Microbiome Informatics Workshop 2019

16S rRNA Sequence Analysis: Intro and Advanced Courses

Center for Microbiome Research

Courses in sequencing and analysis of microbiome data held over two days.

Workshop Summary:

This workshop is designed to be a combination of the theory behind processing sequence data and practical applications of current software in studying the microbiome. The first day of the workshop will cover a primer on sequencing technology, information theory, ecological theory, and statistics as applied to microbial communities. The practical portions will focus on 16S sequence analysis through QIIME2. The goal is to provide a comprehensive survey of sequencing applications and hands-on experience with data analysis.

The second day is intended as an advanced analysis course where participants analyze their own data.

Goals:

By the end of the workshop, participants should:

1. Understand the basics of 16S biology and bioinformatics uses
2. Be able to analyze their own 16S data using QIIME2

Dates:

16S basics/intro course: Tuesday, January 29, 2019 (MFRC 3075)

Advanced 16S: Wednesday, January 30, 2019 (TBRC C2195)

Seminar Coordinators:

Samantha Atkinson [1,2,3] saatkinson@mcw.edu

Orlando 'Landon' DeLeon [1,2] odeleon@mcw.edu

CMR Program Manager

Mary Holtz, PhD[2,4] mlholtz@mcw.edu

Directors:

John Kirby, PhD, Professor and Chair[1,2] jkirby@mcw.edu

Nita Salzman, MD, PhD, Professor[1,2,4] nsalzman@mcw.edu

[1]Dept. of Microbiology and Immunology, Medical College of Wisconsin

[2]Center for Microbiome Research, Medical College of Wisconsin

[3]University of Iowa, Department of Bioinformatics

[4]Dept. of Pediatrics, Medical College of Wisconsin

Schedule:

Prior to the workshop *all participants* need to have QIIME2 installed on their computers, instructions will be provided in the week prior to the workshop.

Day 1 (16S basics): LOCATION MFRC 3075

9-9:30am	Introductions & Sequencing and information theory • DNA Extraction, Library Prep and Sample Considerations • Types of Sequencing Platforms • 16S rDNA gene structure and 16S Sequencing • Analysis Platforms • Taxonomic Assignment Methods • State of the Field for Sequencing Microbiomes
9:30-10:30am	QIIME2 background, Basic command line, Importing data, Quality control, & "OTU" picking (practical) • QIIME2 platform purpose and file/program structure • Using command line to interface with the platform • Importing data to a QIIME2-format • Quality Control of Sequences (PHRED scores and Trimming) • Choosing Operational Taxonomic Units (OTUs)
11-11:15am	Break
11:15-12pm	Statistics, Ecology, what does QIIME tell you • Alpha diversity o What is alpha diversity o Uses for alpha diversity o Metrics for alpha diversity o Interpreting alpha diversity metrics • Beta Diversity o What is beta diversity o Uses for beta diversity o Metrics for beta diversity o Interpreting beta diversity metrics
12-1pm	Lunch: on your own
1-2:15pm	Practical: Ecological profiling of samples
2:15-2:30pm	Break
2:30-4pm	How to ask a question/Practical/remaining questions

Day 2 (Advanced)*: LOCATION TBRC C2195

DATA FORMAT

Day 2 participants are expected to bring their own data. Your data should be in FASTQ formats as CASAVA-ONE-EIGHT, and the sequencing length and type (Paired End or Single) should be known. If your data is not in CASAVA-ONE-EIGHT format, please know the type (FASTA/FASTQ) and formatting of your data.

9-10am	Introductions & Summary of analysis steps • Recap of day 1 • Where we are now in the analysis pipeline
10-10:15am	Break
10:15-12pm	Work on your own data
12-1pm	Lunch
1-2pm	Background advanced topics • Balances o Relative Abundance vs Actual Read Counts o Log Trees o Balances and the resolution they provide o Statistical Models from Balances o Step by step walk through • PICRUST2 (predicted functional assignments) o Limitations and Uses of PICRUST o Interpreting PICRUST o Step-by-step walkthrough
2-4pm	Work on voted option/Continue own data

*Day 2 schedule is a loose schedule, subject to change based on rate that people work/have problems.

Additional Topics We can Discuss on Day 2

- q2-longitudinal (longitudinal & paired sample comparisons)
- MaAsLin: Multivariate Association with Linear Models