

# **What the Taq? The Influence of Different Hi-Fidelity Taq Polymerase on 16S rRNA Sequencing**

Marc A Sze<sup>1</sup> and Patrick D Schloss<sup>1†</sup>

† To whom correspondence should be addressed: [pschloss@umich.edu](mailto:pschloss@umich.edu)

<sup>1</sup> Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI

Co-author e-mails:

- [marcsze@med.umich.edu](mailto:marcsze@med.umich.edu)

1 **Abstract**

2 **Background.**

3 **Methods.**

4 **Results.**

5 **Conclusions.**

## 6 Introduction

## Materials & Methods

**Human and Mock Samples:** A single fecal sample was obtained from 4 individuals who were part of the Enterics Research Investigational Network (ERIN) and the processing and storage of these samples have been published previously (Seekatz et al., 2016). Clinical data and other types of meta data were not utilized or accessed for this study. All samples were extracted using the MOBIO™ PowerMag Microbiome RNA/DNA extraction kit (now Qiagen, MD, USA). The ZymoBIOMICS™ Microbial Community DNA Standard (Zymo, CA, USA) was used in this study and is made up of *Pseudomonas aeruginosa*, *Escherichia coli*, *Salmonella enterica*, *Lactobacillus fermentum*, *Enterococcus faecalis*, *Staphylococcus aureus*, *Listeria monocytogenes*, and *Bacillus subtilis* at equal genomic DNA abundance (<http://www.zymoresearch.com/microbiomics/microbial-standards/zymbiomics-microbial-community-standards>).

**PCR Protocol:** The five different high fidelity (HiFi) Taq DNA polymerase that were tested were AccuPrime™ (ThermoFisher, MA, USA), KAPA HIFI (Roche, IN, USA), Phusion (ThermoFisher, MA, USA), Platinum (ThermoFisher, MA, USA), and Q5 (New England Biolabs, MA, USA). The PCR cycle conditions were the same for every primer (Kozich et al., 2013) ([https://github.com/SchlossLab/MiSeq\\_WetLab\\_SOP/blob/master/MiSeq\\_WetLab\\_SOP\\_v4.md](https://github.com/SchlossLab/MiSeq_WetLab_SOP/blob/master/MiSeq_WetLab_SOP_v4.md)). If the HiFi Taq had a specific activation time that was different then 2 minutes that was used instead. The 30 cycle default was used but the cycle conditions started at 15 and increased by 5 up to 35 cycles and was used for both fecal and mock samples. The fecal PCR consisted of all 4 samples at 15, 20, 25, 30, and 35 cycles for each Taq (total samples = 100). Although, the mock communities also had 4 replicates used for 15, 20, 25, and 35 cycles, 10 replicates were used for 30 cycles for all Taq (total samples = 130). For all the mock community samples there was not enough PCR product at 15 cycles for adequate sequencing.

**Sequence Processing:** The mothur software program was utilized for all sequence processing steps (Schloss et al., 2009). The protocol followed was similar to that used by Kozich, *et al.* (Kozich et al., 2013) ([https://www.mothur.org/wiki/MiSeq\\_SOP](https://www.mothur.org/wiki/MiSeq_SOP)). Two major differences from the stated protocol were the use VSEARCH instead of UCHIME for chimera detection and the use of the OptiClust algorithm instead of average neighbor for OTU generation (Edgar et al., 2011; Rognes et al., 2016; Westcott & Schloss, 2017). Sequence error was determined using the seq.error command on mock samples after chimera removal and classification to the RDP to remove non-bacterial sequences (Schloss et al., 2009; Cole et al., 2013; Rognes et al., 2016).

**Reproducible Methods:** The code and analysis can be found here [https://github.com/SchlossLab/Size\\_PCRSeqEffects\\_XXXX\\_2017](https://github.com/SchlossLab/Size_PCRSeqEffects_XXXX_2017).

## 43 **Results**

44 ***Insert Sub title here***



## 46 **Conclusion**



## 47 **Acknowledgements**

48 The authors would like to thank all the study participants who were a part of each of the  
49 individual studies utilized. We would also like to thank each of the study authors for making  
50 their data available for use. Finally we would like to thank the members of the Schloss lab  
51 for valuable feed back and proof reading during the formulation of this manuscript.

## References

- Cole JR., Wang Q., Fish JA., Chai B., McGarrell DM., Sun Y., Brown CT., Porras-Alfaro A., Kuske CR., Tiedje JM. 2013. Ribosomal database project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Research* 42:D633–D642. DOI: 10.1093/nar/gkt1244.
- Edgar RC., Haas BJ., Clemente JC., Quince C., Knight R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27:2194–2200. DOI: 10.1093/bioinformatics/btr381.
- Kozich JJ., Westcott SL., Baxter NT., Highlander SK., Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq illumina sequencing platform. *Applied and Environmental Microbiology* 79:5112–5120. DOI: 10.1128/aem.01043-13.
- Rognes T., Flouri T., Nichols B., Quince C., Mahé F. 2016. VSEARCH: A versatile open source tool for metagenomics. *PeerJ* 4:e2584. DOI: 10.7717/peerj.2584.
- Schloss PD., Westcott SL., Ryabin T., Hall JR., Hartmann M., Hollister EB., Lesniewski RA., Oakley BB., Parks DH., Robinson CJ., Sahl JW., Stres B., Thallinger GG., Horn DJV., Weber CF. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* 75:7537–7541. DOI: 10.1128/aem.01541-09.
- Seekatz AM., Rao K., Santhosh K., Young VB. 2016. Dynamics of the fecal microbiome in patients with recurrent and nonrecurrent clostridium difficile infection. *Genome Medicine* 8. DOI: 10.1186/s13073-016-0298-8.
- Westcott SL., Schloss PD. 2017. OptiClust, an improved method for assigning

<sup>75</sup> amplicon-based sequence data to operational taxonomic units. *mSphere* 2:e00073–17.  
<sup>76</sup> DOI: 10.1128/mspheredirect.00073-17.

77 **Table 1:**

78 **Table 2:**

79 **Figure 1:** .

80 **Figure 2:** .

81 **Figure 3:** .

82 **Figure 4:** .

83 **Figure 5:** .

84 **Figure 6:** .

85 **Figure 7:** .

86 **Figure S1:** .

87 **Figure S2:** .

88 **Figure S3:** .

89 **Figure S4:** .

90 **Figure S5:** .

91 **Figure S6:** .

92 **Figure S7:** .