DEC 01, 2023

Protocol status: Working
We use this protocol and it's working

Created: Dec 01, 2023

Last Modified: Dec 01, 2023

Nephele DADA2 Workflow

Nandie Elhadidy[1]

[1]North Carolina State University

N   Nandie Elhadidy

ABSTRACT

A protocol outlining the workflow of utilizing DADA2 on Nephele with a provided practice dataset. DADA2 is a pipeline that provides species assignments using Amplicon Sequence Variants or ASVs which compare variants to a predetermined sequence through assumption that the most amplified sequence is the correct one.

Downloading the Dataset

1   To download the practice dataset provided by Nephele, visit:
https://nephele.niaid.nih.gov/user_guide_data/

2   Download the first set of sequences listed as well as the mapping file.

2.1   The reference provided for this set of sequences is listed as: Experimental Microbial Dysbiosis Does Not Promote Disease Progression in SIV-Infected Macaques. NCBI BioProject: PRJNA417022

3   Once downloaded, unzip the downloaded file and ensure all .fastq.gz files are present.

Running DADA2

4   Click the **Home** tab.

5   Choose the **Analyze** option.

6   Under **Amplicon Metagenomics**, select "Start job" option under **DADA2**.

7   Select **Paired End FASTQ**, as this is a paired-end dataset.

8   Click "Upload from my computer" and add all .fastq.gz files from the downloaded dataset. Start upload

9   Select "Next" and upload the corresponding Mapping File.

10   After reaching the Pipeline Selection page, click **Select** and enter a description of the job.

11   Keep Filter and Trim, Merge Pairs, and Analysis settings defaulted. Do not check Ion Torrent Data. Click **Validate and Submit** *Note: This is for following the exact tutorial instructions. If you would like to make changes to these pre-sets, proceed with these changes.*

12   Nephele will send an email upon job completion.

Analysis of the Output

13   From the **My Workspace** tab, hover to the view results icon under **Actions** for the completed DADA2 job.

14   Under **Reported Statistics, Reported Files, Reported Images, Reported Tables,** and **Reported Charts** and explore outputs.

14.1   **Reported files** of note include: **Rooted tree, Rarefaction curve, Alpha diversity, PCoA binomial, BCoA Bray-Curtis,** and **ASV heatmap.**

14.2   **Reported images** of note include: **Quality profiles** of forward and reverse reads

14.3   **Reported tables** of note include: **Read counts by sample** and **Top species table**

14.4   **Reported charts** of note include: **Reads in vs. reads out** and **Track reads**