| | Plotting SARS-CoV-2 Metagenomic Sample Data in Terra using Theiagen's Freyja Plot Workflow | |
|---|---|---|
| | Document TG-FREY-02, Version 3 | |
| | Date: | Workflow Versions: |
| | 4/16/2024 | PHB v2 |

## 1. PURPOSE/SCOPE

To standardize the process of analyzing SARS-COV-2 (SC2) metagenomic data using Theiagen's Freyja Plot workflow in Terra to create a plot visualization of aggregated Freyja data. Acceptable data types include Freyja demixed files generated from the Freyja FASTQ workflow.

## 2. REQUIRED RESOURCES

- Computer
- Internet connection: at least 10 and 5Mbps for download and upload speeds, respectively
- Internet browser
  - Google Chrome, Firefox, or Edge
- Google account
- Terra account, linked to Google account
- Freyja demixed files uploaded to Terra workspace
- Theiagen's Freyja_Plot_PHB workflow in Terra

### REQUIRED WORKFLOW INPUTS FILES

- Freyja demixed files (output from Freyja_FASTQ
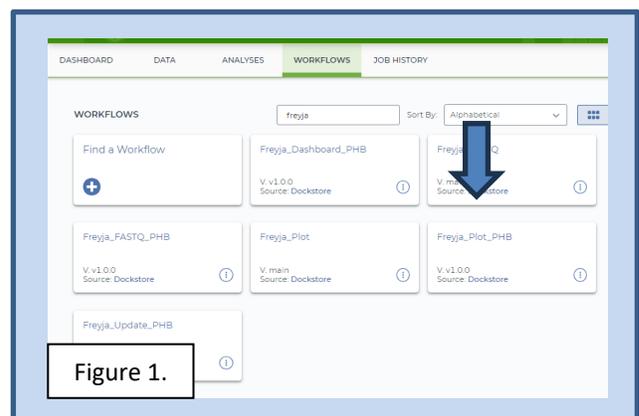- Plot name
- Terra metadata (tsv) file

## 3. RELATED DOCUMENTS

| Document Number | Document Name |
|---|---|
| TG-TER-03 | Uploading Local or SRA NGS Data & Creating a Results Metadata Table in Terra |
| TG-FREY-01 | Running SARS-CoV-2 Metagenomic Samples in Terra using Theiagen's Freyja FASTQ Workflow |

## 4. PROCEDURE

### 4.1 RUNNING THE FREYJA PLOT WORKFLOW

1. Open Terra and navigate to the *workflows* tab within the workspace containing metagenomic SC2 data
2. Select the *Freyja_Plot_PHB* (Fig 1)
3. Choose the latest version of *version 2* in the version dropdown field, or the version used during internal assay validation (Fig 2, a)
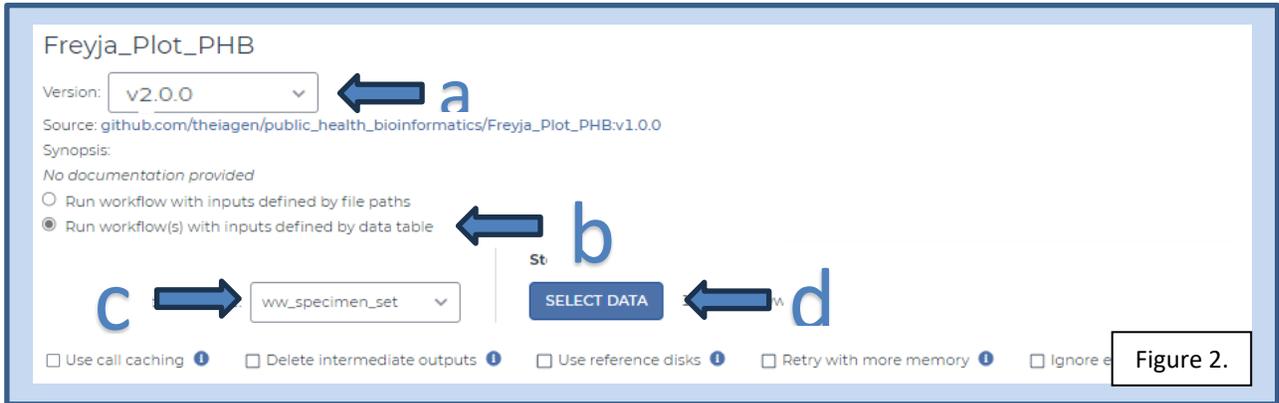4. *Uncheck use call caching* (Fig 2)



Figure 1.

Figure 2.

5. Select the second bullet to *run workflow(s) with inputs defined by data table* (Fig 2, b)

6. Select the relevant SET data table name under the select *root entity type* dropdown (Fig 2, c)

7. Click *select data* (Fig 2, d) and in the pop-up window select the first bullet to *create a new ww_specimen_set from selected ww_specimens* (Fig 3, highlighted text)

    a. *Select the checkbox* for each sample to be included in the analysis (Fig 3, a)

    b. *Optional: name the output file to differentiate it from other files, e.g. FreyjaPlot_LineageTime_YYYYMMDDn (Fig 3, b)*
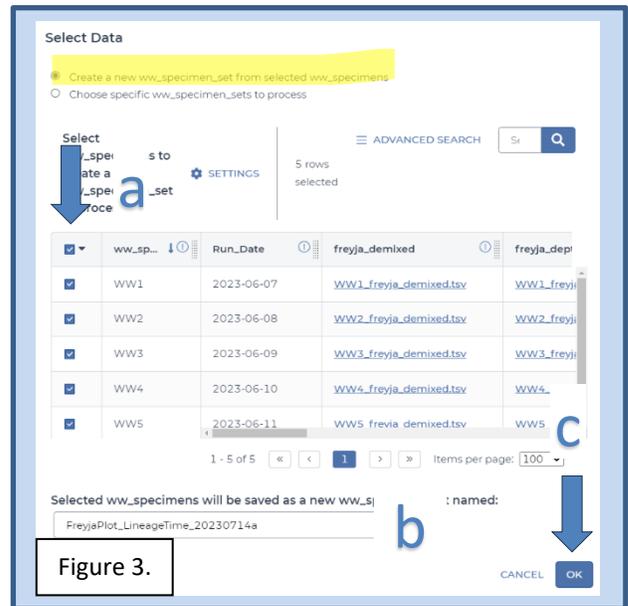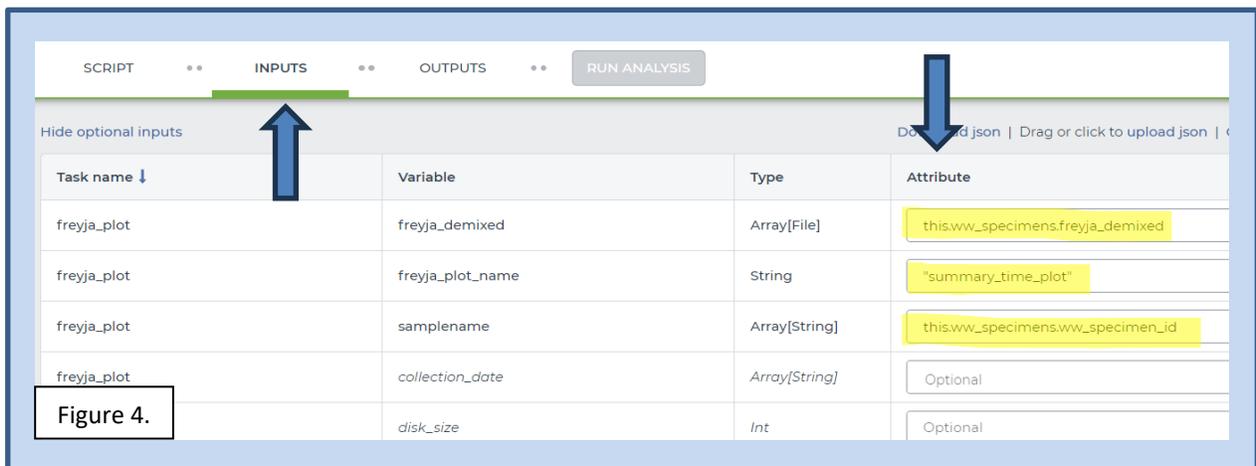


Figure 3.



Figure 4.

   c. Click ok

8. In the *inputs* tab, manually set the first three attributes in the table to the following:
   a. *this.ww_specimens.freyja_demixed*
      i. Notice there is an "s" added to the end of the unique data table name
   b. *"summary_time_plot"*
   c. *this.ww_specimens.ww_specimen_id*

9. *Optional: manually set the following parameters to visualize lineages and time variables (Fig 5):*
   a. *To add lineages to the plot, set the plot_lineages variable to true*
   b. *To plot over time:*
      i. *Set the collection_date variable to this.ww_specimens.Run_Date where:*
         1. *ww_specimen is the unique data table name (add an s to the table name)*
         2. *"Run_Date" is the name of the metadata column containing collection dates*
      ii. *Set the plot_time variable to true*
      iii. *Set the plot_time_interval variable to "D" for days or "MS" for months*

| | | | |
|---|---|---|---|
| freyja_plot | collection_date | Array[String] | this.ww_specimens.Run_Date |
| freyja_plot_task | disk_size | Int | Optional |
| freyja_plot_task | docker | String | Optional |
| freyja_plot_task | mincov | Int | Optional |
| freyja_plot_task | plot_day_window | Int | Optional |
| freyja_plot_task | plot_lineages | Boolean | true |
| freyja_plot_task | plot_time | Boolean | true |
| freyja_plot_task | plot_time_interval | String | "D" |

Figure 5.

10. Specify outputs by clicking on the outputs tab and use defaults (Fig 6)
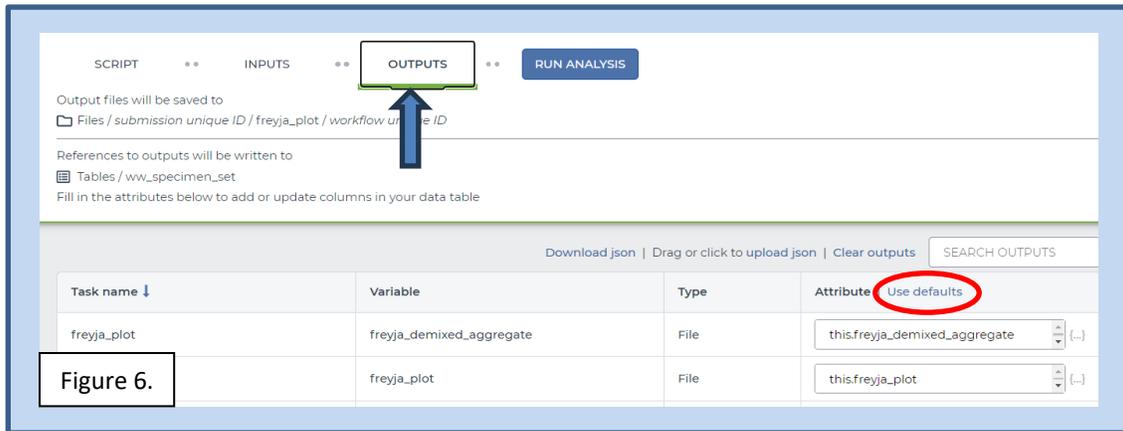11. Click save, run analysis, add comments if desired, and click launch



Figure 6.

## 4.2 VISUALIZING FREYJA DATA PLOTS

1. Navigate to the data tab of the workspace containing metagenomic data and open the pertinent SET data table
2. Open the plot file located in the freyja_plot column for the relevant data set (Fig 7)
   a. If the output file was named as described in section 4.1, step 6b, the data set name in the ww_specimen_set_id column can help differentiate analysis files from one another (Fig 7)
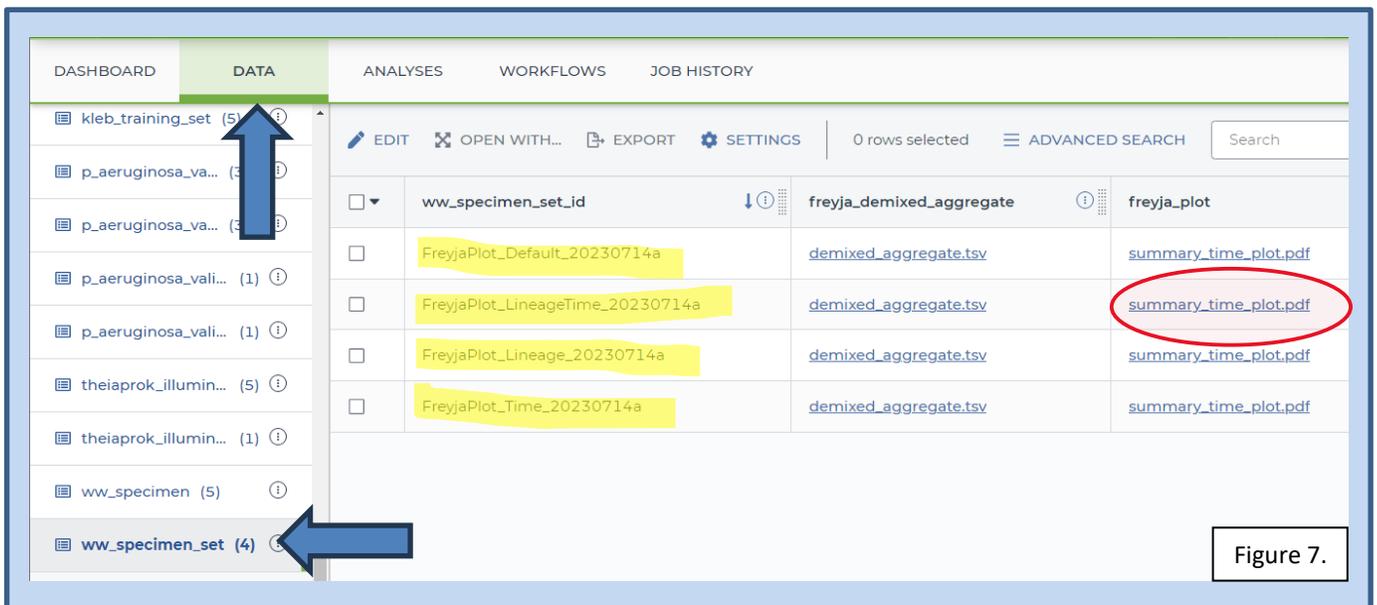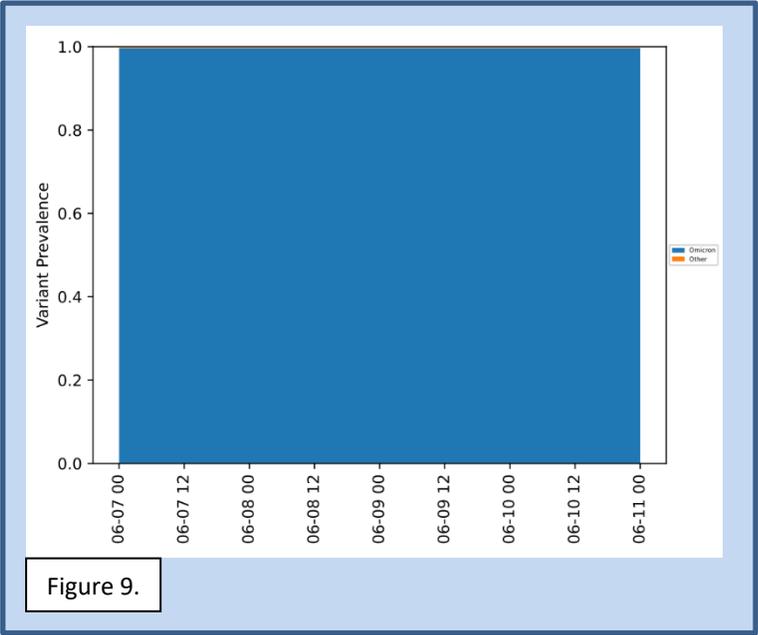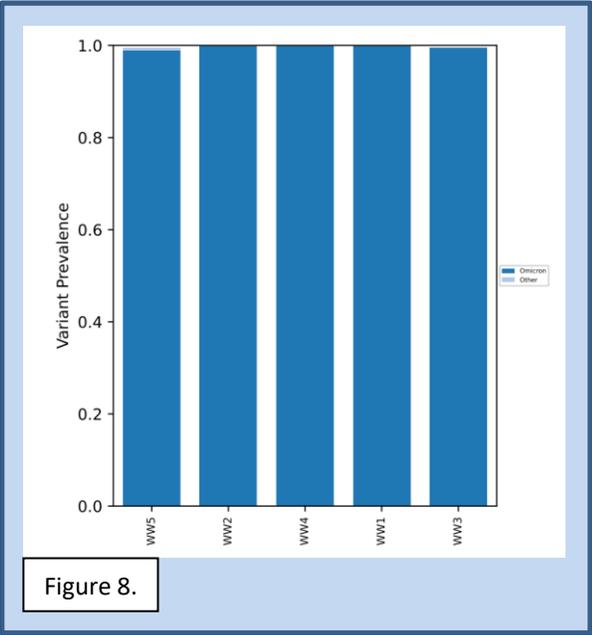


Figure 7.

3. The default plot using summarized outputs (without lineage or time variables) depicts the variant prevalence within each sample (Fig 8)
   a. *This may be useful when assessing different collection sites collected at the same time*



Figure 8.



Figure 9.

4. Plots containing the time variable will display the collection date along the x-axis and the corresponding variant proportions in the y-axis (Fig 9)
   a. *This view can be helpful when analyzing the same site over time*
5. Plots including both time and lineage variables illustrate the collection date on the x-axis and lineage abundance in the y-axis
   a. *Notice by comparing Fig 9 and 10 the drastic difference in resolution of data*
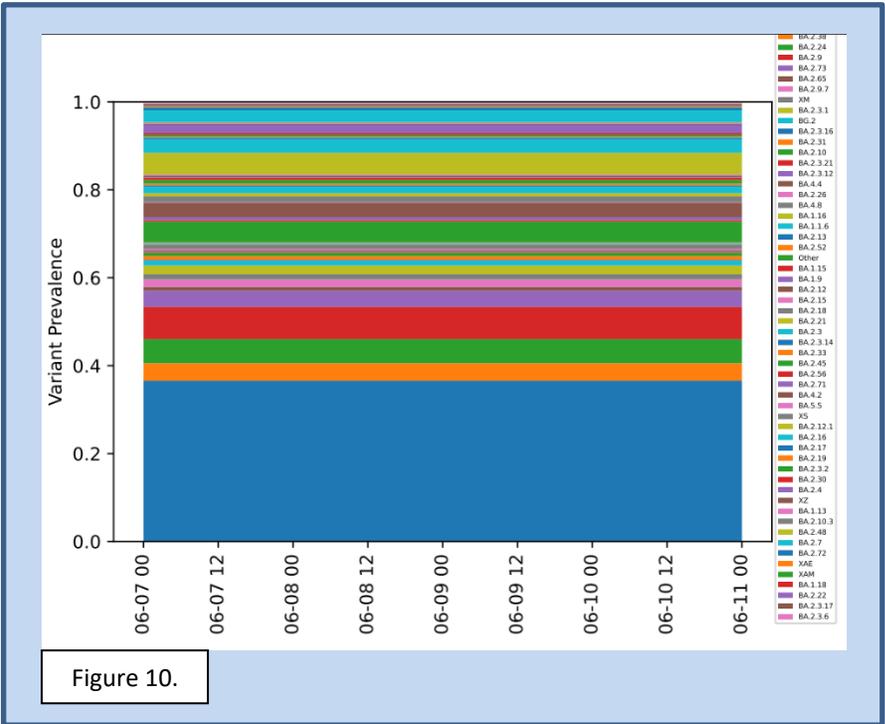


Figure 10.

## 5. QUALITY RECORDS

- freyja_demixed files
- Plot graphs

## 6. TROUBLESHOOTING

- Consult with internal staff familiar with this procedure or contact support@theiagen.com for troubleshooting inquiries
- For document edit requests, contact support@theiagen.com

## 7. LIMITATIONS

1. When creating visualizations from aggregated sample data over time, ensure all samples have been run with Freyja FASTQ using the same curated_lineages and usher_barcodes files
2. Freyja can only be used to analyze SC2 data from Illumina PE sequencing files

## 3. REFERENCES

None

## 4. REVISION HISTORY

| Revision | Version | Release Date |
|---|---|---|
| Document creation | 1 | 8/2023 |
| Corrected 4.1.7.c input attribute value for samplename | 2 | 9/2023 |
| Updating for PHB v2.0.0 release (no major updates); added appendix | 3 | 4/2024 |
| | | |
| | | |

| | Plotting SARS-CoV-2 Metagenomic Sample Data in Terra using Theiagen's Freyja Plot Workflow | |
|---|---|---|
| | Document TG-FREY-02, Version 3 | |
| | Date: | Workflow Versions: |
| | 4/16/2024 | PHB v2 |

## 5. APPENDICES

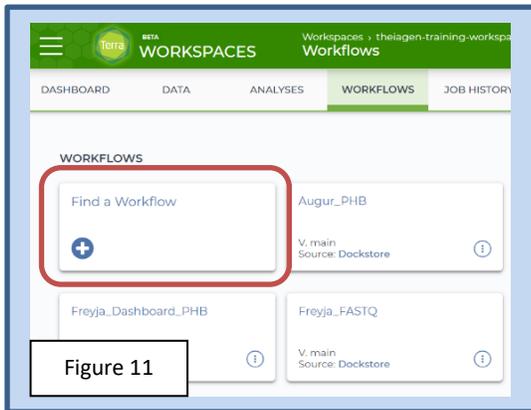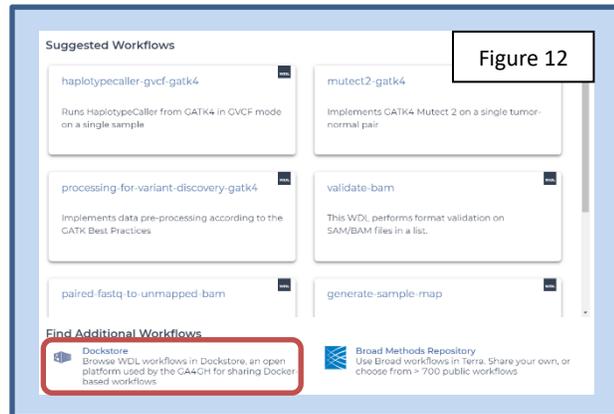### 10.1 IMPORTING THE FREYJA_PLOT_PHB WORKFLOW FROM DOCKSTORE



Figure 11



Figure 12

1. In the *Terra workspace* of interest, open the *workflows* tab and click *find a workflow* (Fig 11)
2. In the pop-up window, click *dockstore* (Fig 12)
3. To find the Theiagen Terra_2_GISAID workflow, type *Freyja_Plot_PHB* in the search bar (Fig 13)
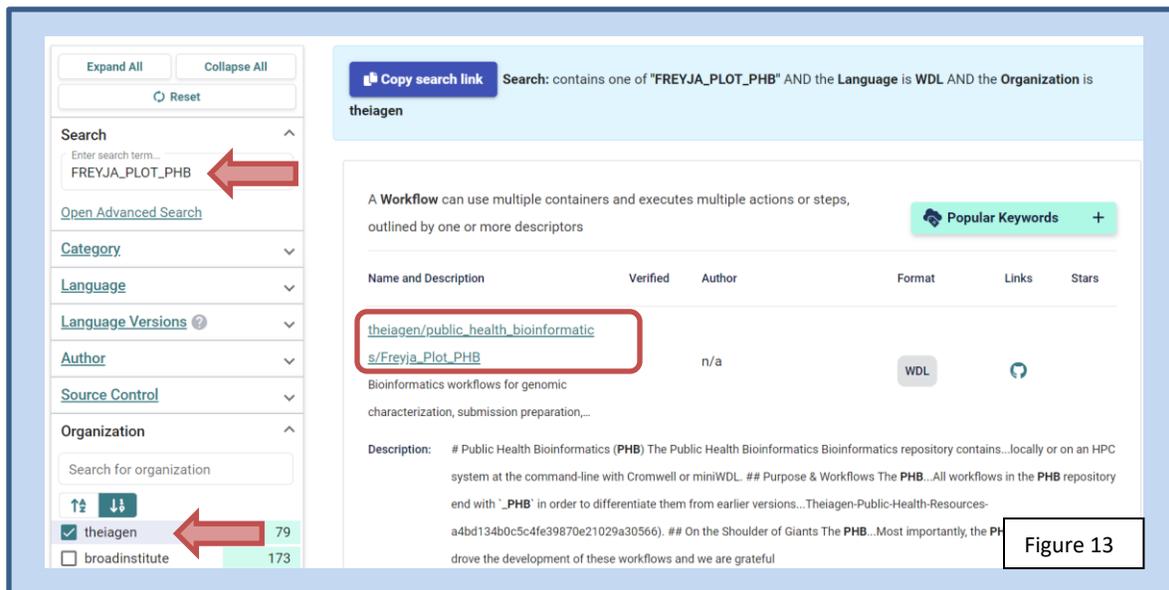4. In the left hand sidebar, scroll down to Organization and select "*theiagen*" (Fig 13)



Figure 13

5. Find the workflow by looking at the file path suffix; click the name to *open the workflow* (Fig 13)
6. Click *Terra* to launch the workflow in Terra (Fig 14)
7. Choose the *destination workspace* in the dropdown and click *import* or create a new workspace (Fig 15)

Figure 14



Figure 15