

Metagenomics in QIIME 2 Using Jupyter Notebooks in VICE

Presented by: Joslynn Lee
April 5, 2019



Today's Focus Forum will cover

- What is CyVerse?
- What is QIIME 2?
- How can I run QIIME 2 on DE's VICE?
- Reproducibility using Jupyter Notebooks
- Demo: QIIME 2 using public data set

*Due to time constraint, this will not cover a full tutorial of QIIME 2



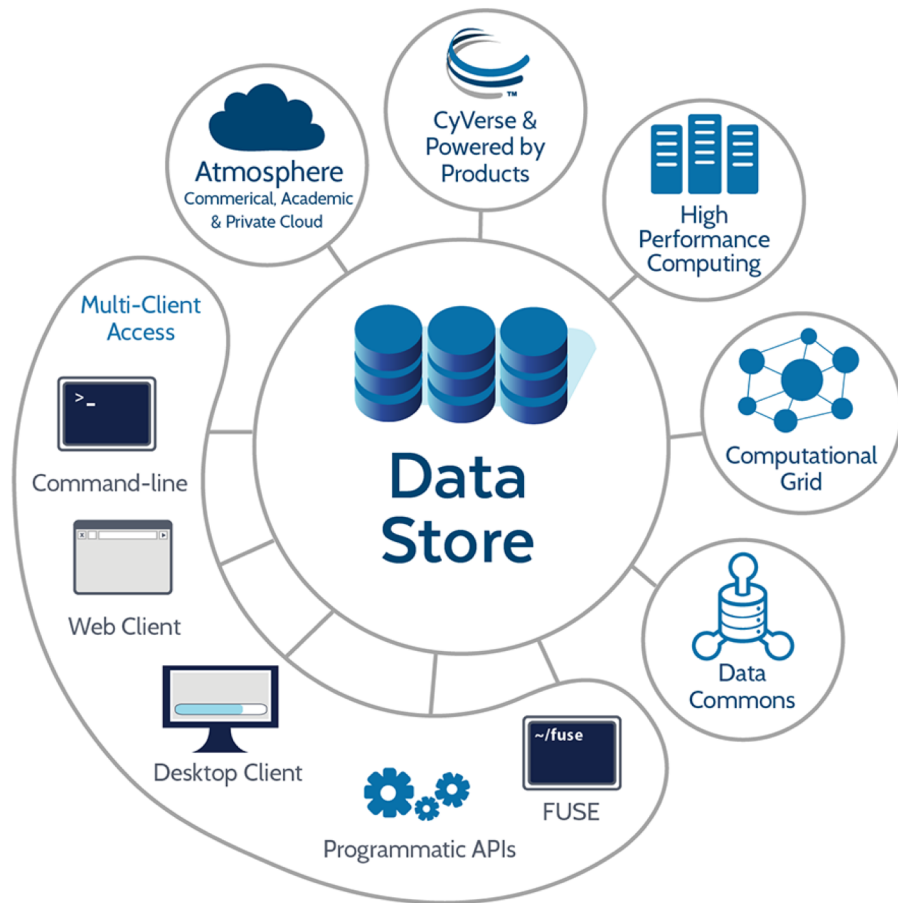
What is CyVerse?

Vision: Transforming Science through Data-Driven Discovery.

Mission: To design, deploy, and expand a national Cyberinfrastructure for Life Sciences research, and to train scientists in its use.



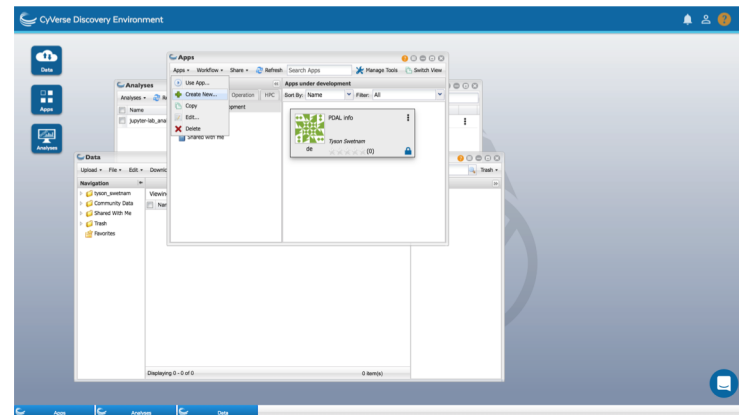
What is CyVerse?



What is the Discovery Environment?

- Easy to use interface for running command line applications
- Seamlessly integrated with data and high performance computing
- User extensible - ready integrate your applications using Docker

de.cyverse.org

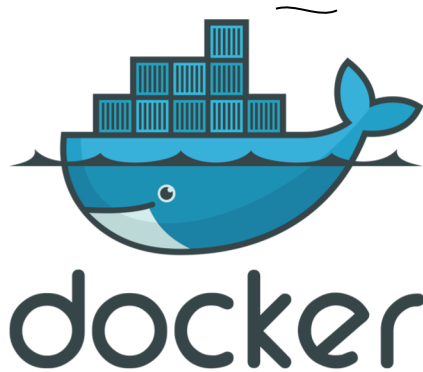


What is VICE?

Visual & Interactive Computing

– Discovery Environment

→ Allows users of the CyVerse Discovery Environment launch web applications packaged into Docker



→ Users access these apps through a secure and authenticated URL (`https://`)



Documentation: <https://learning.cyverse.org/projects/vice/en/latest/index.html>

See previous FF webinar on VICE to learn more: <https://www.youtube.com/watch?v=KpBC0nScfL0>

Performing a microbiome study

Study design
(power, randomization)
(metadata--standard + specific to your study)

Sample collection
(method, shipping/storage)
lab protocols
(include neg/pos controls)

Data processing

Data analysis

For more on these topics, see:

[Conducting a Microbiome Study](#), by Goodrich et al. 2014

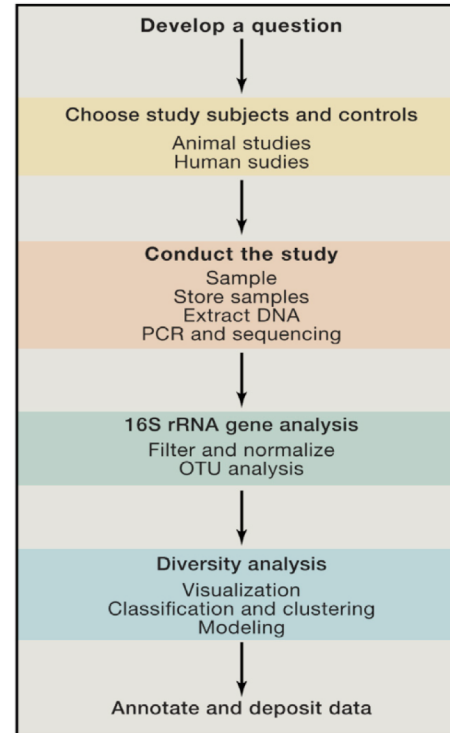
[Reagent Contamination](#), Salter et al. 2014

[Storage effects](#), by Song et al. 2016

[Microbiome Quality Control \(MBQC\)](#), by Sinha et al. 2017

[MIMARKS](#), by Yilmaz et al. 2011

[KatharoSeq low biomass workflow](#), by Minich et al. 2017



Sample Information

Sequence Data

Feature Table

Summary Statistics

Other resources:

[Earth Microbiome Project website](#)

[Human Microbiome Project website](#)

[American Gut Project website](#)





Bioinformatics Workflow

- Quantitative Insights Into Microbial Ecology
- canonically pronounced 'Chime'
- QIIME 2 is a powerful, extensible, and decentralized microbiome analysis package with a focus on data and analysis transparency.
- QIIME 2 enables researchers to start an analysis with raw DNA sequence data and finish with publication-quality figures and statistical results.



<https://qiime2.org>



[NSF Award: 1565100](#)

High-level features (these attract users to system)

- **Latest and greatest microbiome bioinformatics methods and visualizations.**
- **Accessibility** through accurate, detailed, and interesting documentation and well-designed interfaces.
- **A community** of microbiome scientists, developers, and bioinformaticians.



These slides were created and arranged by Greg Caporaso and other members of the QIIME and QIIME 2 development groups.





[NSF Award: 1565100](#)

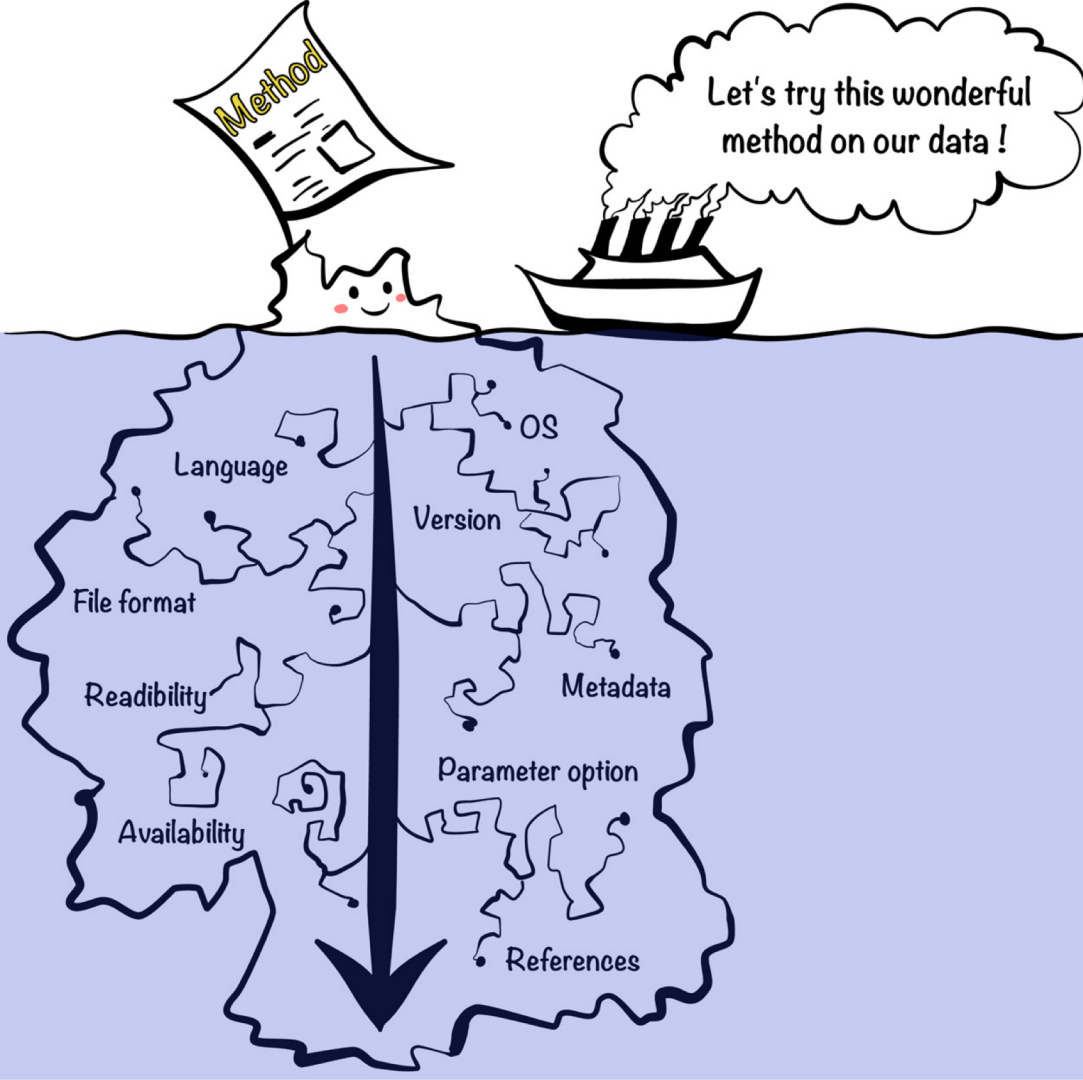
Low-level features (these get our users hooked)

- **Decentralized provenance tracking** automates bioinformatics record keeping facilitating reproducibility.
- **Multiple user interfaces.** The same functionality is accessible through graphical interface, command line interface, and API, which target different types of users.
- **Plugin architecture** allows the software to keep pace with the field. Any developer can create and distribute a QIIME 2 plugin.



These slides were created and arranged by Greg Caporaso and other members of the QIIME and QIIME 2 development groups.





Kim et al. (2018) [Experimenting with reproducibility: a case study of robustness in bioinformatics.](#)

These slides were created and arranged by Greg Caporaso and other members of the QIIME and QIIME 2 development groups.



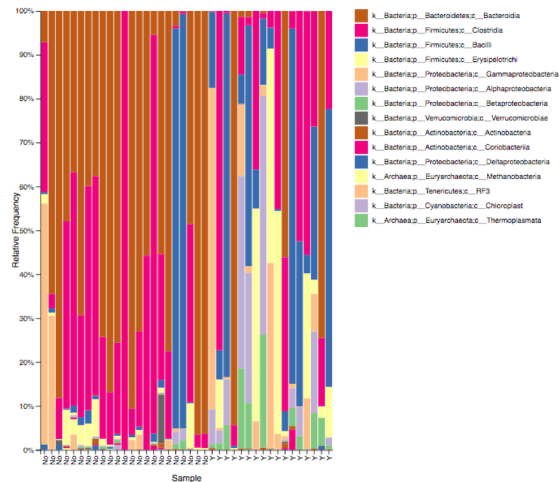
QIIME 2 integrated data provenance ensures reproducibility

qiime2view File: taxa-bar-plots-deblur.qzv Visualization Peek Provenance

Download Taxonomic Level Color Palette Sort Samples By

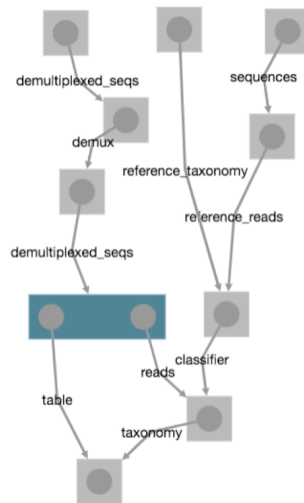
SVG (bars) SVG (legend) CSV Level 3 schemeAccent Infection_in_6 Ascending Relabel X?

Y | k__Bacteria;p__Bacteroidetes;c__Bacteroidia | 74.527%



qiime2view File: taxa-bar-plots-deblur.qzv Visualization Peek Provenance

Provenance Graph Citations



Action Details

- execution:
 - uuid: "42a57f0f-18db-47eb-aa14-4d81618a4b5a"
- runtime:
 - start: 2019-03-27T20:00:55.986Z
 - end: 2019-03-27T20:21:09.860Z
 - duration: "20 minutes, 13 seconds, and 874586 microseconds"
- action:
 - type: "method"
 - plugin: "environment:plugins:deblur"
 - action: "denoise_16S"
- inputs:
 - 0:
 - demultiplexed_seqs: "60168b22-9ac7-477f-b085-09df6a9c2557"
- parameters:
 - 0:
 - trim_length: 300
 - 1:
 - sample_stats: true
 - 2:
 - mean_error: 0.005

This interface can view .qza and .qzv files directly in your browser without uploading to a server. [Click here](#) to learn more.

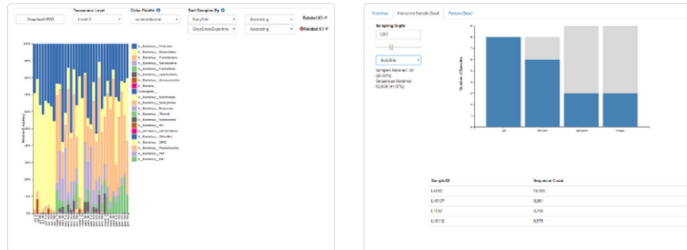
Drag and drop or click here

to view a QIIME 2 Artifact or Visualization (.qza/.qzv) from your computer.

You can also provide a link to a [file on Dropbox](#) or a [file from the web](#).

Gallery

Don't have a QIIME 2 result of your own to view? Try one of these!



q2view: a read-only web interface for viewing results without having QIIME 2 installed

- Reads QIIME 2 outputs (QZAs and QZVs)
- No installation required
- Easily share links using Dropbox with collaborators
- No uploading (your data stays on your computer)!

<https://forum.qiime2.org>

Please read our [Code of Conduct](#) when joining.

all categories ▾ all tags ▾ **Categories** Latest New (6) Unread (1) Top Bookmarks My Posts + Ne

Category Topics Latest

User Support

Post to this category if you need help understanding output produced while running QIIME 2. Examples of this include help understanding plots labels, techniques that are used in QIIME 2, etc. Posts in this category will be triaged by a QIIME 2 Moderator and responded to promptly.

98 / month
1 unread
2 new

F Form of OTU (ASV)
 User Support **queued** feature-table

P Problems with importing data
 Technical Support **import** **queued**

Technical Support

Post to this Category if you are experiencing a technical difficulty while running QIIME 2. Examples of difficulties include installation errors, help deciphering error messages, etc. Posts in this category will be triaged by a QIIME 2 Moderator and responded to promptly.

21 / month
2 new

q2-SCNIC: A tool for making correlation networks, finding modules of observations and summarizing them
Plugins

Community Plugin Support

Post to this category if you have a question about (bug report, technical detail, etc.). Community plugins distributed in the QIIME 2 Core Distribution. Please we are planning on moving away from the notion "Distribution," where all pl...

14 / month

Importing data from ion-torrent run
User Support **import**

General Discussion

Post to this category if you have a general question about science, bioinformatics, or other general question. Examples of posts include study design, etc. Posts in this category will not be triaged by a Moderator and responded to promptly.

Developer Discussion

Post to this category if you are a developer and have an idea, or suggestion. Posts in this category will be triaged by a Moderator and responded to promptly.

Community Contributions

Post to this category if you have a contribution ready. Examples include tutorials, plugins, doc translations. Posts in this category will not be triaged by a QIIME 2 Moderator and responded to promptly.

QIIME2中文帮助文档 (Chinese Manual)

Community Translations **in progress** **initial**

Yong-Xin Liu 20 Jul 17

QIIME 2 2017.7



QIIME 2™ is a next-generation microbiome bioinformatics platform that is extensible, free, open source, and community developed.

[Learn more >](#)

Automatically track your analyses with hierarchical data provenance — no more research on what commands were used.

Interactively explore your data with beautiful visualizations that provide new perspectives.

Easily share results with your team, even those members without QIIME 2 installed.

Plugin-based system — your favorite microbiome methods at its core.

Choose the interface that fits your needs

q2cli the command line interface



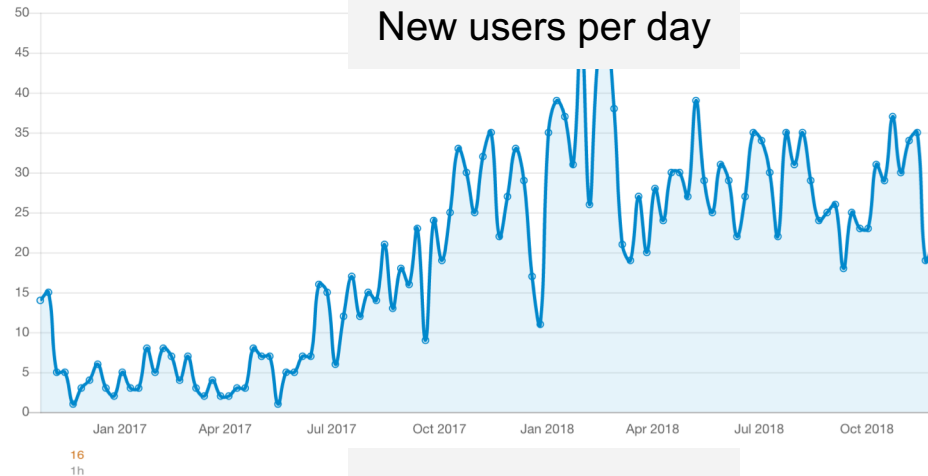
q2hub the graphical user interface



声明: 本文为QIIME2官方帮助文档的中文版, 由中科院遗传发育所刘永鑫博士翻译并监测有效, 文档翻译已获QIIME2团队官方授权. 由于QIIME2更新频繁, 如使用中遇到问题请访问QIIME2官方论坛阅读最新版中文帮助, 如中文翻译没有及时更新, 新闻读英文原版 <https://docs.qiime2.org>

本人只习惯使用命令行模式分析数据, 图形界面和python模式下载使用暂不介绍. 本系列的教程主要以命令行方式

New users per day



User visits per day



Using notebooks to record what you did

nature > toolbox > article

a natureresearch journal

MENU ▾

nature
International journal of science



Search



E-alert



Submit



Login

TOOLBOX · 30 OCTOBER 2018

Why Jupyter is data scientists' computational notebook of choice

An improved architecture and enthusiastic user base are driving uptake of the open-source web tool.

Jeffrey M. Perkel

<https://www.nature.com/articles/d41586-018-07196-1>



New to Project Jupyter?





Live Demonstration

Quick Launch : <https://learning.cyverse.org/projects/vice/en/latest/index.html>

Considerations and Limitations

- QIIME 2 updates – quarterly, consult CyVerse for latest
- Max runtime of 48 hours
- Increasing Cores, GPU, RAM, Duration
- Starting point! Look forwards to future webinar and updates within QIIME 2 shared folder





Questions

email: leej2@hhmi.org

Information about example dataset

- **“Infectious Complications Are Associated With Alterations in the Gut Microbiome in Pediatric Patients With Acute Lymphoblastic Leukemia”**
- <https://www.frontiersin.org/articles/10.3389/fcimb.2019.00028/>
- Find raw FASTQ F/R reads in ‘gut-microbiome’ data folder in :
 /iplant/home/shared/iplantcollaborative/example_data/qiime2
- Find Jupyter Notebooks and metadata file in :
 /iplant/home/shared/iplantcollaborative/example_data/qiime2
- Will update the full QIIME2 analysis by Monday April 8, 2019





CyVerse is supported by the National Science Foundation under
Grants No. DBI-0735191, DBI-1265383, and DBI-1743442.



Cold
Spring
Harbor
Laboratory



How do VICE Apps Work?

