

# INTRODUCTION TO EDGE BIOINFORMATICS PIPELINE

RONALD KIIZA

29<sup>th</sup> June 2023

# EDGE BIOINFORMATICS



- EDGE = **E**mpowering the **D**evelopment of **G**enomic **E**xpertise
- EDGE is a highly integrated and interactive web-based platform .
- Developed for biologists
- Analyse microbial and metagenomics data
- Composed of very many pipelines
- Runs many of the standard analyses required for **viral**, **bacterial**, and **metagenomic** samples.

# EDGE BIOINFORMATICS



- Analysis of some eukaryotes not fully supported
- Developed, maintained and occasionally updated by the
- Updated current version of EDGE is **v2.4.0**



# EDGE METAGENOMICS PIPELINE



Raw  
sequencing data  
(FASTQ files)

Data QC and Host  
Removal

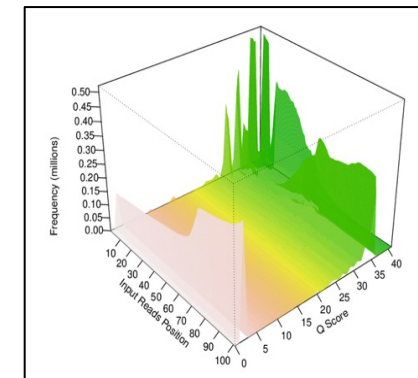
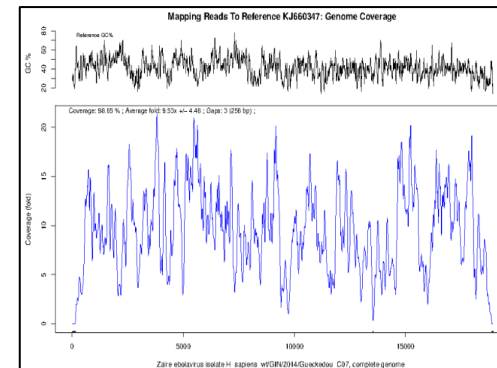
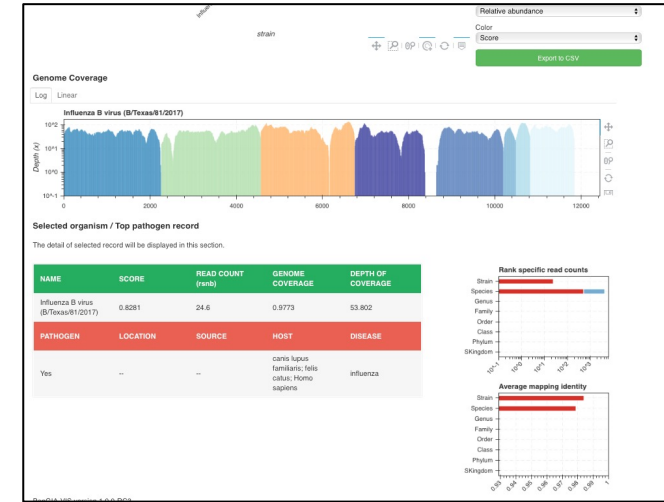
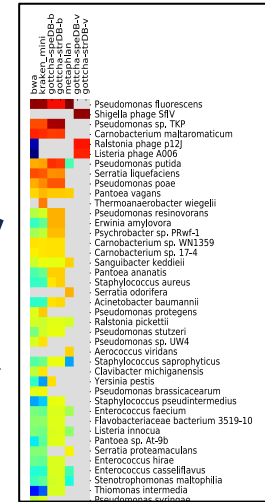
Assembly and  
Annotation

Reference-based  
Analysis

Taxonomic  
Classification

Phylogenetic  
Analysis

PCR Primer  
Analysis



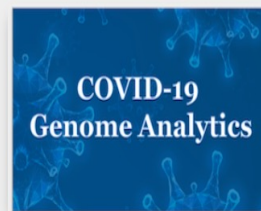


# HOW TO RUN EDGE BIOINFORMATICS

- <https://edgebioinformatics.org/>

## Requirements to run EGDE

- EDGE account
- Project name – **Avoid including spaces in the project name**
- Data – Fastq files, Fasta files (Uploaded), Publicly available data (SRA)
- Strong internet connection (GUI)



**COVID-19 GENOME ANALYTICS**

A platform for COVID-19 analytics



**NEW TO EDGE?**

2.0.0



**RUN EDGE**

2.4.0

EDGE bioinformatics



**RUN QIIME2**

2019.10

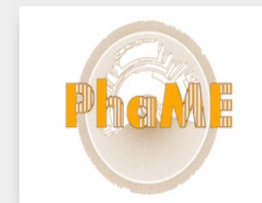
Amplicon Data Analysis, 16s, 18s,  
and Fungal ITS



**RUN PIRET**

0.3.2 beta

A pipeline for Reference based  
Transcriptomics analysis.



**RUN PHAME**

1.0.0

Phylogenetic and Molecular Evolution  
(PhaME) analysis tool.

 Home

Upload Files

Run EDGE

Run Qiime2

Run DETEQT

Run PiReT (BETA)

Job Queue

Reports

Projects


## Input Your Sample

EDGE requires **FASTQ** sequence data files in FASTQ format or **Contigs** sequence data file in FASTA format. EDGE allows both paired-end and single-end sequences.

 **Input Raw Reads** **Batch Project Submission**

## Choose Processes / Analyses

EDGE provides many modules to do various analyses. You can choose to run or skip a specific process. Parameters/options are provided for most of the analyses. You can click here to [turn all on](#), [expand all sections](#) or [close all sections](#).

 **Assembly and Annotation****On** ☒ **Reference-Based Analysis**☐ **Off** **Taxonomy Classification****On** ☒ **Phylogenetic Analysis**☐ **Off** **Gene Family Analysis**☐ **Off** **PCR Primer Analysis**☐ **Off** **JBrowse****On** ☒

Submit

Reset

Run DETEQT

Run PiReT (BETA)

Job Queue

Reports

Projects

Project/Run Name

(required, at 3 but less than 30 characters)

Description

Input Source

Nanopore Reads

Pair-1 FASTQ File

Pair-2 FASTQ File

Single-end FASTQ File

### Select a File

PublicData

- 2019workshop
  - Bacillus
  - Burk
  - ClinicalMeta
    - FQ 1055\_1079-KAPA\_A9\_\_CGGCTATG-TATAGCCT\_\_S28\_R1\_001.fastq.gz**
    - FQ 1055\_1079-KAPA\_A9\_\_CGGCTATG-TATAGCCT\_\_S28\_R2\_001.fastq.gz**
    - FQ 1055\_1080-KAPA\_B9\_\_CGGCTATG-ATAGAGGC\_\_S6\_R1\_001.fastq.gz
    - FQ 1055\_1080-KAPA\_B9\_\_CGGCTATG-ATAGAGGC\_\_S6\_R2\_001.fastq.gz
    - FQ 1055\_1081-KAPA\_E1\_\_ATTACTCG-AGGCGAAG\_\_S23\_R1\_001.fastq.gz
    - FQ 1055\_1081-KAPA\_E1\_\_ATTACTCG-AGGCGAAG\_\_S23\_R2\_001.fastq.gz
    - FQ 1055\_1082-KAPA\_A11\_\_TCTCGCGC-TATAGCCT\_\_S12\_R1\_001.fastq.gz
    - FQ 1055\_1082-KAPA\_A11\_\_TCTCGCGC-TATAGCCT\_\_S12\_R2\_001.fastq.gz
    - FQ 1055\_1125-KAPA\_F9\_\_CGGCTATG-TAATCTTA\_\_S20\_R1\_001.fastq.gz
    - FQ 1055\_1142\_1143-KAPA\_B4\_\_GAGATTCC-ATAGAGGC\_\_S8\_R1\_001.fastq.gz
    - FQ 1055\_1146\_1147-KAPA\_C4\_\_GAGATTCC-CCTATCCT\_\_S9\_R1\_001.fastq.gz

| additional options |

 **Batch Project Submission**



# Current metagenome read-based profiling tools in EDGE

## Keep in mind...

- Make sure the database includes the organisms you're looking for (bacteria, virus, eukaryotes)
- Some tools are faster than others and if you select all, your run will take a very long time to complete

Specific  
F FP

Sensitive  
Few FN

Choose Classification Tools:	
GOTTCHA Bacterial Databases	
GOTTCHA Genus	<input type="checkbox"/>
GOTTCHA Species	<input type="checkbox"/>
GOTTCHA Strain	<input type="checkbox"/>
GOTTCHA Viral Databases	
GOTTCHA Genus	<input type="checkbox"/>
GOTTCHA Species	<input type="checkbox"/>
GOTTCHA Strain	<input type="checkbox"/>
GOTTCHA2 BacteriaViruses Databases	
GOTTCHA2 Species	<input type="checkbox"/>
PanGIA Databases	
PanGIA NCBI_genomes_p_GRCh38	<input type="checkbox"/>
PanGIA NCBI_genomes_p_GRCh38 (orig)	<input type="checkbox"/>
Reads Mapping	
Reads Mapping (BWA against RefSeq)	<input type="checkbox"/>
Other Tools	
MetaPhlAn2	<input type="checkbox"/>
Kraken2	<input type="checkbox"/>
Centrifuge	<input type="checkbox"/>
DIAMOND (Amino acid-based classification)	<input type="checkbox"/>

# Contact information for registering the EDGE accounts

## Contact Us and Citation

Questions? Concerns? Please feel free to email our google group at [edge-users@googlegroups.com](mailto:edge-users@googlegroups.com) or contact a dev team member listed below.

Name	Email
Patrick Chain	<a href="mailto:pchain@lanl.gov">pchain@lanl.gov</a>
Chien-Chi Lo	<a href="mailto:chienchi@lanl.gov">chienchi@lanl.gov</a>
Paul Li	<a href="mailto:po-e@lanl.gov">po-e@lanl.gov</a>
Karen Davenport	<a href="mailto:kwdavenport@lanl.gov">kwdavenport@lanl.gov</a>
Logan Voegtly	<a href="mailto:logan.j.voegtly.ctr@mail.mil">logan.j.voegtly.ctr@mail.mil</a>
Kim Bishop-Lilly	<a href="mailto:kimberly.a.bishop-lilly.ctr@mail.mil">kimberly.a.bishop-lilly.ctr@mail.mil</a>

### 12.1 Citation

Enabling the democratization of the genomics revolution with a fully integrated web-based bioinformatics platform

Po-E Li; Chien-Chi Lo; Joseph J. Anderson; Karen W. Davenport; Kimberly A. Bishop-Lilly; Yan Xu; Sanaa Ahmed; Shihai Feng; Vishwesh P. Mokashi; Patrick S.G. Chain

Nucleic Acids Research 2016;

doi: [10.1093/nar/gkw1027](https://doi.org/10.1093/nar/gkw1027)

# Important documentation for using EDGE

- <https://buildmedia.readthedocs.org/media/pdf/edge/develop/edge.pdf>
- <https://academic.oup.com/nar/article/45/1/67/2572059>

*PRACTICAL DEMO !!!*

*THANK YOU!!!!*