

Microbial Communities Profiling via QIIME 2 and Qiita

CLASS SESSIONS

Monday, June 4, 2018 - Tuesday, June 5, 2018

8:30am – 5:30 pm

Location: Hammer Health Sciences Building, 701 West 168th Street

Directions can be found here: <https://www.cuepisummer.org/contactpage>

INSTRUCTORS

Antonio Gonzalez, PhD

Programmer Analyst

University of California, San Diego, Department of Pediatrics

email: antgonza@gmail.com

Instructor 2: TBD

COURSE DESCRIPTION

This course will provide a theoretical, analytical and practical introduction to QIIME 2 (canonically pronounced 'chime two'), which stands for Quantitative Insights into Microbial Ecology 2, and Qiita (canonically pronounced 'cheetah'), a multiomics and multi-study online tool. QIIME 2 and Qiita are open source software packages for comparison and analysis of microbial communities, primarily based on high-throughput amplicon sequencing data (such as SSU rRNA) generated on a variety of platforms, but also supporting analysis of other types of data (such as shotgun metagenomic, metabolomics or proteomics). The main Qiita deployment (<http://qiita.microbio.me/>) allows users to manage and analyze large studies, their metadata and the multiple data types generated from the same samples. Additionally, it allows users to combine their samples with other published and public studies available in the system. QIIME 2 is a stand-alone environment for the analysis of individual microbiome data sets that can be used on your laptop, university computational resources, and cloud computing resources.

REQUIRED SOFTWARE

Please bring your laptop to the workshop, as a significant portion of the workshop will be hands-on. Any operating system will work, including Windows, Mac OS X, and Linux. You do not need to install any other software besides what is listed below. There will be power outlets and WiFi available to you at the workshop.

Before the workshop, please install the latest version of the Google Chrome web browser. If you don't have Google Chrome installed, [get it here](#). If you already have Google Chrome installed, make sure it is the latest version by [following these instructions](#). Next, [install the Secure Shell](#) app for Google Chrome. Next, try to open [this page](#) to ensure that your browser is configured correctly to view Emperor plots. If you get a notice that the page cannot be opened,

follow the instructions and try the link again. Finally, if you are using Windows or Mac OS X, [install Cyberduck](#).

PREPARING FOR THE WORKSHOP

Students will use the Linux command line extensively during the hands-on portion of the workshop. We highly recommend becoming familiar with the command line prior to the workshop. The following resources are recommended to learn the command line. Depending on previous experience with the command line and level of technical expertise, it may take anywhere from a few hours to a few days to get comfortable using the command line. Thus, we recommend that students start working through this material as soon as possible, including the accompanying exercises.

[Udacity Linux command line course \(free\)](#)

This course is pretty comprehensive and will take a few days to complete. Additionally, it shows users how to set up a virtual machine to run Linux commands in (this step is required for Windows users but can also be performed by Mac OS X and Linux users).

[Software Carpentry Unix Shell introduction](#)

This course is shorter than the Udacity course above, taking about 1.5 hours to complete. If you are on Windows, you will need access to a Linux machine (or virtual machine) to execute commands in. The Udacity course above will show you how to do this, or you can try out this [online Linux emulator](#). If you are using Mac OS X or Linux, you can alternatively use the Terminal program included with your operating system to execute commands in.

UPLOAD YOUR DATASET TO QIITA

If you plan to bring your own dataset to the workshop the requirements are that (1) you upload your sample & preparation templates, (2) and raw data to Qiita (<http://qiita.microbio.me/>). For more information, visit: <http://qiita.microbio.me/static/doc/html/index.html>. If you run into any issue send an email to qiita.help@gmail.com

COURSE LEARNING OBJECTIVES

By the end of the course, participants will be able to:

1. Understand the most recent QIIME2 and Qiita features for microbial community analysis
2. Select the best workflow and parameters to perform the different steps for microbial community analysis
3. Understand and apply on their own datasets different phylogenetic and non-phylogenetic metrics to compare microbial diversity samples
4. Upload and analyze their own datasets using Qiita and compare their studies with other public studies

COURSE READINGS

The following references should give some basic background:

- QIIME2 core concepts. <https://docs.qiime2.org/2017.2/concepts/>
- Gonzalez, Antonio, and Rob Knight. 'Advancing analytical algorithms and pipelines for billions of microbial sequences.' *Current opinion in biotechnology* 23.1 (2012): 64-71.

- Gonzalez, Antonio, et al. 'Our microbial selves: what ecology can teach us.' EMBO reports 12.8 (2011): 775-784.
- Lozupone, Catherine A., and Rob Knight. 'Species divergence and the measurement of microbial diversity.' FEMS microbiology reviews 32.4 (2008): 557-578.
- Kuczynski, Justin, et al. 'Microbial community resemblance methods differ in their ability to detect biologically relevant patterns.' Nature methods 7.10 (2010): 813-819.
- Lozupone, Catherine, and Rob Knight. 'UniFrac: a new phylogenetic method for comparing microbial communities.' Applied and environmental microbiology 71.12 (2005): 8228-8235.
- Bouslimani, Amina, et al. 'Molecular cartography of the human skin surface in 3D.' Proceedings of the National Academy of Sciences 112.17 (2015): E2120-E2129.
- Lozupone, Catherine A., et al. 'Meta-analyses of studies of the human microbiota.' Genome research 23.10 (2013): 1704-1714.

COURSE STRUCTURE

Class meets from 8:30am – 5:30pm, Wednesday, June 28 and Thursday, June 29, 2017 (16 hours total).

COURSE SCHEDULE

Session 1 – Introduction to QIIME2, Qiita and microbiome analysis
Tools background and introduction to the microbiome analysis. Quick overview of the command line interface. Setting up studies in Qiita.
Session 2 – Study design and initial processing steps
Study design and setup with QIIME2 and Qiita. Understand primer selection, picking Operational Taxonomic Units (OTUs) methods, representative set selection, building a phylogenetic tree, assigning taxonomy and finishing with a BIOM table. Process your own studies in Qiita.
Session 3 – Downstream analysis with a BIOM table
Possible analyses with QIIME2 of a BIOM table. Overview of alpha, beta, and taxonomy analyses, their statistical methods and metrics. Tools to generate time series and spatial visualizations of your samples.
Session 4 – Working with Qiita, other tools and visualizations
Analyses in Qiita, combine private and public studies. Other tools: supervised learning, source tracking, II, among others. Review and interpretation of the

methods section of published articles.

Session 5 – Future of QIIME2 and final remarks

Have an overview of what can we expect of the future and open discussion.