

Creating a Dashboard Visualization of SARS-CoV-2 Metagenomic Samples in Terra using Theiagen's Freyja Dashboard Workflow					
Document TG-FREY-03, Version 2					
Date: Workflow Versions:					
4/16/2024	PHB v2				
	Samples in Terra using Theiag Document TG-I Date:				

1. PURPOSE/SCOPE

To standardize the process of aggregating Freyja sample data to generate an interactive visualization using Theiagen's Freyja Dashboard workflow in Terra. Acceptable data types include Freyja demixed output 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
 - \circ Google Chrome, Firefox, or Edge
- Google account
- Terra account, linked to Google account
- Freyja demixed files uploaded to Terra workspace
- Theiagen's Freyja_Dashboard_PHB workflow in Terra

REQUIRED WORKFLOW INPUTS FILES

- Freyja demixed files (output from Freyja FASTQ
- Dashboard name
- Collection dates
- Viral load
- <u>Optional</u>: freyja_dashboard_config json file

3. RELATED DOCUMENTS

Document Number	Document Name
TG-TER-03	Uploading Local or SRA NGS Data & Creating a
19-1EK-05	Results Metadata Table in Terra
TG-FREY-01	Running SARS-CoV-2 Metagenomic Samples in
IG-FRET-OI	Terra using Theiagen's Freyja FASTQ Workflow
TG-FREY-02	Plotting SARS-CoV-2 Metagenomic Sample Data
IG-FRET-UZ	in Terra using Theiagen's Freyja Plot Workflow



Document TG-FREY-03, Version 2

Workflow Versions:

4/16/2024

Date:

PHB v2

4. PROCEDURE

4.1 RUNNING THE FREYJA DASHBOARD WORKFLOW

- Open Terra and navigate to the workflows tab within the workspace containing metagenomic SARS-CoV-2 data
- 2. Select the *Freyja_Dashboard_PHB* workflow (Fig 1)
- Choose the latest version of version 2 in the version dropdown field, or the version used during internal assay validation (Fig 2, a)

DASHBOARD DATA	ANAL	VSES WORKFLOWS	JOB HISTOR	iY	
WORKFLOWS		freyja	Sort	t By: Alphabetical	~
Find a Workflow		Freyja_Dashboard_PH	в	Freyja_FASTQ	
•		V.v1.0.0 Source: Dr	()	V. main Source: Dockstore	()
Freyja_FASTQ_PHB		Freyja_Plo		Freyja_Plot_PHB	
V.v1.0.0 Source: Dockstore	()	V. main Source: Dockstore	1	V.v1.0.0 Source: Dockstore	1
Freyja_Update_PHB					
Figure 1.	()				



- 4. Select the first bullet to <u>run workflow with</u> inputs defined by file paths (Fig 2, b)
- Select the relevant SET data table name under the select root entity type dropdown (Fig 2, c)
- 6. Click select data (Fig 2, d)
- In the pop-up window select the first bullet to <u>create a new ww_specimen_set from</u> <u>selected ww_specimens</u> (Fig 3, highlighted text)
 - a. Select the checkbox for each sample to be included in the analysis (Fig 3, a)

ate	t w imens to e a set to process	s 5 rows selected	⊟ ADVANCED SEARCH	Search	Q
	ww_specimen_id	10	Run_Date	0	viral_load
~	WW1		2023-06-07		19547
¥	WW2		2023-06-08		20486
*	WW3		2023-06-09		17845
~	WW4		2023-06-10		23895
*	WW5		2023-06-11		1589
Freyja	d ww_specimens will be saved as a new Dashbowrd_20230718b gure 3.	1-5 of 5 《	<	ms per page	



•	ntion of SARS-CoV-2 Metagenomic en's Freyja Dashboard Workflow				
Document TG-I	Document TG-FREY-03, Version 2				
Date:	Workflow Versions:				
4/16/2024	PHB v2				

- b. <u>Optional</u>: name the output set name to differentiate it from other workflow runs, e.g. FreyjaDashboard_YYYYMMDDn (Fig 3, b); this populates a row to the SET data table
- c. Click ok

INPUTS ••	OUTPUTS •• RUN ANALYSIS		
			Downlowson Drag or click to upload json Clear inputs SEARCH IN
	Variable	Туре	Attribute
	collection_date	Array[String]	this.ww_specimens.Run_Date
	freyja_dashboard_title	String	"Default_Dashboard"
	freyja_demixed	Array[File]	this.ww_specimens.freyja_demixed
	samplename	Array[String]	this.ww_specimens.ww_specimen_id
	viral_load	Array[String]	this.ww_specimens.viral_load
	dashboard_intro_text	File	Optional
Figure 4.	config	File	"gs://theiagen-public-files-rp/terra/freyja-files/freyja_dash_config.json"

- 8. In the *inputs* tab, set the first five attributes to the following (Fig 5, yellow highlights):
 - a. this.ww_specimens.Run_Date

i.Notice an "s" is added to the end of the data table name: ww_specimen -> ww_specimens ii. "Run_Date" must match the sample data table column containing collection dates

- b. *"Default_Dashboard"*; enter the dashboard name you want to display in quotes (Fig 8)
- c. this.ww_specimens.freyja_demixed
- d. this.ww_specimens.ww_specimen_id
- e. this.ww_specimens.viral_load
 - *i."viral_load" must match the corresponding sample data table column header*
- 9. <u>Optional</u>: configure an additional input setting to create lineage groups if there are too many lineages to be visualized together (Fig 4, orange highlight)
 - a. In the config variable, add <u>"gs://theiagen-public-files-rp/terra/freyja-</u> <u>files/freyja_dash_config.json</u>" to group lineage according to Figure 5; otherwise, create a customized json file specifying how to define lineage groups
- 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



Document TG-FREY-03, Version 2

Workflow Versions:

Date: 4/16/2024

PHB v2

Lineages:	
grp 1:	
name: "grp_1"	
members:	
- "BA.4*"	
- "BA.5*"	
color: "orange"	
grp_2:	
name: "grp_2"	
members:	
- "B.1.617.2*"	
color: "green"	
grp_3:	
name: "grp_3"	
members:	
- "BA.2*"	
color: "blue"	
Figure 5.	

SCRIPT •• INP	UTS •• OUTPUTS •• RUN ANALYS	IS	
Output files will be saved to	eyja_dashboard / work		
References to outputs will be writte Tables / ww_specimen_set	en to		
	r update columns in your data table	Download json Drag	or click to upload json Clear outputs
	r update columns in your data table Variable	Download json Drag	or click to upload json Clear outputs !
Fill in the attributes below to add or			

DASHBOARD	DATA	ANAL	ANALYSES WORKFLOWS JOB HISTORY							
theiaprok_ill	umin	n EDIT	T 🔀 OPEN WITH 🕒 EXPORT 🏟 SETTINGS 0 ro	ows selected						
theiaprok_ill	umin		ww_specimen_set_id	sbhoard 🕕 f						
🗉 ww_specime	en (5)		FreyjaDashboard_20230718b	Dashboard.html						
🗉 ww_specime	en_set (2)		Freyja_Dashboard_PHB_2023-07-18T19-37 Default_1	Dashboard.html <u>f</u>						
Figure 7. E D	DATA ~									



C C	Creating a Dashboard Visualization of SARS-CoV-2 Metagenomic Samples in Terra using Theiagen's Freyja Dashboard Workflow					
Document TG-F	Document TG-FREY-03, Version 2					
Date:	Workflow Versions:					
4/16/2024	PHB v2					

4.2 ANALYZING DASHBOARDS

- 1. Navigate to the *data* tab of the workspace containing metagenomic data and open the pertinent *SET data table*
- Open the *file* located in the freyja_dashboard column for the relevant data set (Fig 7, red oval)
 - a. If the sample set name was specified (section 4.1, step
 6b) the data set name will be indicated in the ww_specimen_set_id column (Fig 7, highlight)
- 3. Different dashboard visualizations are available by clicking on different blue boxes above the graph (Fig 8)
- 4. The variants visualization

(Fig 8) displays the variant prevalence (y-axis) by collection date (x-axis)

- 5. In the *lineages* view (Fig 9), the lineage group prevalence by collection date is shown
 - a. Lineage groups are displayed according to the group definitions provided by the config variable in section 4.1, step 8; modify groups by changing definitions in the json file and re-running the dashboard workflow
- 6. Viral load by collection date may be viewed by selecting the *viral load* tab of the graph (Fig 10)
- 7. <u>Scaled variants</u> (Fig 11) and <u>scaled lineages</u> (Fig 12) tabs depict the variant or lineage abundance viral loads by collection date





Date: 4/16/2024

Workflow Versions:

PHB v2



5. QUALITY RECORDS

- freyja_demixed files
- Dashboard visualizations



Document TG-FREY-03, Version 2

Workflow Versions:

PHB v2

6. TROUBLESHOOTING

- 1. Identify if failed run is due to "too many lineages to show" error
 - a. In *job history*, click on the *failed run* for freyja dashboard (Fig 13)

Date:

4/16/2024

- b. Open *job manager* in the links column within the table by clicking on the three lines (Fig 14)
- c. Click sign in to authenticate sign in; open the log files indicated by the cloud on paper (Fig 15)
- d. Within the code, use the search function (*CtrI+F* or *Cmd+M*) for "*lineage*;" if the error code indicates "*too many lineages to show*," *re-run the workflow* with a configuration file included in the workflow input settings to simplify the number of lineage groups to display in the visualization

i.See section 4.1 step 8 for details.

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

	DASHBOARD	DATA	ANALYSES	WORKFLO	ws	JOB HISTORY					
Γ						Î					
	Submission (click for	details)			Data	entity		No. of Workflows	Status		Submitted 1
	Freyja_Dashboard_F Submitted by kelsey		n.com		Freyj	aDashboard_2023	0718	1	✓ D	one	Jul 18, 2023 1:46 PM
	Freyja_Dashboard_F Submitted by kelsey		n.com		Freyj	a_Dashboard_PHB	_202	1	✓ D	one	Jul 18, 2023 1:38 PM
	Freyja_Dashboard_F Submitted by kelsey				Freyj	a_Dashboard_PHB	_202	1	A D	one	Figure 13.

DASHBOARD DATA	ANALYSES W	ORKFLOWS JO	OB HISTORY						
← Job History • Submission	← Job History > Submission 29d48f2a-6902-4dff-9e01-a78204c6b69								
Workflow Statuses	Workflow Configuration theiagen-training-workspace		op theiagen.co	m	Total Run Cost N/A				
	Data Entity Freyja_Dashboard_PHB_202 ww_specimen_set	Submis 3-07-18T: 29d48f2		01-a7820	Call Caching Enabled				
Comment 🖍 v1.0.0 - call caching - upd					Retry with More Disabled	Memory			
WORKFLOWS INPUTS	OUTPUTS								
Search workflows	Completion	status	~	Downloa	d TSV				
Data Entity 🖡	Last Changed	Status	Run Cost	Message	Workflow ID	Links			
Freyja_Dashboard_PHB	Jul 18, 2023, 12:54 PM	Failed	N/A		f34769c1 📋				
Figure 14.					3	ob Manager			



Document TG-FREY-03, Version 2

Workflow Versions:

4/16/2024

Date:

PHB v2

freyja_dashboard (ID: f34769c1-5161-4a49-a33c- workspace-id: 774455a0-292b-49	88b2581bd3e3		48f2a-6902-4dff-9e01-a	78204c6b695
Status: Failed()	Erro	ors (1 total)		
Tasks: 1 succeeded, 1 failed, 0 currer processed Submitted: Today, 1250 PM Started: Today, 1250 PM Ended: Today, 1254 PM (0h 3m)	rreyja	dashboar	A 🕻 🖿 i	
LIST VIEW	ERRORS	INPUTS	LABELS	TIMING DIAGRAM
Task Name Shard	Failure Message			backend log Projes
freyja_dashboard_task	a78204c6b695/freyja_dash	board/134769c1-5161-4a 455a0-292b-490f-9766	49-a33c-88b2581bd3e3/call -7b83bacbee59/submission	k gs://fc-774455a0-292b-480f-9766-7b83bacbee59/submissions/29448f2a-6902-4dff-9e07- freyja_dashboard_task/stderr for more information. PAPI error code 9. Please check the log fa /29448f2a-6902-4dff-9e01-a78204c6b695/treyja_dashboard/74769c1-5161-4a49-a33c-

call-freyja_dashboard_task/freyja_da	
return _process_result(sub_ctx.command.invol File "/opt/conda/envs/freyja-env/lib/python: return ctx.invoke(self.callback, **ctx.para)	•
Feturn Ctx:InVoke(self.callback, *Ctx.para) File "/opt/conda/envs/freyja-env/lib/python: Peturncallback(*args, **kwargs)	
File "/opt/conda/envs/freyja-env/lib/python3.10/site-packages/freyja/_cli.py", line 307, in dash	
make_dashboard(agg_df, meta_df, thresh, titleText, introText,	
File "/opt/conda/envs/freyja-env/lib/python3.10/site-packages/freyja/utils.py", line 434, in make_dashboard	
color_lin = get_color_scheme(df_ab_lin,	
File "/opt/conda/envs/frevia-env/lib/python3.18/site-packages/frevia/utils.py", line 72, in get_color_scheme	
raise Exception('Too many <mark>lineage</mark> s to show. Useconfig to group.')	
Exception: Too many lineage s to show. Useconfig to group.	
2023/07/10 10:59:10 Starting delocalization.	Figure 16.
2023/07/18 18:53:11 Delocalization script execution started	inguie 10.

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



Date:

4/16/2024

Workflow Versions: PHB v2

8. REFERENCES

None

9. REVISION HISTORY

Revision	Version	Release Date
Document creation	1	8/2023
Minor document revisions for PHB v2.0.0 release; added appendix	2	4/2024



Date:

4/16/2024

Workflow Versions:

PHB v2

10. APPENDICES

IMPORTING FREYJA WORKFLOWS FROM DOCKSTORE 10.1

- 1. In the Terra workspace of interest, open the workflows tab and click find a workflow (Fig 17)
- 2. In the pop-up window, click *dockstore* (Fig 18)

Workspaces > theiagen-training-workspa WORKSPACES Workflows	Suggested Workflows Figure 18.
DASHBOARD DATA ANALYSES WORKFLOWS JOB HISTOR	haplotypecaller-gvcf-gatk4 mutect2-gatk4 Runs HaplotypeCaller from GATK4 in GVCF mode on a single sample Implements GATK4 Mutect 2 on a single tumor- normal pair
Find a Workflow Augur_PHB V. main Source: Dockstore	processing-for-variant-discovery-gatk4 validate-bam State Processing according to the GATK Best Practices SAM/BAM files in a list.
Freyja_Dashboard_PHB Freyja_FASTQ	paired-fastq-to-unmapped-barn generate-sample-map
Figure 17.	Dockstore Eroowse WDL workflows in Dockstore, an open platform used by the CA4CH for sharing Docker- based workflows Erood workflows in Terra. Share your own, or choose from > 700 public workflows

- 3. Type "*Freyja_Dashboard_PHB*" in the search bar (Fig 19)
- 4. In the left hand sidebar, scroll down to Organization and select "theiagen" (Fig 19)
- 5. Find the workflow by looking at the file path suffix; click the name to open the workflow (Fig 19)

Expand All Collapse	All	Copy search link Search: contains	one of "FRE	/JA DASHBOARD PHE	3 " AND the Language is V	VDL AND the O	rganizatio
🗘 Reset		is theiagen					
Search Enter search term FREYJA_DASHBOARD_PHB							
pen Advanced Search		A Workflow can use multiple container outlined by one or more descriptors	s and execute	es multiple actions or s		oular Keywords	; +
Category	~						
anguage	~	Name and Description	Verified	Author	Format	Links	Stars
anguage Versions 🕜	~	theiagen/public_health_bioinformatic					
uthor	~	<u>s/Freyja_Dashboard_PHB</u>		n/a	WDL	0	
ource Control	~	Bioinformatics workflows for genomic characterization, submission preparation,					
Organization	^			blie Lleelth Disinfermation	Disinformation repository	nteine lessliver	
Search for organization		Description: # Public Health Bioinformatic system at the command-line	. ,			,	
theiagen	77	end with `_PHB` in order to di	fferentiate them	n from earlier versionsThe	eiagen-Public-Health-Resourc		
DOH-HNH0303	9	a4bd134b0c5c4fe39870e210	20230566) ##	On the Shoulder of Ciante	The PHR Most importantly	the PHR	gure 19



Workflow Versions: Date: 4/16/2024

PHB v2

- 6. Click *Terra* to launch the workflow in Terra (Fig 20)
- 7. Choose the *destination workspace* in the dropdown and click *import* or create a new workspace (Fig 21)

Dock	store Q Sea	arch 🚓 Organizat	tions 🌘 About	Docs 📃	Forum		Logi	n Registe
C Explore Workf	lows / github.com	n/theiagen/public_heal	th_bioinformatics/Fr	eyja_Dashboard_PHI	}			
st update to this	com/theiagen workflow version: rce repository: 2 ho	-	_bioinformati	cs <mark>/Freyja_Da</mark> s	hboard_PHE	3:main		
Info	Launch	Versions	Files	Tools	DAG	Metrics	Launch with	
Workflow	Information						DNAnexus ⁻	
	w/github.com/thei	agen/public_health_b			-		eLwazi	
	matics workflows f oublic health conce	for genomic characte	rization, submissic	on preparation, and	genomic epidemic	ology of	NHLBI BioData Ca	

Importing from Dockstore			Workflow Name	
github.com/theiagen/public_health_l	aioinformatics/Erevia Dashboa			
B V. main	ploinionnatics/rieyja_Dashboa		Freyja_Dashboard_PHB	
Please note: Dockstore cannot guarantee this Workflow will not change. We advise	that the WDL and Docker image refere you to review the WDL before future ru	enced by	Destination Workspace	9