

Plotting SARS-CoV-2 Metagenomic Sample Data in Terra using Theiagen's Freyja Plot Workflow

Document TG-FREY-02, Version 3 Date:

4/16/2024

PHB v2

Workflow Versions:

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

3. RELATED DOCUMENTS

Document Number Document Name Uploading Local or SRA NGS Data & Creating a TG-TER-03 Results Metadata Table in Terra Running SARS-CoV-2 Metagenomic Samples in TG-FREY-01 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)



REQUIRED WORKFLOW INPUTS FILES

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



Array[String]

Array[String]

Int

freyja_plot

freyja_plot

Figure 4.

samplename

collection_date

disk_size

this.ww_specimens.ww_specimen_id

Optional

Optional



Plotting SARS-CoV-2 Metagenomic Sample Data in Terra using Theiagen's Freyja Plot Workflow

Document TG-FREY-02, Version 3

Date: Workflow Versions:

PHB v2

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

4/16/2024

b. "summary_time_plot"

c. this.ww_specimens.ww_specimen_id

9. <u>Optional</u>: 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
Figure 5.	plot_time_interval	String	"D"

	Plotting SARS-CoV-2 Meta using Theiagen's	genomic Sample Data in Terra Freyja Plot Workflow
	Document TG-	FREY-02, Version 3
	Date:	Workflow Versions:
$\langle \lambda \rangle$	4/16/2024	PHB v2

- 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

SCRIPT •• INF	PUTS •• OUTPUTS •• RUN ANALYS	IS	
Output files will be saved to Files / submission unique ID / f	reyja_plot / workflow ur e ID	_	
References to outputs will be writt Tables / ww_specimen_set Fill in the attributes below to add o	en to		
	Download j:	son Drag or click to up	oload json Clear outputs SEARCH OUTPUTS
Task name 👃	Download j	son Drag or click to up	oload json Clear outputs SEARCH OUTPUTS
Task name ↓ freyja_plot	Download ji Varlable freyja_demixed_aggregate	son Drag or click to up Type File	Attribute Use defaults this.freyja_demixed_aggregate
Task name↓ freyja_plot Figure 6.	Download ji Variable freyja_demixed_aggregate freyja_plot freyja_plot	son Drag or click to up Type File File	Attribute Use defaults this freyja_plot

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)



Plotting SARS-CoV-2 Metag using Theiagen's F	enomic Sample Data in Terra reyja Plot Workflow	
Document TG-F	REY-02, Version 3	
Date:	Workflow Versions:	
4/16/2024	PHB v2	
	Plotting SARS-CoV-2 Metagen using Theiagen's Free Document TG-Free Date: 4/16/2024	Plotting SARS-CoV-2 Metagenomic Sample Data in Terra using Theiagen's Freyja Plot WorkflowDocument TG-FREY-02, Version 3Date:Workflow Versions:4/16/2024PHB v2

- 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





- 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
- 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





Plotting SARS-CoV-2 Metagenomic Sample Data in Terra using Theiagen's Freyja Plot Workflow

Document TG-FREY-02, Version 3

Workflow Versions:

Date: 4/16/2024

PHB v2

5. QUALITY RECORDS

- freyja_demixed files
- Plot graphs

6. TROUBLESHOOTING

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

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 Metag using Theiagen's F	enomic Sample Data in Terra reyja Plot Workflow	
Document TG-F	REY-02, Version 3	
Date:	Workflow Versions:	
4/16/2024	PHB v2	

5. APPENDICES

10.1 IMPORTING THE FREYJA_PLOT_PHB WORKFLOW FROM DOCKSTORE

	ACES	Norkspaces → theiagen- Norkflows	training-workspa
DASHBOARD DATA	ANALYSES	WORKFLOWS	JOB HISTORY
Find a Workflow	A S	ugur_PHB . main ource: Dockstore	
Freyja_Dashboard_PHB		reyja_FASTQ	
Figure 11	(i) s	. main ource: Dockstore	(

- 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)

Ch Prost		🜓 Copy sea	rch link Search: c	ontains one of "FREY	"JA_PLOT_PHB" AND) the Language is WDL AND) the Organization	is
Q Reset		theiagen						
Search	^							
Enter search term								
FREYJA_PLOT_PHB								
Open Advanced Search		A Workflow	can use multiple co	ntainers and execute	s multiple actions or	r steps,	ulaa Kaunuanda	
		outlined by	one or more descrip	tors		- Pop	ular Keywords	+
<u>Category</u>	~							
Language	~	Name and De	scription	Verified	Author	Format	Links St	ars
	•							
Language Versions 🕜	~	theiagen/pu	ublic_health_bioinfor	matic				
Author		s/Frevia Plo	ot PHB		n/a			
Autor	~	Riginformatio	a workflows for gonomi	in .	n/ a	WDL	0	
Source Control	~	Bioinformatic	s worknows for genorin					
Organization	~	characterizati	on, submission prepara	ition,				
organization		Description:	# Public Health Bioin	formatics (PHB) The Pul	blic Health Bioinformati	cs Bioinformatics repository cor	ntainslocally or on a	n HPC
Search for organization			system at the comm	and-line with Cromwell o	r miniWDL, ## Purpose	& Workflows The PHBAll work	flows in the PHB rep	ository
A 13			end with 'PHB' in or	der to differentiate them	from earlier versions	Theiagen-Public-Health-Resourc	-9C-	
					a il al ll (a)	and a set of the set o		
✓ theiagen	79		a4bd134b0c5c4fe39	870e21029a30566). ##	On the Shoulder of Gian	ts The PHBMost importantly,	Eiguro	12

- 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)

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

		ub.com/trielagen/p	ublic_nealtn_blo	informatics/Freyja	a_Plot_PHB			
عithد 👔	ıb.com/theia	agen/public_	health_bio	informatics	/Freyja_Plo	ot_PHB:main	1	☆ (
Last update to	this workflow ve	ersion: 17 hours ag	10					
ast undate to		rv: 39 minutes and	-)					
and apuate to	source repositor	ly. or minuted age						
	source repositor	y. ov minuteo uge						
Info	Launch	Versions	Files	Tools	DAG	Metrics	Launch with	
Info	Launch	Versions	Files	Tools	DAG	Metrics	Launch with	

Importing from Dockstore	Workflow Name
github.com/theiagen/public_health_bioinformatics/Freyja_Plot_PHB V.main	Freyja_Plot_PHB
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.	Destination Workspace
1 version 1.0	Training_demo
<pre>3 import "//tasks/task_versioning.wdl" as versioning port "//tasks/taxon_id/freyja/task_freyja_plot.wdl" as plot gure 15</pre>	Or create a new workspace