	<b>Running Influenza A, H5Nx Metagenomic Samples in Terra using Theiagen's Freyja FASTQ Workflow</b>	
	Document TG-FREY-H5NX, Version 1	
	Date:	Workflow Versions:
	07/08/2025	PHB v3

## 1. PURPOSE/SCOPE

To standardize the process of running Influenza A, H5Nx (H5Nx) metagenomic samples using Theiagen's Freyja FASTQ workflow in Terra to perform lineage deconvolution, abundance determination, and identify coverage metrics. This SOP is specific to Illumina paired end (PE) raw read files. While this SOP can be used for any subtype of H5 influenza, please note that this SOP should NOT be used to run Influenza A, H3N2 samples.

## 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
- Metagenomic Illumina PE raw read files uploaded to Terra workspace
- Theiagen's Freyja\_FASTQ\_PHB Workflow in Terra
  - See the link "Running Freyja on other pathogens" for more details and resources.
  - See Appendix 10.1

### REQUIRED WORKFLOW INPUTS FILES

- Raw Illumina PE read files
- [Primer bed file]
- Reference genome
- [barcodes metadata file]\*


## 3. RELATED DOCUMENTS

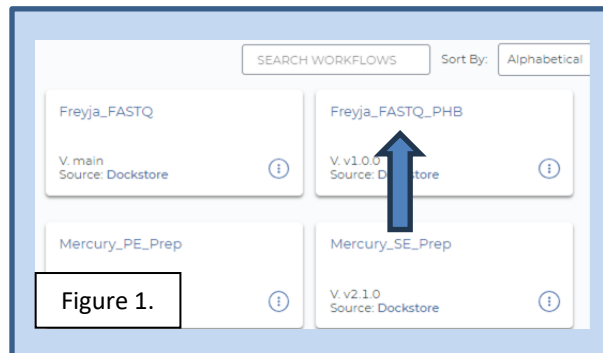
Document Number	Document Name
TG-TER-03	Uploading Local or SRA NGS Data & Creating a Results Metadata Table in Terra

## 4. PROCEDURE

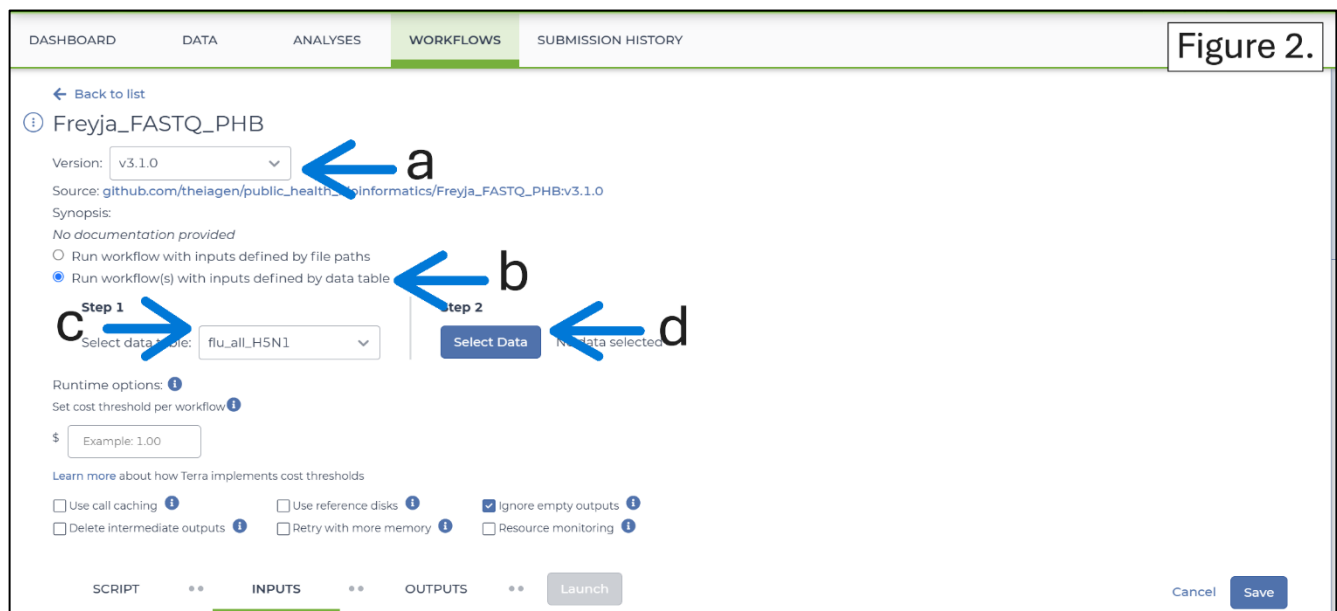
### 4.1 RUNNING THE FREYJA FASTQ WORKFLOW


1. Open Terra and navigate to the **workflows** tab of the workspace containing wastewater data
2. Select the **Freyja\_FASTQ\_PHB** workflow (Fig 1)

	Running Influenza A, H5Nx Metagenomic Samples in Terra using Theiagen's Freyja FASTQ Workflow	
	Document TG-FREY-H5NX, Version 1	
	Date:	Workflow Versions:
	07/08/2025	PHB v3



3. Uncheck **use call caching** (Fig 2)
4. Choose the latest version of **version 3** in the version dropdown field, or the workflow version that was used during internal assay validation (Fig 2, a)
5. Select the second bullet to **run workflow(s) with inputs defined by data table** (Fig 2, b)
6. Select the relevant 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 checkbox** for each sample to be included in the analysis (Fig 3)



	<b>Running Influenza A, H5Nx Metagenomic Samples in Terra using Theiagen's Freyja FASTQ Workflow</b>	
	Document TG-FREY-H5NX, Version 1	
	Date:	Workflow Versions:
	07/08/2025	PHB v3

**Step 1**

Select data table: flu\_all\_H5N1

Runtime options: ⓘ

Set cost threshold per workflow ⓘ

\$ Example: 1.00

Learn more about how Terra implements cost thresholds

☐ Use call caching ⓘ
 ☐ Use reference disks ⓘ
 ☒ Ignore empty outputs ⓘ
 ☐ Delete intermediate outputs ⓘ
 ☐ Retry with more memory ⓘ
 ☐ Resource monitoring ⓘ

SCRIPT   **INPUTS**   OUTPUTS   **Launch**   Cancel   **Save**

**Step 2**

**Select Data**   No data selected

**Figure 4.**

Hide optional inputs   Download json | Drag or click to upload json | Clear inputs   SEARCH INPUTS

Task name ↓	Variable	Type	Input value
freyja_fastq	read1	File	<span>this.read1</span> ⓘ
freyja_fastq	reference_genome	File	<span>workspace.h5n1_reference_genome</span> ⓘ
freyja_fastq	samplename	String	<span>this.flu_all_H5N1_id</span> ⓘ

- a Click the checkbox dropdown and select “all” to select all samples in the data table; if the checkbox at the top is checked, only the first 100 samples in the data table will be selected
- b Additionally, a subset of samples may be chosen using the search bar to filter before selecting the checkbox at the top to only select samples matching the search criteria (Fig 3, highlight)
- c Scroll to the bottom and click **ok**

8. Click on the inputs tab to specify settings (Fig 4)

- a Manually set the first three attributes to the following, respectively
  - i. Reference genome can be found here: <https://github.com/andersen-lab/Freyja-barcodes/tree/main/H5Nx/latest> as “reference.fasta”. This file will need to be downloaded to your computer and uploaded to the workspace data in the Data tab of your Terra workspace (see [appendix 10.2](#) for adding workspace elements and files to Terra).


ii. Unique Terra data table name: this.sample\_id

iii. Raw read1 file: this.read1

1. read2 is further down the page of inputs. If you are supplying Illumina PE reads, you will need to supply this.read2 to the read2 input.

b Scroll down the inputs to *update\_db*. You will either have to provide input to *update\_db* and *freyja\_pathogen* OR provide an input for *freyja\_barcodes* (Fig 5.1 and Fig 5.2).

- i. **Option 1:** set update\_db to **true** and set freyja\_pathogen to “H5NX” (Fig 5.1)
- ii. **Option 2:** Specify the freyja\_barcodes file used to assign H5Nx lineages (Fig 5.2)

	Running Influenza A, H5Nx Metagenomic Samples in Terra	
	using Theiagen's Freyja FASTQ Workflow	
	Document TG-FREY-H5NX, Version 1	
	Date: 07/08/2025	Workflow Versions: PHB v3


1. The H5Nx barcodes file in use by the Andersen lab can be found here as the *barcodes.csv* file: <https://github.com/andersen-lab/Freyja-O/tree/main/H5Nx/latest>. This may be saved in the workflow from prior analysis or copied from the respective Terra workspace files (see [appendix 10.2](#) for details on copying workspace filepaths). It is not necessary to provide a *lineage\_metadata* file to run Freyja\_FASTQ for H5Nx.
- iii. **(Optional):** Provide `primer_bed` file for amplicon sequencing samples (Fig 5.3). Primer bed file: `workspace.[FILENAME]`
  1. For appropriate H5Nx primer sets, ensure primer bed file (.bed file containing the primers used during sequencing) is uploaded to the workspace; it will then be available in the dropdown as `workspace.[FILENAME]`. *If amplicon sequencing was not done, there is no primer bed file; this field can be left blank. Freyja can be run without a primer bed file even for amplicon sequencing, as this is an optional field, but this is not recommended because primers will not be trimmed.*
    - a. See [appendix 10.2](#) for adding workspace elements and files to Terra

freyja	number_bootstraps	Int	Optional	{...}
freyja	update_db	Boolean	true	{...}
freyja_fastq	depth_cutoff	Int	Optional	{...}
freyja_fastq	freyja_barcodes	File	Optional	{...}
freyja_fastq	freyja_lineage_metadata	File	Optional	{...}
freyja_fastq	freyja_pathogen	String	"H5NX"	{...}
freyja_fastq	kraken2_target_organism	String	Optional	

Figure 5.1

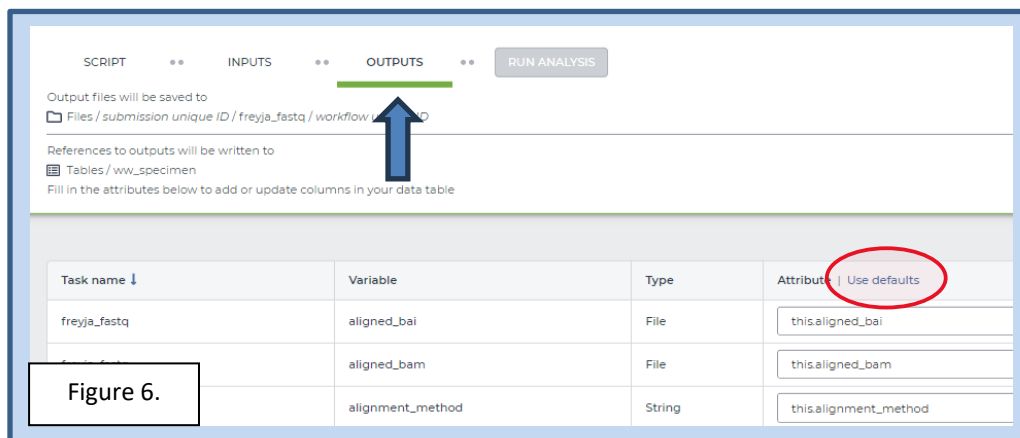
freyja	number_bootstraps	Int	Optional	{...}
freyja	update_db	Boolean	Optional	{...}
freyja_fastq	depth_cutoff	Int	Optional	{...}
freyja_fastq	freyja_barcodes	File	workspace.H5Nx_barcodes	{...}
freyja_fastq	freyja_lineage_metadata	File	Optional	{...}
freyja_fastq	freyja_pathogen	String	Optional	{...}
freyja_fastq	kraken2_target_organism	String	Optional	

Figure 5.2

	<b>Running Influenza A, H5Nx Metagenomic Samples in Terra using Theiagen's Freyja FASTQ Workflow</b>	
	Document TG-FREY-H5NX, Version 1	
	Date:	Workflow Versions:
	07/08/2025	PHB v3

freyja_fastq	freyja_lineage_metadata	File	Optional	Figure 5.3
freyja_fastq	freyja_pathogen	String	Optional	
freyja_fastq	kraken2_target_organism	String	Optional	
freyja_fastq	ont	Boolean	Optional	
freyja_fastq	primer_bed	File	Optional	
freyja_fastq	read2	File	this.read2	

9. Specify outputs by clicking on the **outputs** tab and selecting **Use defaults** (Fig 6)



SCRIPT \*\* INPUTS \*\* **OUTPUTS** \*\* RUN ANALYSIS

Output files will be saved to  
Files / submission unique ID / freyja\_fastq / workflow unique ID

References to outputs will be written to  
Tables / ww\_specimen

Fill in the attributes below to add or update columns in your data table


Task name ↓	Variable	Type	Attribute	Use defaults
freyja_fastq	aligned_bai	File	this.aligned_bai	<input type="checkbox"/>
	aligned_bam	File	this.aligned_bam	<input type="checkbox"/>
	alignment_method	String	this.alignment_method	<input type="checkbox"/>

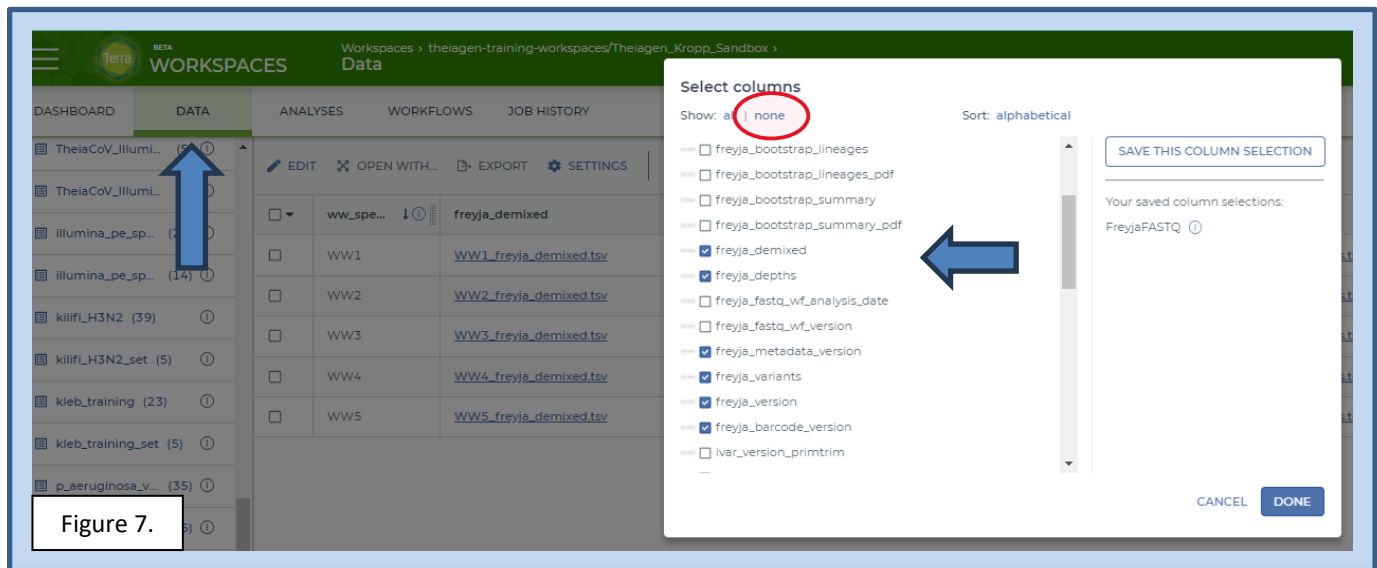
10. Click **save**

11. Launch the workflow by clicking **run analysis**; enter desired comments and click **launch**

## 4.2 DETERMINING LINEAGES, ABUNDANCES, AND COVERAGE METRICS

- In the **data** tab, navigate to the Terra data table containing H5 metagenomic data
- Click **settings** and select **none** to deselect all output columns (Fig 7)
- To simplify the table, select the following outputs:
  - freyja\_barcode\_version**
  - freyja\_demixed**
  - freyja\_depths**
  - freyja\_metadata\_version**
  - freyja\_variants**

	<b>Running Influenza A, H5Nx Metagenomic Samples in Terra using Theiagen's Freyja FASTQ Workflow</b>	
	Document TG-FREY-H5NX, Version 1	
	Date:	Workflow Versions:
	07/08/2025	PHB v3



4. Click on the frejja\_demixed column file to determine the following sample information:
  - a. Lineages identified
  - b. Lineages and relative abundances of lineages
5. Click on the frejja\_variants column file to see all variants identified within the sample
6. Click on the frejja\_depths column file to determine the relative depth of coverage for every variant identified

## 5. QUALITY RECORDS


- Xu, X., et al. (1999). Influenza A virus (A/goose/Guangdong/1/1996 (H5N1)) hemagglutinin (HA) gene, complete cds. NC\_007362.1. NCBI. [https://www.ncbi.nlm.nih.gov/nuccore/NC\\_007362.1](https://www.ncbi.nlm.nih.gov/nuccore/NC_007362.1)
- Workflow version and configuration settings (default and custom inputs)
- Curated lineages and usher barcodes files
- Raw read files
- frejja\_demixed, frejja\_variants, and frejja\_depths tsv output files
- aligned\_bam file for further visualizations

## 6. TROUBLESHOOTING

- Consult with internal staff familiar with this procedure or contact [support@theiagen.com](mailto:support@theiagen.com) for troubleshooting inquiries
- For document edit requests, contact [support@theiagen.com](mailto: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 barcodes file


	Running Influenza A, H5Nx Metagenomic Samples in Terra	
	using Theiagen's Freyja FASTQ Workflow	
	Document TG-FREY-H5NX, Version 1	
	Date: 07/08/2025	Workflow Versions: PHB v3

## 8. REFERENCES

1. Andersen Lab Github. <https://github.com/andersen-lab/Freyja>. Accessed on 07/09/2025..

## 9. REVISION HISTORY

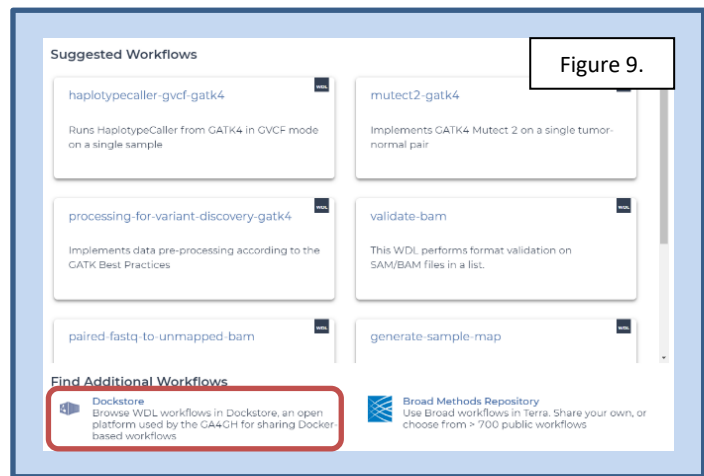
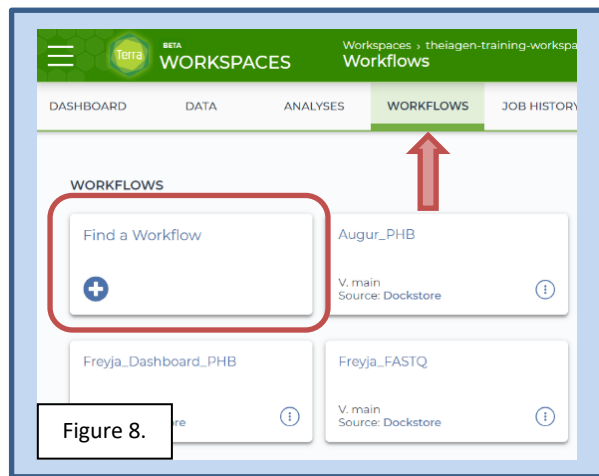
Revision	Version	Release Date
Document creation	1	07/2025

	Running Influenza A, H5Nx Metagenomic Samples in Terra using Theiagen's Freyja FASTQ Workflow	
	Document TG-FREY-H5NX, Version 1	
	Date:	Workflow Versions:
	07/08/2025	PHB v3

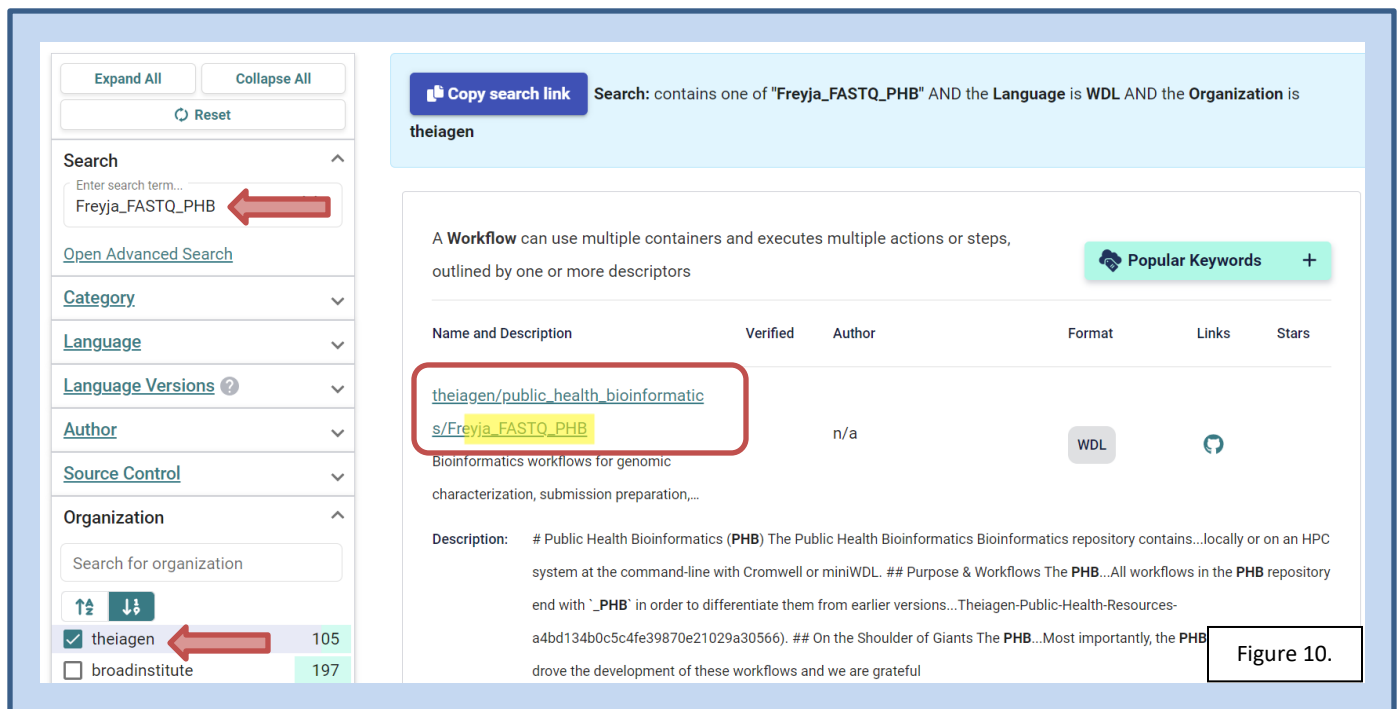
## 10. APPENDICES

### 10.1 IMPORTING FREYJA WORKFLOWS FROM DOCKSTORE


1. In the **Terra workspace** of interest, open the **workflows** tab and click **find a workflow** (Fig 8)
2. In the pop-up window, click **dockstore** (Fig 9)



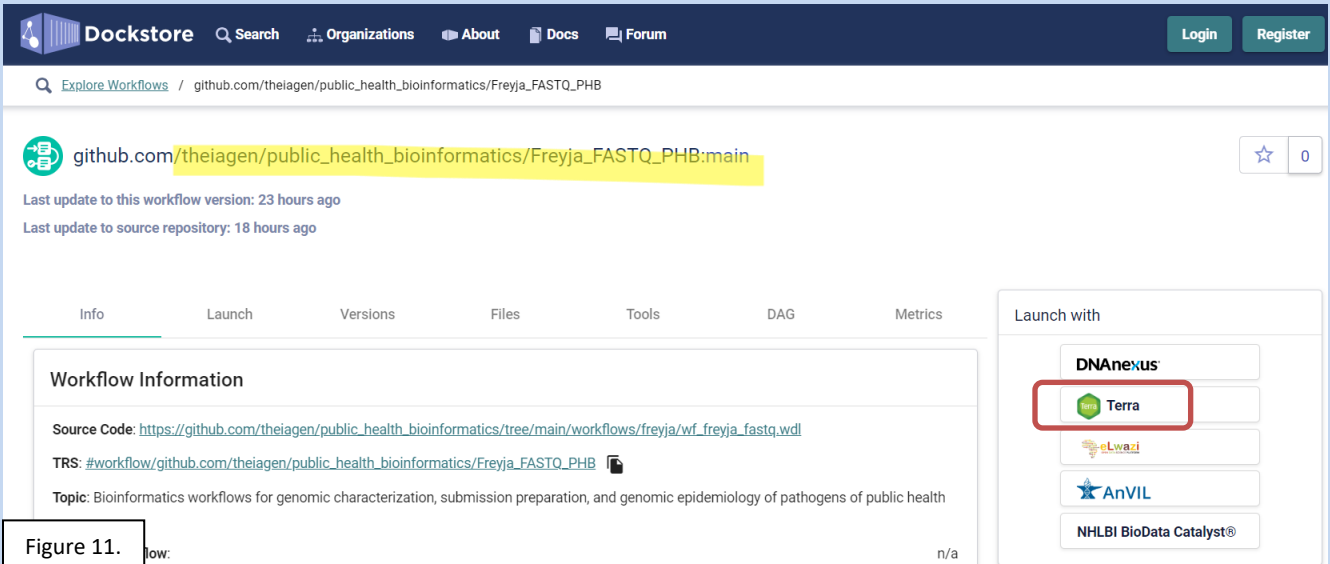
3. To find the Theiagen Freyja FASTQ workflow, type "**Freyja\_FASTQ\_PHB**" in the search bar (Fig 10)
4. In the left hand sidebar, scroll down to Organization and select "**theiagen**" (Fig 10)
5. Find the workflow by looking at the file path suffix; click the name to **open the workflow** (Fig 10)





	Running Influenza A, H5Nx Metagenomic Samples in Terra using Theiagen's Freyja FASTQ Workflow	
	Document TG-FREY-H5NX, Version 1	
	Date:	Workflow Versions:
	07/08/2025	PHB v3

- Click **Terra** to launch the workflow in Terra (Fig 11)
- Choose the **destination workspace** in the dropdown and click **import** or create a new workspace (Fig 12)



**Dockstore** Search Organizations About Docs Forum Login Register

Explore Workflows / github.com/theiagen/public\_health\_bioinformatics/Freyja\_FASTQ\_PHB

github.com/theiagen/public\_health\_bioinformatics/Freyja\_FASTQ\_PHB:main

Last update to this workflow version: 23 hours ago  
Last update to source repository: 18 hours ago

Info Launch Versions Files Tools DAG Metrics

**Workflow Information**

Source Code: [https://github.com/theiagen/public\\_health\\_bioinformatics/tree/main/workflows/freyja/wf\\_freyja\\_fastq.wdl](https://github.com/theiagen/public_health_bioinformatics/tree/main/workflows/freyja/wf_freyja_fastq.wdl)

TRS: [#workflow/github.com/theiagen/public\\_health\\_bioinformatics/Freyja\\_FASTQ\\_PHB](#)

Topic: Bioinformatics workflows for genomic characterization, submission preparation, and genomic epidemiology of pathogens of public health

**Launch with**

DNAexus

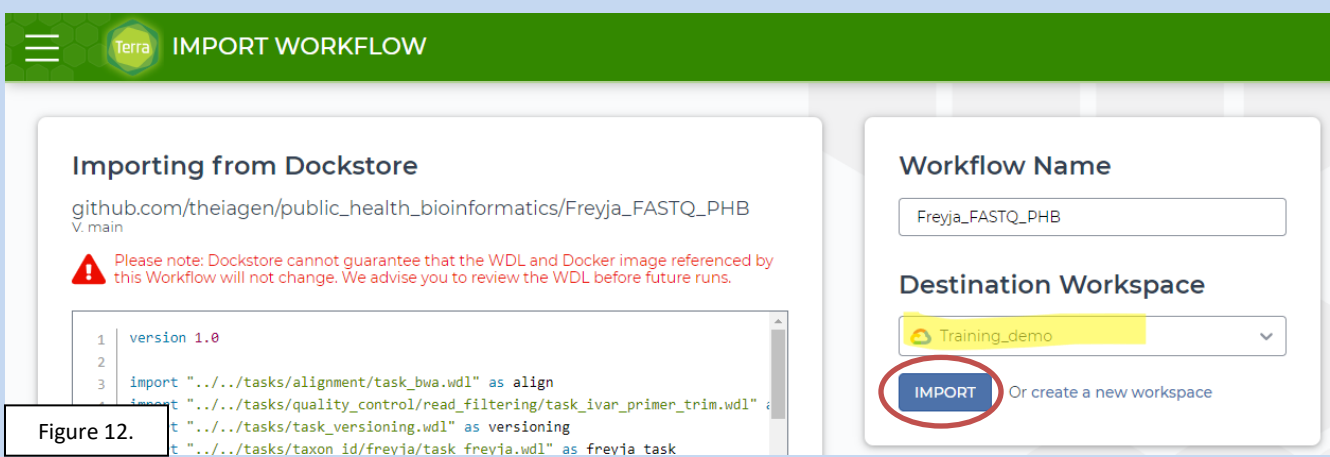
**Terra**

eLwazi

AnVIL

NHLBI BioData Catalyst®

Figure 11.



**Terra IMPORT WORKFLOW**

**Importing from Dockstore**

github.com/theiagen/public\_health\_bioinformatics/Freyja\_FASTQ\_PHB  
V: main

**Please note:** Dockstore cannot guarantee that the WDL and Docker image referenced by this Workflow will not change. We advise you to review the WDL before future runs.

```

1 version 1.0
2
3 import "../tasks/alignment/task_bwa.wdl" as align
4 import "../tasks/quality_control/read_filtering/task_ivar_primer_trim.wdl" as filter
5 import "../tasks/task_versioning.wdl" as versioning
6 import "../tasks/taxon_id/freyja/task_freyja.wdl" as freyja task

```

**Workflow Name**


Freyja\_FASTQ\_PHB

**Destination Workspace**

Training\_demo

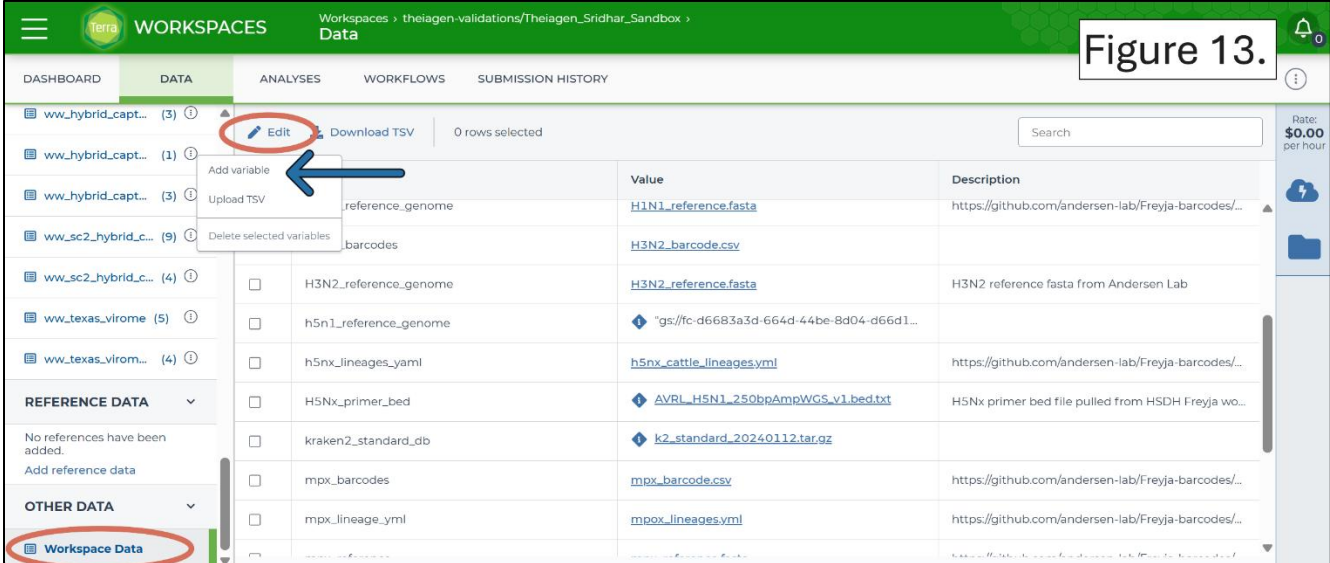
**IMPORT** Or create a new workspace

Figure 12.

	Running Influenza A, H5Nx Metagenomic Samples in Terra	
	using Theiagen's Freyja FASTQ Workflow	
	Document TG-FREY-H5NX, Version 1	
	Date: 07/08/2025	Workflow Versions: PHB v3


## 10.2 ADDING TERRA WORKSPACE DATA ELEMENTS AND FILES

1. Navigate to the **Terra workspace** where analysis will be run
2. To upload local files, open the **Files** tab in the bottom left of the workspace (Fig 13)
  - a. Click **upload** and select the file of interest; ensure the file name does not contain spaces
  - b. Once the upload is complete, **right click** on the file name and click **copy link**
3. Open the **workspace data** tab (Fig 13) and click the **blue plus symbol** in the bottom right (Fig 13)
4. Click in the **key field** and **name the element** being added (Fig 14)
  - a. E.g. to add a primer bed file, the key **SWIFT\_primer\_bed** may be used
5. In the value field, choose **string** as the value type
  - a. **Paste the file path**; the string must start with **"gs://[FILENAME]..."**
    - i. Add a **description** (e.g. updated date/initials), if desired and click the **blue checkmark** (Fig 14)



**Figure 13.**

Key	Value	Description
reference_genome	<a href="#">H1N1_reference.fasta</a>	<a href="https://github.com/andersen-lab/Freyja-barcodes/">https://github.com/andersen-lab/Freyja-barcodes/...</a>
barcodes	<a href="#">H3N2_barcode.csv</a>	
H3N2_reference_genome	<a href="#">H3N2_reference.fasta</a>	H3N2 reference fasta from Andersen Lab
h5n1_reference_genome	<a href="#">"gs://fc-d6683a3d-664d-44be-8d04-d66d1..."</a>	
h5nx_lineages_yaml	<a href="#">h5nx_cattle_lineages.yml</a>	<a href="https://github.com/andersen-lab/Freyja-barcodes/">https://github.com/andersen-lab/Freyja-barcodes/...</a>
H5Nx_primer_bed	<a href="#">AVRL_H5N1_250bpAmpWGS_v1.bed.txt</a>	H5Nx primer bed file pulled from HSDH Freyja wo...
kraken2_standard_db	<a href="#">k2_standard_20240112.tar.gz</a>	
mpx_barcodes	<a href="#">mpx_barcode.csv</a>	<a href="https://github.com/andersen-lab/Freyja-barcodes/">https://github.com/andersen-lab/Freyja-barcodes/...</a>
mpx_lineage_yaml	<a href="#">mpox_lineages.yml</a>	<a href="https://github.com/andersen-lab/Freyja-barcodes/">https://github.com/andersen-lab/Freyja-barcodes/...</a>



**Figure 14.**

Key	Value	Type	Description
H5Nx_barcodes	gs://fc-alf2e(	String	https://github.com/anders...
h5nx_cattle_lineages.yml	<a href="#">h5nx_cattle_lineages.yml</a>		<a href="https://github.com/andersen-lab/Freyja-...">https://github.com/andersen-lab/Freyja-...</a>