

End-to-End Metagenomics: decision making

CINEMA Workshop

July 2023

Is there a “standard” pipeline?

SqueezeMeta, A Highly Portable, Fully Automatic Metagenomic Analysis Pipeline

 Javier Tamames* and  Fernando Puente-Sánchez

Software | [Open Access](#) | Published: 15 September 2018

MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis

[Gherman V. Uritskiy](#), [Jocelyne DiRuggiero](#) & [James Taylor](#)

[Microbiome](#) 6, Article number: 158 (2018) | [Cite this article](#)

Article | [Open Access](#) | Published: 07 June 2022

Metagenomic pipeline for identifying infections among distinct SARS-CoV-2 sequences of concern: study cases from

[Jose Arturo Molina-Mora](#), [Estela Cordero-Laurent](#), [Chacón-Ramírez](#) & [Francisco Duarte-Martínez](#)

[Scientific Reports](#) 12, Article number: 9377 (2022)

Struo: a pipeline for metagenome profilers

[Jacobo de la Cuesta-Zuluaga](#), [Ruth E Ley](#), [Nicholas D Youngblut](#)

[Bioinformatics](#), Volume 36, Issue 7, April 2020, Pages 2314–2315,

Microseek: A Protein-Based Metagenomic Pipeline for Virus Diagnostic and Discovery

by  Philippe Pérot^{1,†},  Thomas Bigot^{2,†},  Sarah Temmam¹,  Béatrice Regnault¹ and  Marc Eloit^{1,3,*}

MetAMOS: a modular and open source metagenomic assembly and analysis pipeline

[Todd J Treangen](#), [Sergey Koren](#), [Daniel D Sommer](#), [Bo Liu](#), [Irina Astrovskaya](#), [Brian Ondov](#), [Aaron E Darling](#), [Adam M Phillippy](#) & [Michael D Ragan](#)

[Genome Biology](#)

Assessment of Common and Emerging Bioinformatics Pipelines for Targeted Metagenomics

Methodology | [Open Access](#) | [Léa Siegwald](#), [Hélène Touzet](#), [Yves Lemoine](#), [David Hot](#), [Christophe Audebert](#), [Ségolène Caboche](#)

[BioMaS: a](#) Published: January 4, 2017 • <https://doi.org/10.1371/journal.pone.0169563>

Review | [Published: 18 October 2021](#)

Music of metagenomics—a review of its applications, analysis pipeline, and associated tools

[Bilal Wajid](#), [Faria Anwar](#), [Imran Wajid](#), [Haseeb Nisar](#), [Sharoze Meraj](#), [Ali Zafar](#), [Mustafa Kamal Al-Shawaqfeh](#), [Ali Riza Ekti](#), [Asia Khatoon](#) & [Jan S. Suchodolski](#)

[Functional & Integrative Genomics](#) 22, 3–26 (2022) | [Cite this article](#)

Review

Ocean viruses: Rigorously evaluating the metagenomic sample-to-sequence pipeline

[Matthew B. Sullivan](#)

with the

[Sergei Kosakovsky Pond](#)^{1,2,6,9}, [Samir Wadhawan](#)^{3,6,7}, [Francesca Chiaromonte](#)⁴, [Guruprasad Ananda](#)^{1,3}, [Wen-Yu Chung](#)^{1,3,8}, [James Taylor](#)^{1,5,9}, [Anton Nekrutenko](#)^{1,3,9} and [The Galaxy Team](#)¹

m: an extensible pipeline for analyzing genomic sequencing experiments

[Louis J. Taylor](#), [Chunyu Zhao](#), [Andrew Connell](#), [Jung-Jin Lee](#), [Bryton Fett](#), [Frederic D. Bushman](#) & [Kyle Bittinger](#)

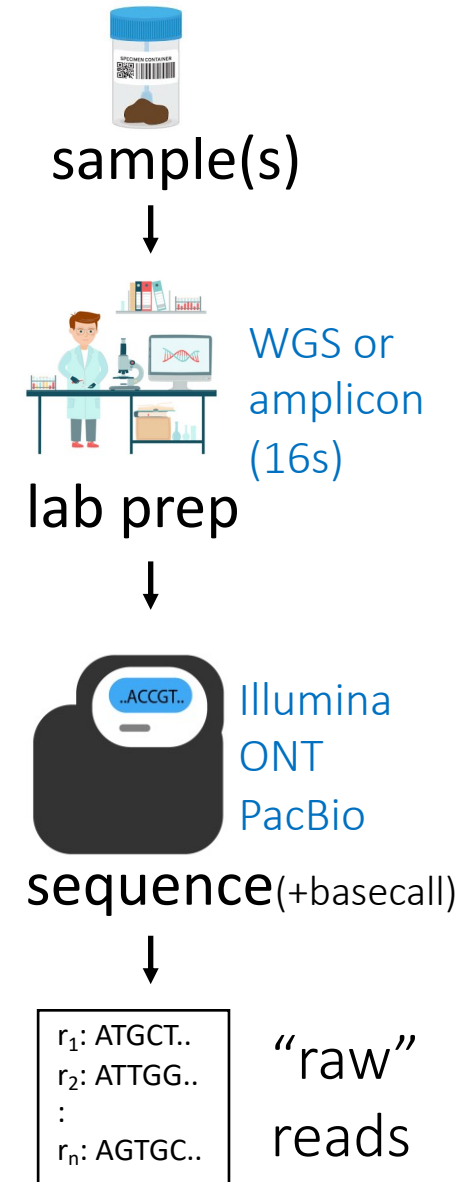
[Microbiome](#) 7, Article number: 46 (2019) | [Cite this article](#)

What questions are you looking to answer?

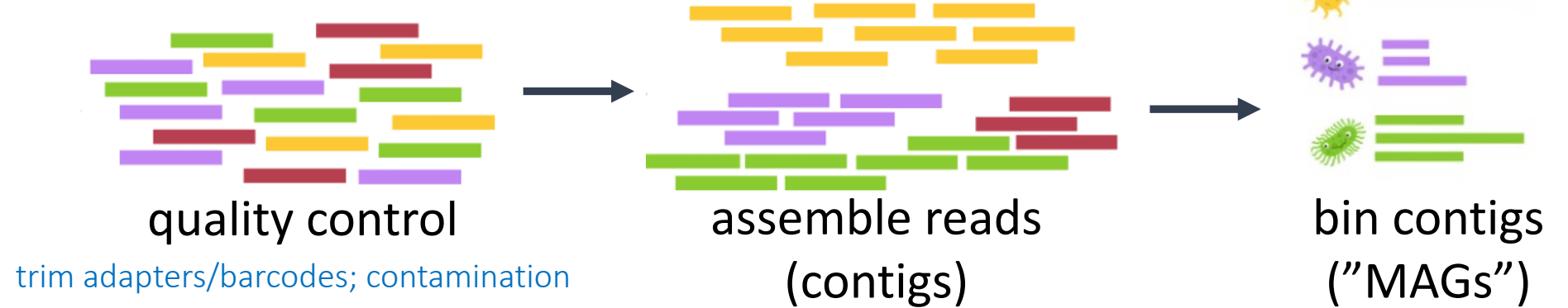
-
-
-
-
-
-
-
-
-
-

Metagenomics pipeline

1. generate reads



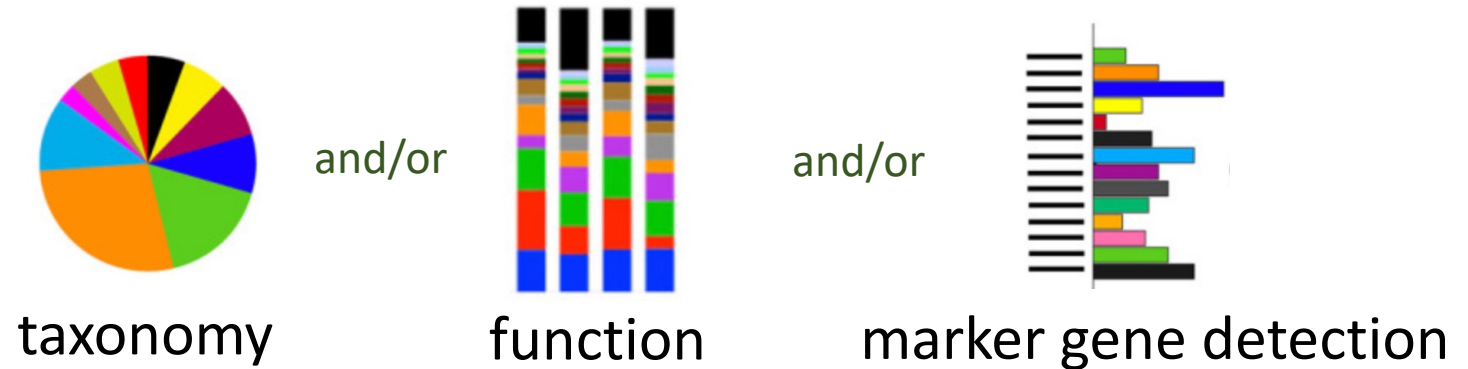
2. clean and sort



3. classify

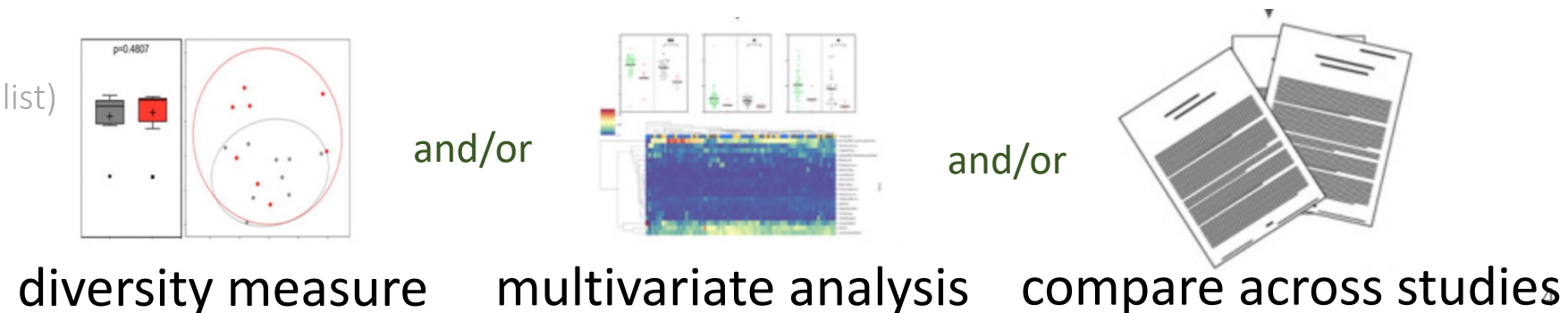
(non-exhaustive list)

input:
reads, contigs, or MAGs



4. stats

(non-exhaustive list)





Long or short reads?



WGS or targeted amplicon (16s)?

Utilizing Rice Virtual Machines

1. Obtain a login from one of your TAs

1. Login from your command line

\$ssh studentX@cinema1.rice.edu

2. Paths to datasets

Zymo Mock (Tuesday): /home/kdc10/datasets/zymo_mock

Waste water contigs (Wednesday): /home/kdc10/datasets/waste_water

Vaginal 16S ONT (Friday): /home/kdc10/datasets/vaginal_16S_ONT