	V	
<b>7</b>	A	
	$\mathbf{\nabla}$	

Creating Static Reference Files for Freyja Analysis in Terra using Theiagen's Freyja Update Workflow			
Document TG-FREY-04, Version 2			
sions:			
S			

#### 1. PURPOSE/SCOPE

To standardize the process of pulling the most up-to-date reference files for curated\_lineages and usher\_barcodes files used as inputs to the Freyja FASTQ workflow for lineage deconvolution and assignment. These files will be copied into the Terra workspace as static files where Freyja\_Update\_PHB is run and can be saved for subsequent Freyja FASTQ analyses to allow samples to be analyzed using the same reference files.

#### 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
- Theiagen's Freyja\_Update\_PHB workflow in Terra

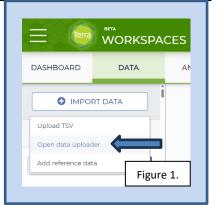
#### 3. RELATED DOCUMENTS

Document Number	Document Name
TG-TER-03	Uploading Local or SRA NGS Data & Creating a
IG-TER-05	Results Metadata Table in Terra
	Running SARS-CoV-2 Metagenomic Samples in
TG-FREY-01	Terra using Theiagen's Freyja FASTQ Workflow

#### 4. PROCEDURE

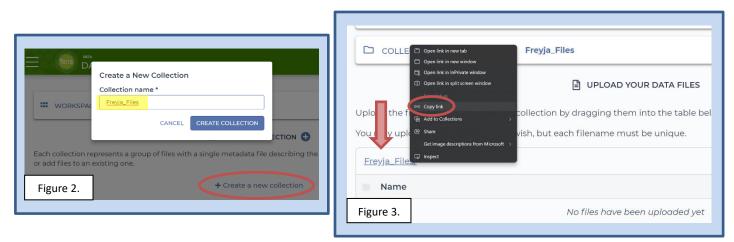
#### 4.1 CREATE A CLOUD LOCATION FOR FREYJA REFERENCE FILES

- If a Freyja\_Files folder has previously been created, but the filepath needs to be copied to enter into the Freyja\_Update workflow inputs, in the Terra data tab, navigate to Files > uploads/ and right click to copy the FreyjaFiles/ folder filepath
- 2. In the *data* tab of the Terra workspace containing metagenomic data, click *import data* and *open data uploader* (Fig 1)
- Click create a new collection (Fig 2) this creates a google bucket storage folder within the Terra workspace where files can be populated



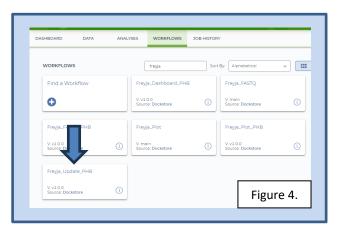
C C	iles for Freyja Analysis in Terra eyja Update Workflow	
Document TG-FREY-04, Version 2		
Date: Workflow Versions:		
4/5/2024	PHB v2	
	using Theiagen's Fr Document TG-I Date:	

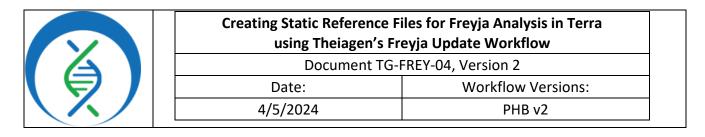
- 4. Name the collection *Freyja\_Files* or something similar and click *create collection* (Fig 2)
- 5. Right click on the *Freyja\_Files folder* and *copy link* (Fig 3); this will be pasted into a Freyja\_Update workflow input field

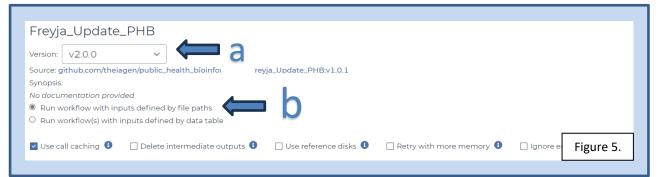


## 4.2 RUNNING THE FREYJA UPDATE WORKFLOW

- 1. Navigate to the *workflows* tab in the workspace containing metagenomic SARS-CoV-2 data
- 2. Select the *Freyja\_Update\_PHB* workflow (Fig 4)
- Choose the latest version of version 2 in the version dropdown field (Fig 5, a)
- Select the first bullet to *run workflow with inputs defined by file paths* (Fig 5, b)
- 5. Uncheck call caching (Fig 5)
- 6. In the *inputs* tab, *paste* in the first attribute field the link copied above in step 4.1, 4:
  - a. "gs://fc-774455a0-292b-490f-9766-7b83bacbee59/uploads/Freyja\_Files/"
    - i.Remove the last "" from the pasted file path so that the first input value reads "gs://fc-774455a0-292b-490f-9766-7b83bacbee59/uploads/Freyja\_Files"
- 7. Click *save*, *run analysis*, *add comments* if desired, and click *launch*
- 8. This will populate files into the Terra workspace folder created in section 4.1 that can be used as Freyja reference files for the Freyja\_FASTQ workflow to ensure continuity across Freyja analysis







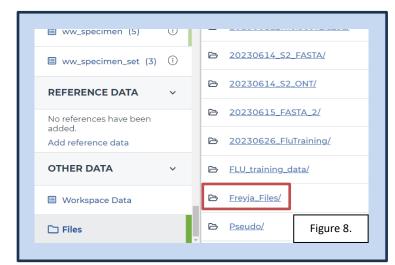
SCRIPT •• INPU	JTS •• OUTPUTS •• RUN AN	IALYSIS	CANCEL SAVE
Hide optional inputs	Download json	Drag or click to upload json	Clear input
Task name ↓	Variable	Туре	Attribute
freyja_update	gcp_uri	String	"gs://fc-774455a0- -292b-490f-9766- 7b83bacbee59/uploads/ Freyja_Files"
freyja_update_refs	disk_size	Int	Optional 🛔 []
Figure 6.	docker	String	Optional

## 4.3 ADDING/UPDATING WORKSPACE DATA ELEMENTS FOR REFERENCE FILES TO USE AS INPUTS FOR FREYJA\_FASTQ\_PHB ANALYSIS

- In the Terra workspace, navigate to the *data tab* (Fig 7)
- Scroll to the bottom of the left-hand sidebar and click *Files* (Fig 7)
- 3. Open the *uploads/* folder (Fig 7)
- 4. Open the *Freyja\_Files* folder created in section 4.1 (Fig 8)
- 5. Open the most recent folder indicated by date, e.g. 2024-04-11 (Fig 9)
- Hover the mouse over the curated\_lineages.json file and click the clipboard to *copy the filepath* or right click on the file name and select copy link address (Fig 10)

DASHBOARD	DATA	ANAL	YSES	WORKFLOWS	JOB HI
🗏 theiaeuk (1)	1	Files			
theiaprok_illumin		🗆 Nar	ne		
theiaprok_illumin	(1) 🕛		TAs/		
theiaprok_ont (2)	()	Ere:	yja Validati	onCriteria.txt	
🗉 ww_specimen (5	) ()	🖻 <u>Ger</u>	nbankFiles	2	
ww_specimen_se	t (3) 🕕		TrackerRe	sults.csv	
REFERENCE DATA	. ~	🖻 Ref	erenceSec	as/	
No references have be added.	een	SD.	PhoenixV	alidationCriteria.txt	i.
Add reference data		🕞 <u>Vali</u>	dationCrit	eria/	
OTHER DATA	~	B gc_	check_tas	k/	
Workspace Data		🕞 sub	missions/		

C C		
Document TG-FREY-04, Version 2		
Date: Workflow Versions:		
4/5/2024	PHB v2	
	using Theiagen's Fr Document TG-I Date:	Date: Workflow Versions:



	WORKSPA	CES Workspaces > theiage Data
DASHBOARD	DATA	ANALYSES WORKFLOWS
🗉 theiaeuk (1)		<u>Files</u> / <u>uploads</u> / <u>Freyja_Files</u>
theiaprok_illur	min (5) 🗊	🗌 Name
🗉 theiaprok_illur	min (1) 🗓	► <u>2023-08-03/</u>
Figure 9.	(2) 🔅	▶ 2024-04-11/

Workspaces > theiagen-training-workspaces Data					
DASHBOARD	DATA	ANALYSES WORKFLOWS JOB HISTORY			
🗉 theiaeuk (1)	(!)	Files / uploads / Freyja_Files / 2024-04-11			
theiaprok_illum	iin (5) 🕕	Name			
theiaprok_illum	in (1) 🛈	c <u>urated_lineages.json</u> Ů			
theiaprok_ont	(2)	update_log.txt			
🗉 ww_specimen	(5) 🔅	usher_barcodes.csv Figure 10.			

- 7. Open the *workspace data* tab in the bottom right of the left-hand sidebar (Fig 11, blue arrow)
- If reference files have never been added as workspace data elements, click the blue plus symbol in the bottom right (Fig 11, red arrow)
  - a. Scroll to the bottom of the table where a new row will be populated
- 9. If reference files were previously added as workspace data elements, hover the mouse over the relevant entry and click the *pencil icon* that will appear to the right to edit the variable
- 10. Click in the *key field* and *name the element* being added (Fig 12)

a. For the Freyja curated\_lineages file, it may be helpful to name the key FreyjaLineageMetadata

- 11. In the value field, choose *string* as the value type and paste the filepath copied above in step 6
  - a. No quotation marks need to be added to this string
- 12. In the description field, it may be helpful to specify Updated YYYY-MM-DD, initials



# **Creating Static Reference Files for Freyja Analysis in Terra** using Theiagen's Freyja Update Workflow

Document TG-FREY-04, Version 2

Workflow Versions:

4/5/2024

Date:

PHB v2

= 🥮 WORKSP/	ACES Data		Figure 11.
DASHBOARD DATA	ANALYSES WORKFLOWS JO	BHISTORY	
🗏 p_aeruginosa_va (35) 🛈		Download TSV   Drag or click to up	pad TSV Search
🖬 p_aeruginosa_vali (1) 🛈	Key	Value	Description
🗏 p aeruginosa vali (1) 🛈	Artic_V3_primer_bed	APTUCOA SATA buttlet bed	
🗉 theiaprok_illumin (5) 🛈	Artic_V4-1_primer_bed	V4-1_nCoV-2021.primer.bed	
🗄 theiaprok_illumin (1) ①	Artic_V4_primer_bed	V4_nCoV_2021_primer.bed	
🗉 ww_specimen (5) 🕕	Midnight_primer_bed	Midnight_Primers_SARS-CoV-2.scheme	
ww_specimen_set (2)	SWIFT_primer_bed	SWIFT_SARS_CoV_2.scheme.bed	Updated 2023-07-05
	freyja_dashboard_config	freyja_dash_config.json	Input 2023-07-18
REFERENCE DATA ~	kraken2 phoenix	k2_standard_08ob_20230605.tar.gz	Updated by Ines on 21/07/2023
No references have been added. Add reference data	nextclade_dataset_tag	5055-0A-50115:00:00X	Updated on 2022-08-12
OTHER DATA	nextclade docker image	nextstrain/nextclade:2.4.0	Updated on 2022-08-12
Workspace Data	golin_docker_image	staphb/pangolin:4.1.2-pdata-1.13	Updated on Y022-08-12
C1 Files	vadr docker image	staphb/vadr:1.4.2	U 2-07-1 ADD VARIABLE +

- 13. Repeat steps 2-12 to copy the most recent or desired usher barcodes filepath and add or update a workspace data element for FreyjaUsherBarcodes
- 14. Proceed to run the Freyja\_FASTQ\_PHB workflow to (re-)analyze SC2 metagenomic samples a. The corresponding workflow inputs for updated reference files will be as follows:
  - i.Freyja\_lineage\_metadata: workspace.FreyjaLineageMetadata *ii.Freyja\_usher\_barcodes:* workspace.FreyjaUsherBarcodes

Artic_V4-1_primer_bed	V4-1_nCoV-2021.primer.bed		Figure 12.
Artic_V4_primer_bed	V4_nCoV-2021.primer.bed		
FreyjaLineageMetadata	gs://fc-58c86846-73: String V	Updated 2024-4-11 kk	⊘ ⊗
FreyjaUsherBarcodes	usher_barcodes.csv	Updated 2024-4-11 kk	
Freyja_ReferenceGenome	<u>nCoV-2019.reference.fasta</u>	MN908947.3	



# Creating Static Reference Files for Freyja Analysis in Terra using Theiagen's Freyja Update Workflow

Document TG-FREY-04, Version 2

Workflow Versions:

PHB v2

#### 5. QUALITY RECORDS

- Freyja curated\_lineages and usher\_barcodes reference files
- Workflow version and configuration settings (default and custom inputs)

Date:

4/5/2024

#### 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

None

#### 8. REFERENCES

None

#### 9. REVISION HISTORY

Revision	Version	Release Date
Document creation	1	8/2023
Updates for PHB v2.0.0 release (no major functionality changes), clarified process to re-use reference files (added section 4.3), updated quality records, added appendix 10.1	2	4/2024



# Creating Static Reference Files for Freyja Analysis in Terra using Theiagen's Freyja Update Workflow

Document TG-FREY-04, Version 2

Workflow Versions:

PHB v2

**10. APPENDICES** 

#### 10.1 IMPORTING FREYJA\_UPDATE\_PHB WORKFLOW FROM DOCKSTORE

Date:

4/5/2024

- 1. In the Terra workspace of interest, open the workflows tab and click find a workflow (Fig 13)
- 2. In the pop-up window, click *dockstore* in the bottom left (Fig 14)

	Workspaces > theiagen-training-workspa Workflows	Suggested Workflows	Figure 14.
	LYSES WORKFLOWS JOB HISTORY	haplotypecaller-gvcf-gatk4 Runs HaplotypeCaller from GATK4 in GVCF mode on a single sample	mutect2-gatk4 Implements CATK4 Mutect 2 on a single tumor- normal pair
Find a Workflow	Augur_PHB V. main Source: Dockstore	processing-for-variant-discovery-gatk4 Implements data pre-processing according to the CATK Best Practices	validate-barn This WDL performs format validation on SAM/BAM files in a list.
Freyja_Dashboard_PHB	Freyja_FASTQ	paired-fastq-to-unmapped-bam	generate-sample-map
Figure 13.	V. main Source: Dockstore	Eind Additional Workflows Dockstore Dockstore Dockstore Dockstore Dockstore, an open platform used by the CA4CH for sharing Docker- based workflows	Broad Methods Repository Use Broad workflows in Terra. Share your own, or choose from > 700 public workflows

- 3. Type "*Freyja\_Update\_PHB*" in the search bar (Fig 15)
- 4. In the left-hand sidebar, scroll down to "Organization" and select "*theiagen*" (Fig 15)
- 5. Find the workflow by looking at the file path suffix; click the name to open the workflow (Fig 15)

Expand All Collapse All	Reset	🕒 Copy sear	rch link Search: contains one	of "Freyja_Update_PHB	" AND the L	anguage is WDL AND	the Organization is t	heiagen	
Search	^								
Enter search term Freyja_Update_PHB	×								
Open Advanced Search		A Workflow can use multiple containers and executes multiple actions or steps, outlined by one or more descriptors					s +		
<u>Category</u>	~								
Language	~	Name and Des	scription	Ve	rified A	uthor	Format	Links	Stars
Language Versions 🕜	~	theiagen/pu	blic_health_bioinformatics/Frey	a_Update_PHB					
Author	~	Bioinformatics	s workflows for genomic characteriza	tion, submission	n,	/a	WDL	0	
Source Control	~	preparation, ar	nd genomic epidemiology of pathoger	ns of public heal					
Organization	^	Description:	# Public Health Bioinformatics (PHI	B) The Public Health Bioinfo	ormatics Bioi	nformatics repository cont	tainslocally or on an HI	PC system at the	a command-
Search for organization			line with Cromwell or miniWDL. ## from earlier versionsThelagen-Pul			1	· -		
12 49			importantly, the <b>PHB</b> user communi	ty drove the development o	of these work	flows and we are grateful			
✓ theiagen	88	Path:	github.com/theiagen/public_health,	_bioinformatics/Freyja_Upd	date_PHB				gure 15.



# Creating Static Reference Files for Freyja Analysis in Terra<br/>using Theiagen's Freyja Update WorkflowDocument TG-FREY-04, Version 2Date:Workflow Versions:4/5/2024PHB v2

6. Click *Terra* to launch the workflow in Terra (Fig 16)

				reyja_Update_PHB			
aithub.	com/theiagen	/public_health	bioinformati	cs/Frevia_Up	date_PHB:ma	in	
	s workflow version:						
update to so	urce repository: 1 ho	our ago					
Info	Launch	Versions	Files	Tools	DAG	Metrics	Launch with
Workflow	Information						DNAnexus
Workflow	Information						DNAnexus <sup>®</sup>
Source Code	https://github.com/	/theiagen/public_hea		_		_update.wdl	
Source Code TRS: <u>#workfl</u>	https://github.com/	/ <u>theiagen/public_hea</u> agen/public_health_t	bioinformatics/Frey	ija_Update_PHB			Terra

 Choose the *destination workspace* in the dropdown and click *import* or create a new workspace (Fig 17)

Importing from Dockstore	Workflow Name
github.com/theiagen/public_health_bioinformatics/Freyja_Update_PHB <sup>v. main</sup>	Freyja_Update_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	C Training_demo
2 3 import "//tasks/taxon_id/freyja/task_freyja_transfer.wdl" as transfer port "//tasks/taxon_id/freyja/task_freyja_update.wdl" as update	IMPORT Or create a new workspace