Microbial Community Analysis Using Pyrosequencing

Jongsik Chun[1] *, Jae-Hak Lee[1], Ok-Sun Kim[1], Hyun-Soo Na[1], Byung-Kwon Kim[1],
Suk-Hwan Yoon[1], Kihyun Lee[1], Seon Joo Ahn[1], You-Seak Go[2], and Jeong-Sun Seo[2]

[1]*School of Biological Sciences and Institute of Microbiology, Seoul National University, Seoul 151-742*
[2]*Macrogen, Inc., Genomic Analysis Division, 10F, World Medridian Center, Seoul 153-781*

DNA sequencing can be used to gain important information on genes, genetic variation and gene function for biological and medical studies. First described by Sanger et al. in 1977, dideoxynucleotide sequencing of DNA has served as the major power horse in modern biology. Even though the technique has constantly improved over years, such as the use of fluorescent dye terminators, capillary electrophoresis and automated process, its low throughput and high cost restrict most of researcher from using DNA sequencing in various biological applications. Recently, several new DNA sequencing methods have been developed or currently under development for commercialization. These techniques, called next generation sequencing (NGS), have ability to process millions of sequence reads in parallel rather than 96/384 at a time, eliminating conventional vector-based cloning and *Escherichia coli*–based amplification stages.

Among the readily available NGS technologies, pyrosequencing, developed by 454 Life Sciences, has proved to be especially valuable in microbial ecological applications, such as whole metagenome shot-gun sequencing and massive sequencing of PCR amplicons of ecological interests. In this presentation, we will focus on the development and critical evaluation of microbial community analysis using pyrosequencing, especially based on currently available Roche GS FLX Titanium machine.

The bioinformatic pipeline (Fig. 1) was developed using JAVA and tested for samples isolated from various environments. We have constructed a hierarchical taxonomic system for bacteria, purely based on 16S rRNA sequences, and used for taxonomic assignment of assembled pyrosequencing reads. The resultant microbial community structures can be viewed and analyzed by the microbial community viewer (MCViewer) program, written in JAVA. Here, we argue that massive sequencing of 16S rRNA gene amplicons generated by Roche GS FLX Titanium is suitable for bacterial community structure analysis.
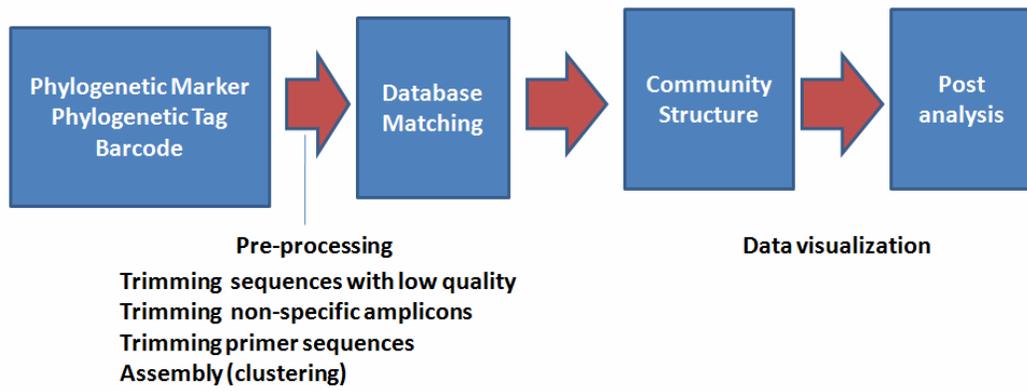
Figure 1. Proposed bioinformatic pipeline for microbial community analysis using pyro-barcodes

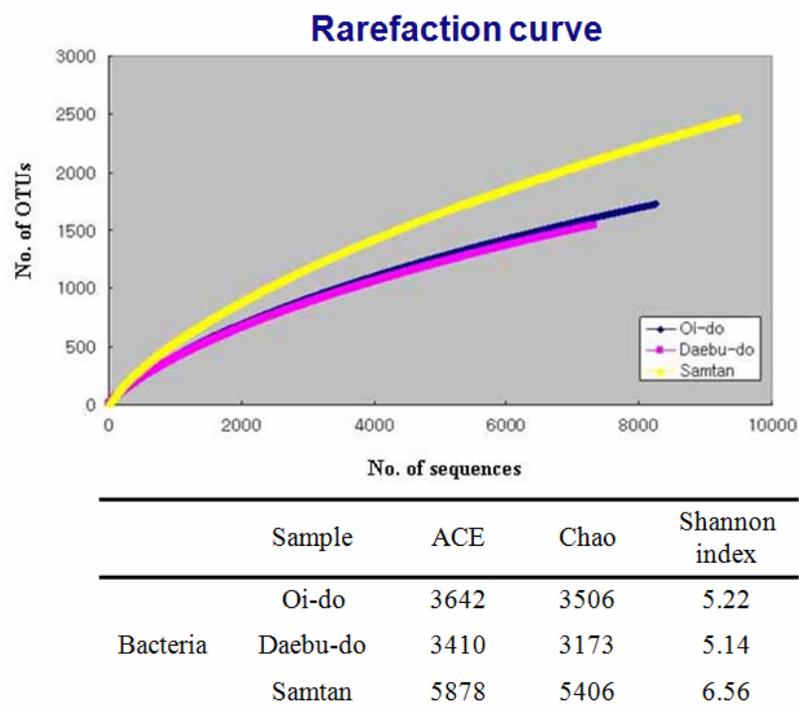|	Sample	ACE	Chao	Shannon index
	Oi-do	3642	3506	5.22
Bacteria	Daebu-do	3410	3173	5.14
	Samtan	5878	5406	6.56

Figure 2. Rarefaction curves of tidal flat sediments and forest soil.