Theiagen Consulting LLC
joel.sevinsky@theiagen.com | kevin.libuit@theiagen.com

SARS-CoV-2 (SC2) Training: Bioinformatics Solutions for Genomic Analysis of SC2 Samples

Session 1 - Terra Onboarding and Configuration

- Introduction to the Terra platform
- Onboarding users to the CDC-Thailand workspace
- Data upload & launching of the Titan_Augur_Prep workflow for genomic epidemiology

Session 2 - Genomic Epidemiology of SC2 Genomes

- Introduction to the bioinformatics methods employed by the Titan workflows for SC2 genomic epidemiology: Titan_Augur_Prep & Titan_Augur_Run
- Submitting a Titan_Augur_Run job for genomic epidemiology analysis of uploaded SC2 assemblies

Session 3 - Assessing SC2 Genomic Epidemiology Outputs

- Phylogenetic tree visualization with the NextStrain web application
- Adding additional layers of metadata to NextStrain visualizations
- Gaining a global perspective with the UCSC UShER algorithm

APHL

Theia gen Genomics
joel.sevinsky@theiagen.com | kevin.libuit@theiagen.com

SARS-CoV-2 (SC2) Pre-Training Materials: Bioinformatics Solutions for Genomic Analysis of SC2 Samples

Virtual Training Workshop Hosted by the Association of Public Health Laboratories (APHL) funded by the Centers for Disease Control and Prevention.

Pre-Assessment Survey

Please complete the pre-assessment survey prior to the first session:
https://www.surveymonkey.com/TBD

Overview

This bioinformatics training workshop will focus on workflows and resources made available on the Terra platform. As such, trainees are encouraged to review the following resources to ensure that they have access to the appropriate compute resources to support bioinformatics analysis using the Terra platform as well as a cursory understanding of Terra and how it applies to bioinformatics solutions for SARS-CoV-2 genomic analysis.

Compute Resource Requirements

Compute Resource	Training Requirements
Internet Browser	Google Chrome preferred, Firefox and Edge acceptable
Minimum Internet Speeds*	10Mbps download; 5Mbps upload
Software Requirements	Microsoft Excel
Terra Access via Google Authentication	Gmail Account Registered to Terra.bio

*Internet speeds can be gauged using the online speedtest resource hosted Ookla available at https://www.speedtest.net/

Introductory Terra Training Videos

Theiagen Genomics has made several introductory Terra training videos available on the Theiagen Genomics YouTube channel. While most of this material will be covered in greater detail during the live training event, trainees are encouraged to watch, at minimum, the first video in this series, Genomic Analysis of SC2, a 10-minute video aimed to provide a cursory understanding of Terra and how it applies to bioinformatics solutions for SARS-CoV-2 genomic analysis.

Step-By-Step Protocols

The Centers for Disease Control and Prevention SARS-CoV-2 Technical Onboarding and Assistance for States Team (CDC TOAST) has released step-by-step protocols for each Terra workflow covered in this training event on the Protocols.io platform. Please note that, while completely free, registration to protocols.io is required to access these resources directly. If registration is not possible, please reach out to the CDC TOAST team directly for assistance: TOAST@CDC.gov

APHL

Technical Details

Additional information, including the technical details and modifiable parameters of each Terra workflow discussed in this training event, is made available through the Public Health Viral Genomics ReadTheDocs page: https://public-health-viral-genomics-theiagen.readthedocs.io/en/latest/

Training Data

A large component of this training workshop will be dedicated to APHL hands-on exercises that will walk trainees through the process of analyzing SARS-CoV-2 genomes on the Terra platform. The data for these exercises have been made available through a publicly accessible GCP Storage Bucket: APHL_Terra_Training_2021.zip--a zipped (compressed) file containing all of the necessary training materials will begin downloading once this link is clicked.

In order to participate in the hands-on bioinformatics exercises, training data must be downloaded, unzipped, and accessible to the trainees before June 8th, 2021. If you experience any issue downloading, decompressing, or accessing these data, please contact us at Support@Theiagen.com. Question or comments about the training, please email Noah Hull, Laboratory Technical Manager, Global Health, APHL (noah.hull@aphl.org).

Theia gen Genomics
joel.sevinsky@theiagen.com | kevin.libuit@theiagen.com

SARS-CoV-2 (SC2) Training:
Bioinformatics Solutions for Genomic Epidemiology of SC2 Samples

Remote Training Workshop Hosted by the Association of Public Health Laboratories (APHL)

Agenda Week One Date: TBD

Live Lecture 45m	**Introduction to the Terra Platform** • *Welcome: Dr. Beth Skaggs (US CDC in Thailand)* • *Meet the event instructor, APHL hosts, and key laboratory personnel* • *Course outline and objectives* • *Introduction to the Terra platform*
Break 5m	**Break**
Tutorial Walkthrough 30m	**Data Upload & Titan_Augur_Prep Job Submission** • *Upload of SC2 assembly files to the Terra platform* • *Create a Terra data table with necessary Augur metadata* • *Submit a Titan_Augur_Prep job to analyze the uploaded data*
Session Wrap Up 10m	**Review of Session Material** • *Summarize lecture content* • *Recap steps performed in the tutorial walkthrough* • *Highlight additional datasets available for continued practice*
Week One Office Hours 60m	**Review of the Terra Platform and the Protocols for Data Upload and Titan_Augur_Prep Job Submission** • *Week 1 material recap, discussion and exercise to reinforce material* • *Significant time will be left open to address trainee questions and expand on specific topics*

Agenda Week Two Date: TBD

Live Lecture 45m	**Genomic Epidemiology of SARS-CoV-2 Genomes** • *Introduction to the bioinformatics methods employed by the Titan workflows for SC2 genomic epidemiology; Titan_Augur_Run & Titan_Augur_Prep* • *Highlight most relevant outputs generated*
Break 5m	**Break**
Tutorial Walkthrough 30m	**Titan_Augur_Run Job Submission** • *Submit job for genomic epidemiology analysis of uploaded SC2 assemblies*

APHL

Theia gen Genomics
joel.sevinsky@theiagen.com | kevin.libuit@theiagen.com

Session Wrap Up 10m	**Review of Session Material** • *Summarize lecture content* • *Recap steps performed in the tutorial walkthrough* • *Highlight additional datasets available for continued practice*
Week Two Office Hours 60m	**Review of Genomic Characterization of SARS-CoV-2 Genomes, the Titan Workflows for Genomic Epidemiology and the Outputs they Generate** • *Week 2 material recap, discussion and exercise to reinforce material* • *Significant time will be left open to address trainee questions and expand on specific topics*

Agenda Week Three Date: TBD

Tutorial Walkthrough 35m	**Assessing the Outputs Generated from the Titan_Augur_Run workflow** • *Phylogenetic tree visualization with the NextStrain web application* • *Adding additional layers of metadata to NextStrain visualizations* • *Gaining a global perspective with the UCSC UShER algorithm*
Break 5m	**Break**
Post-Training Assessment 25m	**Post-Training Assessment** • *Trainees will complete the post-training surveys to gauge the effectiveness of materials presented*
Session Wrap Up 35m	**Review of Course Materials** • *Summary of the Terra platform and its application in public health bioinformatics* • *Recap of the Titan workflows for genomic epidemiology* • *Review of the protocols to run the tools covered in this course* • *Highlight additional datasets available for continued practice*
Week Three Office Hours 60m	**Review of Internationally-Accessible Databases, Required Metadata for Submission, & the Mercury Workflows** • *Week 3 material recap, discussion and exercise to reinforce material* • *Significant time will be left open to address trainee questions and expand on specific topics*

APHL