CHAPTER 2.0: LITERATURE REVIEW

2.1 Bacterial diversity in soil

As we know, soil is the outer covering on earth, which is composed of both organic and inorganic compounds. Therefore, the diversity of microorganisms that survive naturally in soil is much higher than those found in other sources on earth (Atlas and Bartha, 1997). In the studies of soil microbes, Pace and co-workers found that prokaryotic evolution is divided into two major branches, which are bacteria and Archaea (Pace, 1997). Studies have also revealed that 1 g of soil contains more than 2.6 X 10^{29} prokaryotes and may contain up to 4000 different species of bacteria (Whitman *et al*., 1998; Torsvik *et al*., 1990). This indicates that bacteria are the dominant group of microorganisms in soil.

Contributions of soil bacteria in stabilizing the ecosystem are well known (Dubey *et al*., 2006). They are able to catalyze the transformation of organic matter from plants and animals into humus (Whitman *et al*., 1998; O'Donnell *et al*., 2001). Thus, they played essential roles in various biogeochemical cycles processes and nitrogen fixation (Weber *et al*., 2001), which are responsible for the recycling of organic compounds that are essential for sustainable agriculture (Molin and Molin, 1997; Yao *et al*., 2000). Hence, the occurrence of soil and sediment bacteria, and the studies on these bacteria with their ecosystem effects are important.

However, only a small fraction of bacteria (approximately 1 %) could be successfully cultured in the laboratory (Hugenholtz and Pace, 1996; Mclnerney *et al*., 2001), possibly due to the limitations of culture-dependent methods (Torsvik *et al*., 1990; Amann *et al*., 1995). This indicates that the huge bacterial diversity in soil remains unexplored (Hammond, 1995). Therefore, the studies of bacterial diversity in soil are necessary for a better understanding of nature.

2.2 Bacterial diversity studies in the Arctic

Despite being subject to extreme environmental stress, the Arctic and Antarctica provide some living space for numerous organisms that adaptable for freezing (Davey and Clarke, 1991) and desiccated (Warren, 2001) conditions. Studies have shown that bacteria in the Arctic and Antarctica have important roles for both ecosystem processes and global climate feedbacks (Melillo *et al*., 1995; Weintraub and Schimel, 2003; Wallenstein *et al*., 2007).

Whitman *et al*. (1998) had reported that bacteria were the major biomass component in many marine samples. Their genomes encode functions that reflected the details of biogeochemical cycles had been reported (DeLong and Karl, 2005). For instance, marine bacteria play a significant role in the global cycling of carbon (Jorgensen, 1982; Wollast, 1991; Canfield *et al*., 1993), and demineralise the organic matter from primary production that settles to the sea floor (Ravenschlag *et al*., 1999; Wollast, 1991). In general, marine samples have less microbial diversity than in soil samples (Gans *et al*., 2005), and most of the bacterial community found in marine samples were related to sulphate reducing bacteria (Sahm and Berninger, 1998; Ravenschlag *et al*., 1999). Besides that, Ravenschlag *et al*. (2001) had reported that *δ-proteobacteria* and *Bacteroidetes* were the dominant bacteria in Svalbard marine sediments. Decreasing in bacterial diversity with increasing depth was also observed in marine samples, as described in Wadden Sea (Llobet Brossa *et al*., 1998) and Svalbard marine sediments (Sahm *et al*., 1998).

Arctic tundra was reported to have potential source of extensive methane emission due to climatic warming because of its large reservoirs of stored organic carbon (Post *et al*., 1982; Melillo *et al*., 1995; Wartiainen *et al*., 2003; Wallenstein *et al*., 2007). Higher temperatures due to climatic warming would cause an increase in methane production, and thus the abundance of methane-oxidizing bacteria in the arctic soil could be important

regulators to reduce methane emission to the atmosphere (Wartiainen *et al*., 2003). High bacterial diversity was reported in terrestrial samples (Zhou *et al*., 1997). The dominant bacteria in terrestrial samples were *Acidobacteria* (Nemergut *et al*., 2005; Jones *et al*., 2009), *β-proteobacteria* (Zhou *et al*., 1997), *δ-proteobacteria* (Zhou *et al*., 1997; Nemergut *et al*., 2005), *Bacteroidetes* (Nemergut *et al*., 2005), *Clostridia* (Stackebrandt *et al*., 2004) and *Fibrobacteres* (Zhou *et al*., 1997). Kobabe *et al*. (2004) also reported that *β-proteobacteria* and *Bacteroidetes* were dominant in Siberian tundra in their study. Kastovska *et al*. (2005) also reported that *Cyanobacteria* are important oxyphototrophic organisms that inhabiting terrestrial ecosystems both in the Arctic and Antarctica. *Cyanobacteria* play an indispensable role in soil formation, stabilization against wind and water erosion, increasing soil organic matter and nitrogen content and preparing substrata for the subsequent succession stage of organisms (Wynn Williams, 1993; Belnap and Lange, 2001; Kubeckova *et al*., 2002).

Glacier samples may harbor a wide range of bacterial communities, as reported in the Swiss Alps (Sharp *et al*., 1999), Canadian Arctic and Alaska (Skidmore *et al*., 2000), and Svalbard samples (Wadham *et al*., 2004). Mindl *et al*. (2007) had reported that glacial runoff and goose faeces in the Arctic might stimulated and influenced the bacterial production in proglacier lakes and glacier samples. The dominant bacteria in glacier samples were mainly *β-proteobacteria* and *Bacteroidetes* (Skidmore *et al*., 2005). Kastovska *et al*. (2005) reported that rich assemblages of *Cyanobacteria* were also observed in glacier samples in their analysis of microbial community from non-vegetated and plant coverage glacier from Ny-Ålesund. On the other hand, Kastovska *et al*. (2007) also reported glacier samples from southwest Spitsbergen, Svalbard were dominated by *Cyanobacteria*.

The bacterial diversity of lake samples was low, due to phosphorus limitation that as explained by Carlsson and Caron (2001) and Graneli *et al*. (2004). The dominant bacteria in lake samples were *β-proteobacteria* and *Actinobacteria*, as reported by Lindstrom *et al*. (2005) in their analysis of bacterial diversity in Sweden and Ny-Ålesund. However, *Cyanobacteria* were not observed in the lake samples (Lindstrom *et al*., 2005). Although many studies had been conducted for Arctic samples, none of these studies was conducted to compare the bacterial diversity between different type of soils and sediments in the Arctic.

Despite bacterial diversity analysis in samples, several studies on the correlation of Arctic bacterial diversity to their environmental variables were also reported. For instance, Lindstrom *et al*. (2005) had correlated the lake pH with the bacterial diversity and suggesting that different taxa could be selected in acidic and alkaline environments. Fierer and Jackson (2006) had reported low diversity was observed in acidic soils. The lower pH of soils might favor fungal populations (Baath and Anderson, 2003), leading to increased competition between fungal and bacterial populations. Stepanauskas *et al*. (2003) and Yannarell and Triplett (2005) also correlated pH with the bacterial diversity in their samples studied. As for salinity, Crump *et al*. (2004) had analysed the bacterial diversity from three samples varied in salinity (marine, freshwater, and estuarine) using DGGE. Giovannoni and Rappe (2000) also reported the influences of salinity in their marine bacterial communities.

2.3 Methods for bacterial diversity studies

Studies of soil bacterial diversity enable the discovery of relationships between bacterial communities and their surroundings. In the past, several methods have been used to isolate and identify bacteria species from environmental samples. Mainly, these methods

are classified into two approaches, conventional culture-dependent methods and culture-independent molecular approaches.

2.3.1 Conventional culture-dependent methods

In conventional culture-dependent methods, bacterial diversity is studied by the cultivation of bacteria on solid media or agar plate. The development of this method has led to the study of bacteria as individual types according to their nutritional criteria and the morphology examinations through microscope (Torsvik *et al*., 1990). However, this method has many limitations and the results can be quite biased because most of the bacteria cannot be cultured using standard techniques in laboratory (Amann *et al*., 1995).

First of all, there is a big challenge to mimic the growth condition of some bacteria, especially those bacteria living in extreme habitats, such as in the cold and hot regions (Lorenz and Schleper, 2002). Besides that, there is only a mere 0.3 % of bacteria cultivability in soil samples (Torsvik *et al*., 1990). The term "cultivability" was determined as the percentage of cultivable bacteria in comparison with total cell counts (Amann *et al*., 1995). This indicates the conventional culture-dependent methods (media plate) do not reveal the entire bacterial community of a habitat (Handelsman, 2004), but only allow certain bacterial species to grow on the media plates. For instance, *Cyanobacteria* do not or hardly grow on media (Antoon *et al*., 1994). This could introduce a serious bias to bacterial diversity studies (Ferris and Ward, 1997) because bacterial community that obtained from media plates is the one that could grow on the media and is dependent on the media.

There are also difficulties in analyzing the bacterial diversity using conventional culture-dependent methods. As we know, different species of bacteria require different media conditions to grow. There would be a loss of major portions of bacterial communities if the handling techniques (agar plate spread as an example) are not properly

administered (Amann *et al.*, 1995; Ogram and Feng, 1997). Hence, the results given are biased and might not be true to reveal the whole bacterial diversity of a habitat.

As for the observation of bacterial morphology using a microscope in the conventional culture-dependent method, there are also several limitations in describing the taxonomic diversity due to the simple bacterial morphology (Madigan *et al*., 1999). Usually, bacterial communities are the most difficult to phenotypically analyse as compared to other organisms. This might be due to their high level of diversity among the most complex assemblages in the biosphere (Zhou *et al*., 2003). Furthermore, the majority of the viable microscopically bacteria could not be cultured on plates (Staley and Konopka, 1985; Roszak and Colwell, 1987), therefore increasing the difficulty of phenotypical analysis of bacterial diversity.

Conventional culture-dependent methods that employ various media, incubation conditions, and preservation of bacterial colonies on plates, are time-consuming and labour-intensive (Atlas and Bartha, 1997). Therefore, culture-independent molecular approaches, which analyze bacteria based on their DNA, have been introduced to overcome the difficulties and limitations in conventional culture-dependent methods to study bacterial diversity in environmental samples (Steffan and Atlas, 1988).

2.3.2 Culture-independent molecular approaches

Since the last two decades, culture-independent molecular approaches, which analyse bacteria based on their genetic contents, have been widely used to identify bacterial species without the need of bacterial cultivation on agar plate (Amann *et al*., 1995; Hugenholtz and Pace, 1996).

As in culture-independent molecular approach, a universal and highly conserved and small subunit ribosomal ribonucleic acids (SSU rRNA) gene sequence in prokaryotes:

5S rDNA or 16S rDNA, which act as a culture-independent tool, has been studied to determine the bacterial diversity in environmental samples (Stahl *et al*., 1984; Olsen *et al*., 1986; Ward *et al*., 1992; Amann *et al*., 1995).

This approach begins with the lysis of bacterial cells from environmental samples, followed by the cloning of the universal and highly conserved rRNA gene sequences using Polymerase Chain Reaction (PCR). Based on the studies of this specific rRNA gene sequence, phylogenetic trees can be constructed to study the bacterial diversity of environmental samples (Pace, 1997; Hackl *et al*., 2004).

With the advancement of molecular techniques, the development of PCR has overcome the limitations of culture-dependent methods. It is a time saving modified gene cloning method that is able to clone genes in a single test tube by mixing DNA with a set of reagents and carry out the reaction in a thermal cycler (Brown, 2001). Hence, it is an useful and important tool in studying bacterial diversity based on culture-independent molecular approaches. However, the PCR bias, which is usually caused by lysis efficiency of bacterial cells, could affect the results of bacterial diversity studies. Hence, the lysis efficiency of bacterial cells that varies between and within microbial groups should be concerned, to avoid PCR bias (Prosser, 2002). Besides that, the purity (humic acid content) of extracted DNA should be concerned. This is because humic acid contaminated DNA sample could interfere with the reaction of *Taq* DNA polymerase in PCR, hence resulting in a poor PCR product (Tebbe and Vahjen, 1993; Neef *et al*., 1998).

2.4 Extraction of genomic DNA from environmental samples

The culture-independent molecular approaches have become the preferred method of researchers in studying environmental samples, including sea water (Somerville *et al*., 1989), fresh water (Steffan *et al*., 1989), sediments (Steffan *et al*., 1988), and soils (Torsvik

et al., 1990). Techniques in improving the lysis efficiency of bacterial cells from environmental samples, which could extract better purity of DNA to overcome the PCR bias, have been undertaken (Daniel, 2004).

Since soil bacteria exist in or on the surface of soil, the ability to separate these bacterial cells from soil components during cell extraction is vital for studying biodiversity (Trevors, 1998). If the method of cell extraction used is too gentle, only Gram-negative, but not Gram-positive, bacterial cells would be lysed. On the other hand, if the method is too harsh, both Gram-negative and Gram-positive cells may be lysed but their DNA may be degraded and become shared (Wintzingerode *et al*., 1997). Both of these lysis conditions will lead to PCR bias.

Besides that, soil genomic DNA is often co-extracted with humic acids, polyphenolics, and other contaminants. These co-extracted substances will interfere with the reaction of *Taq* polymerase in PCR (Tsai and Olson, 1992; Tebbe and Vahjen, 1993; Neef *et al*., 1998). Therefore, DNA extraction method is the most important criteria to be concerned in order to isolate high quality genomic.

There are two types of DNA extraction method being used to isolate bacterial DNA from environmental samples, which are conventional DNA extraction using chemical lysis (Zhou *et al*., 1996) or physical disruption (Frostegard *et al*., 1999), and DNA extraction kit with the use of spin columns (Yeates *et al*., 1998). The chosen DNA extraction method must be appropriate to the environmental conditions of sample (Zhou *et al*., 1996) because microbial diversity is different in different samples.

Moreover, different methods of DNA extractions will result in different yields of PCR product (Wintzingerode *et al*., 1997), due to DNA extraction which could be influencing the performance of subsequent molecular manipulations (Yeates *et al*., 1998). For instance, high purity of extracted DNA without contaminants (humic acid content) is

essential for a successful PCR amplification (Moreira, 1998). Harsh extraction techniques in conventional DNA extraction methods of physical disruption, such as bead beating, can degrade the bacterial DNA, leading to sharing and smearing problems in PCR detection (Wintzingerode *et al*., 1997; Frostegard *et al*., 1999), although higher DNA yield is extracted (Maarit Niemi *et al*., 2001). On the other hand, a gentler extraction method of chemical lyses can recover only little yield of genomic DNA (Zhou *et al*., 1996; Maarit Niemi *et al*., 2001).

However, both of these conventional DNA extraction methods hardly remove humic acids, which could inhibit subsequent PCR analysis. To solve this problem, DNA extraction kit with spin column is widely used to remove humic acids. The Ultra Clean Soil DNA Isolation Kit (MO Bio Inc., CA) has been reported, to give consistent, clear bands with the most extensive banding patterns after comparing the use of several DNA extraction methods (Maarit Niemi *et al*., 2001).

2.5 The use of 16S rRNA gene fragment in bacterial diversity studies

The first attempts to analyze environmental samples using culture-independent molecular approach was by studying the rRNA gene fragment, which began in the mid-1980s. The rRNA gene fragment is amplified from genomic DNA of environmental samples using PCR, cloned into a vector and then sequenced in order to identify the bacteria species (Stahl *et al*., 1984).

The 5S rRNA gene fragment, with the length of 120 bp, was the first used to characterize microorganisms in environmental samples (Stahl *et al*., 1984; Amann *et al*., 1995). However, the small size of 5S rRNA gene sequence limited bacterial gene information (Madigan *et al*., 1999).

Later, a larger rRNA molecule, 16S rRNA gene fragment with a fragment length of 1500 bp was then used to analyse bacterial diversity. Highly conserved sequences and sufficient sequence variability in 16S rRNA gene fragment have made these molecules an excellent indicator of relatedness in the studies of bacterial diversity (Madigan *et al*., 2000).

Since the year 1990, 16S rRNA gene fragment has discovered many uncultivable bacteria. A phylogenetic tree of bacteria that contains 36 phyla, 13 of which do not have culture representative was discovered (Hugenholtz *et al*., 1998). Besides that, studies have shown that a bacterium whose 16S rRNA gene fragment differs by more than 3% from similar sequences from other organisms, should be considered as a new species (Madigan *et al*., 1999). To date, the rRNA gene fragment has been established as a unique identity molecular "tag" for microbial studies (Janssen, 2006).

2.6 Denaturing Gradient Gel Electrophoresis (DGGE)

Denaturing Gradient Gel Electrophoresis (DGGE) was originally developed to detect point mutations in DNA sequences which have been used to study cancer (Borresen *et al*., 1938; Pellegata *et al*., 1992) and hereditary diseases (Audrezet *et al*., 1994). After Muyzer and co-workers expanded the use of DGGE in studying microbial diversity (Muyzer *et al*., 1993), DGGE has been widely used to examine the changes in microbial communities in Arctic (Ferrari and Hollibaugh, 1999; Sahm *et al*., 1999; Junge *et al*., 2002; Neufeld and Mohn, 2005; Mindl *et al*., 2007; Perreault *et al*., 2007).

Since PCR products from a given reaction are of similar size, conventional separation by agarose gel electrophoresis results only in a single DNA band that is largely non-descriptive. However, DGGE can overcome this limitation by separating PCR products based on sequence differences, which could results in differential denaturing characteristics of the bacteria DNA (Hovig *et al*., 1991).

During DGGE, PCR products come across increasingly higher concentrations of chemical denaturant: urea and formamide, as they migrate through a polyacrylamide gel (Muyzer *et al*., 1993). When reaching a threshold denaturant concentration, the weaker melting domains of the double-stranded PCR product will begin to denature, at which time migration slows dramatically (Muyzer *et al*., 1993).

The differential migration occurs because more denaturant is needed to separate the sequence with higher guanine (G) and cytosine (C) content due to the differences in the number of hydrogen bonds that hold the complementary strand of DNA together (Sheffield *et al*., 1989). As we know, there are three hydrogen bonds between guanine (G) and cytosine (C), but only two between adenine (A) and thymine (T). Thus, as the DNA strands separate, their migration becomes retarded in the polyacrylamide gel, as a distinct and different band.

In order to get sharp bands, the sequence migration is stabilized by adding high GC region to the end of one PCR primer (Sheffield *et al*., 1989). Therefore, the amplified 16S rRNA gene sequence from PCR is preceded to a nested PCR with GC rich primer pair, which the 5'-end of the forward primer contains a 35–40 base pair GC clamp, to prevent a complete strand separation during DGGE.

Theoretically, DGGE can separate DNA up to one base-pair difference (Wu *et al*., 1998; Miller *et al*., 1999). Different sequences of DNA which occur in different bacteria will denature at different denaturant concentrations resulting in a pattern of bands. Each band theoretically represents a different bacterial population that is present in the community whereas, the band intensity theoretically represents the abundance of the bacterial population (Muyzer *et al*., 1995). However, recent study also reported that a gene fragment of closely related bacteria might be separated into two distinct bands (Jackson *et al*., 2001). Besides that, artificial bands that are probably induced by heteroduplex

molecules might also be detected in DGGE profile during analysis of complex DNA (Ferris and Ward, 1997).

Once a DGGE gel is generated, the fingerprints of the DGGE gel acts as a "snapshot" of the whole bacterial community of the sample (Muyzer *et al*., 1995). The fingerprints can then be uploaded into databases in which fingerprint similarity can be assessed statistically by multidimensional scaling (MDS), hierarchical cluster analysis, and diversity index evaluation; to determine microbial structural differences between environments (Boon *et al*., 2002).

Since DGGE does not depend on cultivation, DGGE has revealed greater biodiversity than previously uncovered traditional culture-dependent methods (Wise *et al*., 1999). DGGE has the advantages of being reliable and reproducible (Muyzer, 1999), due to an identical sample that is loaded on a gel can display identical fingerprints (Simpson *et al*., 1999). Multiple samples can be analyzed concurrently, making it possible to follow changes in microbial populations. Meanwhile, DGGE is a time saving molecular and inexpensive approach in studying environmental samples (Nakatsu, 2007), compared to other culture-independent molecular approaches such as the method of cloning and sequencing of multiple plaque samples (Paster *et al*., 2001). Therefore, this approach allows a comparison of different bacteria communities from many studied sites in only one gel and in very short time.

2.7 Statistical analysis of DGGE profiles

There are two types of analysis of DGGE profiles, are which is based on single bands and the other whole fingerprints (Fromin *et al*., 2002). Once the DGGE gel is generated, the banding pattern of the fingerprints is aligned and standardized by a computer-assisted characterization (Rademaker *et al*., 1999), before profile analysis can be

carried out. Markers that are generated in the previous DGGE run which have distinct melting position can be used to normalise the banding pattern from different DGGE profiles (Chong *et al*., 2009a).

Analysis of DGGE profiles that are based on single bands, observed the variation of the presence/absence and the intensity of a single band (Murray *et al*., 1996). The presence of bands can be correlated with several biological and chemical parameters (Widmer *et al*., 2001), such as dependent quantitative variable (comparison of band intensity) to study the abundance of bacteria, and in dependent quantitative variable (comparison of presence/absence band) is able to correlate the bacterial diversity with environmental variables (Fromin *et al*., 2002). A well-defined DGGE band from the gels can also be identified by sequencing (Kowalchuk *et al*., 1997).

Analysis of DGGE profiles that are based on whole fingerprints compare the total number of present band accumulated, which are related to the number of dominant phylotypes or known as diversity (Muller *et al*., 2001). The Shannon diversity index (H') can be calculated to analyses the dominant bacteria between samples (Nubel *et al*., 1999). The index is positively correlated with species richness and evenness, and thus gives more weight per individual to rare species than to common species (Magurran, 1988; Hill *et al*., 2003).

In order to study the similarity of bacterial diversity between samples, the use of Non-metric Multidimensional Scaling (nMDS) plots (Iwamoto *et al*., 2000) and hierarchical cluster analysis (Boon *et al*., 2002) are usually involved. nMDS is an ordination method that transformed the DGGE patterns into a two-dimensional space, with similar objects close to one another and different objects far apart (Fromin *et al*., 2002). Hierarchical cluster analysis targets group samples having similar banding pattern together (Boon *et al*., 2002). In order to study the correlation of bacterial diversity with

environmental variables, The Biota and Environment matching (BEST) procedure was suggested by Clarke and Gorley (2006).

Therefore, the complementation of DGGE profiles with statistical analysis leads to a more precise description and better understanding of microbial diversity and their ecological niches (Fromin *et al*., 2002).