

Microbiome Premature

Jaewoong Lee

Ulsan National Institute of Science and Technology

jwlee230@unist.ac.kr

2020-09-07

Overview

1 Introduction

2 Materials

3 Methods

4 Results

5 Proceedings

References

Ribosomal RNA

Premature

16s rRNA Sequencing

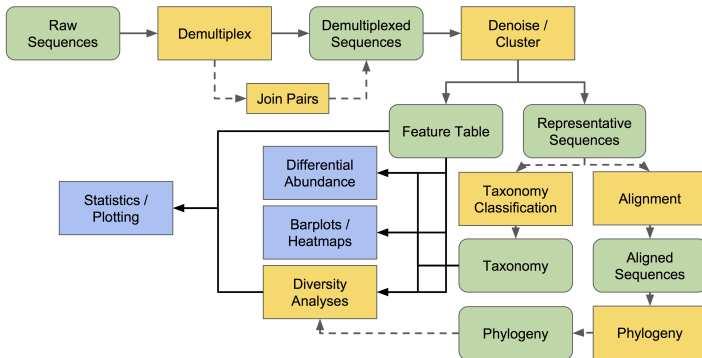


Figure: QIIME 2 workflow

Denoising Techniques

- DADA2 (Callahan et al., 2016)
- Deblur (Amir et al., 2017)

Taxonomy Classification

- Greengenes (DeSantis et al., 2006)
- SILVA (Pruesse et al., 2007; Quast et al., 2012)

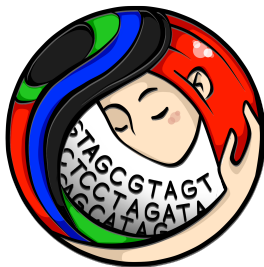


Figure: Mothur

Note: Still in progress

t-distributed Stochastic Neighbor Embedding

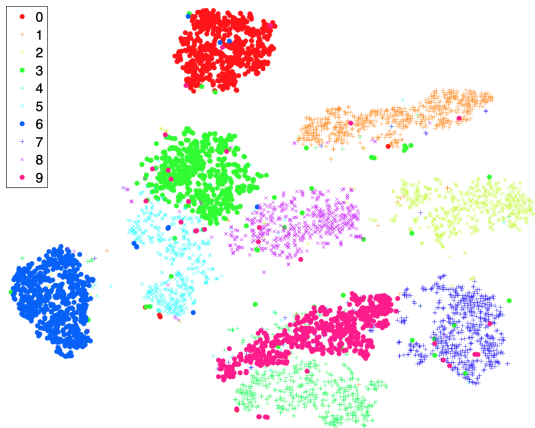
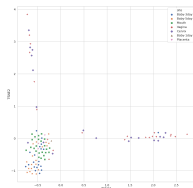
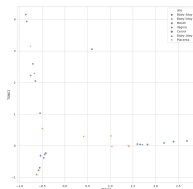
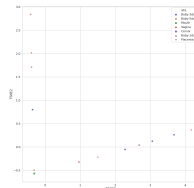


Figure: t-SNE with handwritten data (Maaten & Hinton, 2008)

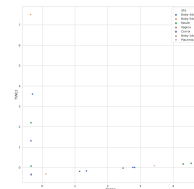
t-SNE with Site



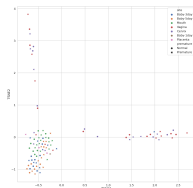
(a) DADA2 + GG (b) DADA2 + SILVA



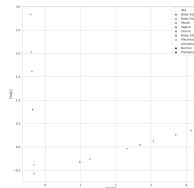
(a) Deblur + GG (b) Deblur + SILVA



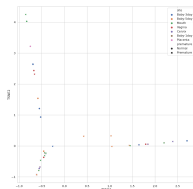
t-SNE with Site + Premature



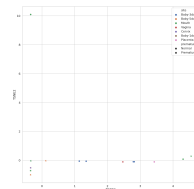
(a) DADA2 + GG



(b) DADA2 + SILVA



(a) Deblur + GG



(b) Deblur + SILVA

Yields I

Requirements I

Expectations I

References I

- Amir, A., McDonald, D., Navas-Molina, J. A., Kopylova, E., Morton, J. T., Xu, Z. Z., ... others (2017). Deblur rapidly resolves single-nucleotide community sequence patterns. *MSystems*, 2(2).
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). Dada2: high-resolution sample inference from illumina amplicon data. *Nature methods*, 13(7), 581–583.
- DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., ... Andersen, G. L. (2006). Greengenes, a chimera-checked 16s rRNA gene database and workbench compatible with arb. *Applied and environmental microbiology*, 72(7), 5069–5072.
- Maaten, L. v. d., & Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov), 2579–2605.

- Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., & Glöckner, F. O. (2007). Silva: a comprehensive online resource for quality checked and aligned ribosomal rna sequence data compatible with arb. *Nucleic acids research*, 35(21), 7188–7196.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... Glöckner, F. O. (2012). The silva ribosomal rna gene database project: improved data processing and web-based tools. *Nucleic acids research*, 41(D1), D590–D596.