile, in RP, none of the parameters were significantly different. In the FSL test, HDL, LDL and HDL/TC ratio was found to be significantly different. The Chinese participants had significantly increased HDL (1.44 ± 0.33 mmol/L) and decreased LDL (3.10 ± 0.92 mmol/L) when compared to the Indians (1.17 ± 0.22 mmol/L and 4.01 ± 0.96 mmol/L, respectively). Meanwhile, the HDL/TC ratio was significantly increased in the Chinese (28.13 ± 7.78%) in comparison to both Malays (23.41 ± 5.75%) and Indians (20.16 ± 4.61%).

Table 4.2 The Characteristics of All Study Participants by Ethnicity, n=90.

Parameters	Malay Mean ± SD or *Median (IQR)* n=30	Chinese Mean ± SD or *Median (IQR)* n=30	Indian Mean ± SD or *Median (IQR)* n=30	*p*-value
Age (years)	48.20 ± 12.89	56.00 ± 12.12	48.97 ± 12.69	**0.033**
Gender (M/F)*	10/20	6/24	10/20	0.421
Height (m)	1.59 ± 0.10	1.60 ± 0.10	1.62 ± 0.11	0.581
Weight (kg)	70.77 ± 14.67	63.57 ± 11.55	75.30 ± 12.91	**0.003**
BMI (kg/m^2)	27.72 ± 4.23	24.94 ± 3.99	29.06 ± 6.06	**0.005**
Diabetic Profile				
FPG (mmol/L)	*5.50 (4.05)*	*5.70 (2.43)*	*5.85 (3.60)*	0.794
HbA1c (%)	*8.30 (3.60)*	*6.70 (2.00)*	*7.20 (2.60)*	0.197
Glucose control (good/poor)*	3/12	6/9	2/13	0.209
Years diagnosed	*3.00 (9.00)*	*3.00 (7.00)*	*6.00 (10.00)*	0.636
Liver function test				
Total protein (g/L)	73.43 ± 4.29	72.73 ± 4.02	73.00 ± 2.69	0.765
Albumin (g/L)	42.33 ± 3.12	42.47 ± 2.61	40.77 ± 2.60	**0.036**
ALP (U/L)	75.00 ± 19.25	77.03 ±14.70	89.20 ± 22.97	**0.011**
ALT (U/L)	*21.00 (21.50)*	*17.50 (11.75)*	*16.00 (8.50)*	0.078
TB (µmol/L)	*10.00 (3.29)*	*11.50 (7.00)*	*9.50 (6.01)*	0.164
Renal profile				
Urea (mmol/L)	*4.15 (1.58)*	*4.50 (1.73)*	*4.05 (2.45)*	0.548
Sodium (mmol/L)	137.70 ± 2.32	138.23 ± 2.31	137.07 ± 2.97	0.214
Potassium (mmol/L)	4.07 ± 0.28	3.96 ± 0.39	4.15 ± 0.36	0.126
Chloride (mmol/L)	101.70 ± 2.65	102.70 ± 2.31	101.80 ± 3.19	0.303
Creatinine (mmol/L)	*65.00 (26.75)*	*58.50 (25.50)*	*65.50 (29.00)*	0.540
Fasting serum lipids				
TC (mmol/L)	5.67 ± 1.24	5.27 ± 0.90	5.86 ± 1.05	0.105
TG (mmol/L)	*1.30 (0.85)*	*1.30 (0.93)*	*1.50 (0.45)*	0.619
HDL (mmol/L)	1.31 ± 0.31	1.44 ± 0.33	1.17 ± 0.22	**0.002**
LDL (mmol/L)	3.60 ± 1.19	3.10 ± 0.92	4.01 ± 0.96	**0.004**
HDL/TC (%)	23.41 ± 5.75	28.13 ± 7.78	20.16 ± 4.61	**<0.001**

Data are expressed as mean ± SD for normal data or median [interquartile range (IQR)] for skewed data, analysed continuous variables using Anova [post-hoc test (Bonferroni)] or Kruskal-Wallis test (pair-wise comparison with Mann-Whitney U test) and *Chi-square test for categorical variables in SPSS, p-value < 0.05 is significant.

M/F, male/female; BMI, body mass index; FPG, fasting plasma glucose; ALP, Alkaline phosphatase; ALT, Alanine aminotransaminase; TB, total bilirubin; TC, total cholesterol; TG, triglyceride, HDL, high-density lipoprotein; LDL, low-density lipoprotein.

Age: Chinese vs Indian *p*=0.033

Weight: Chinese vs Indian *p*=0.002

BMI: Chinese vs Indian *p*=0.004

Albumin: Chinese vs Indian *p*=0.040

ALP: Malay vs Indian *p*=0.016

HDL: Chinese vs Indian *p*=0.001

LDL: Chinese vs Indian *p*=0.003

HDL/TC: Malay vs Chinese *p*=0.012, Chinese vs Indian *p*=<0.001

In Table 4.3, the clinical characteristics of nonDM were compared by ethnicity. The Chinese nonDM (52.87 ± 12.64 years) were older than the Indian nonDM (39.00 ± 8.59 years). Conversely, the Indian nonDM (75.80 ± 16.25 kg) had a higher weight in comparison to Chinese nonDM (60.13 ± 11.10 kg). In LFT, Malay nonDM had significantly increased albumin (43.00 ± 2.59 g/L) and significantly decreased ALP (66.53 ± 18.36 U/L) in comparison to Indian nonDM (40.27 ± 3.31 g/L and 88.47 ± 20.75 U/L, respectively). Meanwhile, there were no significant differences in the diabetic profile and RP test in the nonDM group. In FSL test, Chinese nonDM had significantly increased HDL (1.59 ± 0.30 mmol/L) and HDL/TC ratio (29.97 ± 8.48%) in comparison to both nonDM of Malay (1.42 ± 0.30 mmol/L and 25.32 ± 5.25%) and Indian ethnic groups (1.09 ± 0.13 mmol/L and 19.23 ± 2.85%), respectively.

In Table 4.4, the clinical characteristics of T2DM were compared by ethnicity. The Indian T2DM had a significantly higher BMI (30.01 ± 5.48 kg/m^2) than the T2DM from Chinese ethnic group (25.58 ± 3.43 kg/m^2). In LFT, Malay T2DM had significantly increased ALT [32.00 (24.00) U/L] than both T2DM of Chinese [19.00 (12.00) U/L] and Indian [20.00 (7.00) U/L] ethnic groups. Meanwhile, there were no significant differences found in both diabetic profile and RP test. In FSL test, the Indian T2DM had significantly increased LDL (4.05 ± 1.31 mmol/L) than the Chinese T2DM (2.79 ± 0.85 mmol/L). On the other hand, Indian T2DM had a significantly decreased HDL/TC ratio (21.09 ± 5.83%) in comparison to both T2DM of Malay (21.51 ± 5.76%) and Chinese (26.29 ± 6.79%) ethnic groups, respectively.

Table 4.3 The Characteristics of NonDM Group by Ethnicity, n=45.

Parameters	Malay Mean ± SD or *Median (IQR)* n=15	Chinese Mean ± SD or *Median (IQR)* n=15	Indian Mean ± SD or *Median (IQR)* n=15	*p*-value
Age (years)	43.33 ± 13.22	52.87 ± 12.64	39.00 ± 8.59	**0.007**
Gender (M/F)*	5/10	1/14	4/11	0.188
Height (m)	1.59 ± 0.11	1.57 ± 0.07	1.65 ± 0.10	0.102
Weight (kg)	66.10 ± 11.97	60.13 ± 11.10	75.80 ± 16.25	**0.009**
BMI (kg/m^2)	25.99 ± 2.60	24.31 ± 4.50	28.12 ± 6.65	0.113
Diabetic Profile				
FPG (mmol/L)	5.12 ± 0.41	5.21 ± 0.42	5.31 ± 0.41	0.469
Liver function test				
Total protein (g/L)	73.47 ± 5.42	72.40 ± 3.20	72.40 ± 3.22	0.713
Albumin (g/L)	43.00 ± 2.59	41.47 ± 2.17	40.27 ± 3.31	**0.031**
ALP (U/L)	66.53 ± 18.36	73.47 ± 13.17	88.47 ± 20.75	**0.005**
ALT (U/L)	*18.00 (9.00)*	*15.00 (10.00)*	*16.00 (13.00)*	0.612
TB (µmol/L)	*10.00 (5.00)*	*13.00 (7.00)*	*8.00 (5.00)*	0.113
Renal profile				
Urea (mmol/L)	4.31 ± 1.36	4.11 ± 1.25	3.83 ± 1.25	0.586
Sodium (mmol/L)	138.53 ± 1.88	138.40 ± 2.38	136.60 ± 3.40	0..092
Potassium (mmol/L)	3.99 ± 0.28	3.97 ± 0.24	4.13 ± 0.31	0.238
Chloride (mmol/L)	102.60 ± 2.20	103.47 ± 2.26	102.07 ± 3.24	0.343
Creatinine (mmol/L)	70.40 ± 16.25	60.93 ± 11.50	63.00 ± 15.10	0.180
Fasting serum lipids				
TC (mmol/L)	5.69 ± 0.97	5.50 ± 0.98	5.72 ± 0.49	0.745
TG (mmol/L)	1.07 ± 0.48	1.12 ± 0.60	1.41 ± 0.51	0.171
HDL (mmol/L)	1.42 ± 0.30	1.59 ± 0.30	1.09 ± 0.13	**<0.001**
LDL (mmol/L)	3.77 ± 0.77	3.41 ± 0.90	3.97 ± 0.43	0.114
HDL/TC (%)	25.32 ± 5.25	29.97 ± 8.48	19.23 ± 2.85	**<0.001**

Data are expressed as mean ± SD for normal data or median [interquartile range (IQR)] for skewed data, analysed continuous variables using Anova [post-hoc test (Bonferroni)] or Kruskal-Wallis test and *Chi-square test for categorical variables in SPSS, p-value < 0.05 is significant.

M/F, male/female; BMI, body mass index; FPG, fasting plasma glucose; ALP, Alkaline phosphatase; ALT, Alanine aminotransaminase; TB, total bilirubin; TC, total cholesterol; TG, triglyceride, HDL, high-density lipoprotein; LDL, low-density lipoprotein.

Age: Chinese vs Indian p=0.007

Weight: Chinese vs Indian p=0.007

Albumin: Malay vs Indian p=0.027

ALP: Malay vs Indian p=0.005

HDL: Malay vs Indian p=0.004, Chinese vs Indian p= <0.001

HDL/TC: Malay vs Indian p=0.024, Chinese vs Indian p=<0.001

Table 4.4 The Characteristics of T2DM Group by Ethnicity, n=45.

Parameters	Malay Mean ± SD or *Median (IQR)* n=15	Chinese Mean ± SD or *Median (IQR)* n=15	Indian Mean ± SD or *Median (IQR)* n=15	*p*-value
Age (years)	*54.00 (19.00)*	*60.00 (15.00)*	*61.00 (8.00)*	0.283
Gender (M/F)*	5/10	5/10	6/9	0.908
Height (m)	1.60 ± 0.08	1.62 ± 0.11	1.60 ± 0.11	0.713
Weight (kg)	75.45 ± 15.99	67.00 ± 11.30	74.80 ± 8.99	0.130
BMI (kg/m^2)	29.45 ± 4.89	25.58 ± 3.43	30.01 ± 5.48	**0.026**
Diabetic Profile				
FPG (mmol/L)	9.85 ± 3.67	9.03 ± 4.14	8.87 ± 2.42	0.715
HbA1c (%)	*8.30 (3.60)*	*6.70 (2.00)*	*7.20 (2.60)*	0.197
Glucose control (good/poor)*	3/12	6/9	2/13	0.209
Years diagnosed	*3.00 (9.00)*	*3.00 (7.00)*	*6.00 (10.00)*	0.636
Liver function test				
Total protein (g/L)	73.40 ± 2.95	73.07 ± 4.79	73.60 ± 1.96	0.912
Albumin (g/L)	41.67 ± 3.54	43.47 ± 2.70	41.27 ± 1.58	0.074
ALP (U/L)	83.47 ± 16.65	80.60 ± 15.71	89.93 ± 25.71	0.427
ALT (U/L)	*32.00 (24.00)*	*19.00 (12.00)*	*20.00 (7.00)*	**0.007**
TB (µmol/L)	10.73 ± 3.37	11.20 ± 3.61	10.86 ± 4.66	0.945
Renal profile				
Urea (mmol/L)	*4.2 (1.60)*	*4.70 (2.40)*	*4.90 (3.40)*	0.406
Sodium (mmol/L)	136.87 ± 2.47	138.07 ± 2.31	137.53 ± 2.50	0.408
Potassium (mmol/L)	4.16 ± 0.27	3.96 ± 0.50	4.17 ± 0.42	0.300
Chloride (mmol/L)	100.80 ± 2.83	101.93 ± 2.15	101.53 ± 3.23	0.530
Creatinine (mmol/L)	*64.00 (27.00)*	*73.00 (36.00)*	*68.00 (38.00)*	0.642
Fasting serum lipids				
TC (mmol/L)	5.65 ± 1.50	5.05 ± 0.79	6.00 ± 1.42	0.132
TG (mmol/L)	*1.60 (1.50)*	*1.40 (1.20)*	*1.50 (0.70)*	0.747
HDL (mmol/L)	1.21 ± 0.28	1.30 ± 0.31	1.24 ± 0.27	0.674
LDL (mmol/L)	3.43 ± 1.51	2.79 ± 0.85	4.05 ± 1.31	**0.032**
HDL/TC (%)	21.51 ± 5.76	26.29 ± 6.79	21.09 ± 5.83	**0.046**

Data are expressed as mean ± SD for normal data or median [interquartile range (IQR)] for skewed data, analysed continuous variables using Anova [post-hoc test (Bonferroni)] or Kruskal-Wallis test (Mann-Whitney U test for pairwise comparisons) and *Chi-square test for categorical variables in SPSS, p-value < 0.05 is significant.

M/F, male/female; BMI, body mass index; FPG, fasting plasma glucose; HbA1c, Hemoglobin A1c; ALP, Alkaline phosphatase; ALT, Alanine aminotransaminase; TB, total bilirubin; TC, total cholesterol; TG, triglyceride, HDL, high-density lipoprotein; LDL, low-density lipoprotein.

BMI: Chinese vs Indian *p*=0.039

ALT: Malay vs Chinese p=0.036, Malay vs Indian *p*=0.002

LDL: Chinese vs Indian *p*= 0.027

HDL/TC: Malay vs Indian *p*=0.024, Chinese vs Indian *p*=<0.001

4.3 Genomic DNA Quantification

The gDNA extraction procedure was optimised for bead beating time and speed using two faecal samples. The concentration and quality of gDNA were determined using a spectrophotometer (Implen Nanophotometer, Germany) (Table 4.5) and agarose gel electrophoresis (Figure 4.1, A and B), upon which the bead beating time of 1 minute at 4800rpm was selected.

Overall, gDNA concentrations obtained from the 90 faecal samples (55.53 ± 64.25 ng/uL) were within the required range of more than 20 ng/uL. The purity of gDNA was determined using the ratio of absorbance at 260nm to absorbance at 280nm and 230nm. The absorbance reading of gDNA from the 90 faecal samples obtained at A_{260}/A_{280} ratio (1.90 ± 0.14) was within the required range (1.7 to 1.9). The absorbance reading of the gDNA obtained at the A_{260}/A_{230} ratio (1.11 ± 0.59) was lower than the required range (≥ 1.8) (Table 4.6).

Table 4.5 Assessment of gDNA Concentration and Purity Extracted for Optimisation of Bead Beating Time and Speed.

Well Number	BB Time	Speed (rpm)	Concentration (ng/uL)		DNA Purity (A260/280nm)		DNA Purity (A260/230nm)	
			A	B	A	B	A	B
WELL 0	DNA 1kb LADDER							
WELL 1	No BB	-	43	61	1.686	1.848	1.755	1.371
WELL 2	30s	2500	79	129	1.756	1.817	2.026	1.767
WELL 3	30s	4800	161	155	1.793	1.813	1.911	1.856
WELL 4	1m	2500	112	152	1.828	1.825	1.890	1.772
WELL 5	1m	4800	155	171	1.845	1.829	1.782	1.056
WELL 6	3m	2500	115	193	1.774	1.808	2.101	2.026
WELL 7	3m	4800	188	188	1.808	1.820	2.225	2.060

BB, bead beating time.

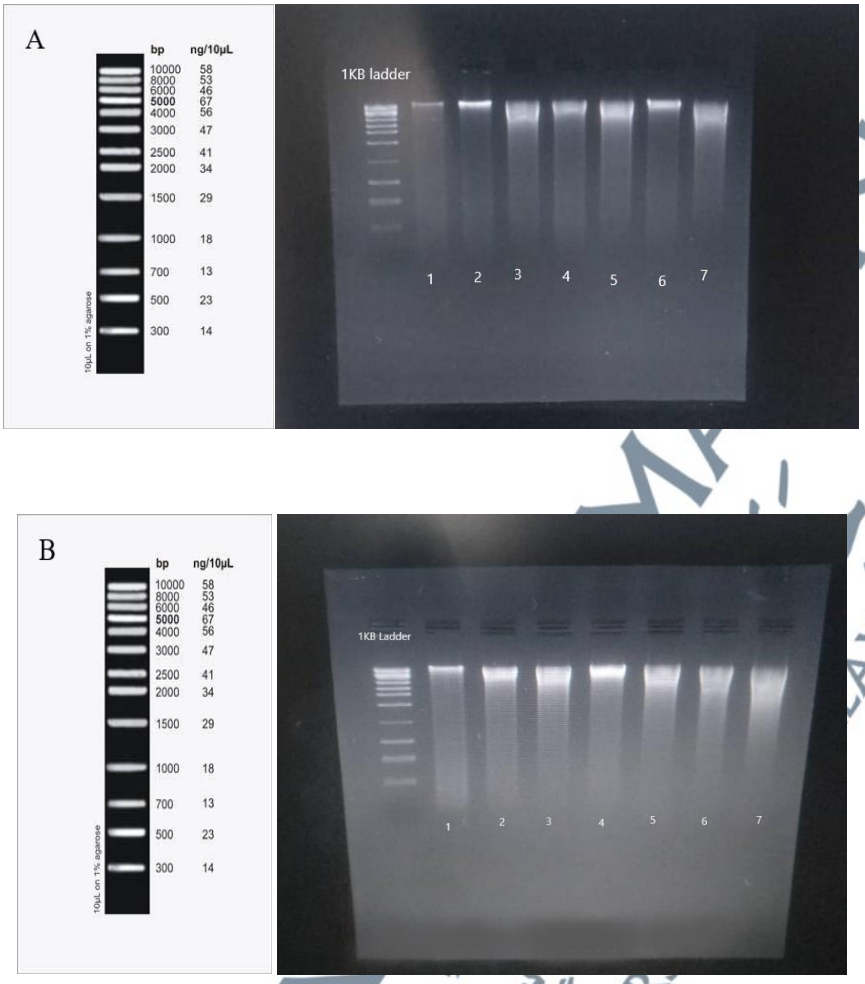


Figure 4.1 Agarose gel Showing gDNA Extracted for Optimisation of Bead Beating Time and Speed

Table 4.6 The gDNA Analysis of All Faecal Samples, n=90.

gDNA Analysis	Mean ± SD n=90 (Required range)
Concentration (ng/uL)	55.53 ± 64.25 (≥ 20)
Purity (A_{260}/A_{280}, nm)	1.90 ± 0.14 (1.70 - 1.90)
Purity (A_{260}/A_{230}, nm)	1.11 ± 0.59 (≤ 1.80)

4.4 Gut Microbiota Analysis

The 90 samples of purified gDNA were sequenced targeting the V3-V4 region of the 16S rDNA using the Illumina MiSeq platform. This gave an output of 1,757,079 raw sequences with 3, 132 ASVs identified using DADA2 across all samples. These ASVs were further used for diversity analysis and taxonomy classification.

4.4.1 Diversity Analysis

The samples were rarefied to capture an even diversity among the study groups. Based on the minimum ASV sequence that could retain most of the samples, the suitable sampling depth was chosen. This set the sampling depth at 13,900 reads/sample which excluded one sample which had the lowest sequence count of 3649 reads/sample (Figure 4.2). This sample was excluded for diversity analysis.

The diversity within samples or alpha diversity was assessed based on species richness (presence or absence of species) using Observed ASV and Chao1 indexes. Pielou's evenness measures species abundance (frequency of species occurrence) while Shannon index measures both richness and abundance. When comparing alpha diversity difference between the three ethnic groups, among the nonDM healthy controls, only Pielou's evenness index showed significant differences, whereby it was significantly lower in the Chinese when compared to the Malay and Indian ethnic groups (p=0.015) (Table 4.7

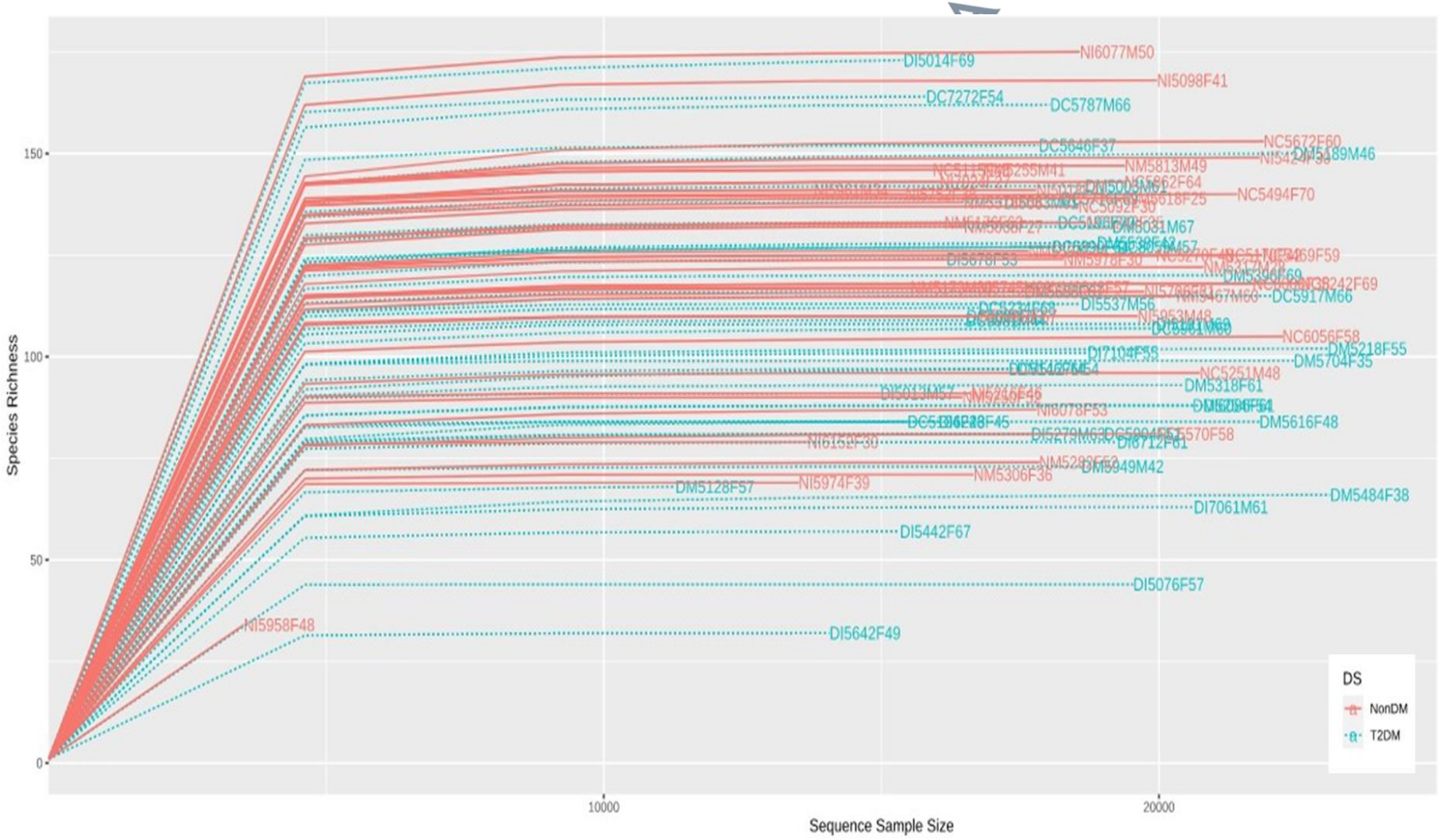


Figure 4.2 Rarefaction Curve in NonDM and T2DM Groups, n=90.

Meanwhile, when comparing nonDM and T2DM groups, all indices used showed that alpha diversity was significantly higher in nonDM in comparison to T2DM (*p*-value < 0.05) (Table 4.8). However, when comparing nonDMs and T2DMs by ethnicity, the alpha diversity of nonDMs was significantly higher than T2DMs in only the Malay (by Pielou's evenness and Shannon indices) and Indian (by all indices) ethnic groups (*p*-value < 0.05) (Table 4.9).

Table 4.7 Bacterial Alpha Diversity in NonDM, n=44.

Diversity Index	NonDM			*p*-value
	Malay n=15	Chinese n=15	Indian n=14	
Observed ASV	138.27 ± 30.07	138.00 ± 22.12	142.71 ± 40.39	0.919
Chao1	138.53 ± 29.99	139.27 ± 22.16	142.93 ± 40.37	0.923
Pielou's evenness[a]	0.83 (0.05)	0.79 (0.02)	0.82 (0.03)	**0.015**
Shannon	5.76 ± 0.46	5.62 ± 0.38	5.83 ± 0.50	0.463

Data are expressed as mean ± SD for normal data or median [interquartile range (IQR)] for skewed data, analysed using Anova [post-hoc test (Bonferroni)] or [a]Kruskal-Wallis test (Mann-Whitney U test for pairwise comparisons) for continuous variables in SPSS, p-value < 0.05 is significant.

Pielou's evenness: Malay vs Chinese *p*=0.037, Chinese vs Indian, *p*=0.010.

Table 4.8 Bacterial Alpha Diversity in NonDM and T2DM, n=89.

Diversity Index	NonDM n=44	T2DM n=45	*p*-value
Observed ASV	139.86 ± 30.84	124.40 ± 36.15	**0.033**
Chao1	140.18 ± 30.81	124.72 ± 36.41	**0.033**
Pielou's evenness[a]	0.82 (0.05)	0.78 (0.07)	**0.001**
Shannon	5.74 ± 0.45	5.31 ± 0.79	**0.002**

Data are expressed as mean ± SD for normal data or median [interquartile range (IQR)] for skewed data, analysed using [a]Student's *t*-test or [a]Mann-Whitney U test for continuous variables in SPSS, p-value < 0.05 is significant.

Table 4.9 Pairwise Comparison of Bacterial Alpha Diversity in NonDM and T2DM by Ethnicity, n=89.

Diversity Index	Malay		*p*-value	Chinese		*p*-value	Indian		*p*-value
	NonDM n=15	T2DM n=15		NonDM n=15	T2DM n=15		NonDM n=14	T2DM n=15	
Observed ASV	138.27 ± 30.07	120.47 ± 28.16	0.105	138.00 ± 22.12	142.67 ± 32.75	0.708	142.71 ± 40.39	110.07 ± 40.58	**0.039**
Chao1	138.53 ± 29.99	120.80 ± 28.16	0.106	139.27 ± 22.16	142.81 ± 32.67	0.730	142.93 ± 40.37	110.53 ± 41.51	**0.043**
Pielou's evenness[a]	0.83 (0.05)	0.77 (0.07)	**0.004**	0.79 (0.02)	0.80 (0.06)	0.412	0.82 (0.03)	0.75 (0.10)	**0.002**
Shannon	5.76 ± 0.46	5.18 ± 0.90	**0.033**	5.62 ± 0.38	5.68 ± 0.48	0.717	5.83 ± 0.50	5.07 ± 0.83	**0.007**

Data are expressed as mean ± SD for normal data or median [interquartile range (IQR)] for skewed data, analysed using [a]Student's t-test or [a]Mann-Whitney U test for continuous variables in SPSS, p-value < 0.05 is significant.

Table 4.10 Beta Diversity of NonDM and T2DM Groups, n=89.

Beta Diversity Indexes	Ethnicity (Malay vs Chinese vs Indian) NonDM n=44		DM status (nonDM vs T2DM) n=89		DM status x Ethnicity					
					Malay (nonDM vs T2DM) n=30		Chinese (nonDM vs T2DM) n=30		Indian (nonDM vs T2DM) n=29	
	F	*p*-value	F	*p*-value	F	*p*-value	F	*p*-value	F	*p*-value
Unweighted UniFrac	1.214	0.138	2.040	**0.006**	1.458	0.061	1.175	0.201	1.617	**0.046**
Weighted UniFrac	1.557	0.123	1.453	0.227	0.559	0.724	1.305	0.242	1.367	0.216
Bray Curtis	1.177	0.077	2.360	**0.001**	1.397	**0.042**	1.132	0.232	0.710	**0.003**

The number of permutations=999; F, pseudo-F, analysed using permutational MANOVA (PERMANOVA) test in QIIME2, p-value < 0.05 is significant.

Beta diversity was measured using Bray Curtis dissimilarity matrix as well as the phylogenetically aware unweighted and weighted UniFrac distance metrics. The significant difference in beta diversity values between the groups was analysed using the PERMANOVA test (Table 4.10). When analysing gut community structure between nonDM of Malay, Chinese and Indian ethnicity, all measures were found to be not significant. This indicates that no difference in community structure was found based on ethnicity (Figure 4.3).

On the other hand, the unweighted Unifrac that analyses community structure based on absence/presence of ASVs was significantly different between T2DM and nonDM groups (*p*-value = 0.006) (Table 4.10). This indicates that a different microbial community is present in nonDM and T2DM groups. When comparing beta diversity by ethnic groups, this significant difference between T2DM and nonDM was only noted in the Indian ethnicity (*p*-value = 0.046) (Table 4.10). However, the weighted Unifrac distance that accounts for differences in community structure based on relative abundance data was found to be not significant across the study groups (Table 4.10). This is probably due to high number of samples having low-abundant ASVs of certain bacterial taxa.

Meanwhile, the Bray-Curtis dissimilarity matrix measures the abundance of dissimilar microbial species found between communities. Overall, the PERMANOVA analysis of the Bray-Curtis dissimilarity matrix was significantly different between nonDM and T2DM groups as well as in Malay and Indian ethnicity (*p*-value = 0.001, 0.042 and 0.003). This was reflected in the PCoA analysis where there was also a distinct clustering of microbial communities between nonDM and T2DM groups

(Figure 4.4) as well as between nonDM and T2DM of Malay and Indian ethnicity (Figure 4.5). This may indicate differences in community composition and/or abundance between nonDM and T2DM respectively.

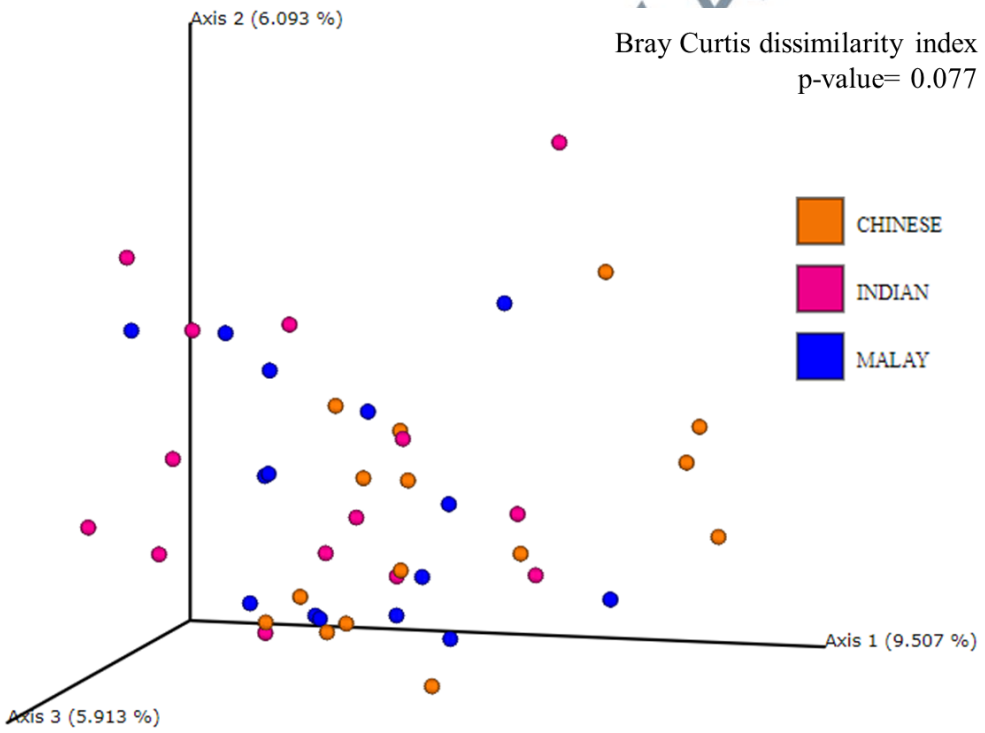


Figure 4.3 Principal Coordinate Analysis (PCoA) based on Bray-Curtis Dissimilarity Matrix between Microbial Communities of NonDM Participants by Ethnicity, n=44.

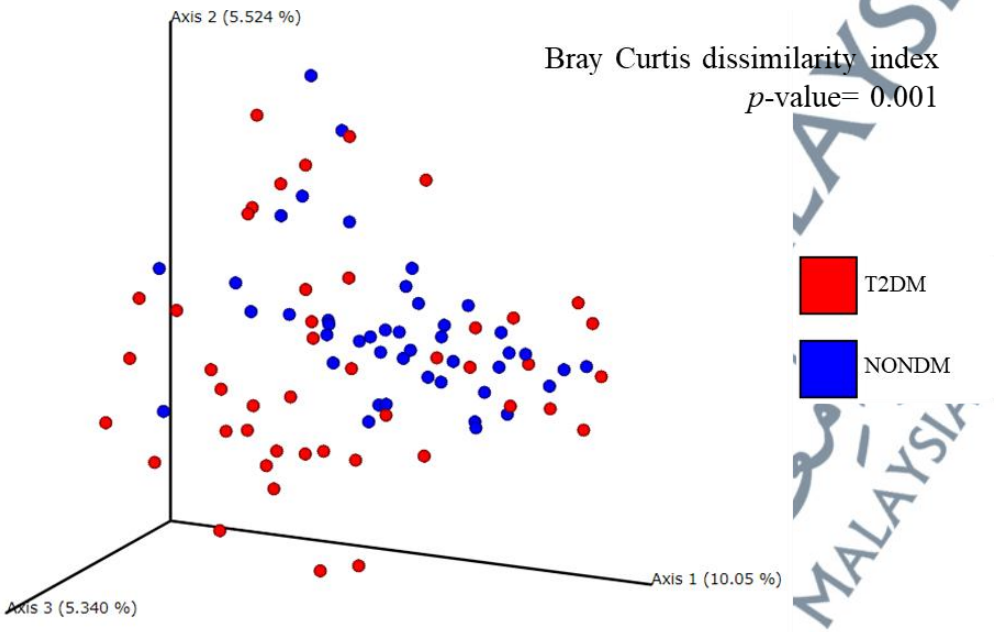


Figure 4.4 Principal Coordinate Analysis (PCoA) based on Bray-Curtis Dissimilarity Matrix between Microbial Communities from NonDM and T2DM Groups, n=89.

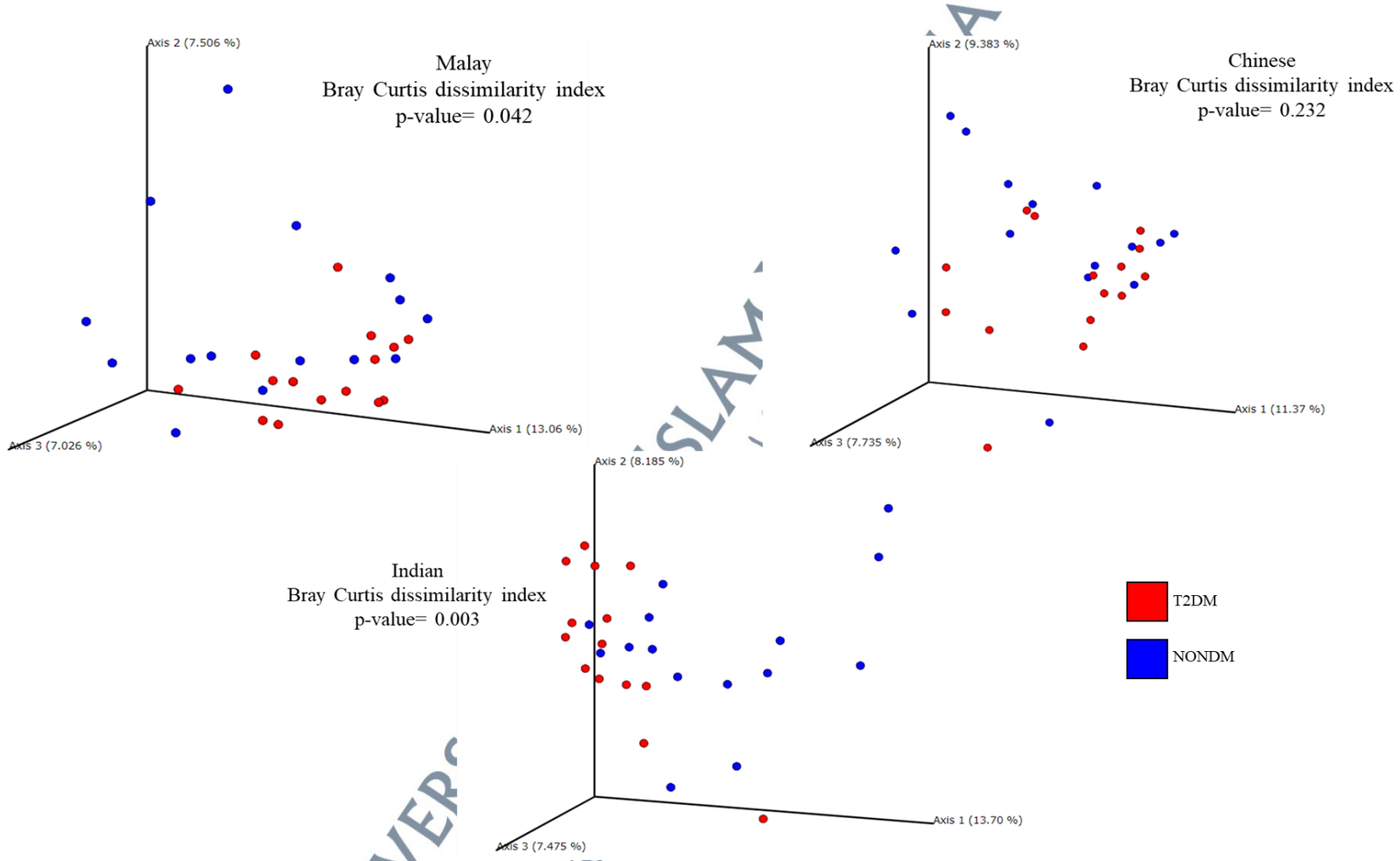


Figure 4.5 Principal Coordinate Analysis (PCoA) based on Bray-Curtis Dissimilarity Matrix between Microbial Communities from NonDM and T2DM Group of Similar Ethnicity, n=89.

4.4.2 Gut Microbiota Composition

In total, 12 phyla and 237 genera were identified across all 90 samples. Figure 4.6 shows the top nine phyla that were prevalent (present) in at least 10% abundance in all 90 samples i.e., *Firmicutes* (70.80%), *Bacteroidetes* (21.60%), *Proteobacteria* (4.52%), *Actinobacteria* (2.09%), *Verrucomicrobia* (0.43%), *Fusobacteria* (0.35%), *Desulfobacteria* (0.12%), *Cyanobacteria* (0.04%) and *Synergistetes* (0.03%). The top 20 genera with more than 1% abundance in all 90 samples were *Blautia* (12.91%), *Bacteroides* (11.73%), *Faecalibacterium* (10.69%), *Prevotella* (8.18%), *Megamonas* (4.71%), *Agathobacter* (3.75%), *Eubacterium* (3.54%), *Ruminococcus* (3.48%), *Escherichia-Shigella* (2.93%), *Subdoligranulum* (2.74%), *Roseburia* (2.52%), *Megasphaera* (2.49%), *Fusicatenibacter* (2.29%), *Lactobacillus* (2.20%), *Dorea* (1.88%), *Catenibacterium* (1.52%), *Anaerostipes* (1.48%), *Bifidobacterium* (1.41%), *Clostridium* (1.17%) and *Holdemanella* (1.06%) as depicted in Figure 4.7.

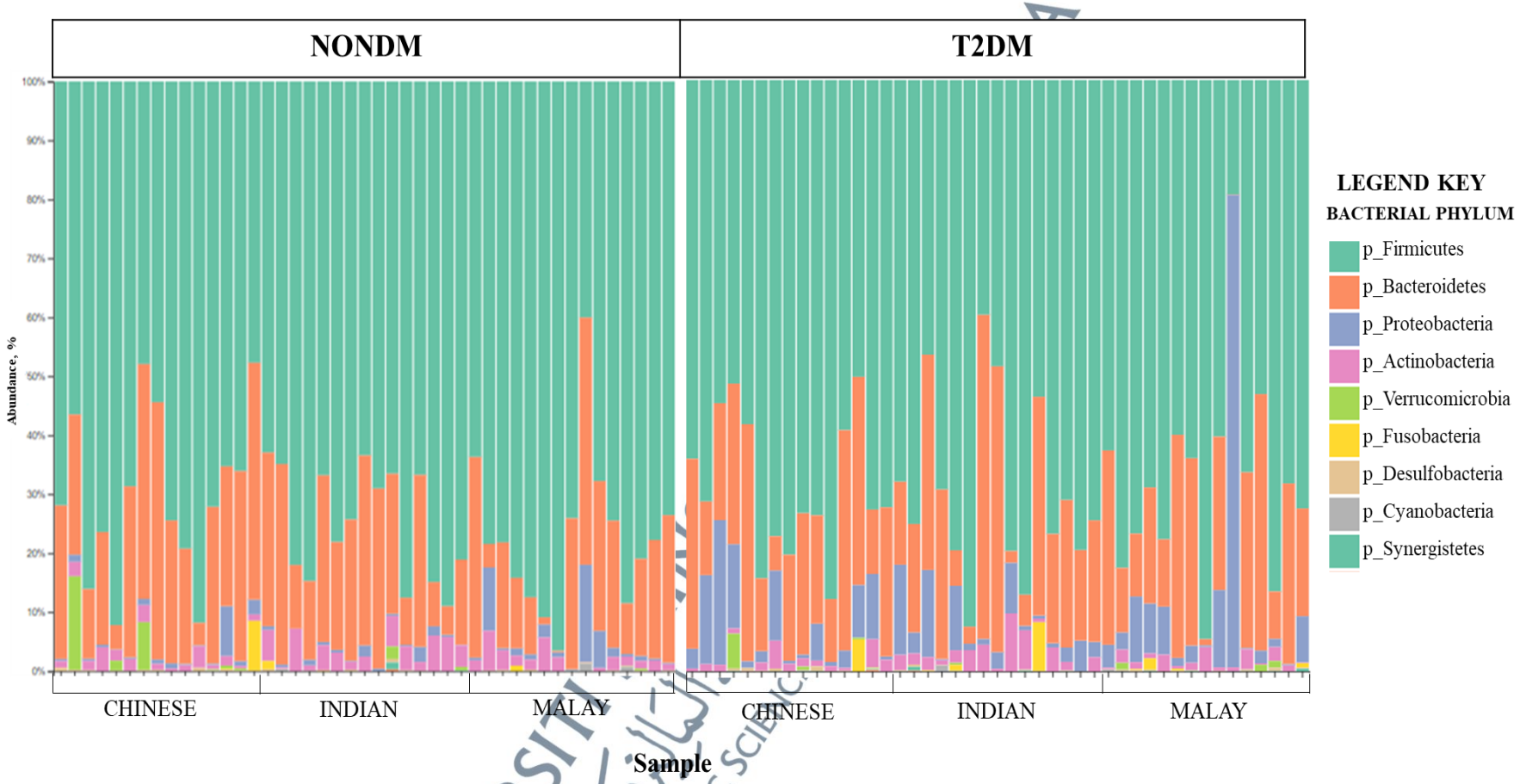


Figure 4.6 The Gut Microbiota Composition at Phylum Level showing the top nine phyla that were present in at least 10% abundance in all 90 samples.

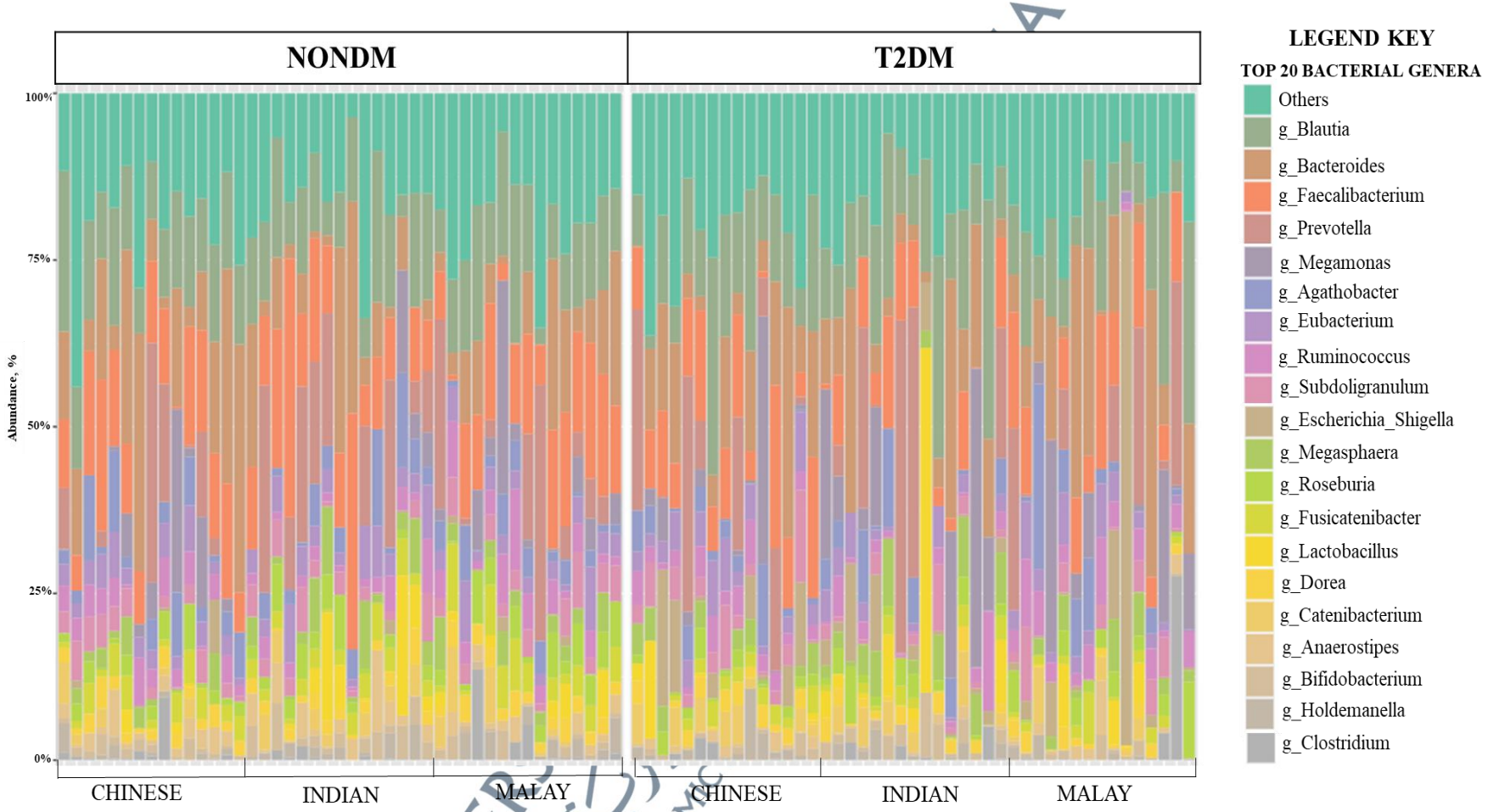


Figure 4.7 The Gut Microbiota Composition at Genus Level showing Top 20 Genera with more than 1% abundance in all 90 samples. All genera with less than 1% abundance were grouped as 'Others'.

4.4.2.1 Gut Microbiota Composition in NonDM

The dominant phyla in nonDM of all ethnic groups were *Firmicutes* (73.75%), *Bacteroidetes* (21.17%), *Actinobacteria* (2.35%) and *Proteobacteria* (1.61%). No significant differences in F/B ratio as well as bacterial phyla abundances were found between the ethnic groups in nonDM ($p < 0.05$) (Table 4.11 and Figure 4.8). Meanwhile, the top 20 genera found in the nonDM group (with abundance of more than 1%) were *Faecalibacterium* (12.77%), *Blautia* (12.35%), *Bacteroides* (11.23%), *Prevotella* (7.88%), *Agathobacter* (4.04%), *Eubacterium* (3.92%), *Ruminococcus* (3.57%), *Megamonas* (3.45%), *Fusicatenibacter* (3.11%), *Subdoligranulum* (3.04%), *Roseburia* (2.91%), *Megasphaera* (2.35%), *Dorea* (2.13%), *Anaerostipes* (2.12%), *Lactobacillus* (1.83%), *Bifidobacterium* (1.57%), *Catenibacterium* (1.37%), *Coprococcus* (1.25%), *Romboutsia* (1.22%) and *Holdemanella* (1.17%). The prevalence of genus *Megamonas* was higher in Malay nonDM (80.00%) when compared to Indian nonDM (33.33%) (p=0.037). On the other hand, the abundance of genus *Bacteroides* was higher in Chinese nonDM (16.12%) in comparison to Indian nonDM (7.79%) (p=0.020). However, the significance of these differences was lost after adjustment for multiple comparisons (FDR<0.05).

Table 4.11 Bacterial Phyla and Genera in NonDM Group Arranged by Abundance, n=45.

Taxa	Bacteria	Prevalence, %						Abundance, %					
		Malay n =15	Chinese n =15	Indian n =15	All nonDM n=45	p-value	FDR	Malay n =15	Chinese n =15	Indian n =15	All nonDM n=45	p-value	FDR
Phylum	*Firmicutes*	100.00	100.00	100.00	100.00	N/A	N/A	76.73	70.38	74.61	73.75	0.222	0.260
	Bacteroidetes	100.00	100.00	100.00	100.00	N/A	N/A	18.02	24.42	20.64	21.17	0.151	0.260
	Actinobacteria	93.33	100.00	93.33	95.60	0.600	0.600	2.03	1.71	3.47	2.35	0.214	0.260
	Proteobacteria	100.00	100.00	93.33	97.80	0.368	0.429	2.91	1.11	0.79	1.61	0.231	0.260
	Verrucomicrobia	20.00	40.00	13.33	24.40	0.217	0.379	0.03	1.59	0.17	0.64	0.153	0.260
	Fusobacteria	6.67	26.67	13.33	15.60	0.314	0.429	0.06	0.63	0.11	0.28	0.339	0.339
	Desulfobacteria	53.33	80.00	46.67	60.00	0.149	0.379	0.07	0.15	0.09	0.11	0.092	0.260
	Cyanobacteria	26.67	6.67	13.33	15.60	0.214	0.379	0.15	0.01	0.03	0.06	0.203	0.260
	Synergistetes	0.00	0.00	13.33	4.40	0.129	0.379	0.00	0.00	0.08	0.02	0.129	0.260
	F/B ratio [median (IQR)]	N/A	N/A	N/A	N/A	N/A	N/A	0.23 (0.25)	0.36 (0.24)	0.32 (0.31)	0.32 (0.30)	0.265	0.265
Genus	*Faecalibacterium* (F)	100.00	100.00	93.33	97.80	0.368	0.632	12.74	12.18	13.51	12.77	0.301	0.700
	Blautia (F)	100.00	100.00	100.00	100.00	N/A	N/A	12.14	13.18	11.57	12.35	0.163	0.619
	Bacteroides[a] (B)	100.00	100.00	100.00	100.00	N/A	N/A	8.93	16.12	7.79	11.23	**0.020**	0.313
	Prevotella (B)	66.67	46.67	60.00	57.80	0.536	0.767	7.30	5.49	11.45	7.88	0.486	0.781
	Agathobacter (F)	100.00	93.33	86.67	93.30	0.351	0.632	3.08	4.90	4.06	4.04	0.477	0.781
	Eubacterium (F)	100.00	100.00	100.00	100.00	N/A	N/A	3.66	3.58	4.64	3.92	0.612	0.816
	Ruminococcus (F)	100.00	100.00	100.00	100.00	N/A	N/A	4.40	3.63	2.57	3.57	0.484	0.781
	Megamonas[b] (F)	80.00	75.00	33.33	57.80	**0.037**	0.577	4.33	4.05	1.75	3.45	0.054	0.466
	Fusicatenibacter (F)	93.33	86.67	100.00	93.30	0.351	0.632	3.79	2.80	2.72	3.11	0.571	0.814
	Subdoligranulum (F)	86.67	86.67	93.33	88.90	0.803	0.874	3.44	2.68	3.04	3.04	0.730	0.880
	Roseburia (F)	100.00	93.33	86.67	93.30	0.351	0.632	3.73	2.05	3.06	2.91	0.309	0.700
	Megasphaera (F)	73.33	66.67	93.33	77.80	0.195	0.632	2.36	1.48	3.39	2.35	0.315	0.700
	Dorea (F)	100.00	100.00	93.33	97.80	0.368	0.632	2.46	2.08	1.82	2.13	0.082	0.548
	Anaerostipes (F)	100.00	100.00	100.00	100.00	N/A	N/A	1.52	2.15	2.76	2.12	0.572	0.814
	Lactobacillus (F)	73.33	46.67	66.67	62.20	0.301	0.632	1.37	0.43	4.05	1.83	0.187	0.649
	Bifidobacterium (A)	80.00	100.00	93.33	91.10	0.153	0.632	1.10	1.21	2.54	1.57	0.098	0.575
	Catenibacterium (F)	40.00	40.00	46.67	42.20	0.915	0.932	1.54	0.94	1.71	1.37	0.817	0.911
	Coprococcus (F)	86.67	73.33	86.67	82.20	0.552	0.767	1.59	1.00	1.18	1.25	0.362	0.700
	Romboutsia (F)	93.33	80.00	86.67	86.70	0.569	0.767	2.16	0.76	0.75	1.22	0.151	0.600
	Holdemanellla (F)	66.67	46.67	60.00	57.80	0.536	0.767	1.63	0.71	1.23	1.17	0.429	0.745

(A), Phylum Actinobacteria; (B), Phylum Bacteroidetes; (F), Phylum Firmicutes. Analysed using Kruskal-Wallis Test (pairwise comparison using Mann-Whitney U test) in MicrobiomeAnalyst, *p*-value and FDR of <0.05 is considered significant difference between nonDM and T2DM. Pairwise comparisons: [a]Chinese vs Indian p=0.010, FDR=0.326; [b]Malay vs Indian p=0.012, FDR=0.534.

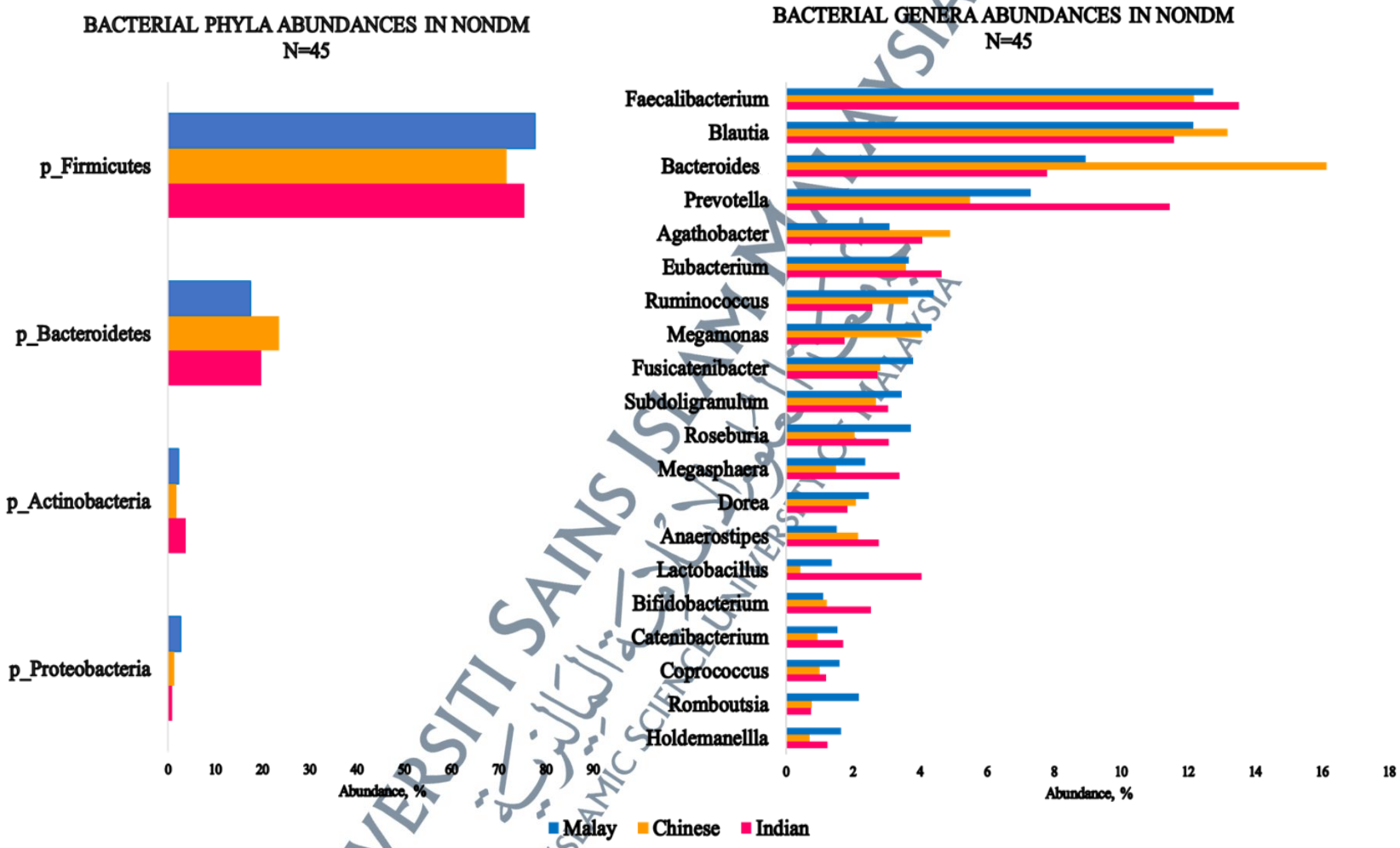


Figure 4.8 Gut Microbiota Abundances in (A) Phylum and (B) Genus Levels in NonDM Group by Ethnicity, n=45. These four phyla account for 99% of all phyla present and the top 20 genera account for 83% of all genera present.

4.4.2.2 Gut Microbiota Composition in T2DM when compared to NonDM by Ethnicity

When comparing T2DM participants with nonDM participants, abundance of only two phyla were found to be significantly different while no significant difference was noted in the F/B ratio. The abundance of phylum *Proteobacteria* was significantly higher (7.32%) in T2DM when compared to nonDM (1.61%) (*p*-value & FDR = <0.001) (Table 4.12 and Figure 4.9). This was reflected within each ethnic group whereby the abundance of *Proteobacteria* was higher in T2DM Malays (9.52%), Chinese (7.23%) and Indians (4.83%) in comparison to the respective nonDM groups (2.91%, 1.11% and 0.79%) (*p*-value = 0.045, 0.001 and 0.002), respectively (Table 4.13 and Figure 4.10).

The prevalence of phylum *Synergistetes* was significantly higher in T2DM (35.56%) when compared to nonDM (4.44%) (p-value=<0.001, FDR=0.002). Although showing low detection levels, the relative abundance of this phylum was significantly higher in T2DM (0.04%) when compared to nonDM (0.02%; p-value=<0.001, FDR= 0.003). This was reflected in the Malay and Chinese ethnic groups whereby the abundance of this phylum was higher in T2DM Malay (0.05%) and Chinese (0.03%), in comparison to nonDM Malay (0.00%) and Chinese (0.00%).

However, after adjustment for multiple corrections, only the difference in abundance of phylum *Proteobacteria* in all participants and specifically among Chinese and Indian ethnicities remained significant (FDR < 0.05).

Table 4.12 Bacterial Phyla and Genera in NonDM and T2DM Arranged by Abundance in All Samples, n=90.

Taxa	Bacteria	Prevalence, %					Abundance, %				
		NonDM n=45	T2DM n=45	All study participants, n=90	*p*-value	FDR	NonDM n=45	T2DM n=45	All study participants, n=90	*p*-value	FDR
Phylum	*Firmicutes*	100.00	100.00	100.00	N/A	N/A	73.75	67.97	70.80	0.212	0.382
	Bacteroidetes	100.00	100.00	100.00	N/A	N/A	21.17	22.02	21.60	0.756	0.756
	Proteobacteria	97.78	100.00	98.89	0.328	0.525	1.61	7.32	4.52	**<0.001**	**<0.001**
	Actinobacteria	95.56	93.33	94.44	0.655	0.655	2.35	1.83	2.09	0.083	0.235
	Verrucomicrobia	24.44	17.78	21.11	0.444	0.525	0.64	0.23	0.43	0.455	0.512
	Fusobacteria	15.56	22.22	18.89	0.425	0.525	0.28	0.42	0.35	0.358	0.460
	Desulfobacteria	60.00	68.89	64.44	0.384	0.525	0.11	0.13	0.12	0.552	0.610
	Cyanobacteria	15.56	8.89	12.22	0.466	0.525	0.06	0.03	0.04	0.320	0.460
	Synergistetes	4.44	35.56	20.00	**<0.001**	**0.002**	0.02	0.04	0.03	**<0.001**	**<0.001**
	F/B ratio [median (IQR)]	N/A	N/A	N/A	N/A	N/A	*0.32 (0.30)*	*0.28 (0.37)*	*0.29 (0.31)*	0.729	0.735
Genus	*Blautia* (F)	100.00	100.00	100.00	N/A	N/A	12.34	13.46	12.91	0.701	0.871
	Bacteroides (B)	100.00	97.80	98.90	0.328	0.480	11.22	12.22	11.73	0.785	0.923
	Faecalibacterium	97.80	86.70	92.20	0.051	0.155	12.75	8.70	10.69	**0.013**	0.173
	Prevotella (B)	57.80	51.10	54.40	0.531	0.630	7.87	8.47	8.18	0.703	0.871
	Megamonas (F)	57.80	55.60	56.70	0.836	0.836	3.45	5.94	4.71	0.705	0.871
	Agathobacter (F)	93.30	75.60	84.40	**0.021**	0.131	4.04	3.47	3.75	0.0683	0.401
	Eubacterium (F)	100.00	97.80	98.90	0.328	0.480	3.92	3.17	3.54	0.077	0.401
	Ruminococcus (F)	100.00	97.80	98.90	0.328	0.480	3.57	3.40	3.48	0.077	0.401
	Escherichia-Shigella (P)	73.30	88.90	81.10	0.062	0.155	0.46	5.32	2.93	**<0.001**	**<0.001**
	Subdoligranulum (F)	88.90	77.80	83.30	0.161	0.341	3.04	2.45	2.74	0.088	0.410
	Roseburia (F)	93.30	80.00	86.70	0.065	0.155	2.91	2.15	2.52	0.083	0.401
	Megasphaera (F)	77.80	71.10	74.40	0.474	0.601	2.34	2.62	2.49	0.836	0.959
	Fusicatenibacter (F)	93.30	80.00	86.70	0.065	0.155	3.10	1.49	2.29	**<0.001**	**0.028**
	Lactobacillus (F)	62.20	53.30	57.80	0.399	0.541	1.83	2.57	2.20	0.560	0.837
	Dorea (F)	97.80	84.40	91.10	**0.028**	0.131	2.12	1.65	1.88	0.116	0.484
	Catenibacterium (F)	42.20	44.40	43.30	0.836	0.836	1.37	1.66	1.52	0.600	0.845
	Anaerostipes (F)	100.00	82.20	91.10	**0.003**	**0.031**	2.12	0.86	1.48	**<0.001**	**0.006**
	Bifidobacterium (A)	91.10	88.90	90.00	0.733	0.819	1.57	1.25	1.41	0.074	0.401
	Clostridium (F)	80.00	51.10	65.50	**0.003**	**0.031**	1.02	1.31	1.17	**0.011**	0.173
	Holdemanella (F)	57.80	46.70	52.20	0.296	0.480	1.17	0.95	1.06	0.313	0.679

A, Phylum Actinobacteria; B, Phylum Bacteroidetes; F, Phylum Firmicutes; P, Phylum Proteobacteria. Analysed using Mann-Whitney U Test in MicrobiomeAnalyst, *p*-value and FDR of <0.05 is considered significant difference between nonDM and T2DM.

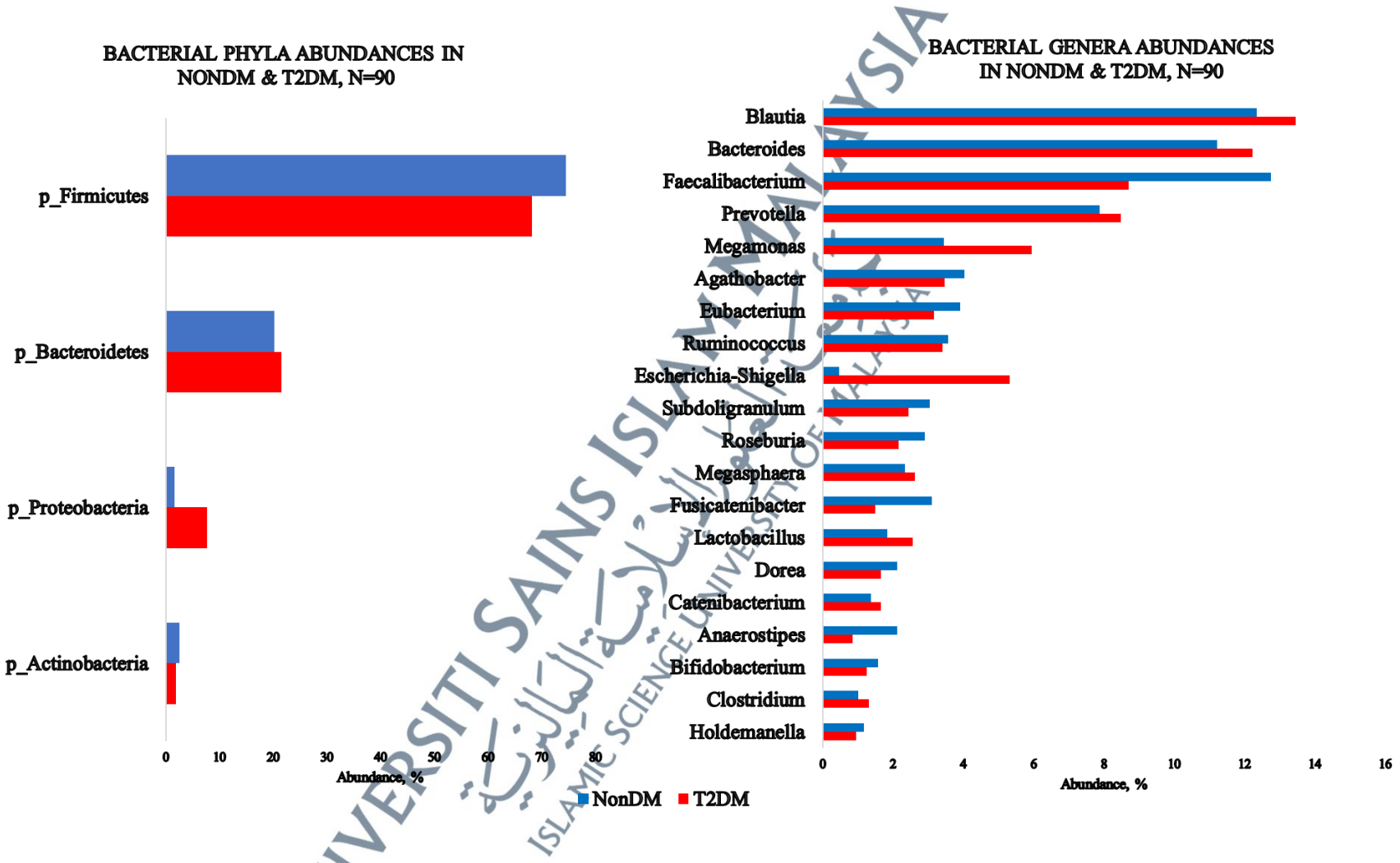


Figure 4.9 Gut Microbiota Composition in (A) Phylum and (B) Genus Levels in NonDM and T2DM Groups, n=90. These four phyla account for 99% of all phyla present and the 20 genera account for 83% of all genera present.

Meanwhile, at the genus level, T2DM had lower prevalence of genera *Anaerostipes* (82.20%), *Dorea* (84.40%), *Agathobacter* (75.60%) and *Clostridium* (51.10%) in comparison to nonDM (100.00%, 97.80%, 93.30% and 80.00%; p-value= 0.003, 0.028, 0.021 and 0.003, respectively). However, after adjustment for multiple corrections, only the difference in prevalence of genera *Anaerostipes* and *Clostridium* remained significant (FDR < 0.05) (Table 4.12).

The relative abundance of genera *Faecalibacterium* (8.70%), *Fusicatenibacter* (1.49%) and *Anaerostipes* (0.86%) were lesser in the T2DM group, when compared to nonDM group (12.75%, 3.10%, and 2.12%; *p*-value = 0.013, <0.001 and <0.001, respectively) (Table 4.12). When looking at the differences in individual ethnic groups, a lower abundance of *Fusicatenibacter* (1.20%) in T2DM when compared to nonDM was only reflected in the Indian ethnicity (2.72%; p-value = 0.023). The genus *Anaerostipes* was lower in abundance in T2DM Chinese (0.90%) and Indian (0.36%) in comparison to their respective nonDMs (2.15% and 2.76%; p-value=0.0014 and 0.004, respectively) (Table 4.13).

On the other hand, genera *Escherichia_Shigella* (5.32%) and *Clostridium* (1.31%) were higher in T2DM when compared to nonDM (0.46% and 1.02%; p-value= <0.001 and 0.011). The increase in genus *Escherichia-Shigella* was reflected within T2DM of Malay (8.13%) and Indian (3.41%) ethnic groups, in comparison to their respective nonDM groups (0.41% and 0.26%; *p*-value = 0.031 and 0.002), respectively.

Additionally, a difference in abundance of a few genera was found in certain ethnic groups but was not reflected in the overall difference between T2DM and nonDM (Table 4.13). The genus *Dorea* was lesser in Malay T2DM (1.30 %), in comparison to the Malay nonDM (2.46 %) (*p*-value = 0.024). Conversely, genus *Clostridium* was higher in Malay T2DM (2.53%) in comparison to Malay nonDM (1.61%) (*p*-value = 0.024). The genus *Bifidobacterium* was lesser in Chinese T2DM (0.64%) when compared to Chinese nonDM (1.21%) (*p*-value = 0.014).

However, only the relative abundance of genera *Escherichia-Shigella*, *Fusicatenibacter* and *Anaerostipes* in overall nonDM and T2DM groups remained significant after adjustment for multiple comparisons (FDR < 0.05).

Table 4.13 Pairwise Comparison of Gut Microbiota Abundances in NonDM and T2DM by Ethnicity, n=90.

Taxa Level	Bacteria	Malay		p-value	FDR	Chinese		p-value	FDR	Indian		p-value	FDR
		nonDM %, n=15	T2DM %, n=15			nonDM %, n=15	T2DM %, n=15			nonDM %, n=15	T2DM %, n=15		
Phylum	*Firmicutes*	76.73	67.13	0.775	0.812	70.38	68.11	0.161	0.270	74.61	68.82	0.901	0.923
	Bacteroidetes	18.02	21.34	0.285	0.418	24.42	22.16	0.305	0.392	20.64	22.70	0.430	0.729
	Proteobacteria	2.91	9.52	**0.045**	0.227	1.11	7.23	**0.001**	**0.013**	0.79	4.83	**0.002**	**0.024**
	Actinobacteria	2.03	1.44	0.389	0.487	1.71	1.34	0.106	0.270	3.47	2.74	0.183	0.729
	Desulfobacteria	0.07	0.11	0.217	0.418	0.15	0.20	0.707	0.707	0.09	0.07	0.511	0.729
	Verrucomicrobia	0.03	0.20	0.812	0.812	1.59	0.49	0.360	0.406	0.17	0.03	0.448	0.729
	Fusobacteria	0.06	0.18	0.085	0.284	0.63	0.42	0.180	0.270	0.11	0.69	0.494	0.729
	Cyanobacteria	0.15	0.02	0.292	0.418	0.01	0.00	N/A	N/A	0.03	0.06	0.604	0.755
	Synergistetes	0.00	0.05	**0.001**	0.083	0.00	0.03	**0.018**	0.081	0.08	0.05	0.385	0.729
Genus	*Blautia*	12.14	12.15	0.870	0.975	13.18	13.19	0.461	0.853	11.57	15.19	0.193	0.589
	Bacteroides	8.93	13.52	0.653	0.817	16.12	11.88	0.285	0.740	7.79	11.17	0.525	0.779
	Faecalibacterium	12.74	8.71	0.198	0.630	12.18	11.46	0.250	0.720	13.51	6.06	0.169	0.589
	Prevotella	7.30	7.55	0.832	0.949	5.49	8.29	0.658	0.879	11.45	9.68	0.876	0.964
	Megamonas	4.33	5.06	0.568	0.792	4.05	3.63	0.832	0.986	1.75	9.15	0.302	0.777
	Agathobacter	3.08	4.09	0.384	0.792	4.90	2.76	0.115	0.584	4.06	3.50	0.721	0.902
	Eubacterium	3.66	3.34	0.389	0.792	3.58	3.97	0.242	0.715	4.64	2.21	0.348	0.777
	Ruminococcus	4.40	4.37	0.795	0.929	3.63	3.12	0.501	0.853	2.57	2.62	0.636	0.829
	Escherichia-Shigella	0.41	8.13	**0.031**	0.465	0.67	4.13	0.130	0.584	0.26	3.41	**0.002**	0.129
	Subdoligranulum	3.44	2.42	0.337	0.784	2.68	3.48	0.934	1.000	3.04	1.49	0.151	0.589
	Roseburia	3.73	2.67	0.229	0.643	2.05	1.48	0.309	0.741	3.06	2.22	0.441	0.777
	Megasphaera	2.36	1.95	0.297	0.738	1.48	2.13	0.627	0.879	3.39	3.84	0.852	0.964
	Fusicatenibacter	3.79	1.50	0.089	0.608	2.80	1.78	0.170	0.584	2.72	1.20	**0.023**	0.525
	Lactobacillus	1.37	1.62	0.610	0.792	0.43	1.19	0.854	0.986	4.05	4.94	0.735	0.909
	Dorea	2.46	1.30	**0.024**	0.465	2.08	1.97	0.547	0.863	1.82	1.73	0.493	0.777
	Catenibacterium	1.54	1.03	0.737	0.901	0.94	2.33	0.169	0.584	1.71	1.71	0.566	0.799
	Anaerostipes	1.52	1.28	0.221	0.637	2.15	0.90	**0.014**	0.584	2.76	0.36	**0.004**	0.132
	Bifidobacterium	1.10	0.78	0.290	0.738	1.21	0.64	**0.014**	0.584	2.54	2.36	0.589	0.814
	Clostridium	1.61	2.53	**0.024**	0.465	0.99	0.45	0.153	0.584	0.39	0.81	0.205	0.589
	Holdemanella	1.63	0.46	0.054	0.592	0.71	1.53	0.431	0.829	1.23	0.94	0.776	0.940

analysed using Mann-Whitney U Test in MicrobiomeAnalyst, p-value and FDR of <0.05 is considered significant.

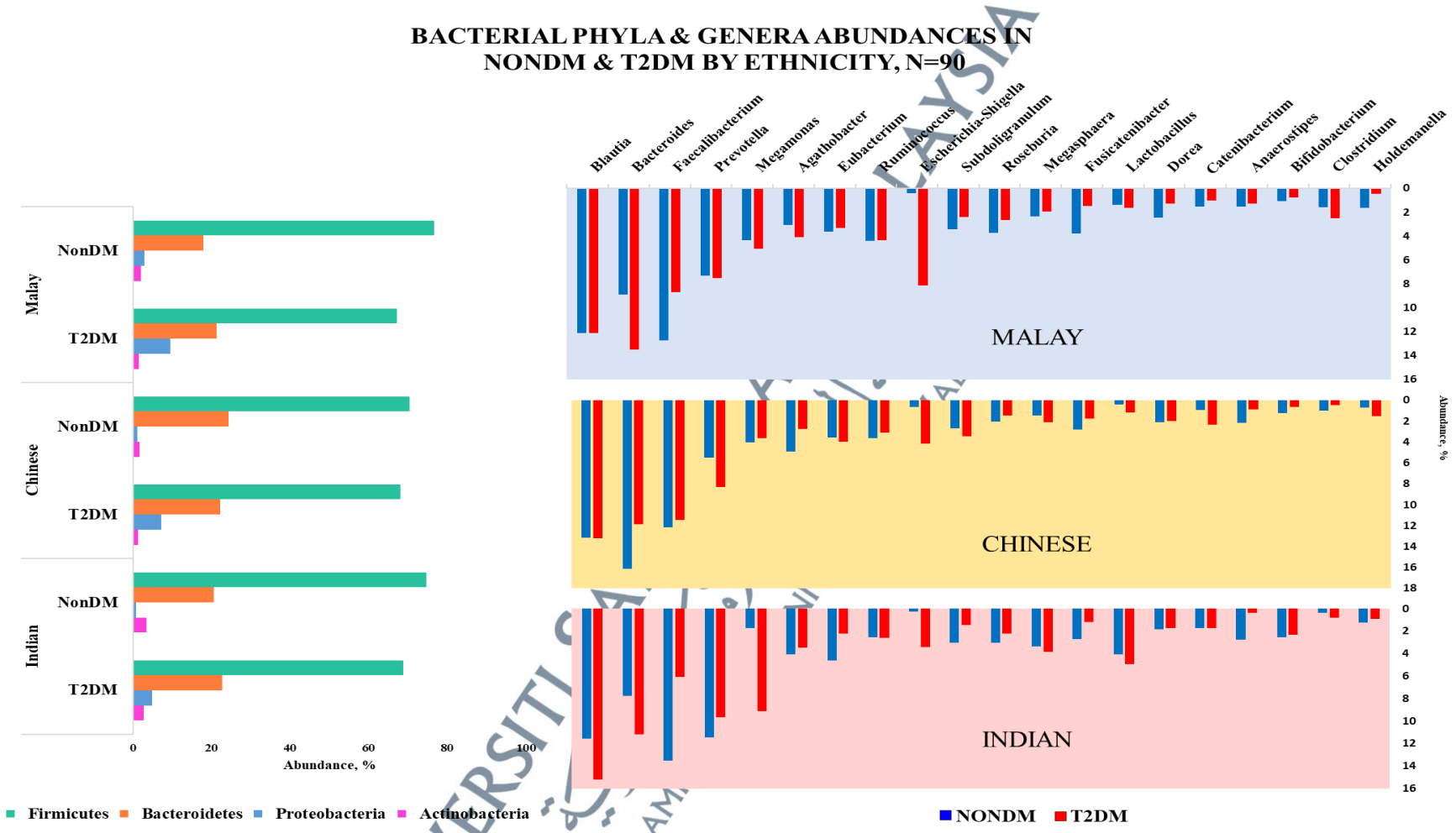


Figure 4.10 Gut Microbiota Composition in (A) Phylum and (B) Genus Levels in NonDM and T2DM by Ethnicity, n=90. These four phyla account for 99% of all phyla present and the top 20 genera account for at least 80% of all genera present.

4.5 Correlation Analysis

4.5.1 Correlation of Top Nine Gut Phyla with Clinical Characteristics

Correlation of bacterial phyla abundance with clinical characteristics (anthropometric, demographic, diabetic profile and biochemical parameters) in all 90 samples are shown in Figure 4.11. Age was seen to correlate positively with phyla *Proteobacteria* and *Desulfobacteria*. Meanwhile, weight was associated negatively with phylum *Desulfobacteria* and BMI was seen to correlate negatively with phylum *Verrucomicrobia*. The FPG was seen to correlate positively with phylum *Proteobacteria*.

The LFT parameters i.e., ALP and ALT correlated positively with phylum *Proteobacteria* while albumin correlated positively with phylum *Bacteroidetes* and negatively with phylum *Actinobacteria*. The RP parameters i.e., urea and creatinine correlated positively with phylum *Proteobacteria.* Meanwhile, chloride correlated positively with phylum *Fusobacteria* and potassium correlated negatively with phylum *Desulfobacteria*. In the FSL test, TG was associated positively with phylum *Proteobacteria* and negatively with phylum *Firmicutes*. The LDL was associated positively with phylum *Actinobacteria* and negatively with phylum *Fusobacteria*.

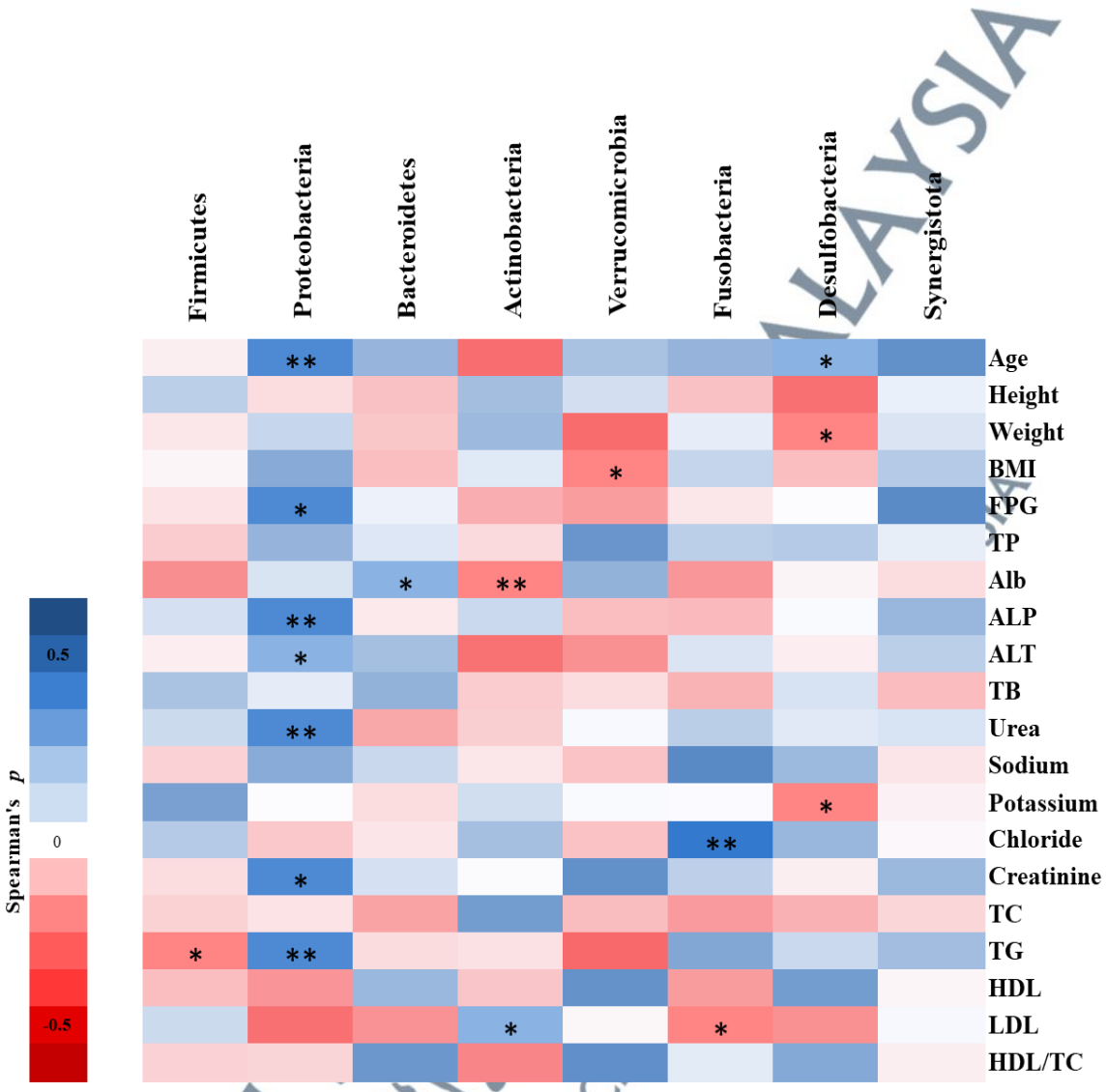


Figure 4.11 Heatmap Showing Correlation of Bacterial Phyla with Clinical Characteristics in All Participants, n=90. Analysed using Spearman Correlation in SPSS, **p*-value <0.05, ***p*-value <0.01.

4.5.2 Correlation of Top 20 Genera with Clinical Characteristics

Correlation between the 20 most abundant genera and clinical characteristics (anthropometric, demographic, diabetic profile and biochemical parameters) in all 90 samples are shown in Figure 4.12. Age was correlated positively with genera *Bacteroides, Escherichia-Shigella* and inversely with genera *Subdoligranulum, Megasphaera* and *Lactobacillus*. Positive associations were seen between height and genus *Fusicatenibacter* as well as weight with *Lactobacillus*. Weight was also negatively associated with the genus *Bacteroides*. BMI was in positive association with genus *Escherichia-Shigella* and in negative association with genera *Faecalibacterium, Fusicatenibacter* and *Anaerostipes*. FPG was correlated negatively with genera *Anaerostipes* and *Clostridium*.

In the LFT, ALP correlated negatively with genera *Faecalibacterium* and *Holdemanella,* ALT correlated negatively with genus *Eubacterium* and TB correlated positively with genus *Agathobacter*. In terms of the RP test, urea was in positive association with genus *Escherichia-Shigella* and negative association with genus *Roseburia*. Positive associations were seen between genus *Roseburia* with chloride and creatinine with genus *Lactobacillus*. In the FSL test, TG was positively associated with genus *Escherichia-Shigella* and negatively associated with genera *Faecalibacterium, Rumincoccus, Fusicatenibacter, Anaerostipes* and *Clostridium*. The genus *Megasphaera* was negatively associated while genus *Clostridium* was positively associated with HDL. The LDL was correlated positively with *Bifidobacterium* and inversely with *Ruminococcus* and *Holdemanella*. The HDL/TC ratio was correlated positively with genera *Bacteroides, Faecalibacterium* and *Ruminococcus.*

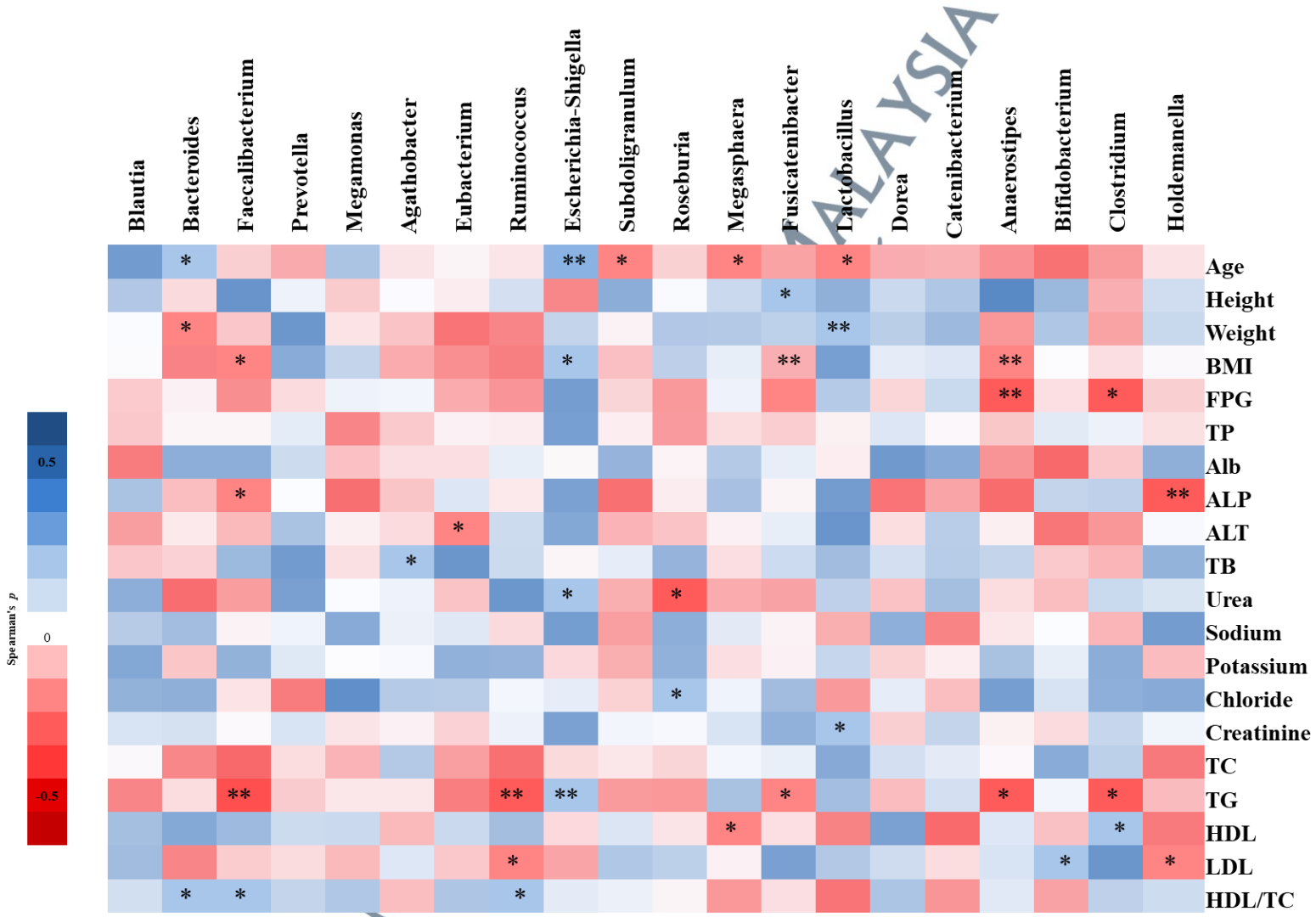


Figure 4.12 Heatmap Showing Correlation of Bacterial Genera with Clinical Characteristics in All Participants, n=90. Analysed using Spearman Correlation in SPSS, * *p*-value <0.05, ***p*-value <0.01.

4.6 Systematic Review

4.6.1 Study Characteristics

Of 3,994 articles identified, 18 studies were included in this systematic review. The characteristics of the study population are summarised in Table 4.14. The 18 observational studies had been conducted between 2013 and 2021. They included 12 case-control (Allin et al., 2018; Bhute et al., 2017; Chen et al., 2019; Gaike et al., 2020; Ghaemi et al., 2020; Lambeth et al., 2015; Li et al., 2020a; Nuli et al., 2019; Wang et al., 2021; Zhang et al., 2013; Zhao et al., 2019; Zhong et al., 2019), four cross-sectional studies (Chávez-Carbajal et al., 2020; Diener et al., 2021; Egshatyan et al., 2016; Wu et al., 2020) and two cohort studies (Ericson et al., 2020; Karlsson et al., 2013). Six studies were conducted in Asia [India (Bhute et al., 2017; Gaike et al., 2020), Iran (Ghaemi et al., 2020), Taiwan (Chen et al., 2019) and China (Li et al., 2020a; Nuli et al., 2019; Wang et al., 2021; Zhang et al., 2013; Zhao et al., 2019; Zhong et al., 2019)], five in Europe [Denmark (Allin et al., 2018), Russia (Egshatyan et al., 2016) and Sweden (Ericson et al., 2020; Karlsson et al., 2013; Wu et al., 2020)] and three in North America [USA (Lambeth et al., 2015) and Mexico (Chávez-Carbajal et al., 2020; Diener et al., 2021)]. These studies compared the gut microbial profiles of either preDM alone (Allin et al., 2018; Ericson et al., 2020), newDM alone(Chen et al., 2019; Li et al., 2020a), preDM and newDM (Egshatyan et al., 2016; Nuli et al., 2019; Wu et al., 2020; Zhang et al., 2013; Zhong et al., 2019), preDM and T2DM (Chávez-Carbajal et al., 2020; Ghaemi et al., 2020; Karlsson et al., 2013; Lambeth et al., 2015; Wang et al., 2021), newDM and T2DM (Bhute et al., 2017; Zhao et al., 2019) or all three preDM, newDM and T2DM (Diener et al., 2021; Gaike et al., 2020) in comparison to gut microbial profile of nonDM.

The 18 studies consisted of 5,489 participants. In 15 of the 18 studies, 43% of the participants were male (n=2, 252) and 57% were female (n= 2,978). Three studies (Bhute et al., 2017; Ghaemi et al., 2020; Zhang et al., 2013) did not specify the participants' gender. The mean age of participants was 50 ± 7.82 years. There was a total of 3,149 participants in the control or non-DM group. The remaining 2,340 participants with varying glucose levels included 1,599 preDM, 406 newDM and 335 T2DM. The pre-DM participants were diagnosed in these studies using either IFG (Allin et al., 2018; Chávez-Carbajal et al., 2020; Ericson et al., 2020; Gaike et al., 2020; Lambeth et al., 2015), IGT (Karlsson et al., 2013), combined glucose intolerance (CGI) (Wu et al., 2020), IFG and/or IGT (Diener et al., 2021; Egshatyan et al., 2016; Ghaemi et al., 2020; Wang et al., 2021; Wu et al., 2020; Zhang et al., 2013; Zhong et al., 2019). New-DM participants were diagnosed by OGTT (Bhute et al., 2017; Diener et al., 2021; Nuli et al., 2019; Wu et al., 2020; Zhang et al., 2013; Zhao et al., 2019; Zhong et al., 2019), HbA1c (Egshatyan et al., 2016; Gaike et al., 2020; Li et al., 2020a; Wu et al., 2020; Zhao et al., 2019) and fasting plasma glucose (FPG) (Chen et al., 2019; Li et al., 2020a; Zhao et al., 2019).

Table 4.14 Study Participant's Characteristics for Systematic Review.

No	First Author, year of publication (Ref)	Type of Study	Country	Sample Size, n	Age group (avg)	Ethnicity	No. of Subjects, n (female/male)				
							preDM			newDM	Controls
							IGT	IFG	CGI		
1	(Allin et al., 2018)	Case-control	Denmark	268	55-68	Danish		134 (53/81)			134 (53/81)
2	(Bhute et al., 2017)	Case-control	India	49	40-60	Indian				14	19
3	(Chávez-Carbajal et al., 2020)	Cross-sectional	Mexico	217	40-63	Mexicans		54 (36/18)			76 (50/26)
4	(Chen et al., 2019)	Case-control	Taiwan	100	20-80	N/A				50 (14/36)	50 (22/28)
5	(Diener et al., 2021)	Cross-sectional	Mexico	430	24-66	Mexicans	42 (29/13)	52 (29/23)	57 (39/18)	48 (31/17)	214 (165/49)
6	(Egshatyan et al., 2016)	Cross-sectional	Russia	97	25-75	Caucasian			25 (18/7)	23 (13/10)	49 (38/11)
7	(Ericson et al., 2020)	Cohort	Sweden	1726	>18	N/A		260 (137/123)			1466 (800/666)
8	(Gaike et al., 2020)	Case-control	India	102	30-60	Indian		17 (11/6)		11 (2/9)	35 (18/17)
9	(Ghaemi et al., 2020)	Case-control	Iran	90	40-60	Iranian			30		30
10	(Karlsson et al., 2013)	Cohort	Sweden	145	70	European	49				43

Avg, average;IGT, impaired glucose tolerance; IFG, Impaired fasting glucose; CGI, combined glucose intolerance.

Table 4.14, continued.

No.	First Author, year of publication (Ref)	Type of Study	Country	Sample Size, n	Age group (avg)	Ethnicity	No. of Subjects, n (female/male)				
							preDM			newDM	Controls
							IGT	IFG	CGI		
11	(Lambeth et al., 2015)	Case-control	USA	49	55-62	Caucasian white, Hispanics, Native Americans		20 (14/6)			15 (10/5)
12	(Li et al., 2020a)	Case-control	China	60	40 -50	N/A				30 (26/4)	30 (26/4)
13	(Nuli et al., 2019)	Case-control	China	60	30-70	Chinese (Uyghur)	20 (8/12)			20 (9/11)	20 (8/12)
14	(Wang et al., 2021)	Case-control	China	126	40-70	Chinese			33 (22/11)		63 (40/23)
15	(Wu et al., 2020)	Cross-sectional	Sweden	1495	50-64	Swedish	310 (172/138)	189 (98/91)	163 (67/96)	104 (44/61)	729 (427/302)
16	(Zhang et al., 2013)	Case-control	China	121	50-55	N/A			64	13	44
17	(Zhao et al., 2019)	Case-control	China	100	40-60	Chinese				16 (9/7)	35 (17/18)
18	(Zhong et al., 2019)	Case-control	China	254	49-75	Chinese			80 (39/41)	77 (44/33)	97 (65/32)

Avg, average;IGT, impaired glucose tolerance; IFG, Impaired fasting glucose; CGI, combined glucose intolerance.

4.6.2 Systematic Review of the Gut Microbiota Composition

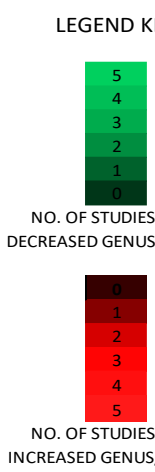
In terms of α-diversity, six of the studies found a statistically significantly lower α-diversity in the disease groups namely preDM (Allin et al., 2018; Chávez-Carbajal et al., 2020), newDM (Bhute et al., 2017; Gaike et al., 2020; Zhang et al., 2013) and in both preDM and newDM (Nuli et al., 2019) in comparison to nonDM. Meanwhile, six studies found a significant difference in β-diversity among preDM (Chávez-Carbajal et al., 2020), newDM (Bhute et al., 2017; Li et al., 2020a; Zhao et al., 2019) and in both preDM and newDM (Nuli et al., 2019; Wu et al., 2020) in comparison to nonDM. Gaike et al. (2020) observed using Principal Coordinate Analysis (PCoA), that bacterial diversity of newDM was distinct from that of nonDM, whereas preDM formed an overlapping cluster with nonDM indicating similarity in bacterial diversity (Gaike et al., 2020).

Eight studies reported that *Bacteroidetes* and *Firmicutes* were the predominant phyla in all groups studied, i.e., preDM and/ or newDM and nonDM (Egshatyan et al., 2016; Karlsson et al., 2013; Lambeth et al., 2015; Li et al., 2020a; Nuli et al., 2019; Zhang et al., 2013; Zhao et al., 2019; Zhong et al., 2019). *Proteobacteria* was reported as the next predominant phylum in five out of these eight studies, (Karlsson et al., 2013; Lambeth et al., 2015; Nuli et al., 2019; Zhao et al., 2019; Zhong et al., 2019). All eighteen studies reported significant differences in gut microbiota composition by microbial taxa in the disease groups i.e., preDM and/or newDM when compared to the nonDM control group. In four studies, a significant increase in the phylum *Firmicutes* along with a significant decrease in phylum *Bacteroidetes* was observed in the newDM group (Bhute et al., 2017; Gaike et al., 2020; Nuli et al., 2019; Zhao et al., 2019). Two of these four studies each found a significant increase (Gaike et al., 2020; Zhao et al.,

2019) or significant decrease (Bhute et al., 2017; Nuli et al., 2019) in *Proteobacteria* respectively. Meanwhile, two other studies reported a significant decrease in phylum *Verrucomicrobia* in the preDM group (Egshatyan et al., 2016; Zhang et al., 2013). Two studies reported increased F/B ratio among newDM (Li et al., 2020a; Zhao et al., 2019) and one study reported an increased F/B ratio among both preDM and newDM (Gaike et al., 2020).

Figure 4.13 depicts findings from all 18 studies on the significantly differing genera/ species belonging to the six predominant gut bacterial phyla in a heat map-like format. When focusing exclusively on changes reported by two or more studies, the composition of a particular genera/ species demonstrated distinct changes in disease groups. The number of *Streptococcus* (Allin et al., 2018; Karlsson et al., 2013; Zhong et al., 2019), *Escherichia* (Diener et al., 2021; Ghaemi et al., 2020; Zhong et al., 2019) and *Veillonella* (Diener et al., 2021; Nuli et al., 2019) in preDM were increased. Similarly, *Lactobacillus* (Bhute et al., 2017; Chen et al., 2019; Gaike et al., 2020) and *Collinsella* (Zhang et al., 2013; Zhong et al., 2019), were increased in newDM. On the other hand, *Faecalibacterium prausnitzii* (Allin et al., 2018; Ghaemi et al., 2020; Karlsson et al., 2013; Wu et al., 2020; Zhong et al., 2019), *Akkermansia* (Allin et al., 2018; Zhang et al., 2013), *Alistipes* (Karlsson et al., 2013; Wu et al., 2020), *Flavonifractor* (Nuli et al., 2019; Wu et al., 2020) and *Roseburia* (Karlsson et al., 2013; Zhong et al., 2019) were decreased in preDM while *Akkermansia* (Gaike et al., 2020; Zhong et al., 2019), *Dialister* (Li et al., 2020a; Zhong et al., 2019), *Haemophilus* (Zhang et al., 2013; Zhong et al., 2019), *Roseburia* (Zhang et al., 2013; Zhong et al., 2019) and *Faecalibacterium* (Bhute et al., 2017; Wu et al., 2020; Zhang et al., 2013) were decreased in newDM. *Bacteroides* (Allin et al., 2018; Ghaemi et al., 2020; Karlsson et al., 2013; Li et al., 2020a; Zhang et al., 2013; Zhao et al., 2019) and *Prevotella* (Bhute

et al., 2017; Egshatyan et al., 2016; Zhang et al., 2013; Zhao et al., 2019) of the phylum *Bacteroidetes*; *Blautia* (Allin et al., 2018; Diener et al., 2021; Egshatyan et al., 2016; Gaike et al., 2020; Zhao et al., 2019), *Eubacterium* (Ericson et al., 2020; Karlsson et al., 2013; Zhang et al., 2013; Zhao et al., 2019), *Clostridium* (Allin et al., 2018; Chen et al., 2019; Karlsson et al., 2013; Wu et al., 2020; Zhong et al., 2019) and *Coprococcus* (Allin et al., 2018; Wu et al., 2020; Zhao et al., 2019; Zhong et al., 2019) of the phylum *Firmicutes* all exhibited changes in both directions.



Phylum	Family	Genus	Decreased in preDM	Decreased in newDM	Increased in preDM	Increased in newDM
Actinobacteria	*Coriobacteriaceae*	*Collinsella*			1	2
	Eggerthellaceae	*Eggerthella*			1	
Bacteroidetes	*Bacteroidaceae*	*Bacteroides*	2	3	2	1
	Barnesiellaceae	*Barnesiella*	1	1		
	Dysgonomonadaceae	*Proteiniphilum*				1
	Prevotellaceae	*Prevotella*		2	2	2
	Rikenellaceae	*Alistipes*	2	1		
	Chrysiogenaceae	*Desulfurispirillum*	1			
Firmicutes	*Eubacteriaceae*	*Eubacterium*	2			2
	Lactobacillaceae	*Lactobacillus*			1	3
	Streptococcaceae	*Streptococcus*	1	1	3	1
	Clostridiaceae	*Clostridium*	4	3	2	1
	Lachnospiraceae	*Anaerostipes*	1	1		
		Blautia	2	2	2	2
		Coprococcus	3	1	1	2
		Dorea	1		1	1
		Lachnospira	1		1	1
		Roseburia	2	2	1	
		Tyzzerella				1
	Peptostreptococcaceae	*Peptostreptococcus*				1
	Oscillospiraceae	*Anaerotruncus*	1			
		Faecalibacterium	5	3	1	1
		Flavonifractor	2			
		Intestinimonas	1			
		Oscillibacter	1	1		
		Pseudoflavonifractor	1	1		
		Ruminiclostridium	1	1		
		Ruminococcus	1	1	1	1
		Sporobacter				1
		Subdoligranulum				1
	Selenomonadaceae	*Megamonas*			1	1
	Acidaminococcaceae	*Phascolarctobacterium*		1		
	Veillonellaceae	*Dialister*	1	2		
		Megasphaera			1	1
		Veillonella			2	1
	Fusobacteriaceae	*Cetobacterium*				1
Fusobacteria	*Sutterellaceae*	*Sutterella*	1		1	1
Proteobacteria	*Desulfovibrionaceae*	*Bilophila*	1			
	Enterobacteriaceae	*Escherichia*			3	1
		Klebsiella	1			
	Pasteurellaceae	*Haemophilus*		2		
	Yersiniaceae	*Serratia*			1	1
Verrucomicrobia	*Akkermansiaceae*	*Akkermansia*	2	2		

Figure 4.13 The Genera/Species found in PreDM and NewDM Groups.

4.7 Conclusion

This chapter focused on the findings obtained from analysing the data collected in this study. Analyses by group were carried out according to ethnicity and diabetes status. When comparing nonDM and T2DM groups, significantly higher alpha diversity was found in nonDMs while beta diversity suggests the presence of low-abundant ASVs and highly dissimilar species between both the groups. This showed that nonDM had a different gut microbiota diversity in comparison to the T2DM group. In comparing T2DM with nonDM by ethnicity, higher abundance of phylum *Proteobacteria*, genus *Escherichia-Shigella* and lower abundance of genera *Fusicatenibacter* and *Anaerostipes* were found in T2DMs. On the other hand, in all study participants, there were various correlations seen between gut bacteria and variables investigated in this study. It is important to note the correlations found between the significantly prevalent/ abundant bacteria and clinical. Phylum *Proteobacteria* correlated positively with age, BMI, ALP, ALT, urea, creatinine and TG while genus *Escherichia-Shigella* correlated positively with age, BMI, urea and TG. On the other hand, genus *Fusicatenibacter* correlated negatively with BMI and TG along with positive correlation with height whereas genus *Anaerostipes* correlated negatively with BMI, FPG and TG. Lastly, gut microbiota diversity and composition of preDM and newDM participants from the systematic review were summarized. These findings will corroborate the discussion in the next chapter.