



# Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

## 1. PURPOSE/SCOPE

To standardize the process of uploading next generation sequencing (NGS) data from BaseSpace into Terra. See [appendix 10.2](#) for details on initial setup to link BaseSpace and Terra accounts. Acceptable data types include Illumina raw sequencing file formats. Read the documentation [here](#).

## 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
- BaseSpace account, linked to Terra account; see [Appendix 10.2](#) for details on linking accounts

### IMPORTANT NOTES

- Metadata column headers and workflow input text indicated in **gray** in this SOP are customizable; **black** is required text
- Terra data table column headers become available as workflow inputs when running workflows, search for them in workflow input dropdowns using the prefix **this.** to filter
- Filter for workspace data and files in workflow input dropdowns using the prefix **workspace.**

## 3. RELATED DOCUMENTS

Document Number	Document Name
None	None

## 4. PROCEDURE

**Initial Setup Required!** Labs using BaseSpace Fetch for the first time must begin by importing the BaseSpace\_Fetch\_PHB workflow and using command line to provide BaseSpace credentials to Terra to pull read files. [Start with appendices 10.1 and 10.2.](#)

### 4.1 DOWNLOAD BASESPACE SAMPLE SHEET AND IMPORT TO TERRA

1. [Login to the BaseSpace](#) account where samples were run at <https://basespace.illumina.com/>
2. In the runs tab, [select the run](#) to be analyzed (Fig 1)
3. Within the files tab, scroll to the bottom and [download the sample sheet](#) (Fig 2 and 3)

STATUS	RUN NAME	AVG Q30	SPP	INSTRUMENT	CREATED
Complete	SEQ002	80.33%	61.26%		
Complete	SEQ001	87.21%	53.46%		

Figure 1.



## Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

Written for Workflow Version:

1/21/2026

PHB v4

The screenshot shows the Nextera DNA CD Indexes - 24 Indexes Tubed analysis page. The top navigation bar includes links for SUMMARY, BIOSAMPLES, SAMPLES, CHARTS, METRICS, INDEXING QC, SAMPLE SHEET, and FILES. Below the navigation is a row of five circular icons representing different types of data or analysis steps. The main content area displays a table with columns for NAME, TYPE, SIZE, PATH, and CREATED. Two files are listed: 'Config' (F...) and 'RunParameters.xml' (xml, 2.89 KB, RunParameters.). At the bottom, a file named 'sampleSheet.csv' (CSV, 978 B, SampleSheet.csv) is highlighted with a red box. A large red arrow points upwards from this highlighted file towards a modal window. The modal, titled 'SampleSheet.csv', contains the CSV file's contents and a 'DOWNLOAD' button, which is also highlighted with a red box.

SEQ002

SUMMARY BIOSAMPLES SAMPLES CHARTS METRICS INDEXING QC SAMPLE SHEET FILES

NAME TYPE SIZE PATH CREATED

Config F...

RunParameters.xml xml 2.89 KB RunParameters.

sampleSheet.csv CSV 978 B SampleSheet.csv

Figure 2.

SampleSheet.csv

[Header]  
Local Run Manager Analysis Id,2002  
Experiment Name,SEQ002  
Date,2020-10-20  
Module,GenerateFASTQ - 2.0.0  
Workflow,GenerateFASTQ  
Library Prep Kit,Nextera DNA CD Indexes - 24 Indexes Tubed  
Description,Chemistry,Amplicon

[Reads]

Figure 3.

4. Open the sample sheet in excel
  5. Click the plus symbol at the bottom of the window to open a new sheet or open a new workbook (Fig 4)

Figure 4.

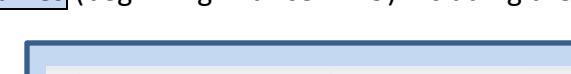
	A	B	C	D	E	F	G	H	I	
1	[Header]									
2	Local Run Manager	2002								
3	Experiment Name	SEQ002								
4	Date	10/20/2020								
5	Module	GenerateFASTQ - 2.0.0								
6	Workflow	GenerateFASTQ								
7	Library Prep Kit	Nexsera DNA CD Indexes- 24 Indexes Tubed								
8	Index Kit	Investigation samples								
9	Chemistry	Amplicon								
10										
11	[Reads]									
12		151								
13		151								
14										
15	[Settings]									
16	adapter	CTGTCTCTTACACATCT								
17										
18	[Data]									
19	Sample_ID	Sample_Name	Description	Index_Plate	Index_Plate_Well	I7_Index_ID	index	I5_Index_ID	index2	Sample_Project
20	2010034999	2010034999-ABC		C	A01	H706	TAGGCATG	H505	CTCCITAC	SEQ002_QC_Salmonella
21	2010015000	2010015000-ABC		C	B01	H707	CTCTCTAC	H505	CTCCITAC	SEQ002_QC_Salmonella
22	2010035001	2010035001-ABC		C	C01	H710	CGAGGGCTG	H505	CTCCITAC	SEQ002_QC_Salmonella
23	2010035002	2010035002-ABC		C	D01	H711	AAGAGGCCA	H505	CTCCITAC	SEQ002_QC_Salmonella
24	NTC-201010	NTC-201010-ABC		C	E01	H714	GCTCATGA	H505	CTCCITAC	SEQ002_QC_Salmonella
25										
26										
27										
28										
	<a href="#">SampleSheet</a>	<a href="#">Sheet1</a>								

6. **Copy the lists of sample IDs and Names** (beginning with cell A19) including the headers “Sample\_ID” and “Sample\_Name” (Fig 4)

7. **Paste** into the new sheet (Fig 5)

8. **Copy + paste the** “Sample\_Project” column into new sheet (Fig 4 and 5)

a. If no “Sample\_Project” listed, use “Experiment Name” for this value



	A	B	C
1	Sample_ID	Sample_Name	Sample_Pro
2	2010034999	2010034999-ABC	SEQ002_QC
3	2010015000	2010015000-ABC	SEQ002_QC
4	2010035001	2010035001-ABC	SEQ002_QC
5	2010035002	2010035002-ABC	SEQ002_QC
6	NTC-201010	NTC-201010-ABC	SEQ002_QC
7			

	A	B	C
1	Sample_ID	Sample_Name	Sample_Project
2	2010034999	2010034999-ABC	SEQ002_QC_Salmonella
3	2010015000	2010015000-ABC	SEQ002_QC_Salmonella
4	2010035001	2010035001-ABC	SEQ002_QC_Salmonella
5	2010035002	2010035002-ABC	SEQ002_QC_Salmonella
6	NTC-201010	NTC-201010-ABC	SEQ002_QC_Salmonella
7			Figure 5.



# Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

9. **Rename the column headers** in the new sheet (Fig 6):

- Column A: **PulseNet\_id** (previously “Sample\_ID”)
  - where **PulseNet** is the unique name of **your** designated Terra data table
- Column B: **basespace\_sample\_name** (previously “Sample\_Name”)
- Column C: **basespace\_collection\_id** (previously “Project\_Name”)

A	B	C
1 PulseNet_id	basespace_sample_name	basespace_collection_id
2 2010034999	2010034999-ABC	SEQ002_QC_Salmonella
3 2010035000	2010035000-ABC	SEQ002_QC_Salmonella
4 2010035001	2010035001-ABC	SEQ002_QC_Salmonella
5 2010035002	2010035002-ABC	SEQ002_QC_Salmonella
6 NTC-201010	NTC-201010-ABC	SEQ002_QC_Salmonella
7		
8		

Figure 6.

10. **Select and copy all cells** containing text, e.g. A1 to C6 (Fig 6)

11. In the Terra workspace where data will be uploaded, click **import data** and **upload tsv** (Fig 7)

12. Click **text import**; **paste into the text box** within the pop-up window; click **start import job** (Fig 8)

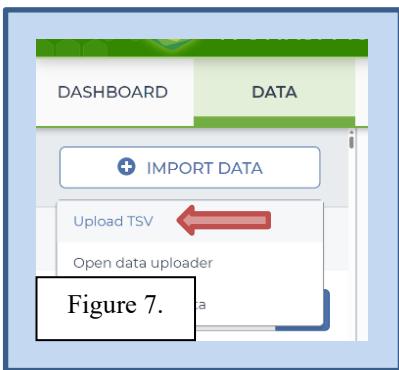


Figure 7.

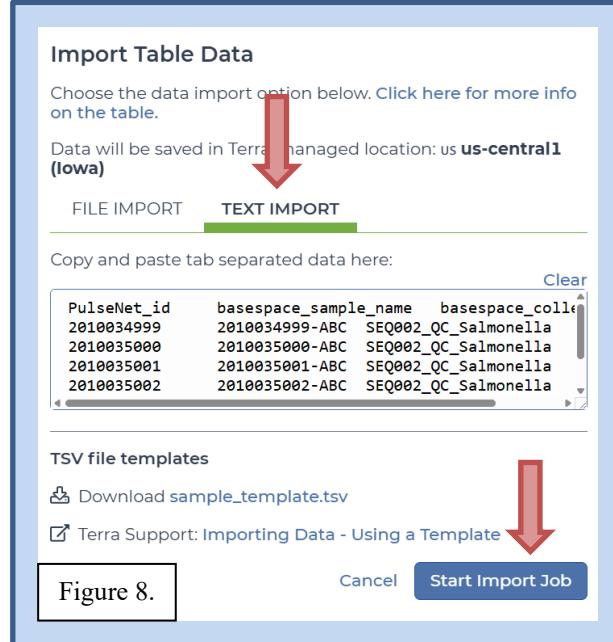


Figure 8.

13. A new data table should now be listed in the tables sidebar or, if adding to an existing table, data should be populated in the relevant table



## Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

### 4.2 RUNNING THE BASESPACE WORKFLOW

- In the Terra workspace where the BaseSpace sample sheet was imported, [open the BaseSpace\\_Fetch\\_PHB workflow](#) within the workflows tab (Fig 9)
  - For instructions on importing the BaseSpace\_Fetch workflow, see [appendix 10.1](#)

Figure 9.

- Choose the **latest** version of the workflow in the [version dropdown field](#) (Fig 10, a)
- Click the second bullet to [run workflow\(s\) with inputs defined by data table](#) (Fig 10, b)
- Select the relevant data table name from the [select data table](#) dropdown, e.g. [PulseNet](#) (Fig 10, c)
- Click [select data](#) (Fig 10, d)

Figure 10.

- In the pop-up window, [choose the samples](#) for which reads should be imported, [modify the sample set name](#) if desired, and click [ok](#) (Fig 11)
  - Optional: to select all samples for sample sets greater than 100, click the down arrow and select all to include all samples in the table
  - Optional: a name will be assigned to this set of samples chosen; for traceability it is useful to modify the job name to include the user's initials, e.g. `BS_FetchPHB_YYYYMMDD[initials]`



# Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

Figure 11.

Select Data

Choose specific PulseNets to process

Select PulseNets to process SETTINGS 5 rows selected

ADVANCED SEARCH Search

Page	basespace_collection_id	basespace_sample_name
All (5)	10034999	SEQ002_QC_Salmonella
None	10035000	SEQ002_QC_Salmonella
<input checked="" type="checkbox"/>	2010035001	SEQ002_QC_Salmonella
<input checked="" type="checkbox"/>	2010035002	SEQ002_QC_Salmonella
<input checked="" type="checkbox"/>	NTC-201010	NTC

1 - 5 of 5 << < > >> Items per page: 100

Selected PulseNets will be saved as a new PulseNet\_set named:  
BS\_FetchPHB\_20230914kk

CANCEL OK

7. Set the first five input settings as follows (Fig 12):

- a. `workspace.access_token`
- b. `this.basespace_collection_id`
- c. `this.basespace_sample_name`
- d. `this.PulseNet_id`
  - i. where `PulseNet` is the unique name of your data table in Terra
- e. `workspace.api_server`

Figure 12.

SCRIPT INPUTS OUTPUTS Launch

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

Task name ↓	Variable	Type	Input value
basespace_fetch	access_token	String	<code>workspace.access_token</code> [...]
basespace_fetch	basespace_collection_id	String	<code>this.basespace_collection_id</code> [...]
basespace_fetch	basespace_sample_name	String	<code>this.basespace_sample_name</code> [...]
basespace_fetch	sample_name	String	<code>this.PulseNet_id</code> [...]
	api_server	String	<code>workspace.api_server</code> [...]



# Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

8. In the outputs tab, click **use defaults** and **save** the workflow (Fig 13)
9. Then click **run analysis**, enter any comments if desired, and click **launch**

Figure 13.

Task name ↓	Variable	Type	Attribute	Use defaults
basespace_fetch	basespace_fetch_analysis_date	String	this.basespace_fetch_analysis_date	<input checked="" type="checkbox"/>
basespace_fetch	basespace_fetch_version	String	this.basespace_fetch_version	<input checked="" type="checkbox"/>
basespace_fetch	read1	File	this.read1	<input checked="" type="checkbox"/>
	read2	File	this.read2	<input checked="" type="checkbox"/>

## 5. QUALITY RECORDS

- Raw read files
- Metadata results table
- Workspace elements and files

## 6. TROUBLESHOOTING

- If the first cell in the metadata table does not start end with `_id`, an error message will prevent file import; adjust the metadata text in cell A1 and re-upload
- If any workflow, input, or output settings are entered incorrectly, the analysis will not run as expected; verify all settings are correct and re-launch analysis
- Terra data table column headers become available as workflow inputs when running workflows; search for them in workflow input dropdowns using the prefix `this`. To filter for workspace data and files in workflow input dropdowns using the prefix `workspace`.
- If analysis fails, navigate to the job history in the workspace and click on the job submission for details; for help resolving run failures, email [support@theiagen.com](mailto:support@theiagen.com)

## 7. LIMITATIONS

None



## Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

### 8. REFERENCES

1. Libuit, Kevin G., Emma L. Doughty, James R. Otieno, Frank Ambrosio, Curtis J. Kapsak, Emily A. Smith, Sage M. Wright, et al. 2023. "Accelerating Bioinformatics Implementation in Public Health." *Microbial Genomics* 9 (7). <https://doi.org/10.1099/mgen.0.001051>
2. Theiagen Genomics [Public Health Bioinformatics Workflow Documentation](#)
3. [Theiagen's BaseSpace Fetch](#) workflow documentation

### 9. REVISION HISTORY

Revision	Version	Release Date
Document Creation	1	9/2023
Aligning with PHB v2.0.0 release; slight format and figure changes	2	4/2024
Edits to align with Terra interface, versioning updates, and links	3	5/2025
Figure, header, and minor formatting edits	4	1/21/2026



# Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

## 10. APPENDICES

### 10.1 IMPORTING THE BASESPACE\_FETCH\_PHB WORKFLOW FROM DOCKSTORE

1. In the **Terra workspace** of interest, open the **workflows** tab and click **find a workflow** (Fig 14)
2. In the pop-up window, click **Dockstore.org** (Fig 15)

Figure 14.

Figure 15.

3. Workflows may be retrieved in two ways:
  - a. Navigate through the organization to see all available workflows (Fig 16)
    - i. Click **organizations**
    - ii. In the search bar type **Theiagen**
    - iii. Click on the **logo**, **view**, or **# collections**
    - iv. Click on collection: **Public Health Bioinformatics (PHB)** (Fig 17)
    - v. Use **ctrl + F** to search for the workflow name, e.g. **basespace** (Fig 18)

Figure 16.

Figure 17.



# Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

Figure 18.

- b. Navigate through the search bar (Fig 19)
  - i. Always select **PHB** versions of Theiagen workflows

Figure 19.

4. Open the workflow (workflow name listed at the end of the file path)
5. Click Terra to launch the workflow in Terra (Figure 20)

Figure 20.



# Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

6. Choose the *destination workspace* in the dropdown (Figure 21)
7. Click *import* or *create a new workspace* (Figure 21)

Figure 21.

Importing from Dockstore

```
github.com/theiagen/public_health_bioinformatics/TheiaCoV_ClearLabs_PHB
V.v1.0.0

⚠ 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.
```

Workflow Name

BaseSpace\_Fetch\_PHB

Destination Workspace

TheiaCoV\_Training\_Demos

IMPORT Or create a new workspace



# Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

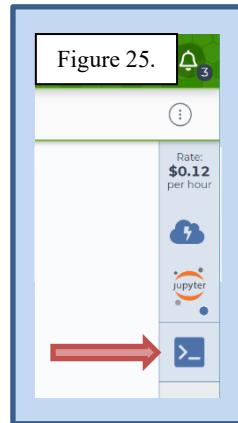
## 10.2 OPEN COMMAND LINE TERMINAL TO PROVIDE BASESPACE CREDENTIALS TO TERRA

\*This setup is up only required once per BaseSpace account. Additionally, the access token and API server keys may be copied into multiple workspaces without having to re-run command line.

Figure 22.

Figure 23.

- In the workspace where BaseSpace reads will be imported, open a cloud environment in Terra by clicking on the **environment configuration** (Fig 22)
- Click on the **Jupyter settings** button (Fig 23)
- Set environment configurations** as follows (Fig 24):
  - Application configuration: **Default: (GATK 4.3.3.0, Python 3.10.11, R 4.3.1)**
    - The version numbers may vary slightly as updates are incorporated into the default application configuration*
  - Cloud compute profile: CPUs **2**, Memory **7.5**
  - Persistent disk: Disk Type **Standard**, Disk Size (GB) **250**
  - Click **create**
- Click on the open terminal button** located in the right sidebar (Fig 25)
- A message will appear that says “Creating cloud environment. You can navigate away and return in 3-5 minutes.”



Jupyter Cloud Environment

A cloud environment consists of application configuration, cloud compute and persistent disk(s).

Running cloud compute cost  
\$0.11 per hr

Paused cloud compute cost  
\$0.01 per hr

Persistent disk cost  
\$10.00 per month

Application configuration

Default: (GATK 4.3.3.0, Python 3.10.11, R 4.3.1)

What's installed on this environment?

Updated: Jul 27, 2023

Version: 2.3.1

Startup script *Optional* Learn more about startup scripts.

URI

Cloud compute profile

CPU

Memory (GB)

7.5

Enable GPUs BETA Learn more about GPU cost and restrictions.

Compute type

Standard VM

Enable autopause Learn more about autopause.

30 minutes of inactivity

Location BETA

us-central1 (Iowa) (default)

Persistent disk

Persistent disks store analysis data. Learn more about persistent disks and where your disk is mounted.

Disk Type

Disk Size (GB)

250

CREATE



# Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

6. Test to see if the terminal is working by typing `ls` and `hit enter`

a. `Verify the output` looks like the following:

```
jupyter@2a9974687b45:~$ ls  
gitignore_global jupyter.log lost+found packages
```

b. *If everything looks okay, proceed with the command line; if not, try closing and re-opening the terminal and re-testing the `ls` command*

7. Enter the following to create a bin: `mkdir ~/bin` and `hit enter`

a. *Ctrl + c and Ctrl + v do not work in this terminal, but right clicking to copy/paste should work*

8. Type the command `ls` and `hit enter` to verify a new bin has been created; you should see "bin" in addition to the following:

```
jupyter@2a9974687b45:~$ ls  
bin gitignore_global jupyter.log lost+found packages
```

9. Enter the command: `wget "https://launch.basespace.illumina.com/CLI/latest/amd64-linux/bs" -O $HOME/bin/bs`

a. *You should get the following returned:*

```
--2023-09-14 22:42:27-- https://launch.basespace.illumina.com/CLI/latest/amd64-linux/bs  
Resolving launch.basespace.illumina.com (launch.basespace.illumina.com) ... 18.160.200.113, 18.160.200.107,  
18.160.200.47, ...  
Connecting to launch.basespace.illumina.com (launch.basespace.illumina.com)|18.160.200.113|:443... connect  
ed.  
HTTP request sent, awaiting response... 200 OK  
Length: 12566528 (12M) [binary/octet-stream]  
Saving to: '/home/jupyter/bin/bs'  
  
/home/jupyter/bin/bs      100%[=====] 11.98M --.-KB/s    in 0.1s  
  
2023-09-14 22:42:28 (112 MB/s) - '/home/jupyter/bin/bs' saved [12566528/12566528]
```

10. Enter the command: `chmod u+x $HOME/bin/bs`

a. *Nothing will be returned for this command*

11. Enter the command: `source ~/.profile`

a. *The following will be returned:* `bash: /home/jupyter/.profile: No such file or directory`

12. Enter the command: `~/bin/bs auth`

a. *The following will be output:*

Please go to this URL to authenticate: <https://basespace.illumina.com/oauth/device?code=...>

b. `Click the link and login` to the BaseSpace account to link to Terra

i. *If the BaseSpace account is already logged in, the login process may be skipped*

c. `Accept the permission request` from Terra in BaseSpace (Fig 26)

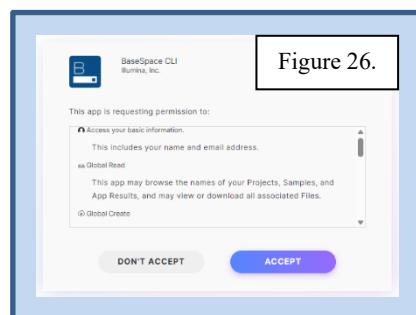


Figure 26.

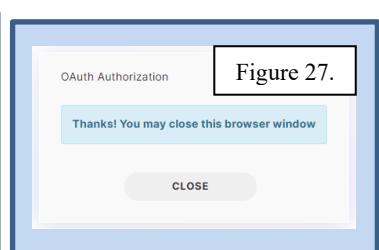


Figure 27.



# Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

13. An authorization message will be received from BaseSpace (Figure 27)
14. Return to the terminal window; a welcome message should be returned for the authorized user similar to the following: Welcome, Kelsey Kropp
15. Verify the basespace access has been added by entering the command `ls -a`
  - a. The following should be returned – notably, the .basespace bin

```
.          bin      .entrypoint.out    .ipython    .local      .welder.log
..          .cache    .gitconfig       .jupyter   lost+found  .wget-hsts
.basespace .config  gitignore global  jupyter.log packages
```
16. Enter the following command: `cat ~/.basespace/default.cfg`
  - a. The following should be returned:

```
apiServer = https://api.basespace.illumina.com
accessToken = XXXXXXXXXXXXXXXXXXXXXXXXX
```
17. Copy the apiServer key <https://api.basespace.illumina.com>; copy by highlighting, right click, copy
18. Open a new Terra window for the relevant workspace
19. Scroll to the bottom of the left sidebar and click on workspace data (Figure 28)

Figure 28.

20. Add a workspace data key by clicking Edit

and Add variable (Fig 29)

21. Specify the following (Fig 30):

- a. Key: `api_server`
- b. Value: paste the apiServer key obtained above in step 16
- c. Description: optional
- d. Click the blue checkmark to save (Fig 30)

Figure 29.

Figure 30.



## Linking BaseSpace and Importing BaseSpace Reads to Terra Workspace

Document TG-TER-04, Version 4

Date:

1/21/2026

Written for Workflow Version:

PHB v4

22. In the terminal window `copy the accessToken` value; **don't include spaces**
23. Return to the Terra workspace and `add a new workspace data key` by clicking *Edit* and *Add variable* (Figure 29)
24. Specify the following, then `click the blue checkmark to save` (Figure 30):
  - a. Key: `access_token`
  - b. Value: `paste the accessToken` value copied from the terminal window obtained in step 16
  - c. Description: optional
25. Continue with [section 4.1](#) to download and import the BaseSpace sample sheet and [section 4.2](#) to run the BaseSpace\_Fetch\_PHB workflow to import sample reads