

NAME OF THIS STUDY

Running title: INSERT RUNNING TITLE HERE

Your Name Here¹, Joeseeph P. Schmo², Sally J. Rivers¹, Patrick D. Schloss^{1†}

† To whom correspondence should be addressed: pschloss@umich.edu

1. Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI 48109

2. Other department contact information

1 Abstract

2 Introduction

3 Results and Discussion

4 **Scaling up.** The advantage of the dual-index approach is that a large number of samples can be
5 sequenced using a number of primers equal to only twice the square root of the number of samples.
6 To fully evaluate this approach, we resequenced the V4 region of 360 samples that were previously
7 described by sequencing the distal end of the V35 region on the 454 GS-FLX Titanium platform
8 (1). In that study, we observed a clear separation between murine fecal samples obtained from 8
9 C57BL/6 mice at 0 to 9 (early) and 141 to 150 (late) days after weaning, and there was significantly
10 less variation between the late samples than the early samples. In addition to the mouse fecal
11 samples, we allocated 2 pairs of indices to resequence our mock community. We generated 2.8
12 million pairs of sequence reads from the 16S rRNA gene with an average coverage of 7,824.2
13 pairs of reads per sample (95% of the samples had more than 2,113.8 pairs of sequences) using a
14 new collection of 8-nt indices (see the supplemental material). Although individual samples were
15 expected to have various amplification efficiencies, analysis of the number of reads per index did
16 not suggest a systematic positive or negative amplification bias that could be attributed to the
17 indices. The combined error rate for the two mock communities was 0.07% before preclustering
18 and 0.01% after ($n = 14,094$ sequences). When we used UCHIME to remove chimeras and rarefied
19 to 5,000 sequences, there was a average of 30.4 OTUs (i.e., 10.4 spurious OTUs). Similar to our
20 previous results, ordination of the mouse fecal samples again showed the separation between the
21 early and late periods and increased stabilization with age (Fig. 4) (Mantel test coefficient, 0.81; P
22 < 0.001). These results clearly indicate that our approach can be scaled to multiplex large numbers
23 of samples.

24 Conclusions

25 Materials and Methods

26 Insert figure legends with the first sentence in bold, for example:

27 **Figure 1. Number of OTUs sampled among bacterial and archaeal 16S rRNA gene**
28 **sequences for different OTU definitions and level of sequencing effort.** Rarefaction curves
29 for different OTU definitions of Bacteria (A) and Archaea (B). Rarefaction curves for the coarse
30 environments in Table 1 for Bacteria (C) and Archaea (D).

31 **References**

- 32 1. **Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA,**
33 **Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Horn DJV, Weber CF.**
34 2009. Introducing mothur: Open-source, platform-independent, community-supported software
35 for describing and comparing microbial communities. *Applied and Environmental Microbiology*
36 **75:7537–7541.** doi:10.1128/aem.01541-09.