Germs everywhere! An Intorduction to Microbiome Bioinformatics. Step by Step Tutorial: Cyverse and Qiime2 for Parkinson's Mouse microbiome tutorial.

Natalia J. Bayona Vásquez

Utah Valley University

This tutorial is a series of steps on how to set up a Cyverse account and to setup analyses for Qiime2.

CyVerse provides life scientists with powerful computational infrastructure to handle huge datasets and complex analyses, thus enabling data-driven discovery.

CyVerse was originally created by the National Science Foundation in 2008 with the name iPlant Collaborative to serve U.S. plant science communities. From its inception, iPlant quickly grew into a mature organization providing powerful resources and offering scientific and technical support services to researchers nationally and internationally. In 2015, iPlant was rebranded to CyVerse to emphasize an expanded mission to serve all life sciences as our cyberinfrastructure is germane to all life sciences disciplines.

Please visit https://cyverse.org/about for more information.

Step 1. Creating a Cyverse Account

Visit the cyverse webpage:https://cyverse.org/ and click on the 'Register' link (Figure 1).

Once you have clicke in the "Register" link there will be a welcom sign on your screen (Figure 2).

You will need to fill all the relevant data in their form (Figure 3).

Step 2. Login to Cyverse and open the Discovery Environment

Once you have set up your Cyverse account, login with your credentials (Figure 4).

In the home page, please LAUNCH your Discovery Environment by clicking the 'launch' link (Figure 5).

Step 3. Open the Apps icon and seacrh for Qiime2 (2018.11)

Once you are in the Dicovery Environment click on the APP icon in the top left of your screen. In the new window search for "Qiime2" applications (Figure 6). Selecte the "JupyterLab-Qiime2-2018.11" application in a purple icon, developed by Upendra Kumar (Figure 6).
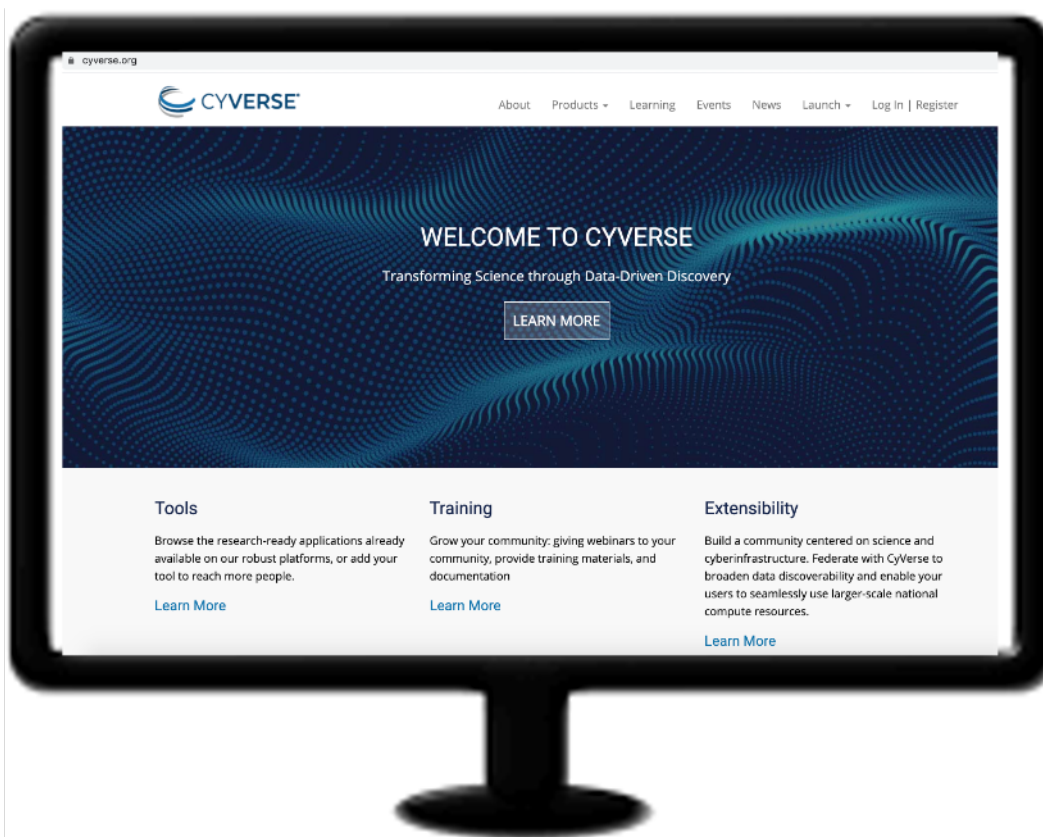
Figure 1: Cyverse webpage

Figure 2: Welcome to Cyverse
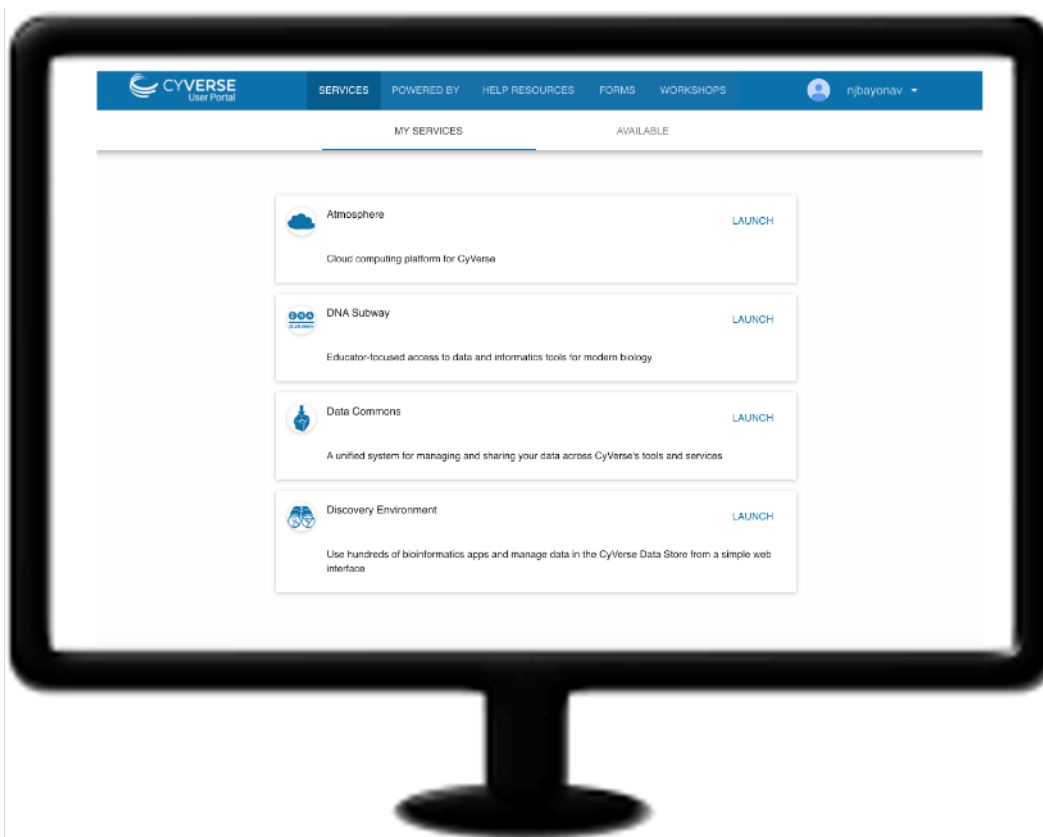
Figure 3: Registration form

Figure 4: Loggin in
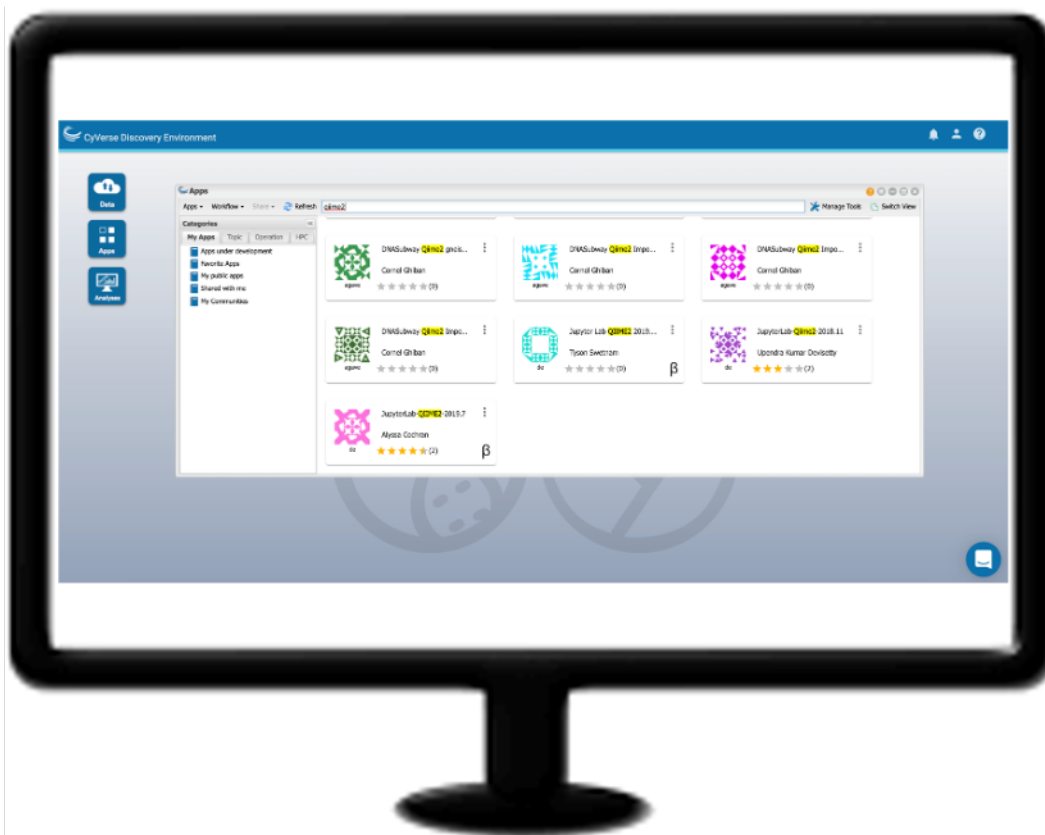
Figure 5: Launch Discoivery Environment

Figure 6: Qiime 2 app

Step 4. Select the input data and output folder

Once the application is opened, you can rename your analysis, select the folder to save the output. Under the "parameters" tab select the input folder, which path is: Shared With Me/njbayonav/Qiime2_Class_2020/Parkiinson_Mouse/ (Figure 7).

NOTE: If you don't have acces to this folder, please send me an email to njbayonav@gmail.com with your cyverse username and I will share it with your. Otherwise you could upload the relevant data specified below, available in my Github tutorial page to your Cyverse Data folder.

Please make sure you have selected the folder (not any file within the folder). Once you have selected the folder clik OK.
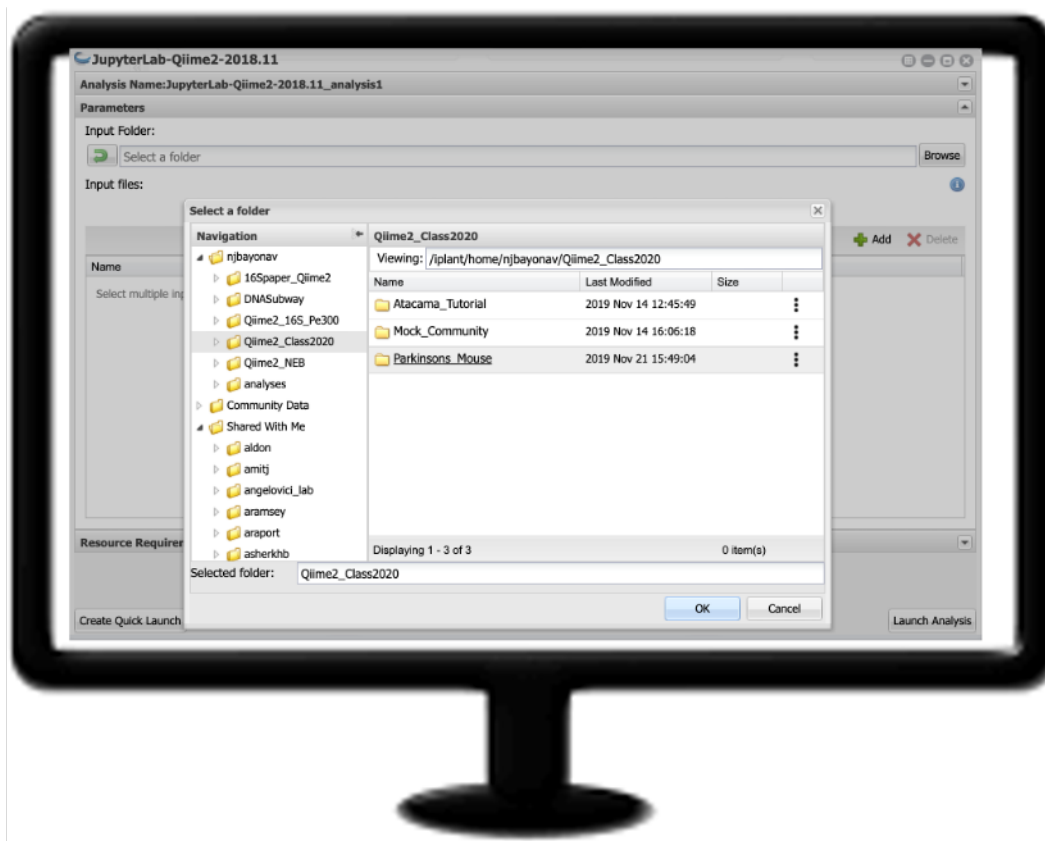

Figure 7: Selecting input folder

Step 5. We are ready to launch the analysis!

Once you have named the analysis and selected your input and output folders, click LAUNCH ANALYSIS. You will get a notification under the bell on the top right of your Discovery environment. You can access to your analysis by clicking the link under notifications (Figure 8) or clicking on the ANALYSES icon.

You will have a new window open in your browser with an icon indicatiing your app is launching (Figure 9).

Once the app is launched, in the new window you will have a left pane with two directories, a central pane with access to creat jupyterbooks, terminal, or R code, among others (Figure 10).

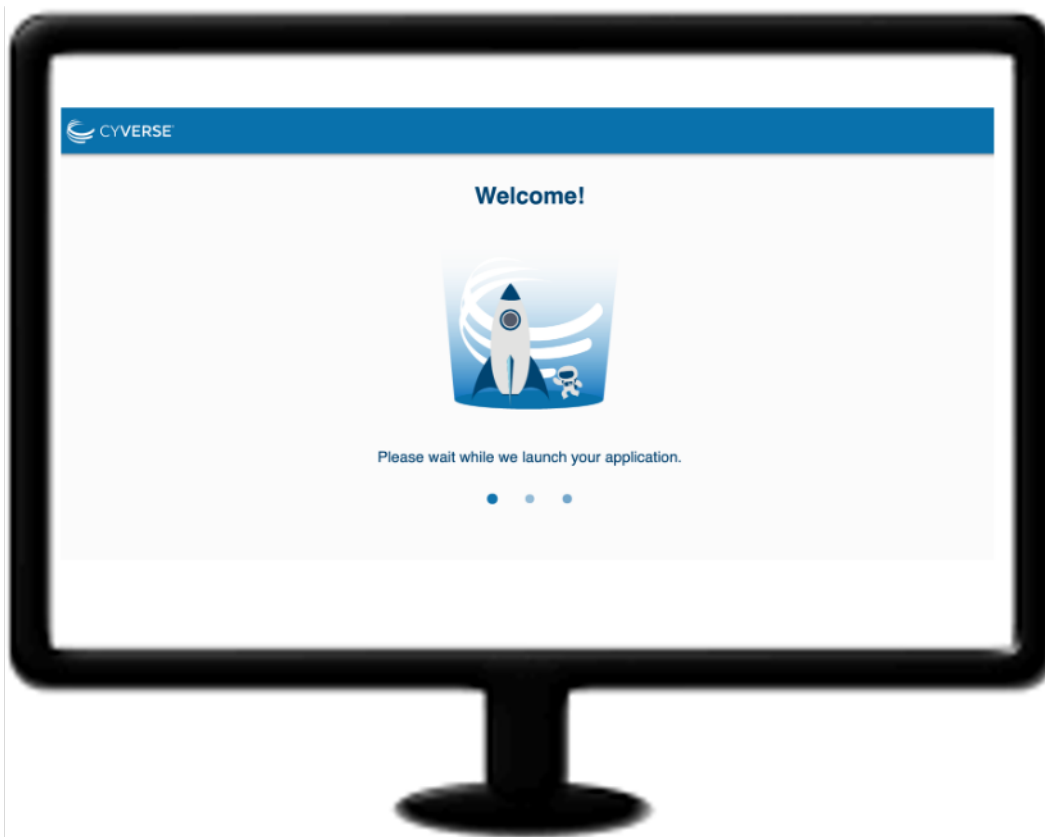Figure 8: Notification of your analysis
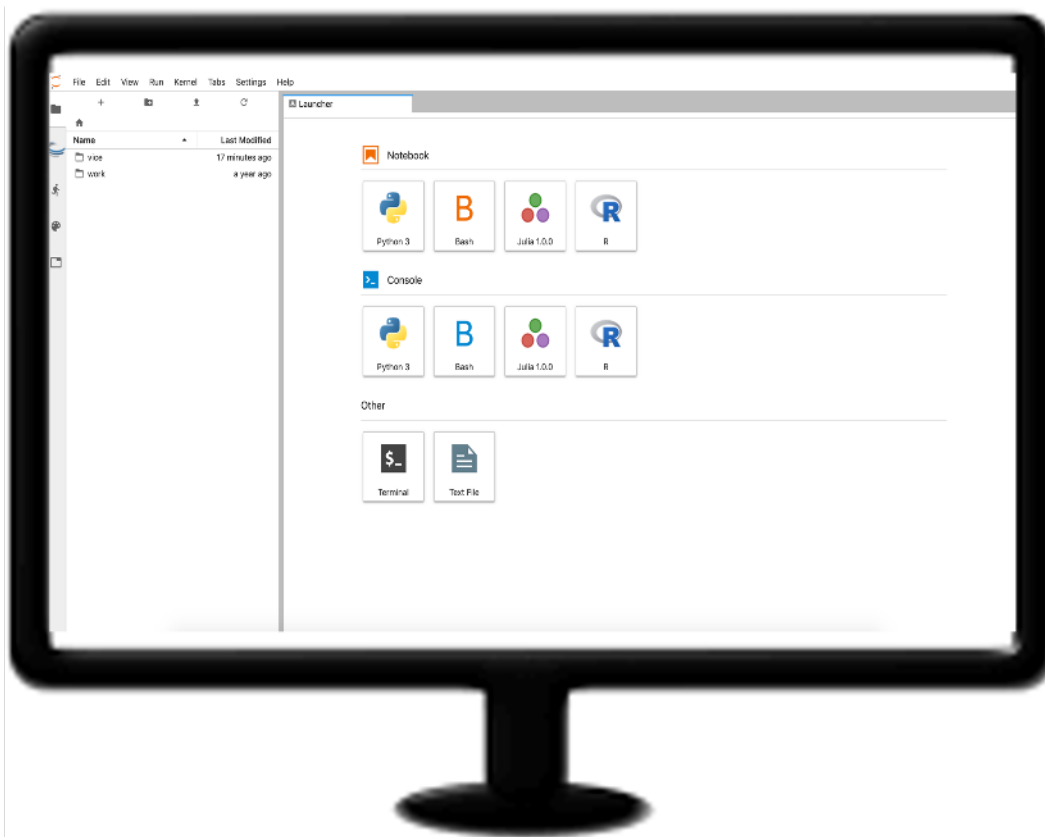
Figure 9: Launch App

Figure 10: Application launched

Within the vice directory, you will find our Parkinson_Mouse folder (Figure 11). Open the Parkinson Mouse folder, here you will be able to locate the following: > 1. fastqs folder: containing the sequencing data needed for the tutorial. > 2. gg_13_8_otus: Greengenes V. 13.8 database to use as referenc for taxonomic assignment. > 3. Qiime2_Parkinson's_Mouse.ipynb: Jupyter notebook to run Qiime2 in the Parkinson's Mouse dataset. > 4. manifest file: Manifest file with sample IDs and read orientation of the sequencing file. > NOTE: This manifest file nees to be updated by the user where the path is replaced with the **absolute** path to the vice folder that contains the fastqs (/home/jovyan/vice/Parkinson_Mouse/fastqs/). > 5. metadata_map_file.tsv: Metadata file containing categorical and numerical information for each sample. > 6. tree.qza: Artifact object of a phylogenetic tree estimated.
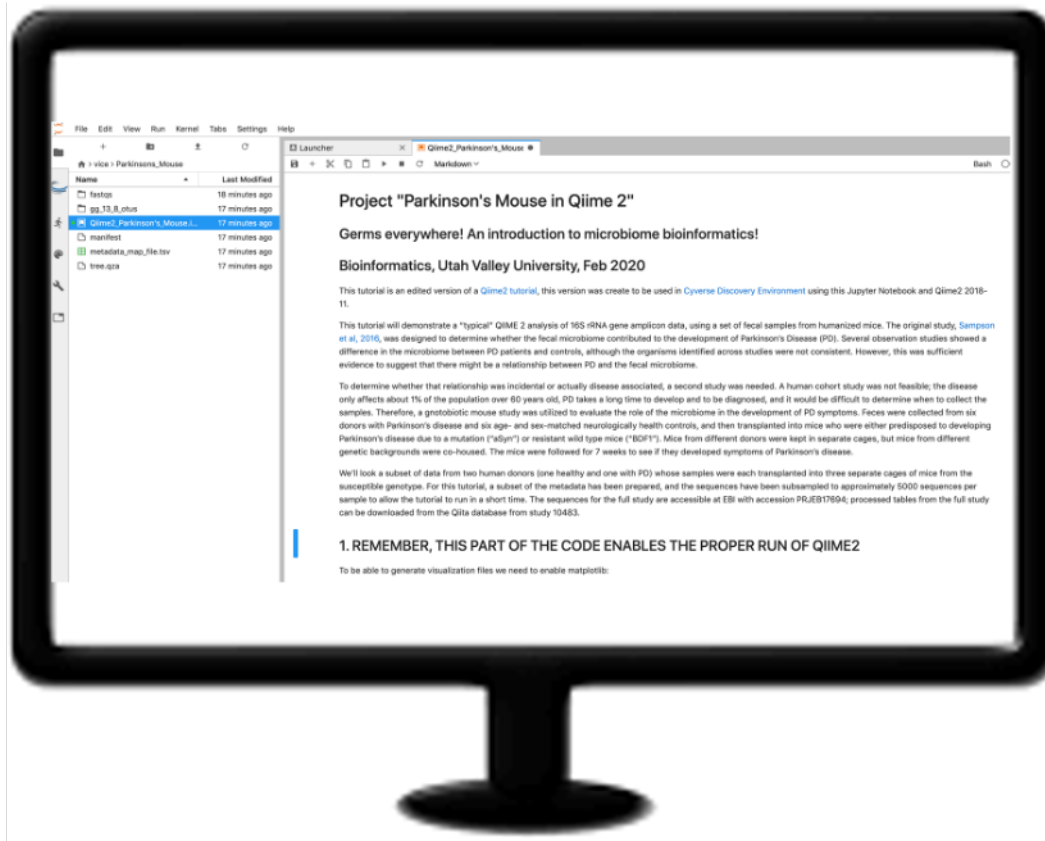
Please open the jupyternotebook file, and we are ready to start! (Figure 11).

Figure 11: Jupyter Notebook