SEA-PHAGES Bioinformatics Workshop Overview

hhmi

Objectives: Overall

- **Characterize and investigate phage genomes** including the details of gene calling, functional assignments, and preparing final submission files **that meet QC requirements**.
- This year's genome is Mycobacterium phage Madiba.
- **Install and use the software** for annotation and analysis of phages. The software includes DNA Master, Phamerator, and other web-based tools.
- **Identify basic phage biology concepts** that underlie bioinformatic investigations.
- **Explore classroom implementation strategies.**

Objectives: Daily: See each day's list

Online Bioinformatics Training Workshop Schedule and Assignments
December 6 – 17, 2021, 11:00AM – 1:00PM

Note: Some of the reading assignments refer to a section in the guide with multiple pages. Participants should read all of the pages in the assigned section.
*Annotation Outline is an outline of much of the rest of the guide and is intended to be a quick reference. You are not responsible for reading all of the linked pages in this section; merely note the organization.

Week 1

Day December	Live Sessions	Pre-live Session Readings	Short Videos	Tasks
6th	• Welcome and Overview	• Introduction • Annotation Outline* • Genome annotation Overview • Decision Trees • Mechanics: Part 1(through "Generating a rough draft…)	• Auto-annotation • BLASTing your genome in DNA Master • Saving your DNA Master file	• Auto-annotate and then BLAST Madiba in DNA Master • Save File
7th	• Genome Overview • Compare genomes using Phamerator • Compare using BLASTN on phagesdb.org	• Guiding principles • Mechanics: Phamerator • Mechanics: Other	• Phagesdb tour • Phamerator overview • GeneMark output • Aragorn output • tRNAscanSE output	• Generate GeneMark S and GeneMark host files • Generate Aragorn and tRNAscanSE outputs. • Phamerator map of Madiba and others
8th	• Call genes 6-12	• Evaluating gene	• Using Starterator	• Call genes

Found on the seaphages.org meeting page

Primary Learning Method

- Learn by doing
  - Gather data
  - Do the work: annotate & analyze the genome
  - Do science like scientists do it. [Translation: new technology changes at a moment's notice]

Workshop Logistics

- Workshop page at seaphages.org
- Try to not work in isolation
- Use the SEA-PHAGES Bioinformatics Guide
- Lab notebook. Take notes while annotating!
- seaphages.org forums: Let's model that at the workshop so it will be easy later!
- Daily objectives and assessments

We Will…

- Use mycobacteriophage Madiba genome to
  - Learn to manipulate the tools we use
    - DNA Master
    - GeneMark, Glimmer
    - BLAST
    - Phamerator
    - HHpred
    - Starterator
  - Practice calling genes and functions
- Investigate other types of genes you will come across
- Produce a **quality** genome annotation
  - Sequence quality control: ✓
  - Annotation quality control: Checklist and files
  - Review to Improve

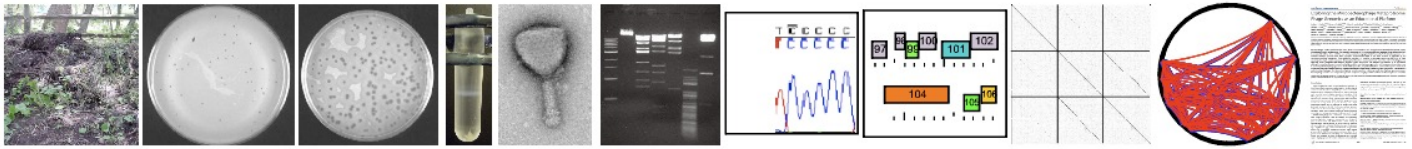
Your learning starts here, but continues.....

- Final file preparations on-line training
  - Time to be announced, late March – April
  - Content will include formatting requirements
  - How to pass preliminary QC
  - Do not hold annotation questions until then, use the forums
- Review to Improve reporting
  - Once your genome is checked, you must submit a comparison report of your annotation with the submitted version
    - How did you do?
  - After you have 3 genomes compared, are you ready to be a trusted submitter?

Important Context

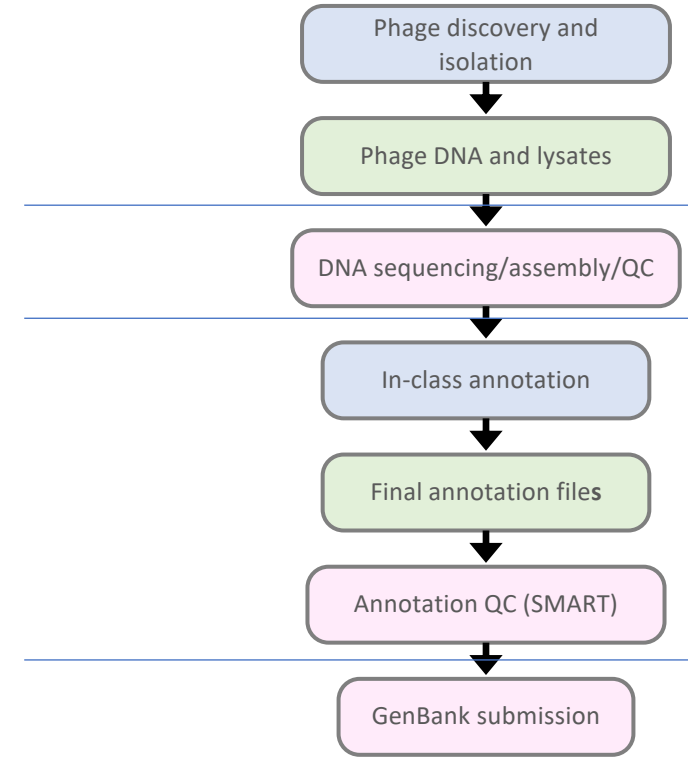
- Biological Information Flow
- Sequencing and Finishing Genomes
- Bioinformatic Information Flow

Concrete → Representational



1 Isolation
2 Purification
3 Amplification
4 Electron Microscopy
5 DNA analysis
6 DNA Sequencing
7 Annotation
8 DNA sequence comparison
9 Comparative analyses
10 Publication

First semester start: Sept. 1



DNA due: Nov. 15

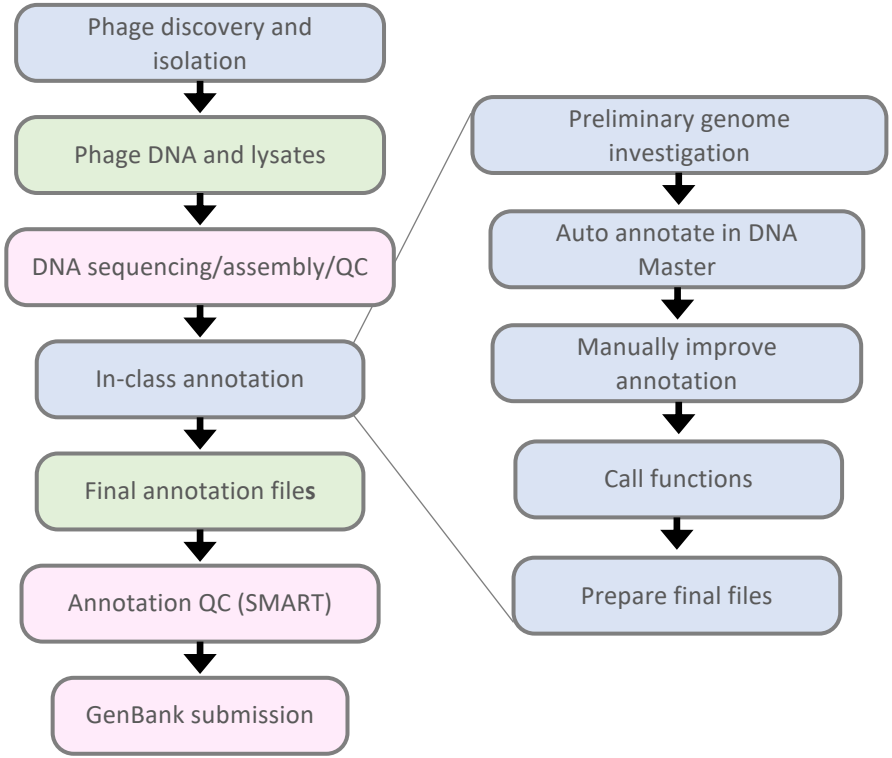
Nov. 15 – Feb. 29

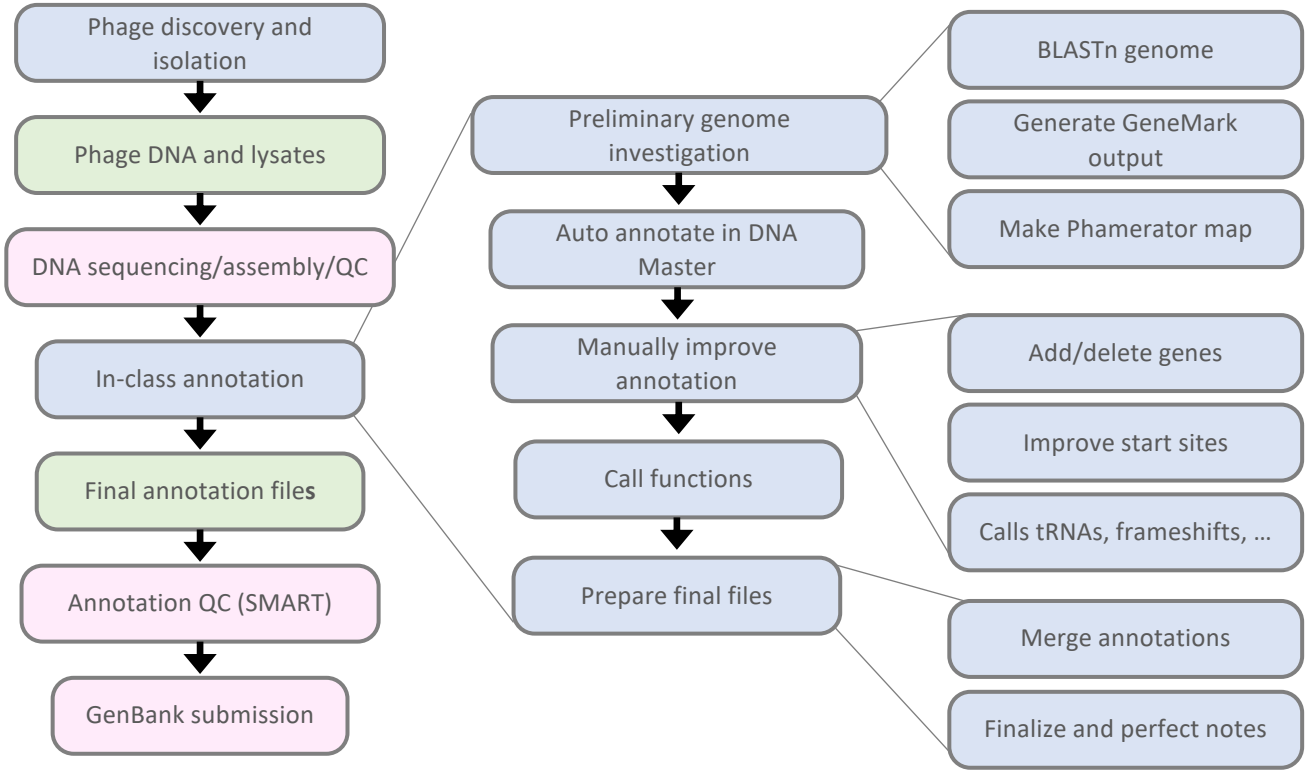
2nd semester start: Jan. 1

Files are due: May 1

Files to GenBank: Sept. 1

First semester start: Sept. 1


Phage discovery and isolation
Phage DNA and lysates
DNA sequencing/assembly/QC
In-class annotation
Final annotation files
Annotation QC (SMART)
GenBank submission
Preliminary genome investigation
Auto annotate in DNA Master
Manually improve annotation
Call functions
Prepare final files


Phage discovery and isolation
Phage DNA and lysates
DNA sequencing/assembly/QC
In-class annotation
Final annotation files
Annotation QC (SMART)
GenBank submission
Preliminary genome investigation
Auto annotate in DNA Master
Manually improve annotation
Call functions
Prepare final files
BLASTn genome
Generate GeneMark output
Make Phamerator map
Add/delete genes
Improve start sites
Calls tRNAs, frameshifts, …
Merge annotations
Finalize and perfect notes