

Periodontitis

Seunghoon Kim

Jaewoong Lee

Semin Lee

2020-11-26

Contents

1	Introduction	3
1.1	Periodontitis	3
1.2	Ribosomal RNA	3
2	Materials	3
2.1	16S rRNA Gene Sequencing	3
3	Methods	3
3.1	QIIME2 Workflow	3
3.1.1	Denoising techniques	3
3.1.2	Taxonomy Classification	3
3.1.3	Rarefaction	3
3.1.4	Alpha-diversity	3
3.1.5	Beta-diversity	3
3.1.6	ANCOM	3
3.2	Python Packages	3
3.2.1	Pandas	3
3.2.2	Scikit-learn	3
3.2.3	Matplotlib	3
3.2.4	Seaborn	3
4	Results	5
5	Discussion	5
6	References	5

List of Tables

List of Figures

1	A Theoretic Overview of QIIME2 Workflow (Bolyen et al., 2019, 2018)	4
2	Denoising Techniques which provided by QIIME2	4
3	Taxonomy Classification which provided by QIIME2	4
4	Example ANCOM Volcano Plot which Provided by QIIME2 (Bolyen et al., 2019, 2018)	6

1 Introduction

1.1 Periodontitis

1.2 Ribosomal RNA

2 Materials

2.1 16S rRNA Gene Sequencing

3 Methods

3.1 QIIME2 Workflow

QIIME2 is a capable, expandable and distributed microbiome analysis package with transparent analysis (Bolyen et al., 2019, 2018). A theoretic overview of QIIME2 workflow is shown as figure 1.

3.1.1 Denoising techniques

There are two denoising techniques provided by QIIME2: DADA2 (Callahan et al., 2016) and Deblur (Amir et al., 2017). Major difference between DADA2 and Deblur is a strategy, the strategy used to divide as different species.

3.1.2 Taxonomy Classification

3.1.3 Rarefaction

3.1.4 Alpha-diversity

3.1.5 Beta-diversity

3.1.6 ANCOM

ANCOM (Analysis of composition of microbiomes) can be used for analyzing the composition of microbiome in multiple populations (Mandal et al., 2015).

3.2 Python Packages

3.2.1 Pandas

Pandas is a Python package of rich data structures and tools for analyzing with structured data sets (McKinney et al., 2011).

3.2.2 Scikit-learn

Scikit-learn grants state-of-the-art implementation of many machine learning algorithms, while controlling an easy-to-use interface tightly integrated the Python code (Pedregosa et al., 2011).

3.2.3 Matplotlib

Matplotlib is a Python graphics package which used for application development, interactive scripting and publication quality image generation (Barrett, Hunter, Miller, Hsu, & Greenfield, 2005). Matplotlib, also, is designed to create simple plots with a few commands (Hunter, 2007).

3.2.4 Seaborn

Seaborn is a Python data visualization package which based on matplotlib, allows a high-level interface for displaying engaging and descriptive statistical graphics (Waskom & the seaborn development team, 2020).

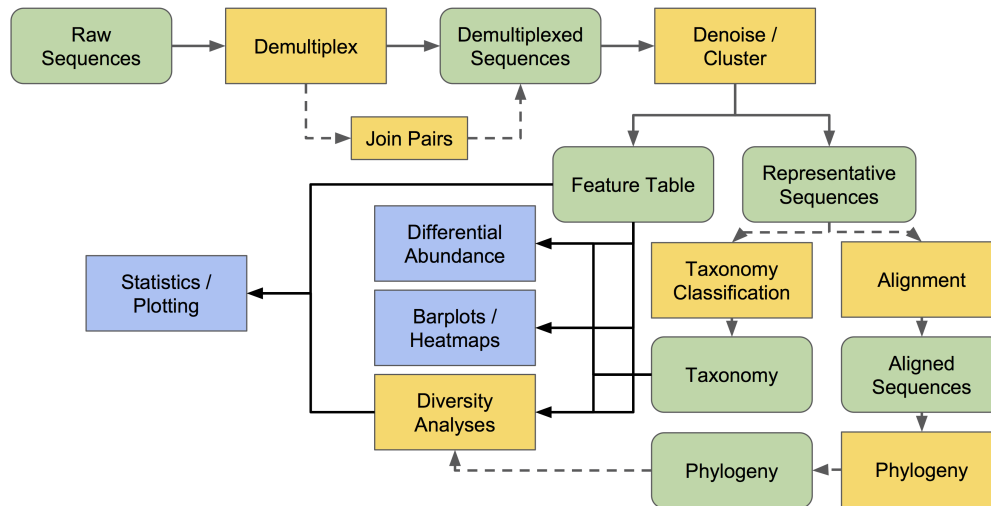


Figure 1: A Theoretic Overview of QIIME2 Workflow (Bolyen et al., 2019, 2018)

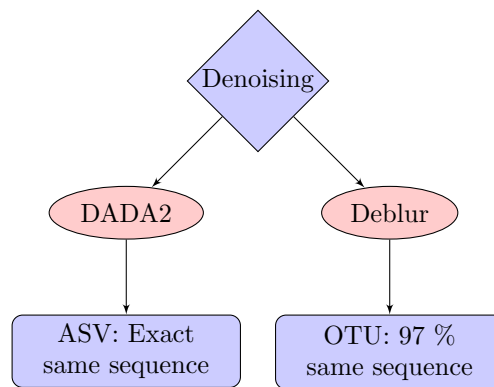


Figure 2: Denoising Techniques which provided by QIIME2

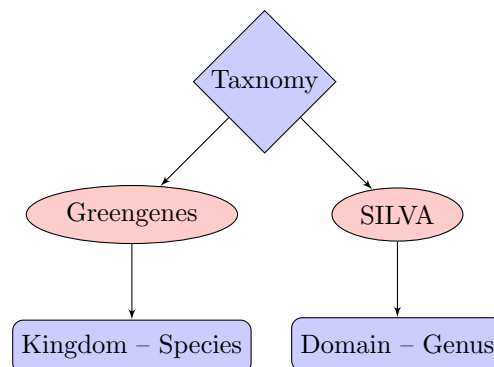


Figure 3: Taxonomy Classification which provided by QIIME2

4 Results

5 Discussion

6 References

- 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).
- Barrett, P., Hunter, J., Miller, J. T., Hsu, J.-C., & Greenfield, P. (2005). matplotlib—a portable python plotting package. In *Astronomical data analysis software and systems xiv* (Vol. 347, p. 91).
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C., Al-Ghalith, G. A., ... others (2018). *Qiime 2: Reproducible, interactive, scalable, and extensible microbiome data science* (Tech. Rep.). PeerJ Preprints.
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., ... others (2019). Reproducible, interactive, scalable and extensible microbiome data science using qiime 2. *Nature biotechnology*, 37(8), 852–857.
- 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.
- Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing in science & engineering*, 9(3), 90–95.
- Mandal, S., Van Treuren, W., White, R. A., Eggesbø, M., Knight, R., & Peddada, S. D. (2015). Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microbial ecology in health and disease*, 26(1), 27663.
- McKinney, W., et al. (2011). pandas: a foundational python library for data analysis and statistics. *Python for High Performance and Scientific Computing*, 14(9).
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... others (2011). Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12, 2825–2830.
- Waskom, M., & the seaborn development team. (2020, September). *mwaskom/seaborn*. Zenodo. Retrieved from <https://doi.org/10.5281/zenodo.592845> doi: 10.5281/zenodo.592845

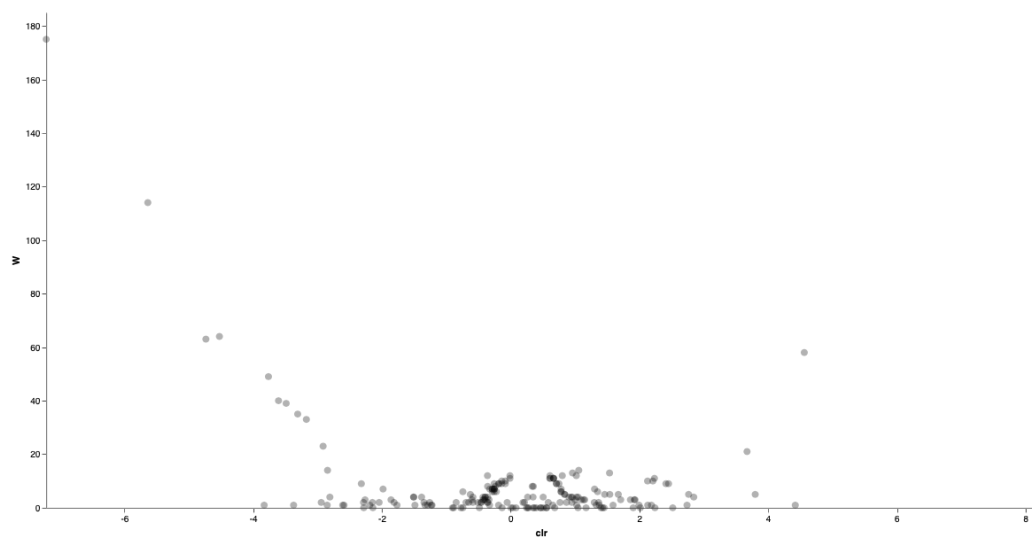


Figure 4: Example ANCOM Volcano Plot which Provided by QIIME2 (Bolyen et al., 2019, 2018)