CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This chapter describes the research methodology used to achieve the objectives involving the study design, study location, study population sampling technique and sample size. Furthermore, ethics approval, data collection and the study procedure to observe the infection of *Blastocystis* in type 2 diabetic mellitus (T2DM) patients and non-diabetic individuals in Malaysia are defined. Lastly, the methods used to analyse the data regarding the infection of *Blastocystis* allied in gut microbiota in both groups were discussed.

3.2 Study design

A comparative cross-sectional study was conducted to evaluate the *Blastocystis* subtypes differentiation in gut microbiota among diabetic patients at the Endocrine Clinic and Primer Clinic, Universiti Kebangsaan Malaysia Medical Centre, Cheras, Kuala Lumpur. There were two groups involved in this study: individuals who were diagnosed with type-2 diabetic mellitus (T2DM) and individuals confirmed to be non-diabetic (NonDM). The stool and blood samples were collected from February 2021 to November 2022 for both groups.

Then, the samples were processed and analysed in the Faculty of Medicine and Health Science Universiti Sains Islam Malaysia (USIM). The following lab tests were performed in the Research Laboratory, USIM:

1. DNA extraction and polymerase chain reaction (PCR)
2. Pro-inflammatory Marker detection by ELISA

The following lab test was performed by Apical Scientific Sdn. Bhd:

1. Next Gene Sequencing (NGS) using 16RNA

3.3 Study site and rationale

This study was carried out at three institutions: USIM, Pusat Perubatan UKM (PPUKM) and Klinik Primer UKM. The lab works were conducted in USIM. PPUKM and Klinik Primer UKM were chosen due to the selection of sample collection because these centres are the primary and tertiary referral hospitals respectively, for diabetic cases in Klang Valley. There were specialists available in these centres with specific times for the appointments. So, we could continuously get potential patients from these centres. Furthermore, both facilities were linked to each other and used a similar system for data accession, diabetic management and care.

3.4 Study population

The study was carried out on individuals aged 18 to 65 years old who were diagnosed diabetic type 2 compared to the non-diabetic group. In this study, participants in the diabetic group were individuals diagnosed with type-2 diabetes mellitus (T2DM). They needed to have been registered at the Endocrine Clinic or Premier Clinic for at least one year and had no history of other chronic diseases, cancer or any immunodeficiency disorders, except for diabetes. Furthermore, they must not take any antibiotics in the three months leading up to their recruitment, as confirmed by a health professional. The diagnosis of diabetes mellitus had been screened according to Malaysian Diabetes Association criteria where the participants were been classified as diabetics if their HbA1c ≥6.3% (≥45 mmol/mol) or if they were currently taking medication for diabetes.

The non-diabetic group (NonDM) was selected from an unselected population undergoing routine health checks at the same health centre and from a community that has the same sociodemographic. They have been identified from the community as non-diabetic group if HbA1c is < 5.7% (<39 mmol/mol), who declared not to have T2DM and if they never receive any diabetic medications. Table 3.1 shows the inclusion and exclusion criteria for both T2DM and NonDM groups.

Table 3.1: Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
• Aged 18-65 years old • Ability to read and write in Malay or English as a subject's information form was provided in these languages **<u>T2DM group</u>** • FBS ≥ 7.0 mmol/L or HbA1c ≥ 6.3% (≥ 45 mmol/mol) within 6 months • Received any diabetic treatment and medication during time of study (Metformin, Glizlzide, Sitagliptin, Pioglitazone, Empagliflozin, Repaglinide, GLP1 analogue, Insulin). **<u>Non-Diabetic group</u>** • No chronic illnesses	• T1DM/GDM/Secondary diabetic • Suffering from any chronic disease and complications; chronic kidney disease, cancer, thyroid, gastrointestinal disorder due to recent acute gastroenteritis (AGE), inflammatory bowel syndrome (IBS), inflammatory bowel disease (IBD) and acute infections • Suffering from any bariatric surgery • Received any antibiotic or probiotic in the past 3 months prior to recruitment • Pregnant and breastfeeding • Received any psychiatric medication • People with disabilities (wheelchair bound)

3.5 Sample size calculation

The sample size was calculated using two proportion values. Using the openEpi software with a power of 80% and confidence interval of 95%. The proportion of the diabetic group (57.14%) and the non-diabetic group (35%) from a pilot test that was conducted in November 2020 at the Endocrine clinic, Pusat Perubatan UKM. The total sample size would be 208 samples and has been divided into 104 samples for each group (Calculation as shown in Appendix 1).

As calculated by Open Epi software, the minimum sample size in this study was 87 per group. However, 20% was added considering there have been a few practical issues during the research such as loss to follow-up, missing data, drop out or refusal to give valid responses to questions and physical measurements (Suresh & Chandrashekara, 2012). So, the absolute sample size would be 104 per armed and the total number of this study would be 208 samples.

Sample size per arm (n) = 87 + 17 (considering 20% dropout of study participants)

Total subjects required= Diabetic group (n=104) + non-Diabetic group (n= 104)

Final sample size (n) = 208 subjects

3.6 Study sampling

In this study, the participants were approached conveniently (Figure 3.1). For the diabetic group, the lists of diabetic patients who were attending appointments and had been registered under the Endocrine Clinic and Premier Clinic, Universiti Kebangsaan Malaysia were accessed. Therefore, within the study period, all patients were invited to join the study conveniently based on their easy accessibility and availability that met the eligibility criteria (Refer Fig. 3.1). For the non-diabetic group, who did not develop any serious disease and were excluded from the diabetic group criteria, had been chosen voluntarily from the daily routine medical check-up, the community around Klang Valley and the area that has a similar sociodemographic to the diabetic group until the targeted sample size were achieved. The poster invitation (Appendix 2) was distributed to invite the volunteers to participate in this study.

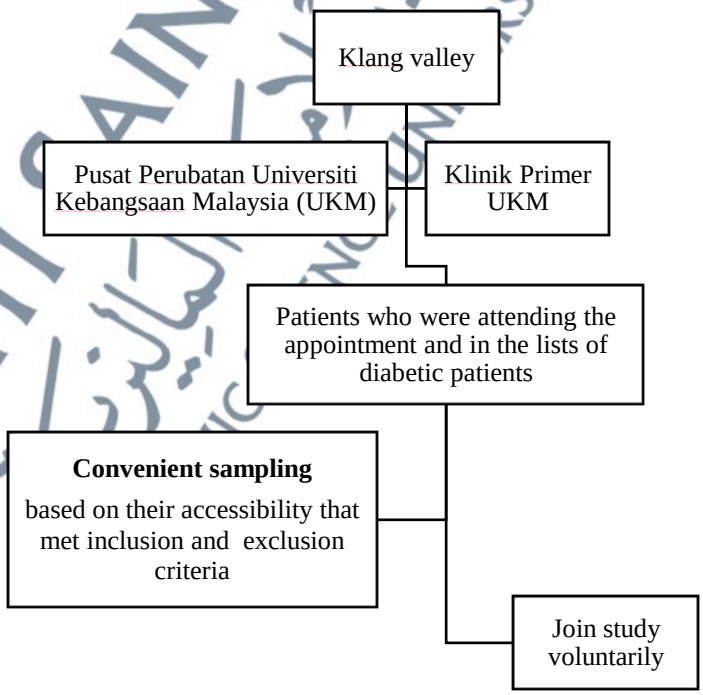

Figure 3.1: Flowchart of convenient sampling

3.7 Methodology flow chart

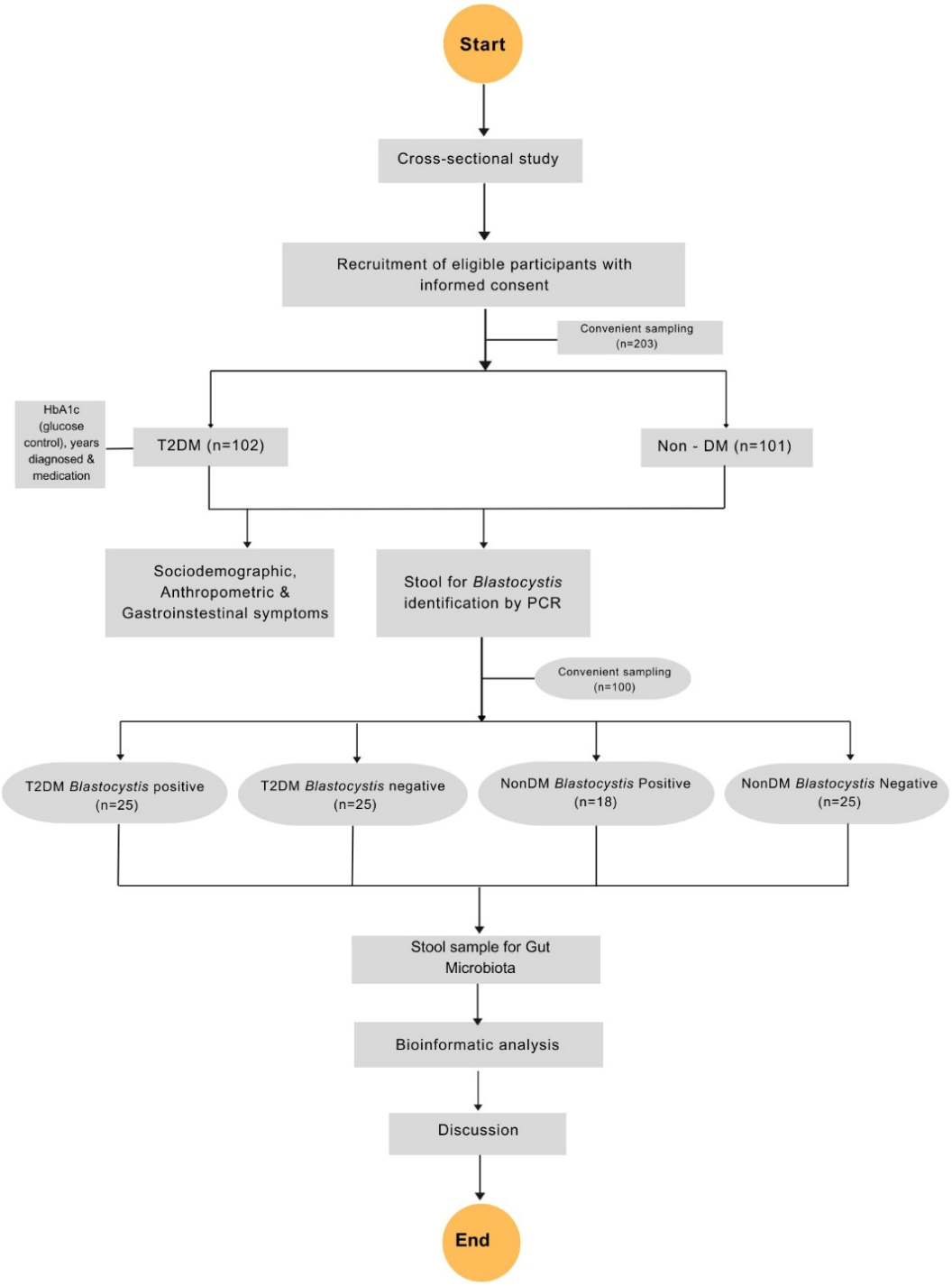

Figure 3.2: Methodology flowchart

3.8 Study procedure

3.8.1 Data and samples collection

Participants were enrolled in this study between February 2021 to November 2022. All the participants were briefed on the study's purpose and design. The consent form and research information sheet were given to all participants before sample collection (refer to Appendix C). The socio-demographic details of the participants were recorded. In addition, the medication intake and clinical health status of diabetic patients were also recorded according to the HUKM database. Then, the labelled capped container with wide mouths was provided to each participant with detailed instructions regarding stool collection to minimise contamination. They were advised to return the sample within 24 hours after collection. The collected samples were transported in an ice box to the laboratory immediately. The fresh stools were separated into several tubes to store at -80 °C for molecular analysis.

For the blood samples, about 10 ml of blood from each participant was drawn in a separator tube (eg, BD Vacutainer® Cat. No. 366430) by using butterfly needles (Terumo, 21'G) and blood was allowed to clot for 30 minutes. The samples were centrifuged for 10 minutes at 1000 rpm (Eppendorf, Centrifuge 5430 R). The serums were transferred immediately into 2ml tubes and were stored at -20°C for further analysis. All the stool and blood specimens were labelled with a code assigned to each participant, written in a combination of a letter ("D" for samples from the Diabetic group, "C" for samples from the non-Diabetic group) and a consecutive number in order of collection.

3.8.2 Stool sample processing: Molecular techniques

3.8.2.1 Detection of *Blastocystis*

a) DNA extraction

The genomic DNA (gDNA) was directly extracted from 203 unfixed stool samples using QIAamp® PowerStool Pro DNA Kit (Qiagen, Germany) according to the manufacturer's protocol. Firstly, PowerBead Pro Tube was spun to ensure that the beads had settled at the bottom. About 250 mg of stool was added to 800 µl of Solution CD1 and the mixture was vortexed. PowerBead Pro Tube was secured horizontally on a Vortex Adapter for 1.5-2 ml tubes and vortexed at maximum speed for 10 minutes. PowerBead Pro Tube was centrifuged at 15,000 x g for 1 minute by using Eppendorf Centrifuge 5430 R (Hamburg, Germany). Then, the supernatant was transferred to a clean 2 ml centrifuge tube and was centrifuged at 15,000 x g for one minute at room temperature. Approximately 700 µl of the supernatant was transferred to a clean 2 ml microcentrifuge tube. 600 µl of Solution CD3 was added and the mixture was vortexed for 5 seconds. 650 µl of the lysate was loaded onto an MB Spin Column and was centrifuged at 15,000 x g for 1 minute. The flow-through was discarded and 650 µl of the lysate was reloaded onto an MB Spin Column. Then, MB Spin Column was recentrifuged at 15,000 x g for 1 minute until the all lysates had passed through the MB Spin Column. The MB Spin Column was placed into a clean 2 ml collection tube.

After that, 500 µl of Solution EA was added to the MB Spin Column and centrifuged at 15,000 x g for 1 minute. The flow-through was discarded and the MB Spin Column was placed back in the same 2 ml collection tube. 500 µl of Solution C5 was added to the MB Spin Column and was centrifuged at 15,000 x g for 1 minute. The

flow through was discarded and the MB Spin Column was placed into a new 2 ml collection tube and was centrifuged at up to 16,000 x g for 2 minutes. The MB Spin Column was placed into a new 1.5 ml elution tube. 50-100 µl of Solution C6 was added to the centre of the white membrane filter membrane and was centrifuged at 15,000 x g for 1 minute. The MB Spin Column was discarded. The final DNA was ready for downstream application and was stored at -20 °C until required for PCR amplification.

b) DNA Quantification

DNA concentration of each extracted DNA was quantified by reading its absorbance at 260 nm by using a nanophotometer (Implen Nanophotometer). Purity was determined by calculation of the ratio of absorbance at 260 nm to the absorbance at 280 nm. Pure DNA showed an A260/A280 ratio of 1.7 to 1.9. the absorbance reading at 260 nm must be within 0.1 to 1.0.

c) PCR Amplification Using Conventional PCR Method

Primer for PCR amplification was prepared accordingly based on protocol (IDT, Singapore). The samples were screened by using a primer complementary to the SSU rDNA sequence from the samples using these primers by Scicluna et al. (2006) which have been described in Table 3.2. Then, the master mix (Promega, USA) was prepared accordingly before PCR amplification.

Table 3.2: RD5 and BhRDr primers

Primer name	Orientation	Target gene	DNA sequence (5'-3')	Length (bp)	Melting temperature
RD5	Forward	SU rRNA (18 S)	ATCTGGTTGATCCT GCCAGT	20	56.2 °C
BhRDr	Reverse	SU rRNA (18 S)	GAGCTTTTTAACTG CAACAACG	22	53.2°C

PCR amplification was performed in a final volume of 25 µl in 0.2 mL PCR tubes using Eppendorf Pro-S 6325 thermal cycler (Hamburg, Germany) by 30 cycles of initial denaturing at 95 °C for two minutes, followed by denaturation at 95 °C for 30 seconds, annealing at 60 °C for 30 seconds, extension at 72 °C for one minute and an additional cycle of five minutes chain elongation at 72 °C. The PCR products were separated in 1.5% agarose gel. Positive control was obtained from a clinical specimen from a pilot study at Endocrine Clinic HUKM which has been confirmed as *Blastocystis*-positive by IVC and PCR, meanwhile, the negative control was nuclease-free water.

d) PCR Product Analysis Using Gel Electrophoresis

Agarose solution (1.5%) was prepared by adding 0.75 grams of agarose (Simply, USA) to a 250 ml conical flask containing 50 ml 1X TAE buffer (1st Base, Singapore). The agarose was heated in a microwave oven (Panasonic, Malaysia) at high temperature for one minute and 30 seconds. It was left to cool to approximately 50 °C for 10 minutes. Subsequently, 5 µl of Ultra GelRed Stain (Vazyme, China) was added and mixed by swirling. The agarose solution was poured into a gel mould fitted with the appropriate comb for moulding the agarose gel wells. The gel was left at room temperature to fully solidify for 30 minutes and was transferred to a gel tank containing 1X TAE buffer. Then, an aliquot of one µl of 6X gel loading dye (1st Base, Singapore) was mixed by repeated pipetting with five µl product of each PCR product on a small sheet of clean parafilm (PARAFILM® USA). Each mixture was transferred with a micropipette and was loaded into each respective well of the agarose gel. As a reference for molecular size, five µl of the 100-basepair DNA ladder (GeneDireX, USA) was added to the first well, which has been usually designed as a reference. The electrophoresis was performed at 90 V (Major Science, California) for 90 minutes and the gel was visualized

using Gel & Blot Imager- UVP ChemStudio PLUS (Analytic Jena, Thailand). Gels were photographed using a camera to fix the transilluminator.

e) DNA sequencing

The samples were sequenced by a commercial company with the Applied Biosystems DNA Sequencer instruments. Then, sequence chromatograms were viewed using CodonCode Aligner (CodonCode Corporation, USA). Forward and reverse sequences were edited, and manually aligned and the consensus was created for each sample.

f) Identification of *Blastocystis* Subtypes and Alleles

All the partial SSU rDNA sequences obtained from sequencing were compared with previously published sequences from GenBank™ using BLASTN program (http://www.ncbi.nlm.nih.gov/blast). Identification of the alleles was performed by sequence query into the "*Blastocystis* ST (18S) and Multi Locus Sequence Typing (MLST) https://pubmlst.org/.

3.8.2.2 Detection of Microbiota: Metagenomics of Gut Microbiota

a) Genomic extraction

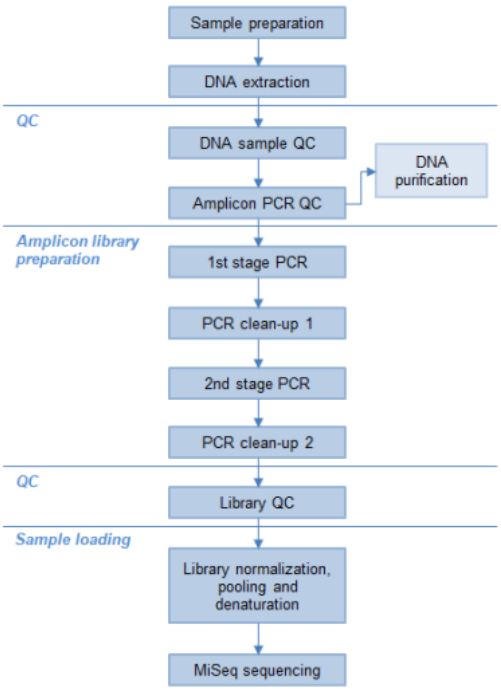

Figure 3.3: Flowchart for 16S Amplicon Library Preparation Workflow

For Next Generation Sequencing (NGS), 100 samples were selected from an initial 203 based on two criteria; positive detection of *Blastocystis* and the quality of extracted DNA. Out of the 203 samples, 26 were identified as *Blastocystis*-positive, and only 25 samples from these individuals with type 2 diabetes mellitus (T2DM), were chosen for NGS. In the non-diabetic (NonDM) group, only 18 samples were available and were used for comparison. Additionally, two *Blastocystis*-negative groups, one from T2DM and one from NonDM individuals, followed the *Blastocystis*-positive groups in the analysis. Despite the unequal sample sizes, this approach was necessary due to the limited availability of *Blastocystis*-positive cases in that group. The sample

was further analysed by using an amplicon library preparation workflow as shown in Figure 3.3.

Firstly, stool samples were thawed and transferred into a 2ml PowerBead Pro Tube. The genomic DNA (gDNA) of 200 mg of each homogenized stool sample was extracted by using QIAamp® PowerStool Pro DNA Kit (Qiagen, Germany) according to the manufacturer's protocol. Then, the DNA purity was measured using Implen Nanophotometer and was performed agarose gel electrophoresis monitoring the DNA quality before submitting the samples to outsources according to the gel electrophoresis protocol with some modifications.

b) DNA Quality control and Library Construction for Next Gene Sequencing (NGS)

A total of 100 samples of purified gDNA were submitted to Apical Scientific Sdn. Bhd. for 16S rDNA sequencing. The following gDNA quality control (QC) along with 16S sequencing were performed by the service provider. As shown in Figure 3.3, the quality of gDNA was first monitored after the extraction step. According to the protocol, aliquot of 1µl diluted gDNA was run on 1% TAE agarose gel at 100V for 60 minutes. This was run along a 1kb ladder and positive control (50 ng of bacterial DNA). The concentration of gDNA was again measured using a spectrophotometer (Implen NanoPhotometer® N60/N50) and fluorometric quantification using iQuant™ Broad Range dsDNA Quantification Kit. Then, the purified gDNA that passed the DNA sample quality check was sequenced for Bacterial 16 rRNA amplification of the V4 region to profile gut microbiota composition by using PCR primer pairs sequence in Table 3.3. All the PCR reactions were carried out with REDiant 2X PCR Master Mix (1st Base, Singapore).

Table 3.3: V4 region profile of Bacterial 16 rRNA amplification

Region	Amplicon size (bp)[a]	Name	Direction	Primer sequence (5'-3')
V4	292	515F	Forward	CCTACGGGNGGCWGCAG
		806RB	Reverse	GACTACHVGGGTATCTAATCC

[a]Mean amplicon size shown for hypervariable region amplified plus PCR primers

All the samples that passed the amplicon PCR quality check were eligible for amplicon library preparation. There were two stages of PCR for library construction based on Figure 3.3. In the first stage of PCR, the bacterial 16S rRNA gene of selected regions (16S V4) was amplified using locus-specific primers with overhang adapters (Forward overhang: 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG and Reverse overhang: 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG) according to Illumina's 16S metagenomic library preparation guideline using KOD-Multi & Epi® (Toyobo). Then, dual indices were attached to the amplicon PCR using Illumina Nextera XT Index kit v2 (Illumina, San Diego, CA, USA) following the manufacturer's protocols. The quality of the libraries was measured using Agilent Bioanalyzer 2100 System by Agilent DNA 1000 Kit (Agilent Technologies, USA, Santa Clara) and qPCR by JetSeq™ Library Quantification Lo-ROX Kit/fluorometric quantification by Helixyte Green™ Quantifying Reagent (Bioline Reagent Ltd). The libraries that passed the library QC were then normalised and pooled following the Illumina protocol before being subjected to next-generation sequencing (NGS) on the Illumina MiSeq platform with 300 PE (San Diego, CA, USA).

c) Bioinformatic analysis

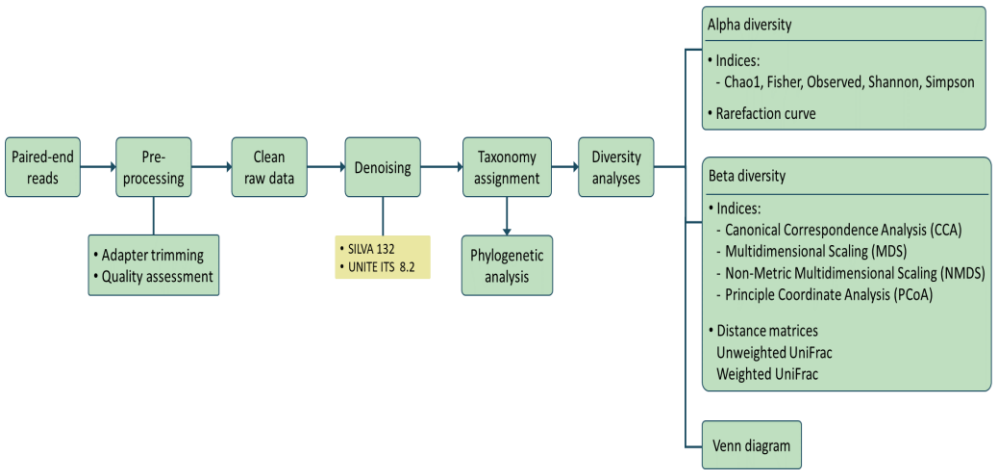

Figure 3.4: 16S Amplicon Sequencing Analysis Pipeline

According to Figure 3.4, sequencing of the region in 16S rRNA was performed using a pair-end (PE) Illumina Miseq platform that generates 300bp raw reads. From the raw reads, sequence adapters and low-quality reads were removed using BBDuk (version 38.76). Then, the raw reads were aligned and merged using the Quantitative Insights Into Microbial Ecology (QIIME2) analysis program (version 2019.10) (Bolyen et al., 2019).

Following that, the Divisive Amplicon Denoising Algorithm 2 (DADA2) pipeline (version 1.14) was used to denoise, to remove low-quality regions and chimeric errors to obtain an amplicon sequence variant (ASV) (Nearing et al., 2018). For this, duplications were eliminated by combining identical sequencing reads into single sequences. Chimeras were removed before taxonomic allocation. The quality control of the DADA2 output was analysed and visualised using MultiQC before downstream analysis to ensure data accuracy and reliability (Figure 3.4).

Next, the taxonomic assignment was completed by using DADA2 pipeline that comes with a naïve Bayesian classifier to classify large sequences across multiple ranks from Kingdom to Genus. It compares a set of taxonomically assigned sequences provided from formatted reference fasta files databases such as SILVA for ribosomal rRNA database and makes individual taxonomic assignments (Cellahan, 2020). SILVA database was used to analyse the sequence variant (ASV) reads with recommended parameters at a 97% similarity threshold (Xue et al., 2018). The classification level of phylum up to the classification level of species and the distribution histogram of the relative abundance was generated accordingly.

Then, a phylogenetic tree was generated by first performing multiple sequence alignment of the sequence using the mafft and FastTree algorithms. This pipeline was carried out in QIIME2. Finally, R phyloseq package was used to carry out the analysis of microbial diversity. The alpha diversity was measured using Observed, Chao1, Fisher, Shannon and Simpson. Alpha rarefaction curve was plotted after samples were rarefied to 20,000 sequences per sample. Beta diversity was computed using distance metrics ordinated weighted UniFrac and unweighted UniFrac by analysing using PCoA (Principal Coordinate Analysis) to show the presence or absence of dissimilarity of the ASVs in between samples.

3.8.2.3 Pro-inflammatory Marker Detection: Solid Phase Sandwich ELISA (Enzyme-Linked Immunosorbent Assay) for IL-6 and IL-1β Quantification

The levels of IL-6 and IL-1β in the serum were determined using Elabscience Human (IL-6 and IL-1β) ELISA Kit and were performed according to the manufacturer's protocols (Elabscience Biotechnology Inc, USA). Briefly, all the reagents including wash buffer, standard working solution, biotinylated detection Ab working solution, HRP Conjugate working solution were prepared accordingly. Then, the wells for diluted standard, blank and sample were determined. 100 µl each dilution of standard, blank and sample was added into the appropriate wells. The plate was covered with the sealer provided in the kit and was incubated for 90 minutes at 37°C. After that, the liquid was decanted from each well and was added immediately with 100 µl of Biotinylated Detection Ab working solution to each well. The plate was covered with new sealer and was incubated for 1 hour at 37°C. The solution was decanted from each well.

Then, 350 µl of wash buffer was added to each well. The solution was soaked for 1 minute, was aspirated from each well and was pat dried against clean absorbent paper. This step was repeated three times. Further after washing the plate, 100 of HRP Conjugate working solution was added to each well. The plate was covered with a new sealer and was incubated for 30 minutes at 37°C. The solution was decanted from each well and the wash step by using a wash buffer was repeated 5 times. After done washing step, 90 µl of Substrate Reagent was added to each well. The plate was covered with a new sealer and was incubated for about 15 minutes at 37°C. Then, 50 µl of stop solution was added to each well. The stop solution changes the colour from blue to yellow. The microwell absorbances were read at 450nm by using Infinite® 200 microplate readers (Tecan Trading AG, Switzerland) to determine the optical density (OD value).

Cytokine levels were determined with the help of standard curve and concentration were expressed as pg/mL.

3.9 Statistical analysis

Data was analysed using SPSS software (Statistical Package for the Social Sciences) program for Window version 25 (SPSS, Chicago, IL, USA). Data entry was cross-checked by independent individuals to ensure correct data entry. Prevalence rates of *Blastocystis* in diabetic patients and non-diabetic groups were determined by descriptive analysis. The continuous data was expressed as mean ± standard deviation (SD) for normally distributed data and median (interquartile range, IQR) for non-normally distributed (skewed) data and the qualitative variables were summarized in frequencies and proportions. For all continuous values, normality assumptions were evaluated using the Kalmogorov-Smirnov and Shapiro Wilk. Pearson's chi-squared test and Fisher's Exact tests were used to assess the possible association between *Blastocystis* subtype infection and the different studied variables.

A student's t-test (normal data) and Mann Whitney U Test (non-normal data) were conducted to measure the mean difference in alpha diversity and richness between T2DM and NonDM in the presence of *Blastocystis* and without *Blastocystis* using values derived from the diversity index by using SPSS. ANOVA (Analysis of Variance) test for normal data and Kruskal Wallis test for non-normal data were also conducted for comparisons involving more than two groups with Bonferroni correction. Correspondingly, the beta diversity measure between study groups was analysed by using per Mutational Multivariate Analysis of Variance (PERMANOVA) test by using

MicrobiomeAnalyst (version 2.0). Probability (p) values $p < 0.05$ were considered to indicate statistically significant differences.

Additionally, the taxonomic data was expressed in prevalence (%) and relative abundance unit by using SPSS and the bar chart was illustrated by using Excel. The significant differences in bacterial taxa among the *Blastocystis* infection and *Blastocystis* free groups in both T2DM and non-diabetic were determined using the Kruskal-Wallis test and Mann-Whitney U test for comparison between the two groups. Finally, the Spearman Correlation test was conducted for the association between *Blastocystis* infection, and microbial composition with pro-inflammatory cytokines (IL-6, IL1-β) by using SPSS. The heatmap correlation was illustrated by Phyton. The *p*-values of gut microbiota composition were corrected using the False Discovery Rate (FDR) by using the Benjamin-Hochberg method in Phyton script (version 3.90). Any taxon showing significant differences (adjusted *p*-value < 0.05) between different conditions was associated with the condition being observed, like T2DM.

3.10 Ethics and protection of subjects

3.10.1 Approval of Institutional Review Board

The study protocol which involved stool and blood collection from the diabetic patients and non-diabetic participants was subjected to approval by USIM(USIM/JKEP/2020-82) and UKM (UKM PPI/111/8/JEP-2019-848) ethics committee (Refer Appendix D).

3.10.2 Patient's information Sheet and Consent

Before stool and blood collection, the subjects were explained briefly the procedures, potential risks and benefits of the study. They were informed that their identity and personal were kept strictly confidential. This was done by labelling each sample with a specific ID instead of the study participant's name. During the meetings, the aim of the study was described to patients or their guardians and their participation is voluntary and therefore they could withdraw from the study at any point in time without giving any reason. After they agreed to participate, their concerns were obtained in written form (sign). The Patient Information Sheet and Consent Form were available in English and Bahasa Melayu. The participants were also informed that the related data would be enclosed for them if they were interested to know.

Any symptomatic diabetic patients and healthy controls infected with *Blastocystis* with severe cases were referred to treating for the treatment according to Centres for Disease Control and Prevention (CDC) recommendations.

3.10.3 Management of sample collection during the COVID pandemic

Standard operating procedures (SOP) were adhered to by both researchers and participants, including the use of respirators or three-ply surgical masks, social distance throughout participant screening and sample collection within participants, and interview sessions. All the researchers wore personal protective equipment (PPE) in the clinical setting. Any participants who exhibited COVID-19 symptoms during the screening process were excluded from the study. However, participants who had been infected by COVID-19 during the specified timeframe were allowed to participate and provide samples after completing their isolation sessions.

3.11 Conclusion

In conclusion, this chapter has provided a detailed overview of the research methodology employed to investigate the prevalence of *Blastocystis* infection among T2DM and non-diabetic groups in Malaysia, along with its association with gut microbiota. By thoroughly discussing the study design, location, population, sampling method and sample size, a methodology framework for data collection, sample and data processing by using a series of lab procedures was recognised. Lastly, the data of these procedures were subjected to data analysis. The results and findings of this study are discussed in the next chapter.