



Terra, Zoom, & Public Notion Pages:

Virtual Resources for Global Public Health
Bioinformatics Workforce Development

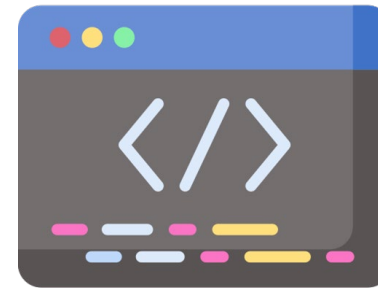
Kevin G. Libuit, MS, Theiagen Genomics

APHL IDLC – March 13, 2023

Innovations in Bioinformatics

Maturation of Software Development Practices

- Significant progress towards standardized and scalable solutions for performing pathogen genomics in public health
 - StaPH-B Docker Builds
 - Use of workflow management systems
 - Nextflow, WDL, Snakemake
 - Growing adoption of cloud compute platforms
 - GCP, Azure, AWS



**Interoperable and
Reproducible Pipelines**



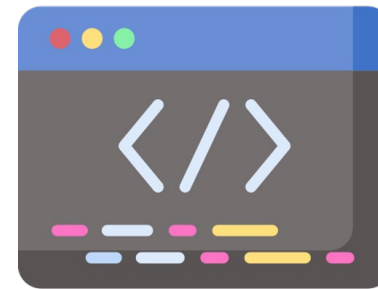
**Scalable
Compute Resources**

Innovations in Bioinformatics

The Challenge is Accessibility

- Use typically requires a **highly-technical practitioner** and sophisticated compute infrastructure
 - Proficiency at command line interface (CLI)
 - System administrators for complex cloud back-ends
- Trainings to develop these advanced practitioners are necessary, but difficult to deploy rapidly and at scale

Urgent need to **bridge the gap** between these technical solutions and the public health workforce



**Interoperable and
Reproducible Pipelines**



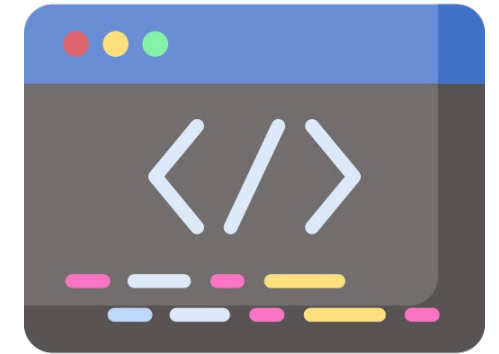
**Scalable
Compute Resources**

The Terra Platform

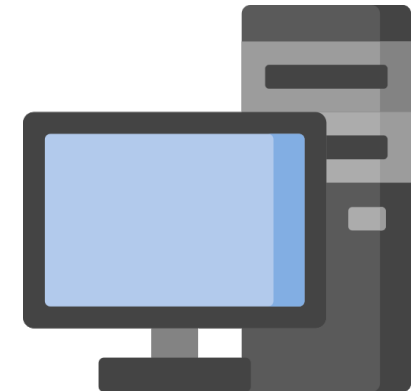
Terra is a **web application** that connects users to bioinformatics workflows & dynamic cloud computing resources through a **graphical user interface** (GUI)



Provides **immediate access** to bioinformatics solutions for PHLs



Interoperable and Reproducible Pipelines



Scalable Compute Resources

Terra: Data-Centric User Interface

The screenshot displays the Terra Data-Centric User Interface. The top navigation bar is green and contains the Terra logo, the word "BETA", and the text "WORKSPACES". Below this, a breadcrumb trail shows "Workspaces > theiagen-demos/MGen-Theiagen-2023 > Data". The main navigation bar is white and includes tabs for "DASHBOARD", "DATA" (which is highlighted), "ANALYSES", "WORKFLOWS", and "JOB HISTORY".

Below the navigation bar, there is a section for "DATA". On the left, there is a sidebar with a search bar labeled "Search all tables" and a magnifying glass icon. Below the search bar, there are two sections: "TABLES" and "REFERENCE DATA". The "TABLES" section shows a list of tables, including "case_study_sample (30)" and "case_study_sample... (3)". The "REFERENCE DATA" section shows a message: "No references have been added. Add reference data".

The main content area displays a table with the following columns: "case_study_sampl...", "gambit_predicted_ta...", "assembly_length", "n50_value", and "seqsero2_predicted_ser...". The table contains 10 rows of data, all of which are "Salmonella enterica" samples. The first row is highlighted. Above the table, there are buttons for "EDIT", "OPEN WITH...", "EXPORT", and "SETTINGS". To the right of these buttons, it says "0 rows selected". To the far right, there is a link for "ADVANCED SEARCH".

case_study_sampl...	gambit_predicted_ta...	assembly_length	n50_value	seqsero2_predicted_ser...
SRR500494	Salmonella enterica	4667688	77246	Bareilly
SRR500493	Salmonella enterica	4728993	233810	Bareilly
SRR498444	Salmonella enterica	4745922	59044	Bareilly
SRR498442	Salmonella enterica	4835798	58727	Bareilly
SRR498436	Salmonella enterica	4749845	98190	Bareilly
SRR498434	Salmonella enterica	4755902	144214	Bareilly
SRR498433	Salmonella enterica	4759145	148538	Bareilly
SRR498431	Salmonella enterica			

User interface that is **accessible** to the public health workforce

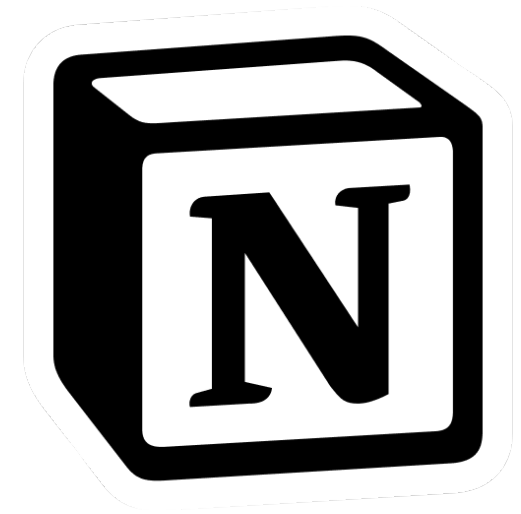
Model for Rapid Bioinformatics Workforce Development



**GUI Bioinformatics
Web Applications**



**Virtual Lectures &
Office Hours**



**Open Access to
Training Materials**

Zoom: Virtual Lectures and Office Hours

Virtual Meeting Platform for PHLs

- **Global familiarity** to the Zoom platform and interface
- Allowed trainings in **multiple languages simultaneously**
- **Recorded** sessions



Formal Lectures



Weekly Office Hours

Zoom: Virtual Lectures and Office Hours



Engagement with **virtual cohort** through active polling, open discussions, and post-lecture quizzes

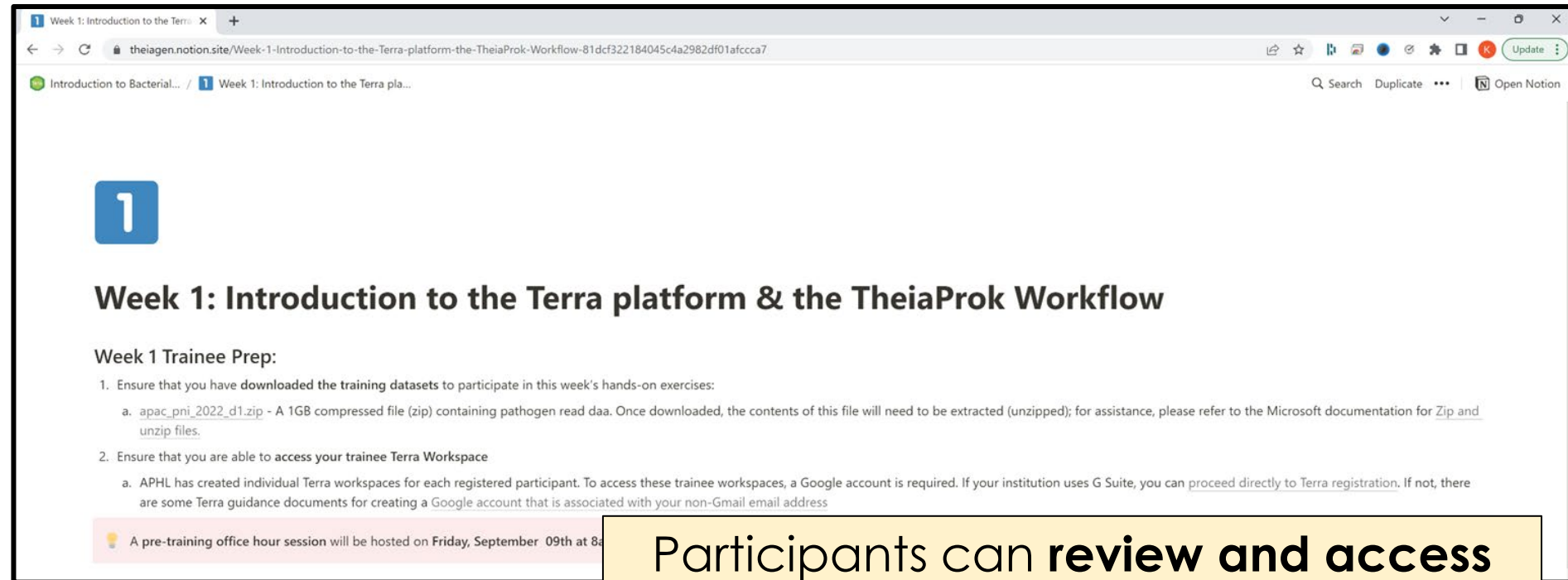
Introduce **hands-on exercises** as **early** in the trainings as possible



Notion: Open Access to Training Materials

Notion is a **note-taking** application that facilitates rapid creation of **publicly-accessible** webpages

- Solution for **easily distributing** of training materials



Participants can **review and access** training material **at their convenience**

Model for Rapid Bioinformatics Workforce Development

Training Public Health Scientists on Critical Technologies

- **Develop trainings** for use of the Terra platform to perform pathogen genomics
- **Deliver of materials** through Zoom to allow training of scientists in the global public health community
- **Enable open and asynchronous access** to training materials through public Notion pages

We have applied this **framework** to train an estimated **800 participants** representing over **40 countries** worldwide



Workforce Development Must Include Solutions for Providing Ongoing Support

Technical Support for Global Training Cohort



**Time Zones and
Availability**

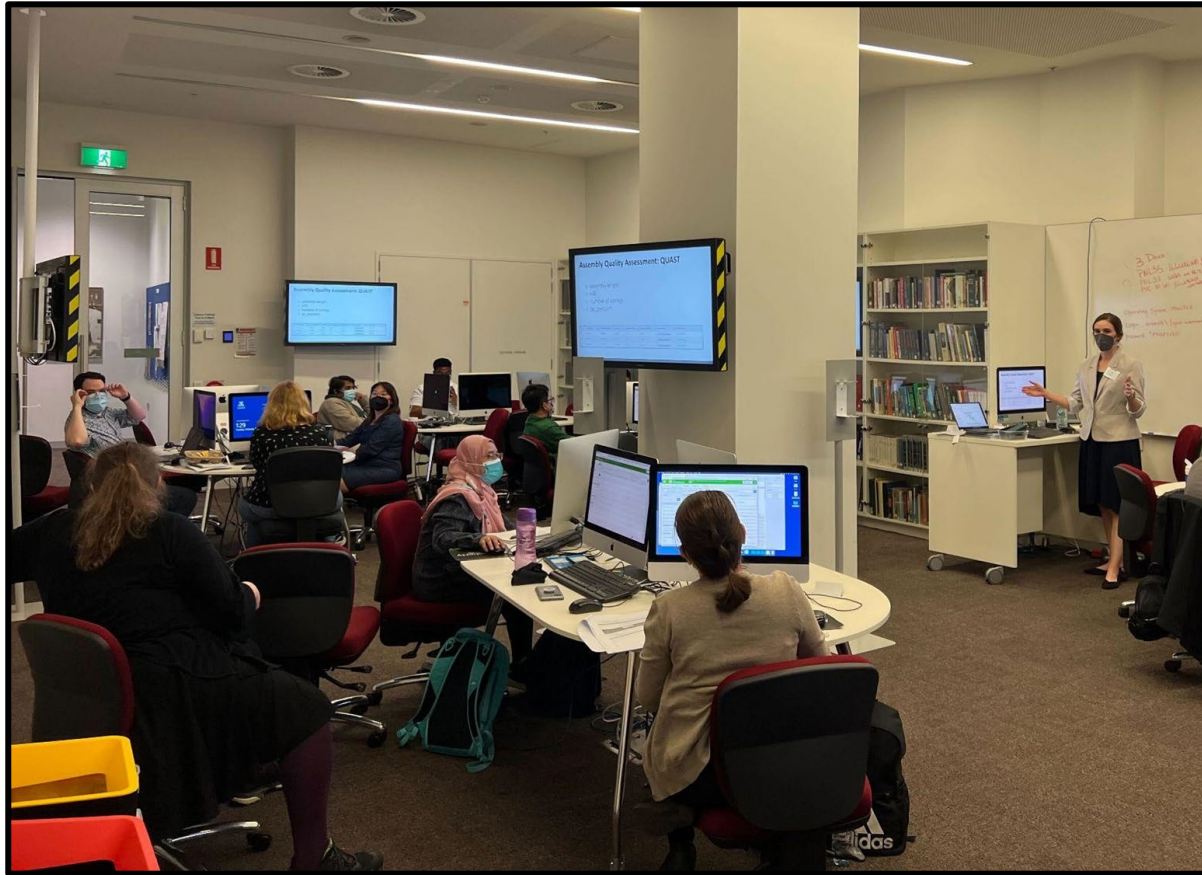


**Forums of
Communication**



**Follow-Up
Site Visit**

Follow-Up Site Visits



Opportunity to **reinforce** materials that were delivered virtually

Follow-Up Site Visits



Site visits also help to strengthen trainer-trainee **relationships**

Acknowledgements



**Global Community of Public Health and
Pathogen Genomics Scientists**

