Advancing Science with DNA Sequence

PyroTag: Pyrosequencing as a Powerful Tool to Study the Diversity of Microbial Community

Feng Chen, Anna Engelbrektson, Natasha Zvenigorodsky, Howard Ochman (University of Arizona), and Philip Hugenholt

DOE JOINT GENOME INSTITUTE
US DEPARTMENT OF ENERGY
OFFICE OF SCIENCE

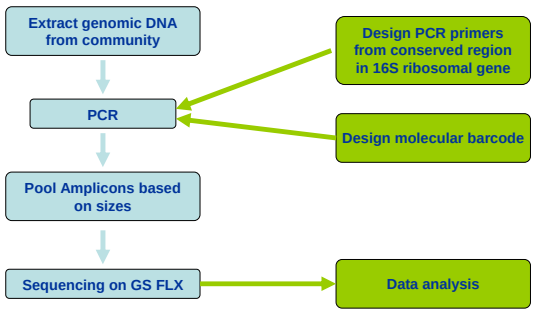
Introduction

At JGI, major efforts have been spent on using the 16S RNA sequencing to study the composition of bacterial community. Currently, 16S rRNA genes from a microbial community are amplified with conserved PCR primers and the mixed amplicon is cloned and sequenced using traditional Sanger sequencing method. Using Roche GS FLX amplicon sequencing method, individual amplicon can be sequenced directly on the platform, producing hundreds of thousand ~250 bp tags of the full length 16S rRNA gene. We call this pipeline PyroTag. The benefit of the PyroTag technology includes higher resolution and less bias than traditional PCR and cloning based sequencing. The disadvantage is that the read length is short as we used FLX reagent kit for the study.

The aim for this study is to systematically test the effect of amplicon length and 16S rRNA regions targeted on diversity coverage of complex microbial community. The goal is to develop a standard PyroTag pipeline to provide to broad range JGI users.

Molecular barcodes have been used to distinguish different amplicon lengths and different targeted 16S regions. In this study, we use the bacterial community from termite gut as the targeted environmental sample. Strategy and detailed experiment design as well as preliminary results will be presented.

Experiment Design



Contact information:
Feng Chen, Ph.D.
US DOE Joint Genome Institute
2800 Mitchell Dr. B400
Walnut Creek, CA 94598
925 296 5733
fchen@lbl.gov

This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program and the by the University of California, Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231 and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396.

Method

Genomic DNA extraction:

- ❑ Termite hind guts were collected on the site (Costa Rica)
- ❑ Community genomic DNA was extracted by regular phenol-chloroform extraction method
- ❑ Same pool of genomic DNA was used for all 16S region PCR amplifications
- ❑ Diluted genomic DNA (5ng/ul) was used for PCR

PCR and pooling:

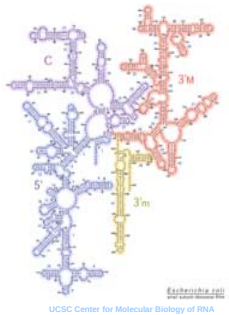
- ❑ 25 cycles of optimized PCR reaction
- ❑ Standard PCR clean-up
- ❑ PCR product quantification
- ❑ Same number of molecules were pooled
  - Similar length or different length

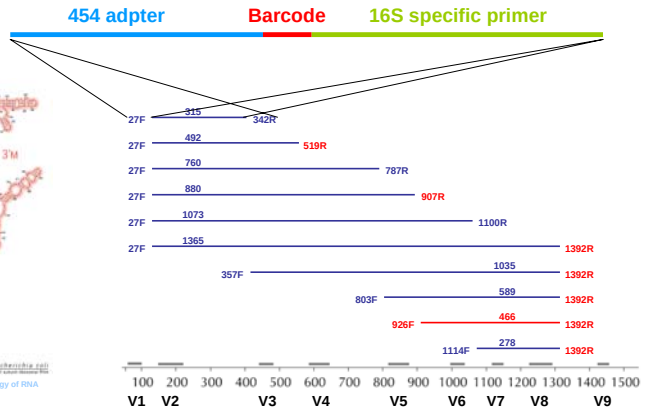
454 sequencing:

- ❑ Standard Amplicon sequencing on GS FLX

PCR primer and molecular barcode design:

- ❑ Unmodified 454 A and B primer sequence
- ❑ Molecular barcode after TCAG key
- ❑ 16S specific forward and reverse primer after barcode





Data Analysis pipeline (under development):

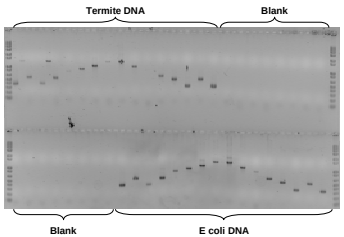
Raw data ⇨ QC filtering ⇨ Assembly ⇨ Sort by sample specific tags ⇨ Locate PCR primers ⇨ Locate high Quality span ⇨ Recruit set of contigs to carry forward ⇨

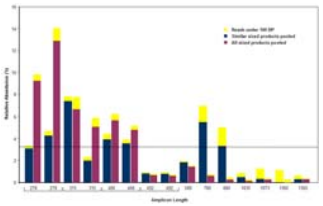
Dereplicate ⇨ Choose proxy for each cluster ⇨ Classification ⇨ Tree construction ⇨ Intra-sample diversity discovery ⇨ Inter-sample diversity discovery

Results

PCR product size validation:

All amplicon sizes worked well under our PCR condition





- ❑ emPCR biased against long fragments when different sizes were mixed
- ❑ Amplicon sizes up to 900 bp worked well
- ❑ Similar sized fragments can be pooled in emPCR stage

Conclusion

PyroTag is a powerful approach to study bacterial community diversity. It tremendously reduces the workload and provides high resolution. It is also less biased than the traditional cloning/sequencing approach. Molecular barcodes provide a effective way to study multiple communities in one experiment setting. Different hypervariable regions can be targeted separately or together although the amplicon size limit seems to be at 900 bp.

We are testing the Titanium platform for this purpose. The long read length of the platform should provide more flexibility.

Office of Science
U.S. DEPARTMENT OF ENERGY
University of California Lawrence Livermore National Laboratory
Los Alamos NATIONAL LABORATORY