3.1 Study Sites

Three sites were selected to elucidate the soil bacterial compositions of Antarctic barren fell field sites from different locations on Antarctic Peninsula (Fig. 3.1.a) (Table 3.1). The first location is in Rothera Point (67^{o} 34'S, 68^{o} 08'W), Adelaide Island (Fig. 3.1.b). This island has high glacial mountains and the annual temperature ranges between 5℃ in summers to -20℃ in winters (NERC-BAS, 2010). The sampling was done near the Rothera Research Station belonging to the British Antarctic Survey (BAS).

Table 3.1
Study sites specifications and soil characteristic

Location	GPS	Air Temp.	Soil Temp.	Description
Rothera (RO) (Bonner Lab)	S67° 34.235' W68° 07.662'	3.9℃	6.6℃	Compact soil (high human activities)
Viking Valley (VV)	S71° 49.989' W68° 20.632'	4.7℃	12.3℃	Stony bare rock
Léonie Island (LE)	S67° 35.703' W68° 20.673'	5.3℃	11.1℃	Scattered rock with *Dechampsia* and moss; possible former melt snow run off

The second chosen site is Viking Valley (71^{o} 50'S, 68^{o} 21'W) (Fig. 3.1.c) on north-eastern side of Mars Glacier, Alexander Island (Australian Antarctic Devision, 2010). This island is on the western coast of Antarctic Peninsula and was once completely isolated by ice (Siegmund and Hall, 2000). The area is known for its constant permafrost and is entirely barren besides small amounts of stream algae and very little lichen and moss (Convey and Smith, 1997).

The last site is Léonie Island, an island (67^{o} 36'S, 68^{o} 21'W) on Northern Marguerite Bay (Fig. 3.1.b). Based on field survey information given by Convey and Smith (1997), this island has a steep height and about more than half of it is covered by ice throughout the year. Some parts of the island are prone to radiation from the neighbouring glaciers on Adelaide Island. Water availability during summer supports vegetation on a few sites while other parts are majorly harboured by lichens. Overall, the vegetation is quite diverse and dynamic considering the latitude of the island.

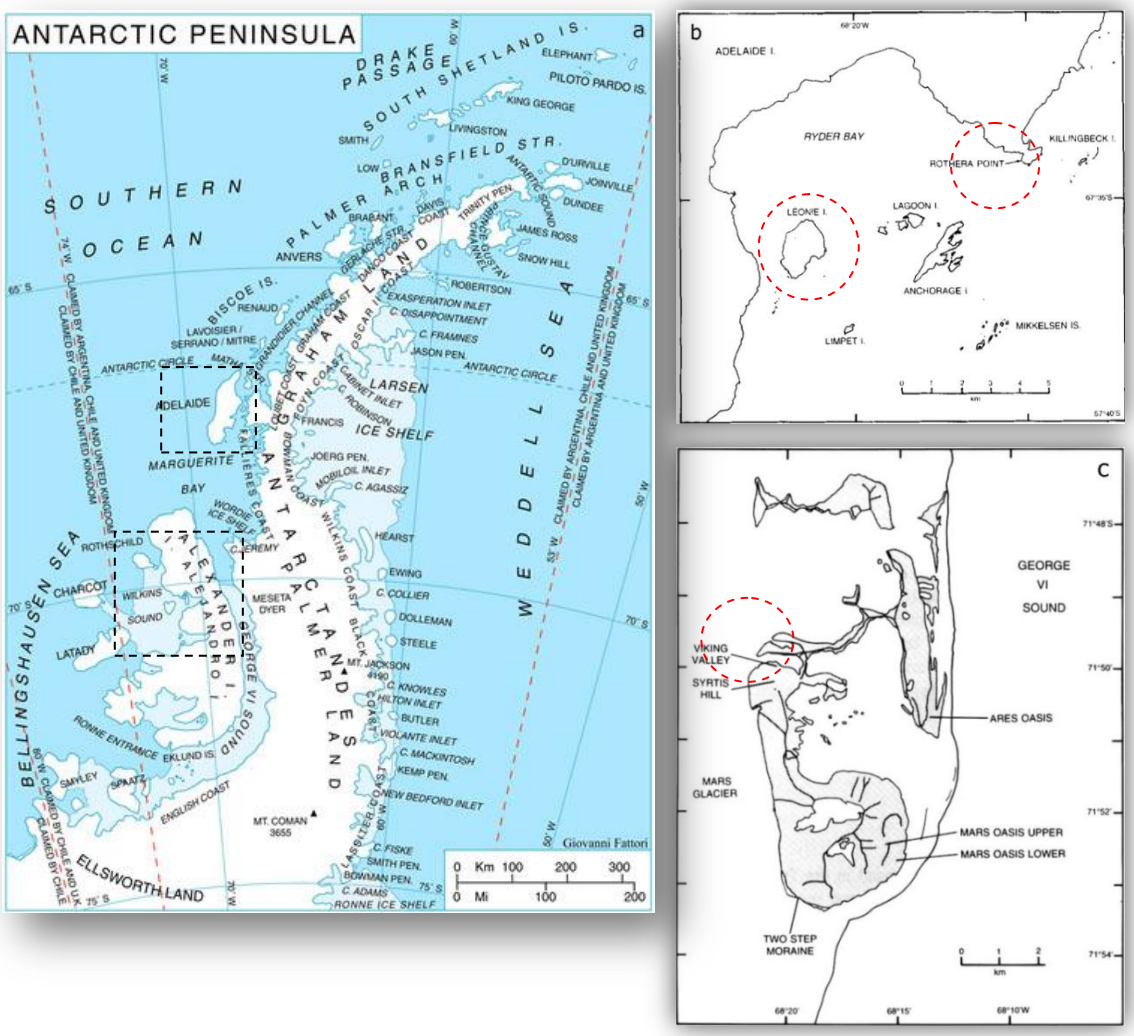


Figure 3.1 a. Map of Antarctic Peninsula showing two main Islands where study sites are located **b.** Map of Rothera Point (on Adelaide Island) and Léonie Island (obtained from Convey and Smith, 1997) **c.** Map of Viking Valley on Alexander Island (Convey and Smith, 1997)

3.2 Soil DNA Extraction

Total genomic DNA was extracted from each soil sample at the Rothera Research Station within 8 hours of collection. The soil collection and DNA extraction were carried out by Chun-Wie Chong and Yuh-Shan Goh during their expedition to Rothera in the austral summer of 2008/2009. DNA extraction was carried out using UltraClean™ Soil DNA Isolation Kit (MoBio Inc., USA) and the extracted DNA for each site was stored at -20℃, and transported back to Malaysia. The rest of the studies were conducted at the National Antarctic Research Centre of Malaysia, University of Malaya.

3.2.1 DNA concentration

The original DNA concentration for each sample was measured by spectrophotometry using a cuvette, UVette 220-1600 nm (Eppendorf AG, Germany), at 600 nm. This was done to indicate the amount of dilution required for each sample to be used as PCR template. For this purpose, the spectrophotometer was set on dsDNA setting option with dilution factor of 2% v/v (DNA to dsH_2O ratio). The machine was blanked with 98 µl of distilled water before 2 µl of DNA sample was added and the optical density (OD) measured.

3.3 PCR Amplification of 16S rRNA Gene

The soil bacterial compositions of the selected locations were analyzed using 16S rRNA genes cloning method. For each site, duplicate clone libraries were made to ensure reproducibility of the results. From the soil total genomic DNA, the 16S rDNA fragment was amplified by polymerase chain reaction (PCR) using 27F (5′-AGA GTT TGA TCM TGG CTC AG-3′) and 1492R (5′-TAC GGT TAC CTT GTT AGG ACT T-

3′) primers (Newberry *et al.*, 2004). The PCR was carried out in a reaction mixture of 1× *Taq* buffer [200 mM Tris–HCl (pH 8.4), 500 mM KCl], 1.5 mM $MgCl_2$, 0.5 µM of each primer, 0.25 mM of each dNTP, and 3 U of *Taq* DNA polymerase (invitrogen, USA). For convenience purpose, a master mix was prepared including all reagents and dH_2O to a final volume of 30 µl for every 50 µl reaction and aliquot into each tube accordingly (Table 3.2).

Table 3.2
Master mix composition for different number of PCR reactions

Reagents	**Volume (µl)**				
	1 reaction	**2 reactions**	**3 reactions**	**4 reactions**	**5 reactions**
50 mM $MgCl_2$	1.5	3	4.5	6	7.5
10× Buffer	5	10	15	20	25
10 mM dNTP Mixture	0.6	1.2	1.8	2.4	3
5 mM Primers (each)	5	10	15	20	25
5 U/µl *Taq* DNA Polymerase	0.6	1.2	1.8	2.4	3
dH_2O	12.3	24.6	36.9	49.2	61.5
Total	30	60	90	120	150

The amount of template required for each PCR reaction was optimized in the following manner. Each sample of the total genomic DNA was diluted before its use as PCR template. Several trials with different volumes of the diluted sample were carried out to decide the sufficient amount of DNA template for best results. In general, dilution factors up to 50 times were required to obtain detectable amplification in the subsequent agarose gel electrophoresis. A negative control with no template was set up to monitor possible reagent contamination.

The thermal cycles of the PCR protocol optimized for 27F and 1492R primers were set in the thermal cycler (Bio-Rad, USA) as follows: initial denaturation for 2 min at 95℃ followed by 30 cycles of denaturation for 30 sec at 94℃, annealing for 30 sec at 52℃ and elongation for 1.5 min (+1 sec each round) at 72℃, and final elongation for 5 min at 72℃.

3.4 Agarose Gel Electrophoresis

The resultant PCR products were examined by gel electrophoresis. First, a 1.0 % (w/v) agarose gel (band separation resolution 0.2-25 kb) was prepared in single strength ultra pure grade TAE (0.040 M Tris-Acetate and 0.001 M EDTA) buffer (1stBASE) by heating periodically in a microwave oven. The molten gel mixture was casted in a tray containing the well comb for 20 min until set.

The comb was removed from the set gel carefully without breaking the wells and the tray containing the gel was placed in the corresponding gel tank (SCIE-PLAS, UK). The gel tank was filled with 1× TAE buffer until the wells were completely covered. At this point, the wells were loaded with 5 µl of PCR product mixed with 1 µl of 6× loading dye and the current was run at 90-100 volts for about 40 min. One of the wells was loaded with 5 µl of HyperLadderI (BioLine, London, UK) as molecular weight marker to indicate the sizes of the DNA fragments (Fig. 3.2).

Once the running time was over, the gel was stained in 2.5 µg/ml ethidium bromide (Promega, USA) for 5 min and destained in water for 10 min. The gel was then viewed on a UV transilluminator (Syngene Bio Imaging, England) and the images were captured and saved for later reference.

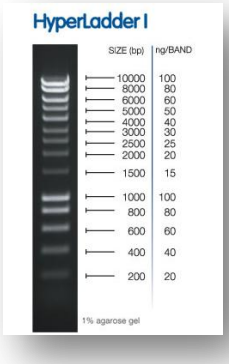


Figure 3.2 Five microlitres HyperLadder I/lane, 1% molecular biology grade agarose in 1× TAE stained with ethidium bromide (*retrieved from: http://www.bioline.com/h_ladderguide.asp, on May 19, 2010*)

3.5 PCR Product Purification

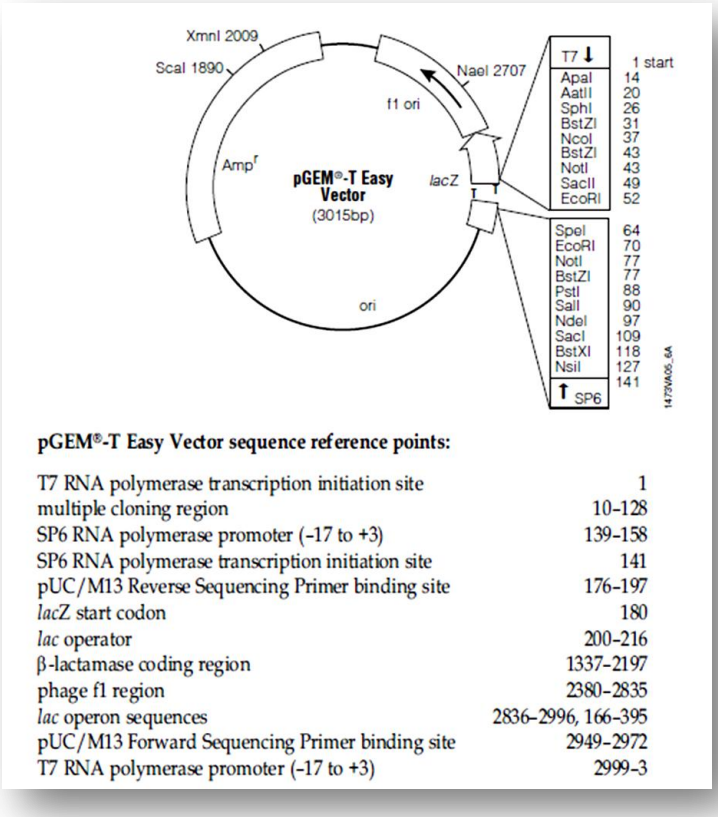
DNA template for cloning and sequencing were purified using the MEGAquick-spin™ PCR and Agarose Gel DNA Extraction System (iNtRON Biotechnology, Korea). A total of 100 µl of each DNA sample amplified previously was transferred to a sterile microcentrifuge tube (Eppendorf AG, Germany).

According to the manufacturer's protocol, a 5× volume of DNA binding BNL buffer was added to the tube and mixed by vortex. The mixture was centrifuged briefly and transferred to a spin column placed inside a collection tube. The spin column assembly was centrifuged for 1 min at 13,000 rpm to bind the DNA to the membrane. The flow-through was discarded and the spin column placed back in the collection tube for next step.

The DNA was then washed with 700 µl of washing buffer and centrifuged at 13,000 rpm for 1 min. The flow-through was discarded by decanting and an extra drying step with the same centrifugation force was carried out to drain the remaining washing buffer. For optimal purification quality, samples were washed one more time with 500 µl of washing buffer before spin-drying. The spin column was then moved to a clean 1.5 ml microcentrifuge tube and labelled accordingly. After 1 min incubation in room temperature, 35-40 µl of elution buffer was added to the membrane and centrifuged. The purified DNA was observed by agarose gel electrophoresis, its concentration verified by spectrophotometry (as explained in section 3.2.1) and was stored in -20℃ freezer for later use.

3.6 16S rRNA Gene Cloning

The isolated 16S rRNA gene from each site was cloned into pGEM®-T Easy Vector System (Fig. 3.3) (Promega, Madison, WI, USA) and competent cells were transformed with this vector. For this purpose, a pure culture of Top10 *E. coli* was grown in a 250 ml Erlenmeyer flask containing 100 ml autoclaved Luria Bertani (LB) broth (stopped with sterile cotton and gauze and covered with aluminium foil) for 16 hours in 37℃ with 160 rpm shaking. After primary incubation period, 1 ml of this culture was transferred to a new flask with fresh LB broth under aseptic conditions and incubated for 3 hours in 37℃ with 180 rpm shaking. At this point, the cells were ready for the cloning procedure (refer to section 3.6.3.1).



pGEM®-T Easy Vector sequence reference points:

T7 RNA polymerase transcription initiation site	1
multiple cloning region	10–128
SP6 RNA polymerase promoter (–17 to +3)	139–158
SP6 RNA polymerase transcription initiation site	141
pUC/M13 Reverse Sequencing Primer binding site	176–197
*lac*Z start codon	180
lac operator	200–216
β-lactamase coding region	1337–2197
phage f1 region	2380–2835
lac operon sequences	2836–2996, 166–395
pUC/M13 Forward Sequencing Primer binding site	2949–2972
T7 RNA polymerase promoter (–17 to +3)	2999–3

Figure 3.3 pGEM®-T Easy Vector map and sequence reference points showing multiple cloning sites (Technical Manual for pGEM®-T and pGEM®-T Easy Vector Systems; *retrieved from: http://www.promega.com/tbs/tm042/tm042.pdf, on May 30, 2010*)

3.6.1 Media preparation

3.6.1.1 Selective Luria Bertani (LB) agar

At the beginning of each cloning session, required amount of LB agar was prepared in a Scott bottle according to the following proportions:

i. Bacto®-tryptone 10 g;
ii. Bacto®-yeast extract 5 g;
iii. NaCl 10 g;
iv. Agar 20 g; and distilled water to the volume of 1000 ml (the same formula minus the agar was used in LB broth preparation).

The medium was autoclaved and set aside to cool down for addition of antibiotics, but not too cool to start coagulating. For achieving optimal blue colony production, 40 mg/ml of X-Gal and 0.1 M IPTG were added to the LB agar along with 100 µg/ml Ampicillin as selective agent. The components were mixed by rolling the bottle gently a few times and poured into sterile plates to a depth of about 1 cm. Once the agar was set, the plates were stored upside down in the 4℃ fridge until later use.

3.6.1.2 Super Optimal Broth with Catabolite Repression (SOC) medium

For preparing SOC media, 1000 ml of LB broth containing 20 g Bacto®-tryptone, 5 g Bacto®-yeast extract, and 0.5 g NaCl was weighed into a Scott bottle. Separately, 0.25 M KCL, 2 M $MgCl_2$ and 1 M glucose were prepared in different universal bottles. All the solutions together with the LB broth were then autoclaved and cooled down. Once cool, the chemicals were added to the LB broth to a final concentration of 2.5 mM KCL, 10 mM $MgCl_2$, and 20 mM glucose (Table 3.3) and mixed well by shaking.

Table 3.3
SOC media chemical components in 1000ml LB broth

Reagents	Volume (ml)
0.25M KCL	10
2M $MgCl_2$	5
1M Glucose	20

3.6.2 Ligation

3.6.2.1 Insert:vector ratio calculation

In order to clone the gene into the vector, first the amount of insert DNA needed was calculated using the following equation (Formula 1):

Formula 1: $x \text{ ng insert} = \frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{insert: vector molar ratio}$

With the given information of 50 ng of 3.0 kb vector, an optimized 2:1 ratio of insert to vector, and the band size of 1.5 kb of DNA insert, the amount of insert required was calculated to be 50.06 ng. According to the OD reading of each purified DNA, the necessary volume of sample DNA for the ligation reaction was calculated as given below (Formula 2) and used in the ligation mixture.

Formula 2: $x\ \mu l\ sample\ DNA = \frac{50.06\ ng}{DNA\ concentration\ ng/\mu l}$

3.6.2.2 Cloning PCR product with pGEM®-T Easy Vector System

The ligation process was carried out in 0.2 ml low DNA-binding tubes under sterile conditions. One positive control for the ligation reaction was set up each time as well as the sample reaction according to the protocol given by company as listed in Table 3.4.

Table 3.4
Ligation reaction set up for cloning PCR product with pGEM®-T Easy Vector System

Reagents	**Sample reaction (µl)**	**Positive control (µl)**
2× Rapid Ligation Buffer	5	5
pGEM®-T Easy Vector (50 ng)	1	1
PCR Product	X*	-
Control Insert DNA	-	2
T4 DNA Ligase (3 Weiss units/µl)	1	1
dH_2O to a final volume of	10	10

*according to calculation
(Promega Technical Manual)

After all reagents were added and mixed by pipetting, the reactions were incubated for 1 hour at room temperature and later, stored at -20℃ for next part of the procedure.

3.6.3 Transformation

3.6.3.1 Cold $CaCl_2$ treatment

Competent cells were made from fresh culture of Top10 *E. coli* as mentioned earlier (see section 3.6). One millilitre of these cells were transferred to a chilled 1.5 ml microcentrifuge tube and placed back on ice for 10 min. The tube was then soft-spin in cold centrifuge (4℃) for 3 min at 4,200 rpm to spin down the cells.

Once the supernatant was removed without disturbing the pellet, 1 ml of cold 0.1 M $CaCl_2$ (previously prepared and autoclaved) was used to resuspend the pellet. This was done by gently pipetting the entire mixture while the tube was kept on ice. The tube was centrifuged again and this step was repeated one more time for maximum competency.

After the second centrifugation and removing the supernatant, the pellet was resuspended in 500 µl of cold 0.1 M $CaCl_2$ for the last time and the tube was buried

deep inside the ice for 30 min. In the meantime, the ligation mix prepared earlier was thawed on ice and centrifuged briefly to collect everything at the bottom of tube.

At this point, 100 µl of the competent cells was added to each of four labelled 1.5 ml microcentrifuge tubes chilled on ice earlier. Four microlitres of the ligation mixture containing the DNA insert was mixed gently with the competent cells of the corresponding microcentrifuge tube. The same was done with the positive control. As a control for transformation efficiency, 1 µl of uncut blue plasmid was mixed with the cells, and as a negative control, 4 µl of dH_2O was added to the last tube. All tubes were put back deep in the ice for 30 min.

3.6.3.2 Heat shock

In order to complete the transformation process, the cells were given a heat shock in a 42℃ water bath for exactly 90 sec. The tubes were then taken out and immediately chilled on ice for about 5 min.

3.6.3.3 Pre-incubation

Eight hundred microlitres of SOC medium (room temperature) was added to each tube and the cells were allowed to grow in 37℃ incubator with 180 rpm shaking for about one hour. The tubes were then centrifuged for 3 min at 10,800 rpm to pellet the cells.

3.6.3.4 Plating

From each tube, 800-850 µl of the supernatant was drained and the pellet was resuspended in the remaining fluid. The entire mixture was spread on the selective agar plates (room temperature) using a hockey-stick spreader under aseptic environment. The plates were labeled accordingly and incubated overnight in 37℃ after sealing properly with parafilm.

3.6.4 Screening

After 14-16 hours incubation, the plates were examined for positive insert by blue/white screening. In case of successful cloning outcome, the colonies were further inspected for possession of desired DNA fragment. For this reason, about 100 white colonies from sample plate were picked with sterile toothpicks under aseptic conditions and transferred to labelled falcon tubes containing 20 ml autoclaved LB broth and grown overnight in 37℃ with 160 rpm shaking.

3.7 Detection of Positive Insert

3.7.1 SP6 and T7 promoter PCR amplification

All the selected colonies were grown and screened for the correct DNA insert by colony PCR amplifications using PCR primers SP6 (5'-GAT TTA GGT GAC ACT ATA GAA-3') and T7 (5'-TAA TAC GAC TCA CAT TAG GG-3') (1stBASE, Malaysia). The PCR reactions used were as described as in section 3.3.

Final PCR reaction contained 1 µl of liquid culture of the cloned cells prepared earlier (see section 3.6.4) as template, 15 µl of master mix (as explained in section 3.3) and 9 µl of dH_2O. The PCR protocol optimized for SP6 and T7 was set as: initial lysis of cells and DNA denaturation for 10 min at 95℃, followed by 25 cycles of denaturation for 30 sec at 94℃, annealing for 30 sec at 60℃ and elongation for 2 min at 72℃, and one final elongation for 10 min at 72℃. The products were then inspected by gel electrophoresis to find positive inserts.

3.7.2 Glycerol stock

The samples showing the right size band were considered cloned and stored in -80℃ freezer for later use. To ensure maximum vitality of the cells, a glycerol stock of the cell cultures was made. Five hundred microlitres of each cloned culture was mixed well with 500 µl of sterile 50% glycerol in a 1.5 ml microcentrifuge tube, labelled accordingly and sealed with parafilm before being stored in -80℃ freezer. In the meantime, the PCR products showing the correct DNA insertion were selected for RFLP process.

3.8 Restriction Fragment Length Polymorphism (RFLP)

3.8.1 Restriction enzyme digestion with HaeIII and HhaI

The PCR products from clone libraries were screened for similarity via RFLP. Two restriction enzymes (RE) were used to digest the amplified DNA insert: *Hae*III and *Hha*I (Promega, Madison, WI, USA). Following the instruction given by the manufacturer (Table 3.5), an optimized reaction was set up.

Table 3.5
Restriction enzyme reaction preparation

Reagents	**Volume (µl)**
RE 10× Buffer C (100 mM TrisHCl (pH 7.9), 500 mM NaCl, 100 mM $MgCl_2$ and 10 mM DTT at 37℃)	2
Acetylated BSA (Bovine Serum Albumin) , 10 µg/µl	0.2
DNA Substrate, 1 µg/µl	5, 10, or 15*
RE *Hae*III, 10 u/µl	0.5
RE *Hha*I, 10 u/µl	0.5
Sterile dH_2O to final volume of	20

*depending on DNA concentration detected by agarose gel electrophoresis

As with other PCR reactions, a master mix of all reagents minus the DNA for wanted number of reactions was prepared and aliquot to separate tubes. The final RE reaction mixture consisted of 15, 10, or 5 µl of master mix along with 5, 10, or 15 µl of substrate DNA respectively. The volumes were decided according to the quality of DNA as detected by band intensity in an agarose gel electrophoresis. Additionally, one positive control was set up in the same way using 50 ng of lambda DNA to monitor the activity of RE enzymes.

For optimal enzymatic activity, the tubes were incubated in thermal cycler set on 37℃ for 4 hours, followed by denaturation of the enzymes in 65℃ for 20 min. At the end, the RFLP patterns were examined in 1.5% (w/v) agarose gel ran for 45 min at 90 V. Gel pictures were used as reference for categorizing fingerprint patterns from the clones of all three sites.

3.8.2 RFLP analysis

Fingerprint profiles were compared for similarity and grouped into unique categories using the program GelComparII (version 6.0 Applied Maths, Kortrijk, Belgium). A 14 day evaluation version was generously provided by the company. All the RFLP gel pictures were uploaded to the software and normalized based on the DNA weight marker. The bands were specified automatically according to band intensity and number of bands present. Further manual configurations of bands were done where necessary.

The individual patterns from each clone library were grouped together in a fingerprint type comparison by calculating a cluster analysis using Dice similarity coefficient with 7% optimization and a band matching tolerance of 2% (automatically set by the software). An overall comparison of all patterns was performed to find the overlapping fingerprints between the three sites and composite a final list of unique

phylotypes for sequencing. In case of ambiguities, fingerprints were matched manually if possible according to the original gel pictures.

3.9 Sequencing

The unique RFLP patterns (phylotypes) found earlier were indicative of possible different bacteria. Therefore, the corresponding 16S rRNA gene was amplified one last time to be sequenced. For this purpose, the remaining product from primary colony PCR was diluted 25× and used as template in 16S rDNA PCR amplification. In this method, 2 µl of diluted PCR product was added in the reaction mixture with 27F and 1492R primers (refer to section 3.3). After satisfactory band detection, the PCR product was purified and the concentration obtained by spectrophotometry. The samples were then sent for single pass DNA sequencing to 1stBASE Commercial Laboratories, Malaysia, along with 10 mM forward primer 27F.

3.10 Data Analysis

3.10.1 Phylogenetic analysis

The sequence results were uploaded to NCBI (National Center for Biotechnology Information, www.ncbi.nlm.nih.gov) website and a BLAST (Basic Local Alignment Search Tool) search was performed to find the closest match with highest scores to known sequences in GenBank. For identification purposes, matching sequences with highest scores were selected from GenBank along with the first known source (i.e. with higher level of classification). Supplementary sequences with lower scores were occasionally selected to better interpret the phylogenetic relationships of the individuals. Sequences chosen from GenBank together with the subject sequences were

exported to the BioEdit software version 7.0.5.3 (Hall, 1999) to generate a multiple sequence alignment using ClustalW function. The alignment was further manipulated and edited manually by removing gaps and omitting ambiguities.

The trimmed sequences were then checked for similarity using the online program FastGroupII from http://biome.sdsu.edu/fastgroup/ website (Yu *et al.*, 2006). The criteria were set for PSI (percentage sequence identity) with gaps of 97% similarity. Any two or more sequences grouped together were considered as one and treated under the same OTU (operational taxonomic unit) throughout the calculations. All the remaining sequences were checked for chimeras using the online version of Pintail tool (http://www.bioinformatics-toolkit.org/Web-Pintail/) from Bioinformatics Toolkit website (Ashelford, 2007).

Partial 16S rRNA gene sequences along with sequences obtained from GenBank were uploaded to the online Ribosomal Database Project II (RDP Release 10). Primarily all the sequences were checked for phylogenetic affiliations in the Classifier tool based on naïve Bayesian rRNA classification (Wang *et al.*, 2007). Further inspection of taxonomic affiliations of obtained sequences was enabled by the construction of a phylogenetic tree using MEGA4 software (Tamura *et al.*, 2007). The program uses Neighbor-Joining method to infer evolutionary history (Saitou and Nei, 1987). The evolutionary distances are computed using the Maximum Composite Likelihood method (Tamura *et al.*, 2004) and are in the units of the number of base substitutions per site. All positions containing gaps and missing data are automatically eliminated from the dataset by the program (complete deletion option).

3.10.2 Statistical analysis

Using the clone library data, once confirmed by sequencing, the richness and diversity of phylotypes were estimated to characterize bacterial community composition. For this reason, phylotype abundances were separately recorded for each clone library based on the sampled clones. Clone library coverage was determined using Good's coverage (Kemp and Aller, 2004). Clone libraries were compared to each other for similarity using PAST software package (Hammer, 2001). Alpha diversity indices (Shannon's index *H'*, Simpson's diversity index *D*, and Pielou's evenness *J'*) were calculated from the same program. For a similarity comparison between the three sites, Sorensen's similarity index was calculated. Also, to evaluate the level of differentiation between the sites, a non-metric multidimensional scaling (nMDS) was constructed by grouping the data according to observed phylotypes in each clone library using Bray-Curtis similarity measure.

The list of the phylotypes observed in each site along with their abundance was used to develop a data set for phylotype richness estimation using EstimateS software version 8.2.0 (Colwell, 2009). Each clone was considered a sample and each phylotype was treated as a species, in a format where rows represent species and columns represent samples. All the calculations were carried out by the software with 50 runs of sample randomization without replacement. Total phylotype richness was determined for each clone library using S_{Chao1} estimator (Kemp and Aller, 2004).