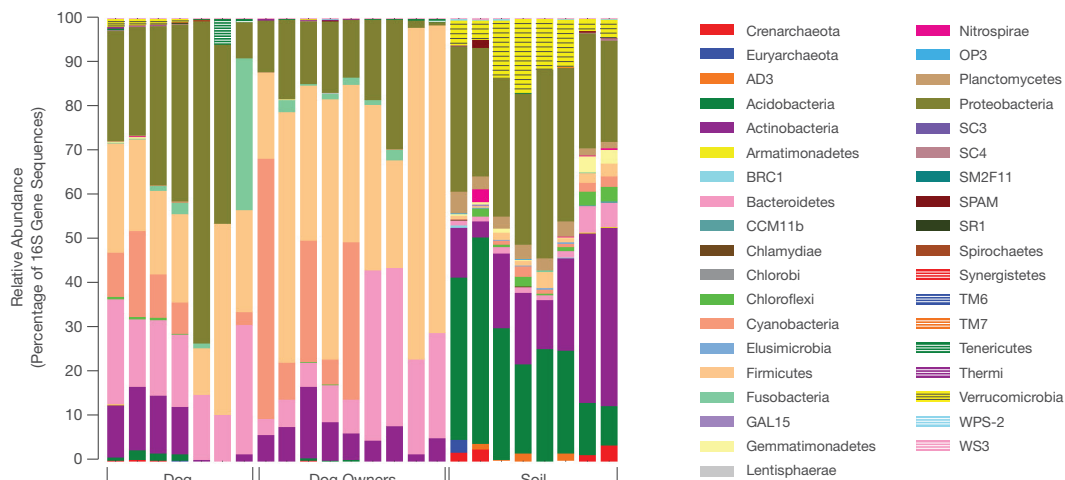


Figure 2: QIIME Taxon Assignment at Phylum Level



To examine whether known differences in the 16S sequence between microbial communities could be reproduced using massively parallel sequencing, qseq files for read 1 were analyzed using QIIME. QIIME taxon assignment at phylum level for samples taken from various anatomic sites of a dog, the dog owners, and soil. Taxonomy assignments were made with QIIME using the Greengenes taxonomy.

24 different barcodes and combined into a single library for sequencing on the MiSeq system.

Sequencing on the MiSeq System

The sample containing 24 pooled barcoded samples was loaded onto the MiSeq reagent cartridge, and then onto the instrument along with the flow cell. Automated cluster generation and paired-end sequencing with a 13-cycle index read was carried out without any further user intervention, taking 28 hours.

Data Analysis

Primary analysis (image analysis, basecalling) was performed on the MiSeq instrument. Quality filtered qseq files were analyzed offline using QIIME³. QIIME is designed to take users from their raw sequence data through to publication-quality graphics, including providing supporting analyses such as quality filtering of reads, demultiplexing, operational taxonomic unit (OTU) picking, taxonomy assignment, and alpha and beta diversity analyses. Figure 2 shows a phylum-level taxonomic summary based on read 1 (i.e., the 5' read) from the 24 samples from various environments, including dog, the dog's human owners, and soil samples. These samples were also concurrently run on the HiSeq 2000 system, and all data were highly reproducible across different flow cell lanes and across sequencing platforms² (data not shown).

Conclusions

With a simple multiplexing strategy and high data yield, the MiSeq system is ideally suited to microbial profiling, enabling rapid turnaround from sample to answer. The V4 region of the 16S ribosomal RNA from various microbial populations was sequenced, allowing phylum-level identification. Using open source tools such as QIIME, complex community analysis can be carried out, and

publication-ready data available in a matter of days after sample acquisition. The MiSeq system has the capacity to accommodate a greater number of samples than those presented in this study, as well as multiple 16S variable regions, permitting deeper genomic scrutiny of larger metagenomic populations.

References

1. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, et al. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 108:4516–4522.
2. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, et al. (2011) Manuscript in preparation.
3. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman, FD, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7(5):335–336.
4. Desantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72(7): 5069–5072.
5. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, et al. (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucl Acids Res* 37:D141–D145.
6. Liu Z, Lozupone C, Hamady M, Bushman FD, Knight R. (2007) Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucl Acids Res* 35:18.
7. Liu Z, DeSantis TZ, Andersen GL, Knight R. (2008) Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. *Nucl Acids Res* 36:18.
8. http://qiime.org/tutorials/processing_illumina_data.html

Learn More

Go to www.illumina.com/miseq to learn more about the next revolution in personal sequencing.

ILLUMINA, INC. • 9885 Towne Centre Drive, San Diego, CA 92121 USA • 1.800.809.4566 toll-free • 1.858.202.4566 tel • techsupport@illumina.com • illumina.com

FOR RESEARCH USE ONLY

© 2011 Illumina, Inc. All rights reserved. Illumina, illuminaDx, BeadArray, BeadExpress, cBot, CSPro, DASL, DesignStudio, Eco, Genetic Energy, GALx, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, Sentrix, Solexa, TruSeq, VeraCode, the pumpkin orange color, and the Genetic Energy streaming bases design are trademarks or registered trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners. Pub. No. 770-2011-013 Current as of 18 August 2011

