An Efficient Method for Sequencing Microbial Communities

Jessica Kuo, Sergey Nuzhdin, Joseph P. Dunham, Maren Friesen
Department of Molecular and Computational Biology
University of Southern California

Introduction

High-throughput DNA sequencing generates large volumes of data at a low cost, essential for genetic studies.

Metagenomic and evolutionary genomic studies currently target species diversity and interactions within microecosystems, but require thousands of samples to be sequenced at once.

High-throughput sequencing determined the diversity of bacteria found in the human gut

Our newly developed technology increases the cost efficiency of sequencing. Modified barcodes ligate to both PCR amplicons and Illumina adaptors, making fewer primers necessary, and yet allowing for thousands of samples to be sequenced simultaneously as needed for community characterization.

To test the limits of the technology, a bacterial community containing functional genes found in soil was designed. After decreasing the copies of genome with functional genes proportionate to the total, we expect to see that same decrease in the number of sequence reads.

Microbiome Design

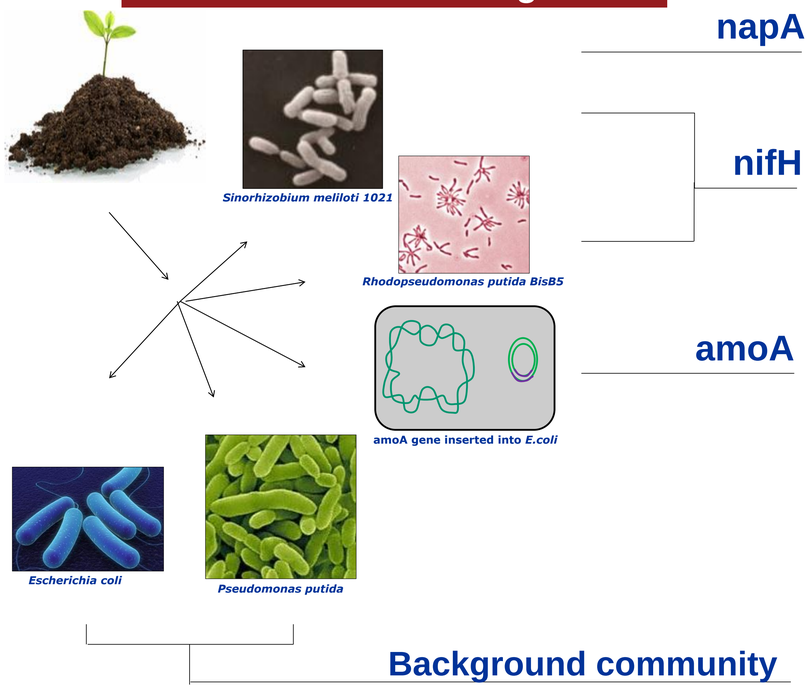
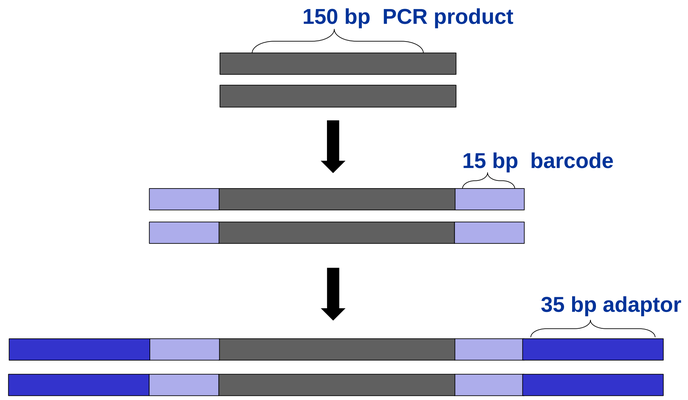


Figure 1. Bacterial community consists of functional genes nifH, napA, and amoA. Bacteria without the genes were added to normalize sequences with 16S reads.

The Technology

Traditional barcoding methods requires a primer to be designed with a unique barcode within. By contrast, our method allows our 8 base barcode to be added directly to a pool of amplicons by the "A" overhang generated by DNA polymerase.

In the same step, Illumina adaptors ligate to the bases at the end of the barcode.



Methods and Materials

In soil, the functional genes nitrogenase reductase (nifH), nitrate reductase (napA), and ammonia monooxygenase (amoA) are genes critical for the nitrogen cycle. Universal primers for the three functional genes were designed, including the 16S rRNA gene used to normalize reads.

Serial dilutions of the bacterial community ranged from 10^0 to 10^{-5} to test at which dilution the technology is unable to detect reads for the functional genes.

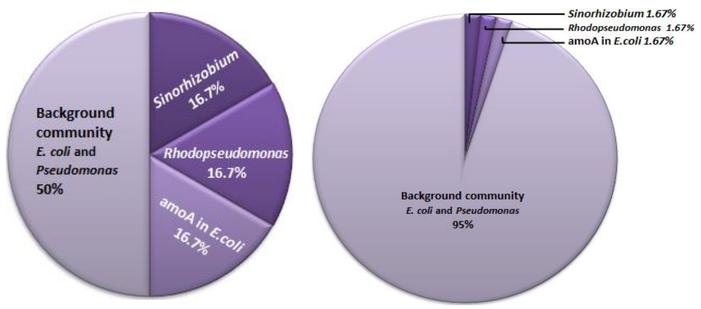


Figure 2: Relative genome proportions for two dilutions. On left shows microbial community with functional genes at 10^0, and on right shows at 10^{-1}.

Results

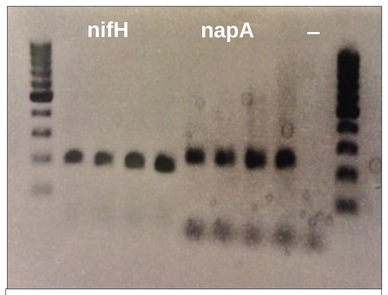


Figure 3: PCR of nifH and napA with 4 soil samples and a negative control

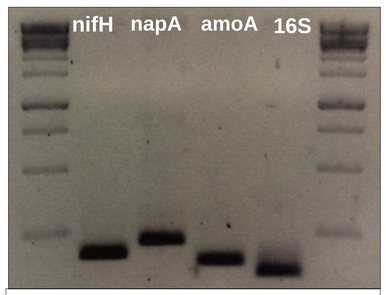


Figure 4: PCR on microbial community with functional genes at 10^{-1}

Preliminary PCR testing demonstrates that the universal primers work in soil and the bacterial community at 60 µM. Between PCRs at 10^0 and 10^{-4}, the total ng of nifH product dropped from 148 total ng / reaction to just 7 ng / reaction.

The 10^0 pool was selected for sequencing testing, and the final library concentration was 13 ng/µL in 26 µL.

Sequenced samples show that 94% of samples contain the dogtags, but data analysis is still ongoing.

Conclusion

Libraries for all dilutions will be sequenced soon and data will be analyzed for the breadth of the technology.

With our new sequencing technology utilizing dogtags:

- Thousands of samples can be pooled into a single Illumina sequencing lane
- Fewer primers used for high-throughput sequencing
- Applies to real environments such as soil
- Allows the characterization of communities for population genetics studies involving structure and interspecies interactions

Future experiments include using the technology to distinguish between closely related bacterial strains of Sinorhizobium and applications to natural samples.

Special thanks to the Nuzhdin Lab, USC Women in Science and Engineering, and the Rose Hill Foundation for supporting and funding this project.