


Omics Logic Online Summer Programs 2020

			
<i>Program</i>	Omics Logic Basics Online	Bioinformatics for Infectious Diseases	SARS-COV2: Genomic Data Analysis
Length	3 months	2 months	1 month
Live sessions	NA	Yes	Yes
Number of Live Group Sessions	NA	13	10
Online Chat Assistance	Yes	Yes	Yes
Expert Q&A	NA	Yes	Yes
One-one-one support	NA	Yes	NA
Bioinformatics Project	NA	Yes	Yes
Access to Educational Platform	asynchronous* online courses	asynchronous* online courses	asynchronous* online courses
Access to Server	Demo	Educational	Educational
Certificate	Course certificates	Program Certificate	Program Certificate

* Asynchronous courses are available for registered members only

** Live sessions and expert support are offered online via ZOOM, scheduled in advance

2020 Summer Bioinformatics Programs Details:

1. OmicsLogic Online

This program is best suited for students interested to learn about various -omics technologies and how bioinformatics is used in biotechnology, healthcare, agriculture and basic research. Program access provides access to all the [asynchronous* online courses](#) (basic course certificates only):

1. **Introduction to bioinformatics (1 course)**
2. **Genomics (3 courses)**
3. **Transcriptomics (4 courses)**
4. **Metagenomics (3 courses)**
5. **Epigenomics (4 courses)**
6. **Biomedical Data Science (3 courses)**

Program Length: 3 months, certification: individual course

Each course includes quizzes and some courses include practical assignments that can be completed before program end to receive certificates of completion. A passing quiz grade is required for completion. Assignments are checked for pass/no pass and can be resubmitted after feedback.

Level of difficulty: intermediate, great for a flexible schedule and independent learning.

2. COVID-19 Genomics

Guided online program** designed to address the Basics of COVID-19. The program prepares beginners to start the very basics of viral genomics using COVID-19 and work on COVID-19 Bioinformatics project prepared by Experts. The program provides training about the novel coronavirus genomics and understand how genomic data analysis tools can help identify specific viral strains, understand multiple sequence alignment, phylogenetic analysis and the significance of mutations in the context of viral protein structure and function.

Program Length: 1 month, certification: Program and individual course

Includes access to introductory [asynchronous online courses](#) and [project](#) based on current research publications on SARS-COV-2 in various domains (oncology, neurodegenerative diseases, agriculture and biotechnology) using public domain data. Course material is suggested based on assessment given to participants at the start of the course.

Many of the included asynchronous courses include quizzes. Some courses include practical assignments that can be completed within this 3-month period to receive certificates of completion. A passing quiz grade, completion of online assignments and participation of 30% of the sessions is required for completion. Student activity is recorded and analyzed to resolve technical issues and generate points that indicate proficiency in acquired skills and comprehension of covered material.

Level of difficulty: Basic & Intermediate, great for generating interest in Beginners, perfect for development of critical thinking, problem solving and data-driven research for COVID-19.

3. Bioinformatics for Infectious Diseases (Viral and Bacterial Infections)

Guided online program** designed to provide comprehensive training in next generation sequencing data techniques for viral and bacterial genomes (i.e. viral and metagenomic amplicon sequencing) as well as experience using analysis tools such as multiple sequence alignment (MSA: ClustalW), phylogenetic evolutionary analysis (BEAST), functional annotation (NCBI Virus) and protein function (PDB, ChimeraX) as well as big data analysis approaches like GWAS. Includes an introduction to host response (RNA-Seq) using publicly available data on Ebola, Coronaviruses, EV-D68 and Tuberculosis.

Program Length: 2 months, certification: Program and individual course

Includes access to select asynchronous online courses and projects on viral and bacterial infectious diseases based on current research publications and public domain data.

Many of the included asynchronous courses include quizzes. Some courses include practical assignments that can be completed before program end to receive certificates of completion. A passing quiz grade, completion of online assignments and participation of 30% of the sessions is required for completion. Student activity is recorded and analyzed to resolve technical issues and generate points that indicate proficiency in acquired skills and comprehension of covered material.

Level of difficulty: intermediate, great for an in-depth review of different types of data and medical topics, perfect for a mix of guided and independent learning.

* Asynchronous courses are available for registered members only

** Live sessions and expert support are offered online via ZOOM, scheduled in advance

SARS-COV-2: Genomic Data Analysis

Using genomics to understand the COVID-19 pandemic (high school students and undergraduate college students interested in research as well as “citizen scientists”)



OMICSLOGIC
BIOINFORMATICS TRAINING

STAY@HOME&LEARN
ABOUT THE VIRUS

In this program, you will learn about genomic data, where to find it and how to analyze it. By analyzing publicly available data you will learn about the **COVID-19** pandemic origin, evolution, cell entry and replication.

JOIN TODAY!

INTRODUCTORY COURSE

ANALYSIS OF SARS-COV-2 GENOMIC DATA

1. Learn where to find appropriate genomic data for analysis
2. Analyze big data using intuitive cloud based tools
3. Read COVID-19 publications with understanding
4. Understand key principles of vaccine and drug design

© Pine Biotech 2020. All rights reserved.

Biorix PLATFORM PINE.BIO TAUBER

Program Topics:

The 2019–20 coronavirus pandemic has had far-reaching consequences lethal for many and disruptive to almost everyone around the world. Many of us are now interested to contribute our efforts to study, research and understand how we **COVID-19** can be stopped and what can be done to prevent such pandemics in the future. The Pine Biotech team along with our collaborators at the Tauber Bioinformatics Research Center, Louisiana Biomedical Research Network and Georgetown University Medical Center has been actively working to provide computational resources and training materials to scientists, researchers and students all over the globe who want

to learn more about this virus. As a citizen scientist, you too can be part of this global initiative and contribute by understanding what data is available and how it can be studied using **bioinformatics**.

In this program, you will learn about the novel coronavirus genomics and understand how genomic data analysis tools can help identify specific viral strains, understand how they differ from previously studied viruses by using such tools as multiple sequence alignment, phylogenetic analysis and characterization of mutations in the context of viral protein structure and function. We will apply these analytical methods to publicly available SARS-COV2 genomic data, learning about the virus emergence, spread and disease pathology. As a result, you will have the chance to perform your own analysis and develop a research question related to this pandemic.

This online program includes: Instructor support via online interactive sessions, tracking of student progress via online course and pipeline activity, access to online educational resources and analysis tools for self-paced learning, access to bioinformatics experts for guidance and problem-solving as well as educational use of analytical pipelines. SARS-COV-2: Genomic Data Analysis program is supported by our cloud-based tools that enables all program participants to analyze complex genomic data generated during this pandemic without having to invest in high-end computational infrastructure.

Registration Details: Online Coursework and 10 ZOOM sessions (total of 15 hours), Duration: 30 days training, followed by an optional research license for independent research.

Program Outcomes:

- A. Learn where to find appropriate genomic data for analysis
- B. Analyze publicly available data using intuitive cloud-based tools
- C. Learn to read research publications on this topic with understanding
- D. Understand key principles of vaccine & antiviral drug design

SARS-COV2 Genomics Program Syllabus

1. What is Next-generation sequencing?

Learn about Next Generation Sequencing techniques used to detect viral genomes in clinical samples and generate genomic data about pathogens and infected cells: short reads, genomic sequences, file formats and laboratory techniques for data preparation.

Associated online course/resource: Introduction to Metagenomics, Metagenomics 1

2. Looking for viral genomes in a sample (part 1)

Using NGS data to find a new pathogen: bioinformatics pipelines and processing steps to structure genomic data using BowTie, HiSat and STAR, annotation of identified sequences using the NCBI virus reference database.

Associated online course/resource: Spodoptera frugiperda and the contamination of biologics production (<https://edu.t-bio.info/projects/spodoptera-frugiperda-fall-armyworm/>)

3. Identifying a novel virus genome

Extracting reads that did not map to the host genome from FASTQ files after alignment, visualizing how they align to specific viral genomes from a database of viral genomes. Understanding genomic variation in short reads.

Associated online course/resource: Introduction to Genomics

4. Multiple Sequence Alignment and Phylogeny

How closely related are different viruses? Comparing genomic sequences (Multiple Sequence Alignment) and finding a consensus sequence (ShoRAH) in FASTQ files or between multiple FASTA files available on NCBI.

Associated online course/resource: Introduction to Metagenomics

5. Finding reliable genomic data on COVID-19 on NCBI

Public resources where SARS-COV-2 data is made available. Types of databases, access control and utilization. Finding the right genomic sequences, using NCBI alignment to check for quality and preparing data for analysis.

Associated online course/resource: NCBI Virus Database

6. Phylogenetic Analysis and Probable Origins of SARS-COV2

Studying evolutionary analysis of viral genomes – comparison of variable regions and identification of genomic variation. Mutations, conservation, and viral evolution. Modeling relationships between sequences based on probability (phylogeny, evolution and conservation)

Associated online course/resource: Genomics 1

7. Hands-on session: using found data to prepare and run a bioinformatics pipeline

Bioinformatics – a step-by-step overview of a pipeline that is used to align sequences, translate trinucleotide segments into amino acids and use the alignment for phylogenetic analysis using BEAST.

8. Q&A and DISCUSSION of pipeline results

Different Workflows: what to do if we have FASTA or FASTQ files? Which databases to use: Detection of viral genomes by mapping on databases. Interpretation of Phylogenetic Analysis: Evolutionary relationships between genomes, evolutionary time and specific changes of interest.

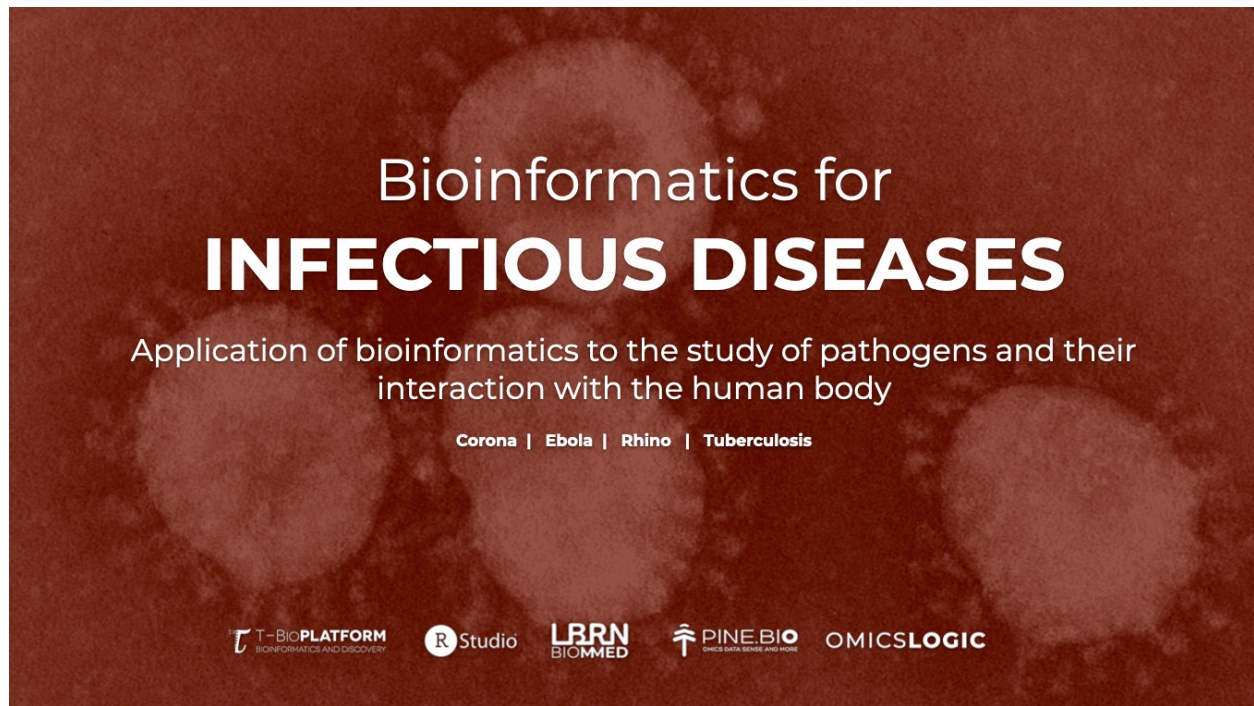
9. Discussion about Program topics and literature review

Utilization of NGS genomics to study viral diseases and understand emergence and evolution of pandemic outbreaks. Zoonotic spillover events and cross-species viral adaptation. Analysis of viral genomic data to explore potential SARS-COV2 origins, and transmission by finding and analyzing public domain data.

10. Bioinformatics in vaccine research and antiviral drug design

In conclusion, we will review how recent examples from the current and previous outbreaks of viral infections (i.e. Ebola 2014-2016) reveal the significance of widely accessible and accurate bioinformatics tools as well as the significance of bioinformatics skills for data-driven life science research enabled by high throughput data and rapid molecular sequencing. The direct outcome of evolutionary studies, functional annotation and publicly accessible genomic data enables the translation of research findings into vaccine and antiviral drug applications.

Bioinformatics for Infectious Diseases



This program is dedicated to the study of viral diversity and its role in epidemic infectious diseases that keep re-emerging, including zoonotic spillover, transmission between humans and the process of viral and bacterial disease development. Participants will get a chance to learn about bioinformatics and analyze genomic data by applying various analysis approaches to study viral genomes. As a result, you will learn to understand relationships between viral strains and haplotypes, find differences in sequence data and see the implications for drug and vaccine design. This program will provide opportunities to practice analyzing data to gain hands-on experience with curated datasets from public domain collections, guided by experts with bioinformatics experience and knowledge about virology.

The topics we will cover include:

- Finding genomic data from epidemic outbreaks and research projects (i.e. genomic sequences and NGS data)
- phylogenetics and Multiple Sequence Alignment (MSA)
- Downstream analysis of genomic data (differential mutations, data mining and association with phenotype)
- GWAS studies for viral and bacterial genomes
- Variation mapping on protein structures

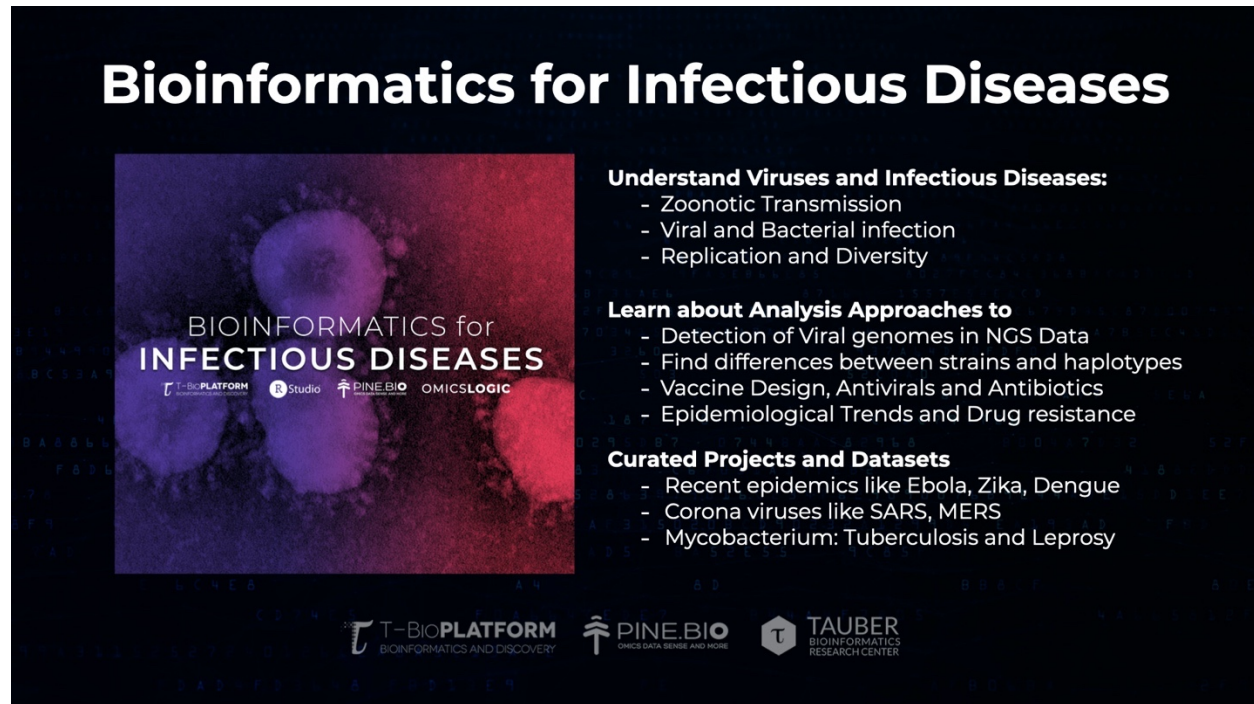
Bioinformatics for Infectious Diseases is an online training program designed for biologists, clinicians and students that are interested in virology and immunology and would like to learn about the use of bioinformatics and big data for infectious disease research, diagnostics as well as drug and vaccine development.

In this program, you will:

1. Understand Viral Diversity and its role in epidemic Infectious Diseases that keep re-emerging, including zoonotic transmission, transmission between humans and the process of viral and bacterial disease development
2. Learn about Bioinformatics Analysis Approaches to study viral genomes and understand relationships between viral strains and haplotypes, finding differences in sequence data and seeing the implications for drug and vaccine design
3. Gain hands-on experience by analyzing Curated Datasets from public domain collections, guided by experts with bioinformatics experience and knowledge about virology.

Program Topics:

- Finding genomic data from epidemic outbreaks and research projects (i.e. genomic sequences and NGS data)
- phylogenetics and Multiple Sequence Alignment (MSA)
- Downstream analysis of genomic data (differential mutations, data mining and association with phenotype)
- GWAS studies for viral and bacterial genomes
- Variation mapping on protein structures



Bioinformatics for Infectious Diseases

Understand Viruses and Infectious Diseases:

- Zoonotic Transmission
- Viral and Bacterial infection
- Replication and Diversity

Learn about Analysis Approaches to

- Detection of Viral genomes in NGS Data
- Find differences between strains and haplotypes
- Vaccine Design, Antivirals and Antibiotics
- Epidemiological Trends and Drug resistance

Curated Projects and Datasets

- Recent epidemics like Ebola, Zika, Dengue
- Corona viruses like SARS, MERS
- Mycobacterium: Tuberculosis and Leprosy

Logos: T-BioPLATFORM, RStudio, PINE.BIO, OMICSLOGIC, TAUBER BIOINFORMATICS RESEARCH CENTER

Bioinformatics for Infectious Diseases Program Syllabus:

Part1: INTRODUCTION:

Introduction to bioinformatics basics and data sources:

1. **Next-generation sequencing: viral genomes in host transcriptome**
 - Overview of NGS: reads, sequences, file formats
 - alignment, annotation and non-mapped reads
 - Alignment to databases of viral genomes

2. Multiple Sequence Alignment and Phylogeny

- Comparing sequences (Multiple Sequence Alignment)
- Finding a consensus sequence
- Identifying relationships between sequences (phylogeny, conservation)

3. Hands-on session, preparing and running your pipeline:

multiple sequence alignment of viral genomes and building a phylogenetic tree

- Finding full genome sequences and preparing FASTA files
- Selecting appropriate genomic sequences
- Preparing a full pipeline of MSA and Phylogeny

4. Q&A and DISCUSSION of pipeline results:

- Workflows: what to do if we have FASTA/FASTQ files?
- Which databases to use: Detection of viral genomes by mapping on databases
- Interpretation of Phylogenetic Analysis: Evolutionary relationships between genomes, evolutionary time

Part 2: Genomics and Virology:

How can bioinformatics be used to study pandemics?

5. From Infection to Pandemic: viral adaptation

- Hosts and origins
- Transmission
- Cell entry and tissue tropism

6. Symptom severity: Disease progression and outcomes

- Viral proteins
- Replication
- Immune evasion

7. Hands-on project discussion

EXAMPLE: the origin of human infection with MERS, SARS, and SARS-2 pandemics

8. Rate of Mutation - mutation variant types

- Point mutations, substitutions, insertions
- Mutation types (synonymous/nonsynonymous; sense/missense)
- Mutation rate and fitness (frequency, entropy, conservation)

9. Mutation Annotation & Significance for analysis

- Codon/amino-acid and chemical properties
- Location on genome and protein function

10. Host-pathogen interaction

- Protein-protein interaction and host response
- Immune response (adaptive, innate)
- Vaccine design: factors for consideration

CONCLUSION and OUTCOMES:

In this program, we will learn about important principles of bioinformatics in application to virology, including:

Use of bioinformatics in virology

- Methods of analysis
- Databases and references
- Raw data types and repositories

Important factors for vaccine design

- Exposed parts of viral proteins
- Vaccine types: protein-based, virus-based
- Novel approaches to rapid-response vaccines: interference

Important factors for antivirals

- Prevention of cell entry
- Inhibition of replication
- Toxicity, specificity
- Solubility, permeability

Case studies we will utilize in this program:

- Coronaviruses and the recent COVID-19 epidemic
- EBOLA outbreaks over the last decade: emerging diseases
- Flu and other respiratory disorders: challenges with vaccines and antivirals
- Tuberculosis: bacterial chronic diseases and antibiotic resistance

Bioinformatics methods:

Alignment, annotation, characterization of mutations, GWAS, phylogenetic analysis, RNA-seq