

Getting Started with Partek Flow

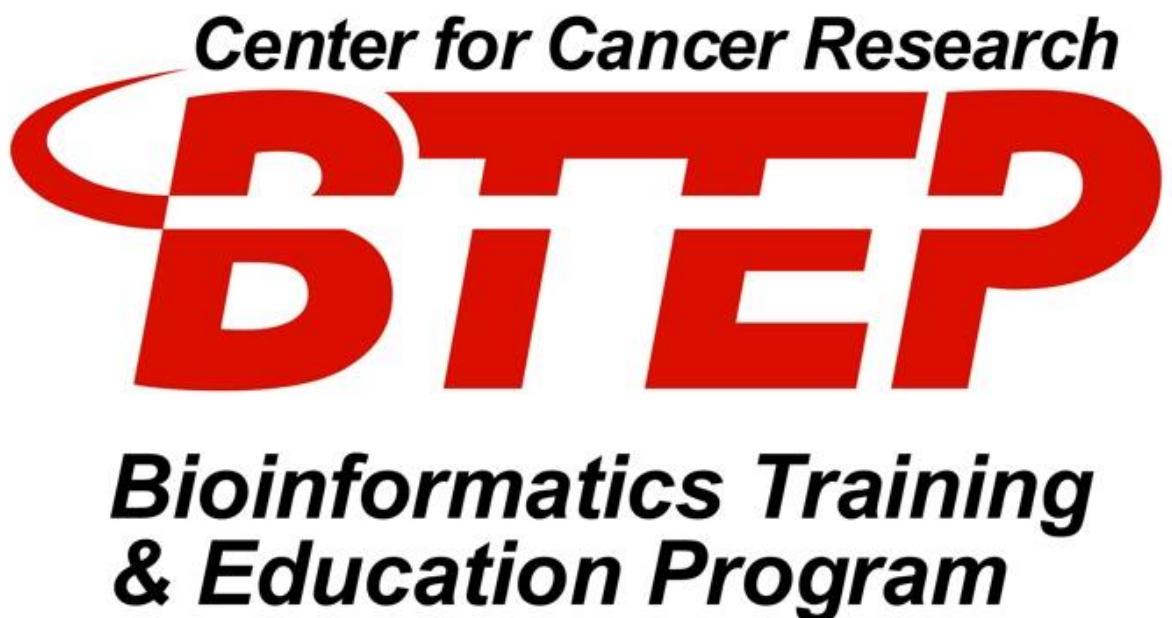


Table of Contents

Class overview

● Getting Started with Partek Flow at NIH	4
● Learning Objectives	4
● What is Partek Flow?	4
● Instructions for Obtaining Access to Partek Flow	4
● Overview of the Partek Flow User Interface	5

Class overview slides

Tips on Biowulf for Partek Flow users

● Tips on Biowulf for Partek Flow Users	13
● Signing onto Biowulf	13
● Changing into the User's Data Folder	13

Transfer data to Partek Flow server using Globus

● Transferring Data to the NIH Partek Flow Server Using Globus	15
● Globus	15
● Step 1: Logging into Globus	15
● Step 2: Setting up a Globus Endpoint to the Partek Flow Server	20
● Step 3: Downloading Data from the NCI CCR SF DME	26

Transfer data to Partek Flow server using command line

- Transferring Data to the NIH Partek Flow Server Using Command Line 30
- Copy from User's Biowulf data Folder to Partek Flow uploads Folder 30
- scp from Personal Computer to the Partek Flow uploads Folder. 31

Transfer data to Partek Flow server using Flow web GUI

- Transferring Data to the NIH Partek Flow Server Using the Web Tool 33
- Using the Partek Flow web tool to transfer data 33

Import data to Partek Flow projects

- Importing Data to the Partek Flow Projects 42
- Data import into Partek Flow projects 42

Partek Flow Bioinformatics Biowulf Globus Data transfer from NCI CCR Sequencing Facility

As far as the documentation goes, it should say that the licenses are NCI only. People can send a message if they want to ask beyond that.

Getting Started with Partek Flow at NIH

Learning Objectives

After this class, participants will

- Know how to acquire access to Partek Flow
- Become familiar with approaches for transferring data to the NIH Partek Flow server such as Globus
- Be able to import data into a Partek Flow project

What is Partek Flow?

Partek Flow is a point-and-click software and is suitable for those who wish to avoid the steep learning curve associated with analyzing sequencing data through command line and/or code. It enables the analysis of high dimensional multi-omics sequencing data including DNA, RNA, single cell RNA, ATAC/ChIP, and spatial transcriptomics. At NIH, Partek Flow is hosted on the Biowulf High Performance Computing (HPC) cluster. Researchers interact with Partek Flow through a web browser using a URL supplied by Biowulf after a Biowulf and Partek Flow account has been established. This enables investigators to take advantage of the compute power offered by HPC while using a graphical user interface to construct a sequencing data analysis workflow. Partek Flow also enables the creation of publication quality visualizations.

Note

NCI holds a license for Partek Flow.

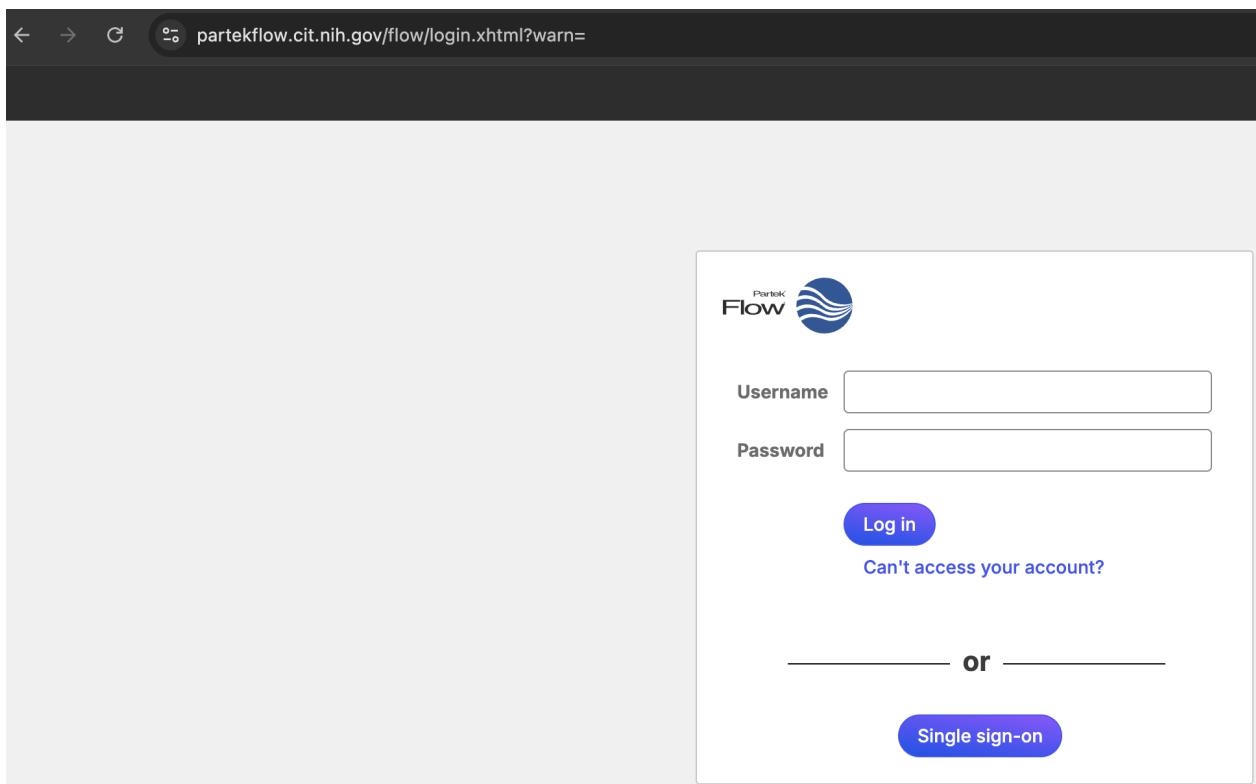
Instructions for Obtaining Access to Partek Flow

NCI researchers can find instructions for accessing Partek Flow at <https://bioinformatics.ccr.cancer.gov/btep/partek-flow-bulk-and-single-cell-rna-seq-data-analysis/> (<https://bioinformatics.ccr.cancer.gov/btep/partek-flow-bulk-and-single-cell-rna-seq-data-analysis/>). Here is what is needed:

- Biowulf (The High Performance Computing cluster) account — [see here](#) for information about how to obtain a HPC account (<https://hpc.nih.gov/docs/accounts.html>).

- /data directory on Biowulf with enough disk space to hold their Partek Flow files — please fill out [this online form](https://hpc.nih.gov/dashboard/storage_request.php) (https://hpc.nih.gov/dashboard/storage_request.php) if you require more disk space.
- Partek Flow account — please contact staff@hpc.nih.gov.

Once these steps have been accomplished, use <https://partekflow.cit.nih.gov/flow> (<https://partekflow.cit.nih.gov/flow>) to sign onto the NIH Partek Flow server. NIH single sign-in enables users to authenticate and connect to the Partek Flow server. Alternatively, users can log on by supplying username and password. However, the password may be different from the one in which the user signs in to Biowulf with as HPC staff allow for the researcher to set up something different upon Partek Flow account creation. The username though, is the user's NIH username.



Overview of the Partek Flow User Interface

Upon logging in, users will be taken to Partek Flow's main page. In this page users can access

- Help
- View a list of existing projects
- Create new project
- Transfer files (this does not use Globus)
- Access user profile

The screenshot shows the Partek Flow Home page. At the top, there are navigation links: Queue, Projects, and Help. On the right, there is a user profile icon and a search bar. Below the navigation, there are two buttons: '+ New project' and 'Transfer files'. A search bar with the placeholder 'Search project names and descriptions...' is followed by a magnifying glass icon. The main area displays a table of projects. The columns are: Project name (with a sort arrow), Owner (with a sort arrow), Your role (with a sort arrow), Last modified (with a sort arrow), Size (with a sort arrow), and Actions. A single project row is shown: 'project1' by 'username' (Project owner), last modified '27 Jun 2024, 12:56 PM', size '453.43 MB', and an 'Actions' button. Below the table, a message says 'Projects will be listed here'. A dropdown menu for 'Optional columns' is open, showing 'Project name ↑↓', 'Owner ↑↓', 'Your role ↑↓', 'Last modified ↑↓', 'Size ↑↓', and 'Actions'. The 'Last modified ↑↓' option is highlighted with a red box. A 'Rows per page' dropdown is set to '10'. A search bar with the placeholder 'Search by -- Select --' and a dropdown menu showing 'No criteria applied' are also visible on the right.

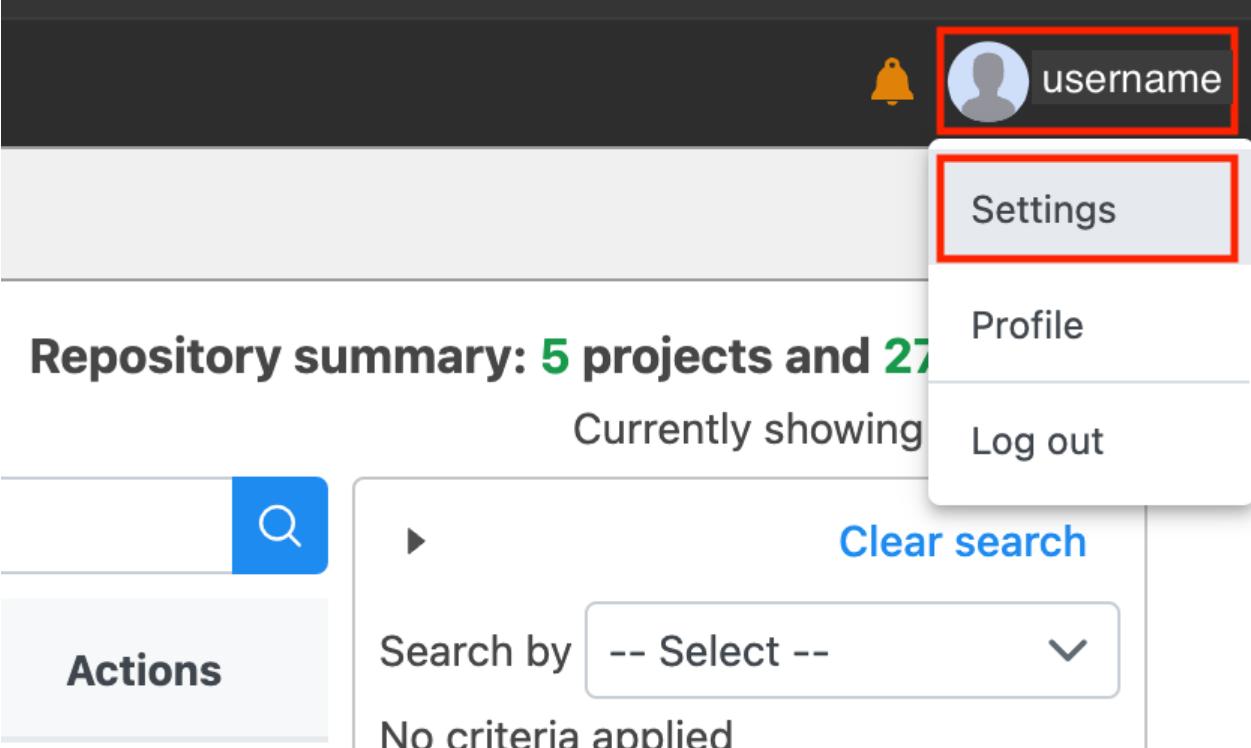
Users can access help documentation, how-to videos and information for contacting support under the "Help" menu. The email for Partek support is support@partek.com.

The screenshot shows the Partek Flow Home page with the 'Help' menu open. The 'Help' menu is a dropdown with the following options: 'How-to videos', 'FAQ', 'Documentation', 'Release notes', 'Contact us', and 'About'. The 'Documentation' option is highlighted with a red box. The rest of the page is identical to the first screenshot, showing the project list and search interface.

Where applicable, users can find links to how-to video and documentation for specific tasks, such as importing data.

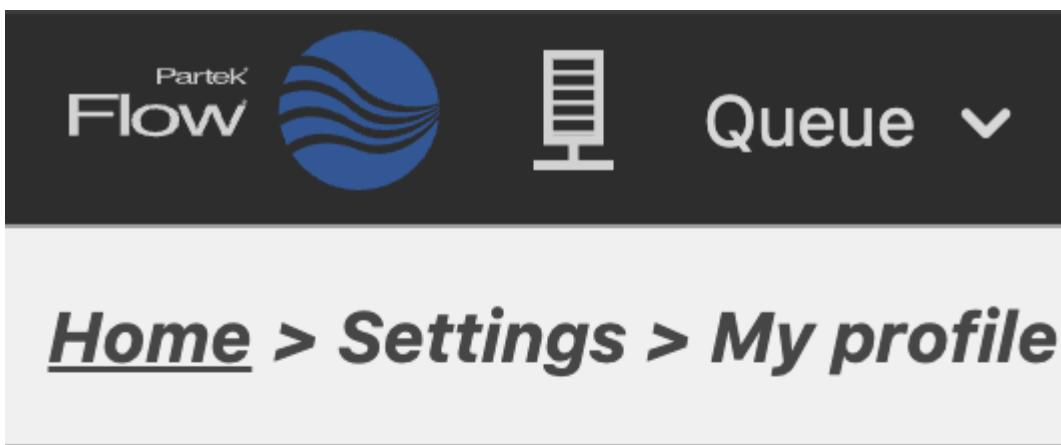
The screenshot shows the Partek Flow project settings page for 'project1a'. The top navigation bar includes 'Queue', 'Projects', and 'Help'. The main content area shows the project owner information: 'Home > project1a (Project owner)'. Below this, there are tabs: 'Analyses' (highlighted with a blue underline), 'Metadata', 'Log', 'Project settings', 'Notebook', 'Data viewer', and 'Attachments'. A large callout box on the left contains the following text: 'Welcome to your project.' with a blue circular icon containing a white plus sign. Below it, the text says: 'You'll need samples before you can run an analysis. Click the blue button on the left then choose your assay and file type.' and 'For more help please watch this [video](#) or see our [documentation](#)'. The 'video' and 'documentation' links are highlighted with red boxes.

The settings section is divided into five tabs, which are expandable to reveal more options.



The screenshot shows the Partek Flow interface. At the top right, there is a user profile icon with a bell notification and the text "username". A dropdown menu is open from this icon, containing "Settings" (which is highlighted with a red box), "Profile", and "Log out". Below the profile area, the text "Repository summary: 5 projects and 27" is visible. The main content area shows a search bar with a magnifying glass icon, the text "Currently showing", and a "Clear search" button. A dropdown menu for "Search by" is open, showing the option "-- Select --". The text "No criteria applied" is displayed below the search controls. On the left, there is a sidebar with a "Actions" tab.

Click on the user profile tab at the top right to reveal the option for changing settings.



The screenshot shows the Partek Flow software interface. At the top left is the "Partek Flow" logo with a blue circular icon. To its right is a "Queue" button with a downward arrow. Below the logo is a breadcrumb path: **Home > Settings > *My profile***.

- > Personal**
- > System**
- > Components**
- > Access**
- > Usage**

User profiles can be viewed and edited under "My Profile" in the "Personal" settings section. Note the information regarding the user's Biowulf /data folder storage usage and quota.

Display name username

Username username

Email first.last@nih.gov

Password

Disk usage 103.16 GB (3% of 3.00 TB quota)

Users can access an example bulk and single cell RNA sequencing dataset under the "System" setting tab.

An important feature in the settings is that users can manage genomic references and annotations by clicking on "Library files" under "Components".

Personal

- My profile
- My preferences

System

- System information

Components

Library files

- Lists
- Pipelines
- Filters
- Option sets
- Installed tasks
- Data repository

Access

Usage

- System queue
- System resources

Genomic library files

Assembly Aegilops tauschii - Aegilops tauschii

Reference files *i*

Add reference files

Library file	Owner	Ignore	Actions
Reference sequence	owner	<input type="checkbox"/>	⋮

Reference aligner indexes *i*

Add reference aligner indexes

Library file	Owner	Ignore	Actions
Bowtie index	owner	<input type="checkbox"/>	⋮
Bowtie 2 index	owner	<input type="checkbox"/>	⋮
HISAT2 2.1.0 index	owner	<input type="checkbox"/>	⋮

Users can view "Systems resources" under the "Usage" settings. Here, users are informed of how many tasks are running and waiting in the queue. A visualization of Biowulf Partek Flow worker nodes is also provided.

The screenshot shows the 'System resources' page of the Partek Flow web interface. On the left, a sidebar lists 'Personal', 'System', 'Components', 'Access', 'Usage' (which is expanded), 'System queue', and 'System resources' (which is selected and highlighted in blue). The main content area has three sections: 'Queue status' (with 'Running tasks' 1 and 'Waiting tasks' 0, and a note that estimates are continuously updated), 'Licensing' (showing 48 out of 50 available worker licenses), and 'Active workers' (a table listing three workers: 'hpcpartek.cit.nih.gov' (Server), 'cn2287' (Worker), and 'cn3168' (Worker)). Each worker row includes a 'CPU usage' and 'Memory usage' chart, and a 'Machine capacity' and 'Uptime' column. Below the table is a legend: a blue square for 'Worker usage', a yellow square for 'System usage', and a grey square with a triangle for 'Worker limit'.

Name	CPU usage	Memory usage	Machine capacity	Uptime	Type
hpcpartek.cit.nih.gov			12 cores 63 GB RAM		Server
cn2287			56 cores 252 GB RAM	1 day	Worker
cn3168			56 cores 252 GB RAM	07:55:54	Worker

Legend: Worker usage (blue), System usage (yellow), Worker limit (grey with triangle)

Class overview slides

|

Tips on Biowulf for Partek Flow Users

Even though Partek Flow runs on Biowulf, the NIH Unix-based high performance computing cluster, users do not need to be experts in command line to use this software. However, there are somethings that are useful to know because Partek Flow projects, input, and output are all stored on Biowulf.

Signing onto Biowulf

To sign onto Biowulf, open a Terminal if using MacOS or Command Prompt if using Windows 10 or above. Once opened, at the prompt, do the following to sign on. The breakdown of the command is as follows.

- `ssh`: Connect to remote computer (ie. Biowulf)
- `user`: User's NIH username
- `@biowulf.nih.gov`: The remote computer to connect to (ie. NIH Biowulf)

```
ssh user@biowulf.nih.gov
```

Users will then be asked to enter a password for signing onto Biowulf. This is the password that the user utilizes to sign onto NIH applications. The password will not appear as the user is typing but keep typing and hit enter when done.

```
user@biowulf.nih.gov's password:
```

Changing into the User's Data Folder

Upon signing onto Biowulf, users will land in the `/home/user` directory (again replace `user` with user's NIH username). This directory only has 16 GB of storage space and cannot be increased, thus users will need to perform analyses in the `/data/user` folder. Further, a folder called `/PartekFlow` in the `/data/user` directory is created when the Partek Flow account is activated for the user.

To check which folder on Biowulf a user is currently in, use the `pwd` command.

Upon signing onto Biowulf:

```
pwd
```

```
/home/user
```

To change into the user's `data` folder do the following.

```
cd /data/user
```

Biowulf staff will create a folder called `PartekFlow` in the user's `/data/user` directory upon activating a Partek Flow account. This folder contains all Partek Flow projects, inputs, outputs, and data Uploads.

List the contents of the `/data/user` folder.

```
ls /data/user
```

```
PartekFlow
```

Change into the `PartekFlow` folder.

```
cd /data/user/PartekFlow
```

The `PartekFlow` folder houses project folders (only if users have projects), which contain input and output as well a folder called `uploads` that stores data that has been loaded to the NIH Partek Flow server.

```
Project_RNAsequencing_1234
uploads
```

Note

Most likely, Partek Flow users will not have to use the command line to do anything Partek Flow related as tasks such as creating projects, deleting projects, uploading files, etc. can be done in the software or using Globus.

Transferring Data to the NIH Partek Flow Server Using Globus

Globus

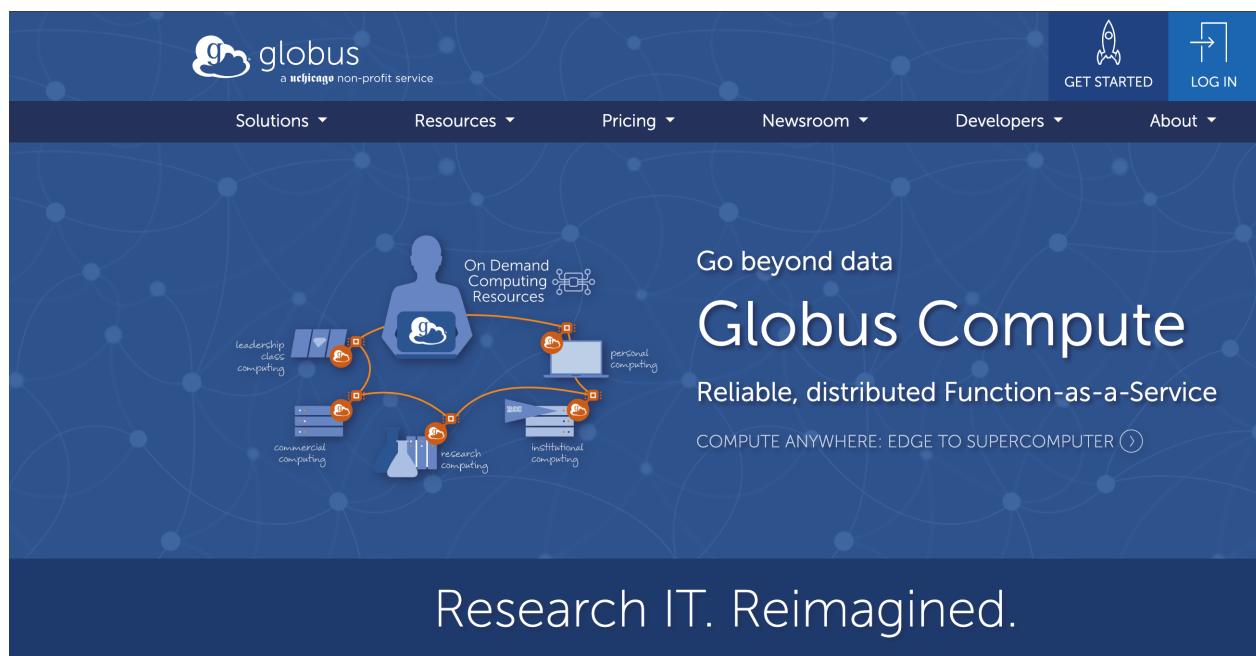
NCI CCR researchers will likely use the NCI CCR Sequencing Facility (or SF) (https://bioinformatics.ccr.cancer.gov/docs/resources-for-bioinformatics/raw_data_from_cores/) for sequencing projects. Data generated from the NCI CCR SF will be stored within its Data Management Environment (DME). Researchers can use Globus to transfer data from NCI CCR SF DME to the NIH Partek Flow server and the steps for accomplishing this are described below.

Note

The staff at Biowulf has created detailed documents for Globus, which can be found at <https://hpc.nih.gov/docs/globus/setup.php> (<https://hpc.nih.gov/docs/globus/setup.php>).

Step 1: Logging into Globus

Goto <https://www.globus.org> (<https://www.globus.org>) to log in by clicking on the "LOG IN" icon at the top right corner of the page.



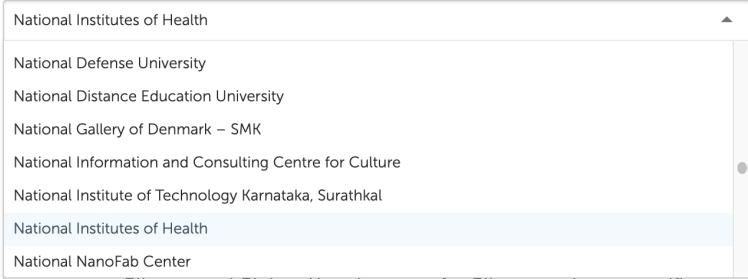
In the next page, select organizational affiliation from the drop down menu (in this example it is National Institutes of Health).



Log in to use Globus Web App

Use your existing organizational login

e.g., university, national lab, facility, project



A screenshot of a dropdown menu titled 'Use your existing organizational login'. The menu lists various institutions, with 'National Institutes of Health' highlighted. Other listed institutions include National Defense University, National Distance Education University, National Gallery of Denmark – SMK, National Information and Consulting Centre for Culture, National Institute of Technology Karnataka, Surathkal, National Institutes of Health, and National NanoFab Center.

Click on "Continue" when the organizational affiliation has been selected.



Log in to use Globus Web App

Use your existing organizational login

e.g., university, national lab, facility, project



A screenshot of the same dropdown menu as the previous image, but with the selection 'National Institutes of Health' highlighted by a red box.

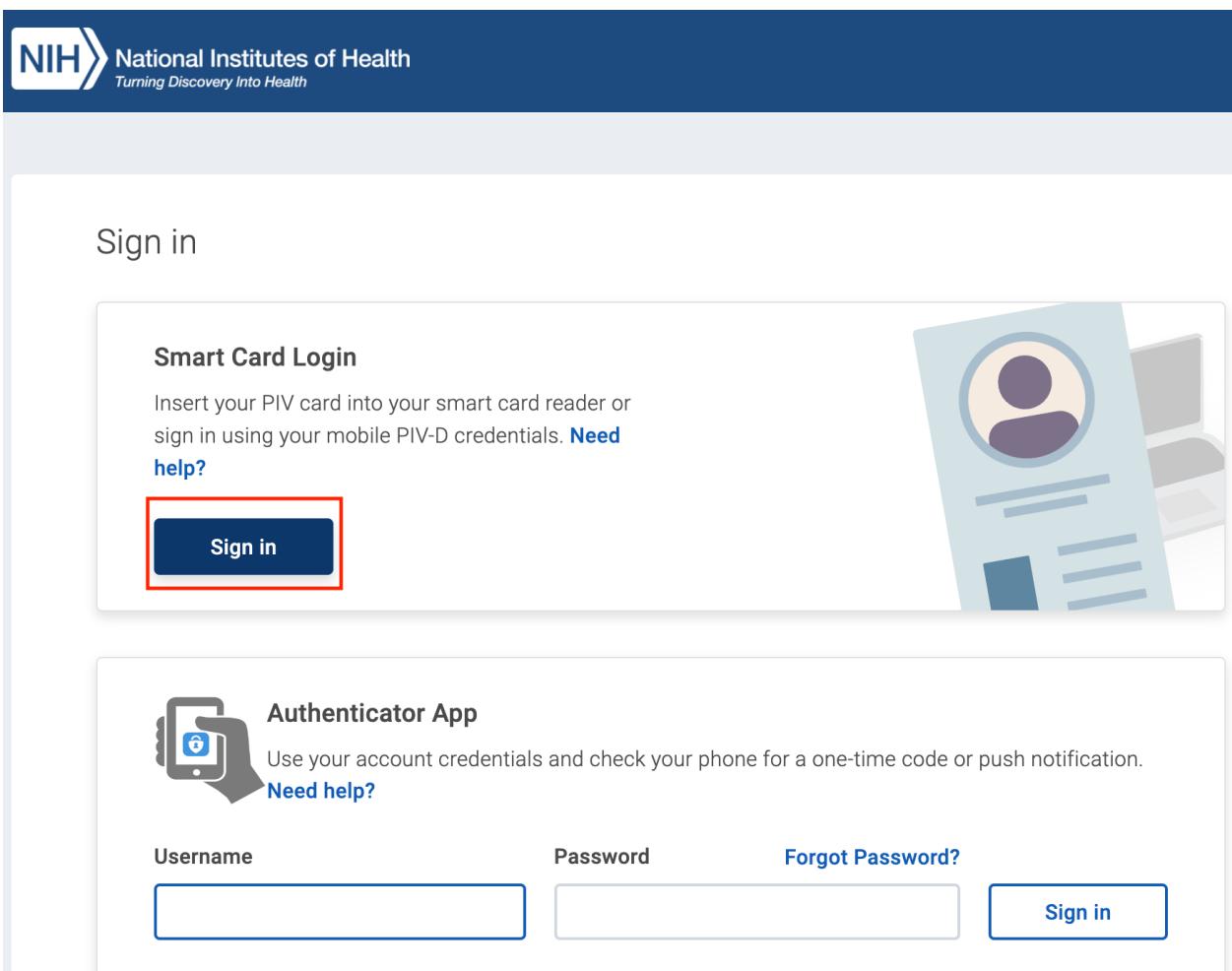
By selecting Continue, you agree to Globus [terms of service](#) and [privacy policy](#).

→ **Continue**



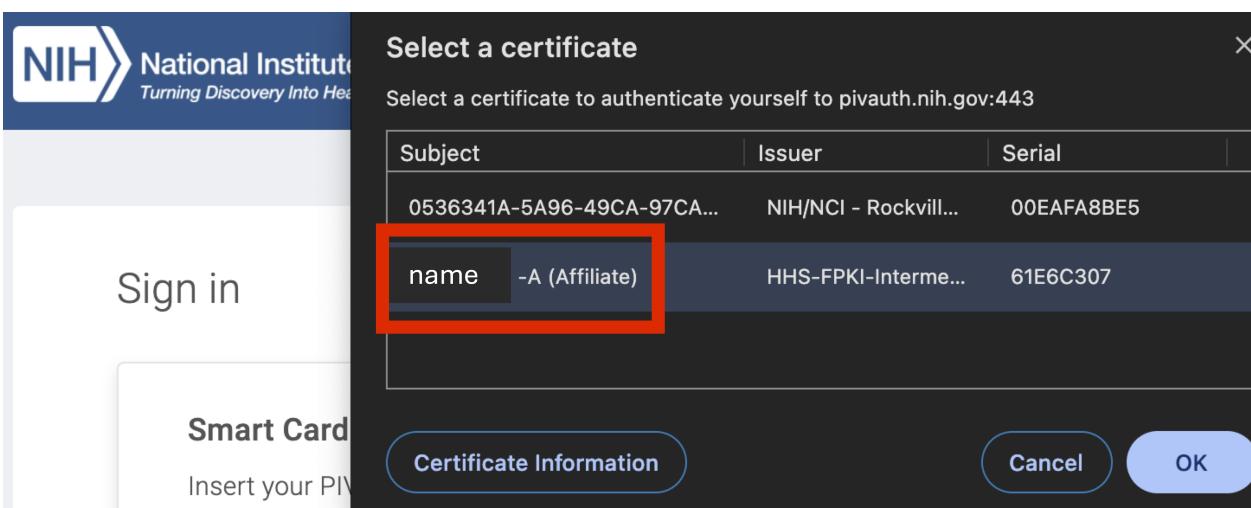
Globus uses CILogon to enable you to Log In from this organization. By clicking Continue, you agree to the [CILogon privacy policy](#) and you agree to share your username, email address, and affiliation with CILogon and Globus. You also agree for CILogon to issue a certificate that allows Globus to act on your behalf.

Subsequently, users will be taken to the NIH authentication page. Click on "Sign in" to authenticate using PIV card.



The image shows two screenshots of the NIH login interface. The top screenshot is titled 'Smart Card Login' and shows a button labeled 'Sign in' with a red box around it. To the right is an icon of a card with a person's profile. The bottom screenshot is titled 'Authenticator App' and shows fields for 'Username' and 'Password' with a 'Forgot Password?' link. It also has a 'Sign in' button and an icon of a smartphone with a lock. Both screenshots are set against a light gray background with the NIH logo at the top.

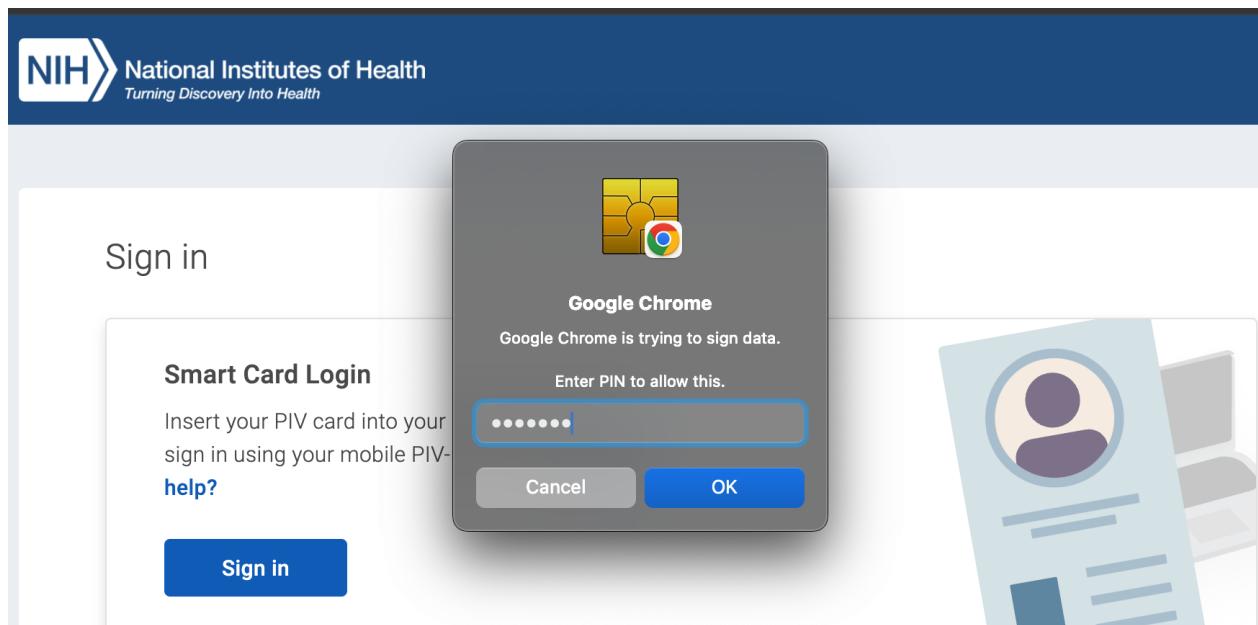
Next, select the appropriate PIV card certificate (usually the one with the user's name followed by "- A (Affiliate)") and a pop-up will appear to take the user's NIH pin.



The image shows a 'Select a certificate' pop-up window. It has a title 'Select a certificate' and a sub-instruction 'Select a certificate to authenticate yourself to pivauth.nih.gov:443'. The window lists certificates in a table format:

Subject	Issuer	Serial
0536341A-5A96-49CA-97CA...	NIH/NCI - Rockvill...	00EAFA8BE5
name -A (Affiliate)	HHS-FPKI-Interme...	61E6C307

The 'name -A (Affiliate)' row is highlighted with a red box. At the bottom of the window are buttons for 'Certificate Information', 'Cancel', and 'OK'.



In the next page, scroll to the bottom and click "I Agree".



Here's the information to be released

First name: First

Last name: Last

E-mail address: First.Last@nih.gov

NetID: userid@nih.gov

The privacy policy of the service you're connecting to details things like why the service requires this information, how long the information will be retained, who the information will be shared with, etc. In general, the information is needed to facilitate your access, it will not be shared, and it will be retained for a limited time. You can review the service's privacy policy here:

<https://www.cilogon.org/privacy>



Do you agree to release the information listed above to this service?

Please select your sharing preference from the options below and click on the I Agree button.

Ask me again at next login

- I agree to send my information this time, but I want to be prompted again the next time I access a service that requests this information.

Ask me again if information to be provided to this service changes

- I agree to send the information listed above to this service now and in the future but I want to be prompted if the information the service requires changes.

Send now and in the future - Do not ask me again

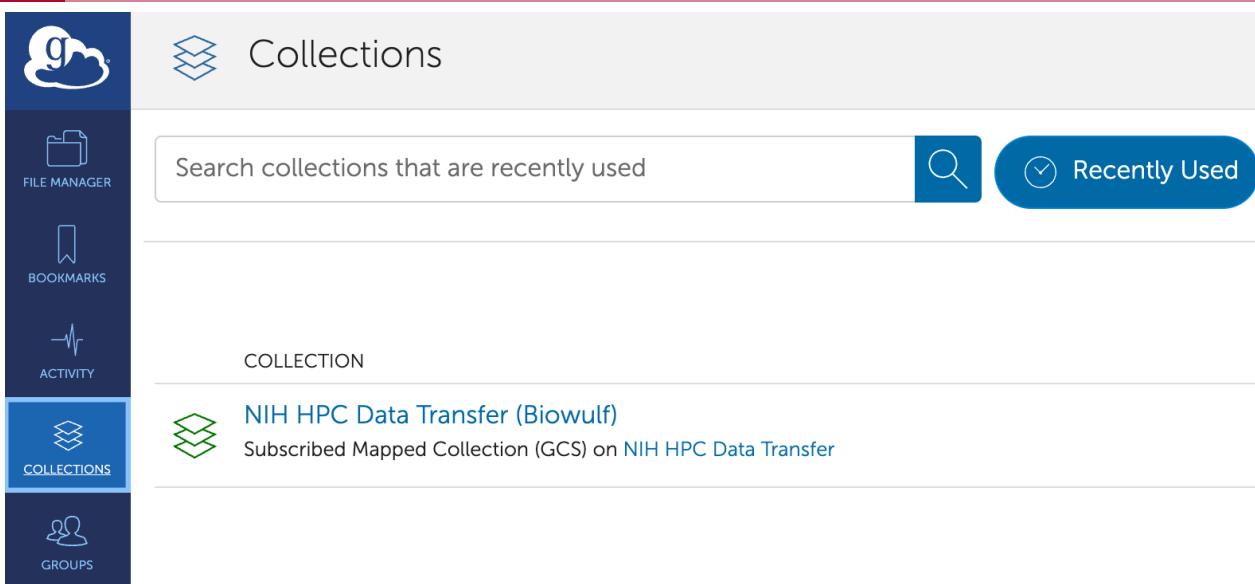
- I agree to release the information listed above to this service and to any service that asks for the same information. I want to be prompted if a service asks for different information.

I Do Not Agree

I Agree



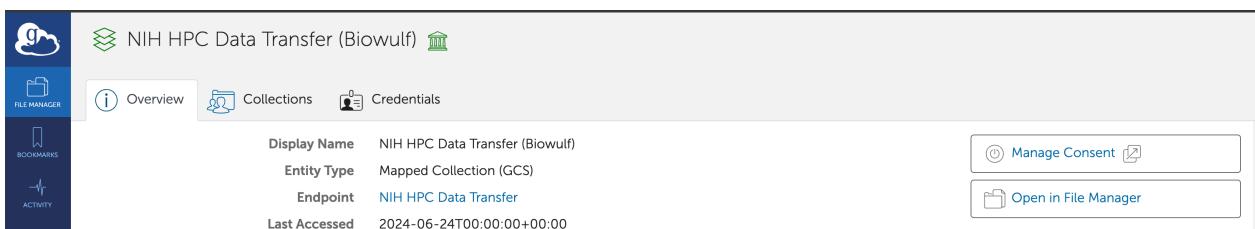
Users will then be brought to the Globus interface where file transfers are managed. Click on "COLLECTIONS" to see the recently used Globus data transfer endpoints here. The "NIH HPC Data Transfer (Biowulf)" (<https://hpc.nih.gov/docs/globus/setup.php#endpoints>) endpoint points to the content available on the cluster.



The screenshot shows the Globus Collections interface. On the left is a sidebar with icons for FILE MANAGER, BOOKMARKS, ACTIVITY, COLLECTIONS (which is selected and highlighted in blue), and GROUPS. The main area is titled 'Collections' and contains a search bar with the placeholder 'Search collections that are recently used' and a 'Recently Used' button. Below the search bar is a section titled 'COLLECTION' with a single entry: 'NIH HPC Data Transfer (Biowulf)'. The entry includes the description 'Subscribed Mapped Collection (GCS) on NIH HPC Data Transfer'.

Step 2: Setting up a Globus Endpoint to the Partek Flow Server

Click on the "NIH HPC Data Transfer (Biowulf)" endpoint and then "Open in File Manager". This will take users to their Biowulf /home directory. Username is the Biowulf user name for the specific user.



The screenshot shows the details page for the 'NIH HPC Data Transfer (Biowulf)' endpoint. The top navigation bar includes 'Overview', 'Collections' (selected), and 'Credentials'. Below the navigation are the following details:

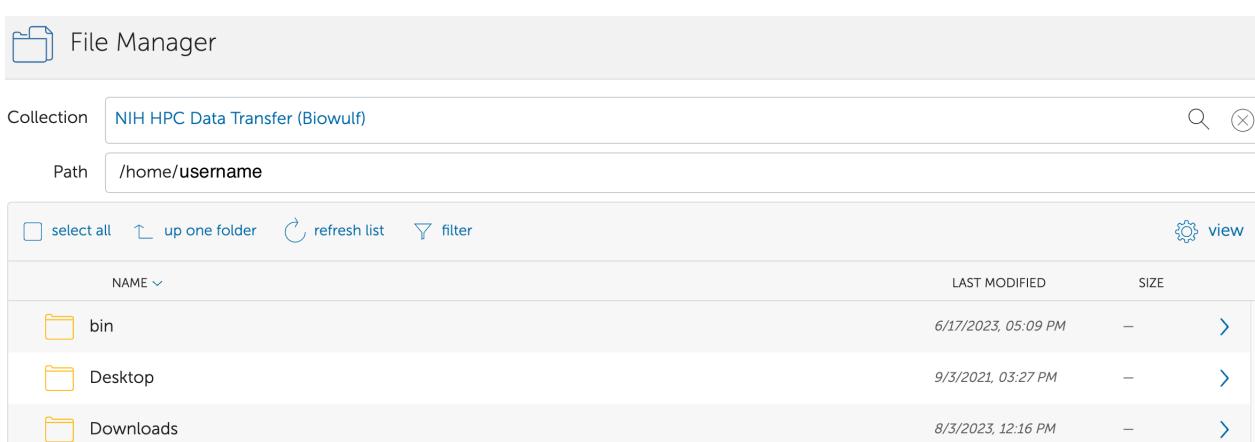
- Display Name: NIH HPC Data Transfer (Biowulf)
- Entity Type: Mapped Collection (GCS)
- Endpoint: NIH HPC Data Transfer
- Last Accessed: 2024-06-24T00:00:00+00:00

On the right side, there are two buttons: 'Manage Consent' and 'Open in File Manager'.

Recall

The user's Biowulf /home directory is not suitable for analyzing data. To conduct analysis, use the /data directory.

To goto the user's /data/directory, replace /home/username in the box labeled "Path" with /data/username.

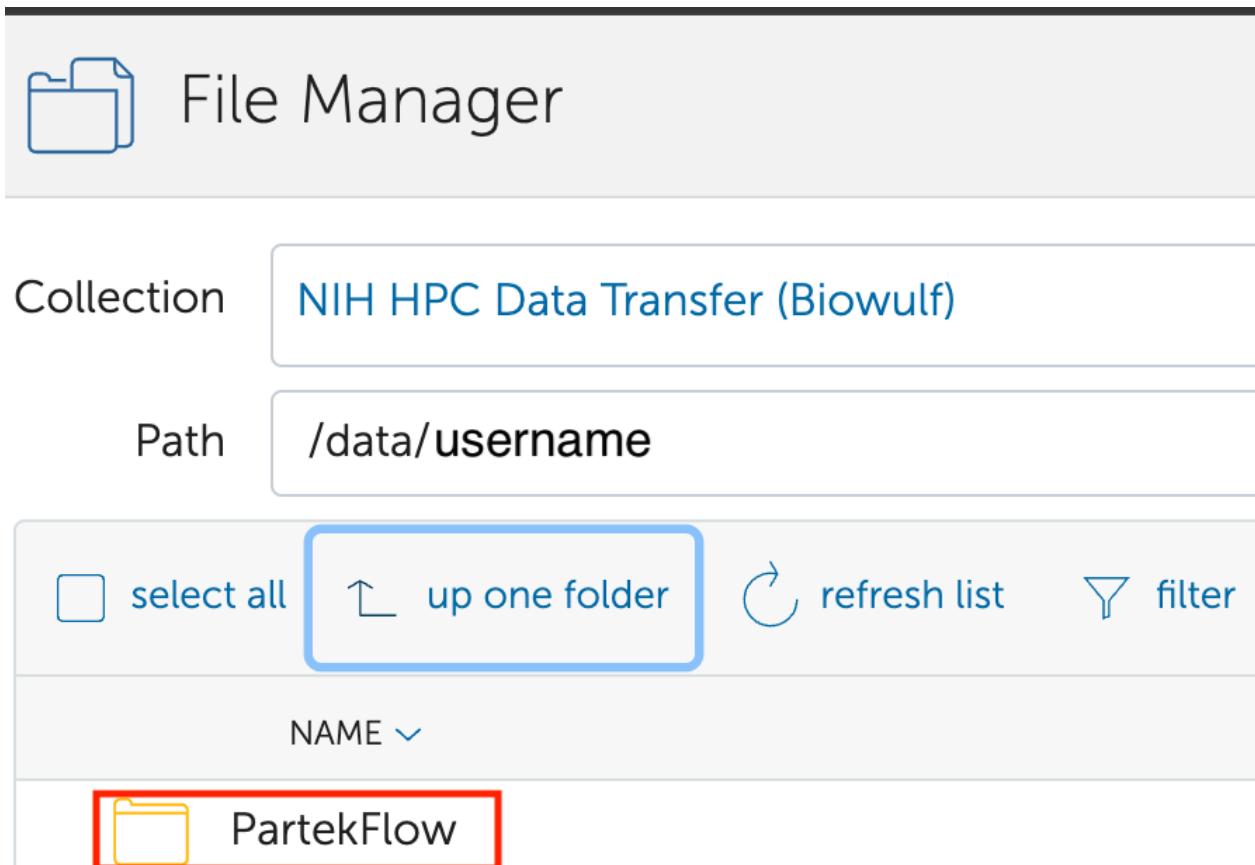


The screenshot shows the Globus File Manager interface. The top navigation bar includes 'File Manager' and a search bar. Below the navigation, the 'Collection' dropdown is set to 'NIH HPC Data Transfer (Biowulf)' and the 'Path' dropdown is set to '/home/username'. The main area shows a list of files and folders in the '/home/username' directory:

NAME	LAST MODIFIED	SIZE
bin	6/17/2023, 05:09 PM	-
Desktop	9/3/2021, 03:27 PM	-
Downloads	8/3/2023, 12:16 PM	-

At the bottom of the interface are buttons for 'select all', 'up one folder', 'refresh list', 'filter', and 'view'.

After switching to the user specific /data directory, find and click into the PartekFlow folder. Recall that the PartekFlow folder will exist only if the user has contacted Biowulf staff about activating a Partek Flow account.



Collection NIH HPC Data Transfer (Biowulf)

Path /data/username

select all up one folder refresh list filter

NAME ▾

PartekFlow

Next, click on "New Folder" to make a folder called "globus" to store data uploaded to the Partek Flow server via Globus.



Collection NIH HPC Data Transfer (Biowulf)

Path /data/username/PartekFlow

select all up one folder refresh list filter view ⋮

NAME ▾

NAME	LAST MODIFIED	SIZE	⋮
globus	6/24/2024, 07:54 PM	–	Share Transfer or Sync to... New Folder

Then, click into the globus folder and create one named example_data_transfer.

The screenshot shows the Globus File Manager interface. At the top, it displays the collection 'NIH HPC Data Transfer (Biowulf)' and the path '/data/username/PartekFlow/globus'. A folder named 'example_data_transfer' is selected. In the context menu on the right, the 'New Folder' option is highlighted with a red box.

Go back to the "NIH HPC Data Transfer (Biowulf)" endpoint and click on the "COLLECTIONS" tab after the `example_data_transfer` folder has been created. From there, click on "Add Guest Collection".

The screenshot shows the NIH HPC Data Transfer (Biowulf) endpoint. The 'Collections' tab is selected. The 'Add Guest Collection' button is highlighted with a red box.

In the "Add Guest Collection" menu, click "Browse" to select the folder in which the endpoint will reference (ie. `/data/username/PartekFlow/globus/example_data_transfer`). Enter a display name and description for the endpoint and then "Create Collection" when ready.

Create new Guest Collection

You are creating a guest collection on NIH HPC Data Transfer (Biowulf) to share data

Directory:

This folder is empty.

Display Name: example data transfer

Description: Endpoint for transferring data from the NCI CCR Sequencing Facility Data Management Environment to the NIH Partek Flow server.

Keywords:

Force encryption on transfers to and from this collection

[view more fields](#)

The user will then be taken to a page for setting up sharing between the endpoint and the location where the data is (ie. NCI CCR Sequencing Facility DME).

example data transfer

Overview Permissions Roles

Shared With

Mapped Collection: NIH HPC Data Transfer (Biowulf)
Base Path: /data/wuz8/PartekFlow/globus/example_data_transfer/

USER OR GROUP	CREATED	EXPIRATION	READ	WRITE
Permissions granted by role				

Leave the entry in the box labeled "Path" as "/". Make sure to mark the "Write" permission box because the NCI CCR Sequencing Facility DME has to write the data into this Globus endpoint. Then click "Select a Group" to choose the group in which to share /data/username/PartekFlow/globus/example_data_transfer with.

EXAMPLE DATA TRANSFER

Add Permissions - Share With

Path / [Browse](#)

Share With user - share with specific individuals group - make data accessible to members of a group all users - make data accessible to all logged in users of Globus public (anonymous) - make data accessible to everyone

Group [Select a Group](#)

Permissions read write

[Add Permission](#) [Cancel](#)

In the drop down menu, select "HPCDME-PROD-App-Accts-Pool-FNLCR" to return to the "Add Permissions" page. Hit "Add Permission" when ready.

EXAMPLE DATA TRANSFER

Select a Group

Search by group name or select one below

[HPCDME-PROD-App-Accts-Pool-CCBR](#)

[HPCDME-PROD-App-Accts-Pool-HiTIF](#)
Pool of Globus app accounts for the NIH NCI DOC known as HiTIF

[HPCDME-DEV-App-Accts-Pool-FNLCR](#)
HPC DME Globus Group for FNLCR

[HPCDME-PROD-App-Accts-Pool-FNLCR](#)
Pool of Globus app accounts for the NIH NCI DOC of the Sequencing Facility at FNLCR

Click "Done" to finish the adding permission process.

EXAMPLE DATA TRANSFER

Add Permissions - Share With

Path: /

Share With:

- user - share with specific individuals
- group - make data accessible to members of a group
- all users - make data accessible to all logged in users of Globus
- public (anonymous) - make data accessible to everyone

Group: HPCDME-PROD-App-Accts-Pool-FNLCR

Permissions:

read write

Permissions Added

Read and write permissions were added for HPCDME-PROD-App-Accts-Pool-FNLCR on path: /

The user will then be returned to the "example data transfer" endpoint and see that this endpoint has been shared with "HPCDME-PROD-App-Accts-Pool-FNLCR".

example data transfer

Overview Permissions Roles

Shared With

Mapped Collection: NIH HPC Data Transfer (Biowulf)
Base Path: /data/wuz8/PartekFlow/globus/example_data_transfer/

USER OR GROUP	CREATED	EXPIRATION	READ	WRITE
Permissions granted by role				
Path: /	6/24/2024, 09:34 PM	never expires	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
HPCDME-PROD-App-Accts-Pool-FNLCR				

Click on the "Overview" tab and scroll to the bottom of the page. Take note of the UUID, which tells Sequencing Facility DME where to send data. The UUID for each Globus endpoint will be different.

Step 3: Downloading Data from the NCI CCR SF DME

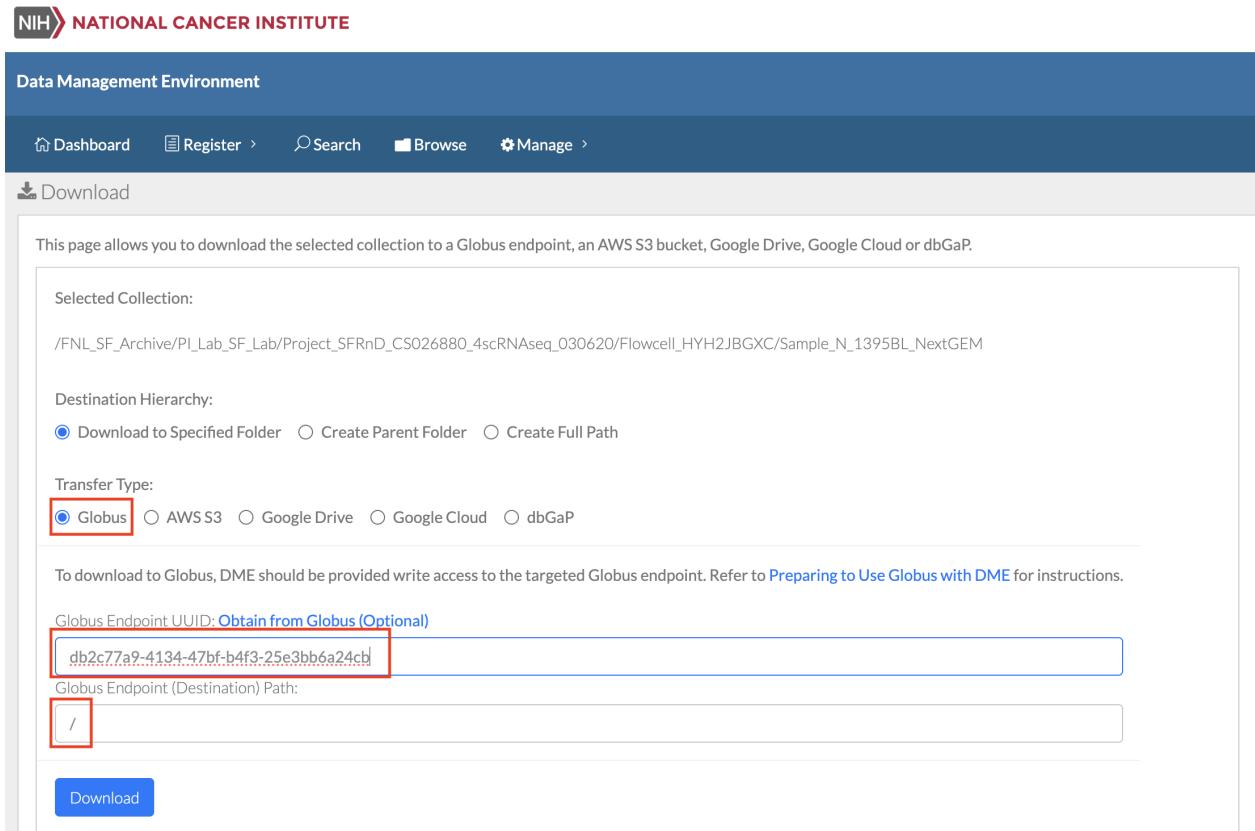
Copy the link to the data provided by the NCI CCR SF and sign in with user specific NIH credentials and the page below will be shown. Users can download all data or browse through their data.

This example will click on the tab for browsing data and download FASTQ files in the folder labeled "Sample_N_1395BL_NextGEM". To download, just click on the "down arrow" corresponding to this folder in the column labeled "Download".



The screenshot shows the NIH Data Management Environment (DME) interface. The top navigation bar includes the NIH logo, 'NATIONAL CANCER INSTITUTE', and the 'Data Management Environment' title. The top right shows the user 'Joe Wu', help link, and version '3.2.1'. Below the header, there are tabs for 'Dashboard', 'Register', 'Search', 'Browse' (which is selected), and 'Manage'. The main content area shows a file tree on the left and a list of files on the right. The file tree includes 'FNL_SF_Archive', 'PI_Lab_SF_Lab', and 'Project_SFRnD_CS026880_4scRNAseq_030620'. The list of files under 'Project_SFRnD_CS026880_4scRNAseq_030620' includes 'Flowcell_HYH2JBGXC' and several sub-folders like 'AggregatedDatasets_aggregate.tar', 'N_1395BL_NextGEM_count.tar', 'T_1395_NextGEM_count.tar', and 'T_N_NextGEM_Hash_GEX_A10_count.tar'. On the right, a table lists files with columns for 'File Name', 'File Size', 'Last Updated', and 'Download'. The 'Download' column for the 'Sample_N_1395BL_NextGEM' file is highlighted with a red box.

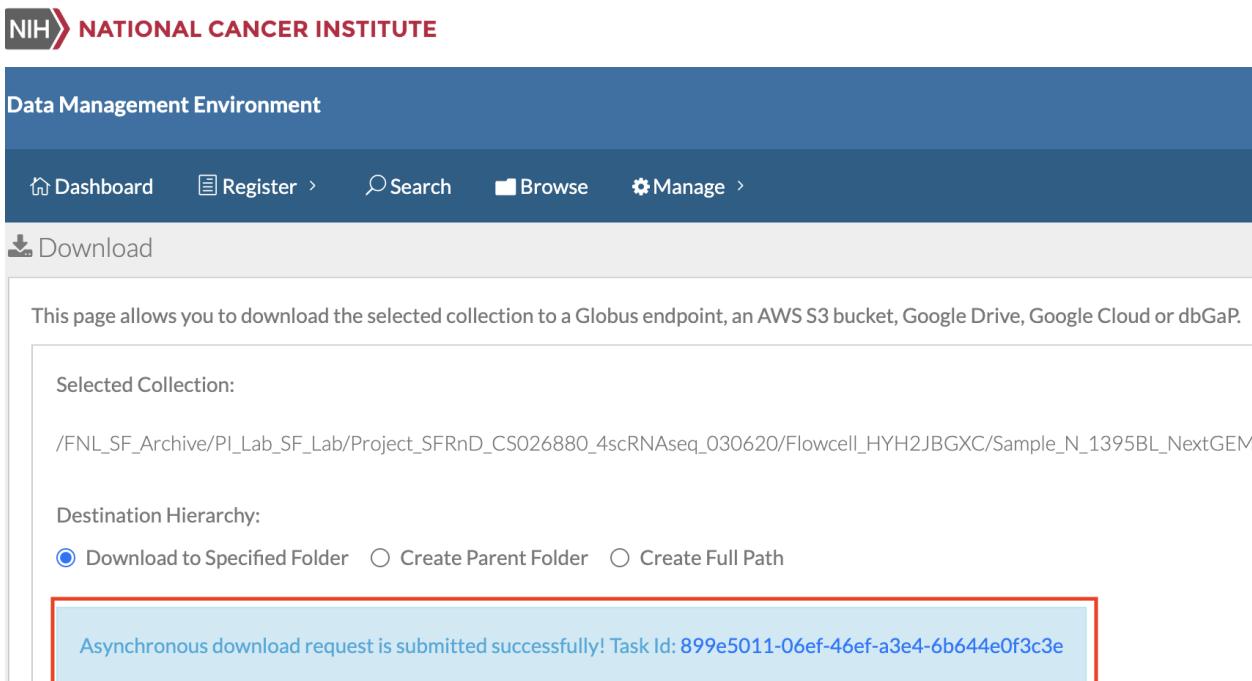
After clicking on the download arrow, users will be taken to the dialogue page shown below. Be sure to select Globus for the "Transfer Type", go back to Globus and copy then paste the endpoint UUID (this will ensure that the data gets transferred to the right place), and finally, leave the path as "/" as it was set when creating the "example data transfer" Globus Endpoint. When ready, hit "Download".



The screenshot shows the 'Download' dialog box from the NIH Data Management Environment (DME). The title bar says 'Download'. The dialog contains the following fields:

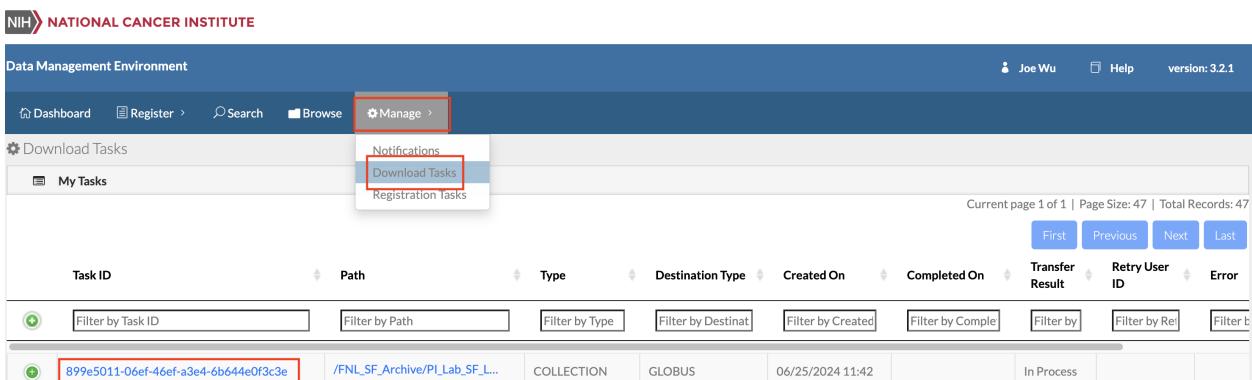
- Selected Collection:** /FNL_SF_Archive/PI_Lab_SF_Lab/Project_SFRnD_CS026880_4scRNAseq_030620/Flowcell_HYH2JBGXC/Sample_N_1395BL_NextGEM
- Destination Hierarchy:**
 - Download to Specified Folder
 - Create Parent Folder
 - Create Full Path
- Transfer Type:**
 - Globus
 - AWS S3
 - Google Drive
 - Google Cloud
 - dbGaP
- Globus Endpoint UUID:** [Obtain from Globus \(Optional\)](#) (The text 'db2c77a9-4134-47bf-b4f3-25e3bb6a24cb' is highlighted with a red box)
- Globus Endpoint (Destination) Path:** / (The path '/' is highlighted with a red box)
- Download** button

If all goes well, users will see a message indicating that the data transfer request has been submitted successfully.



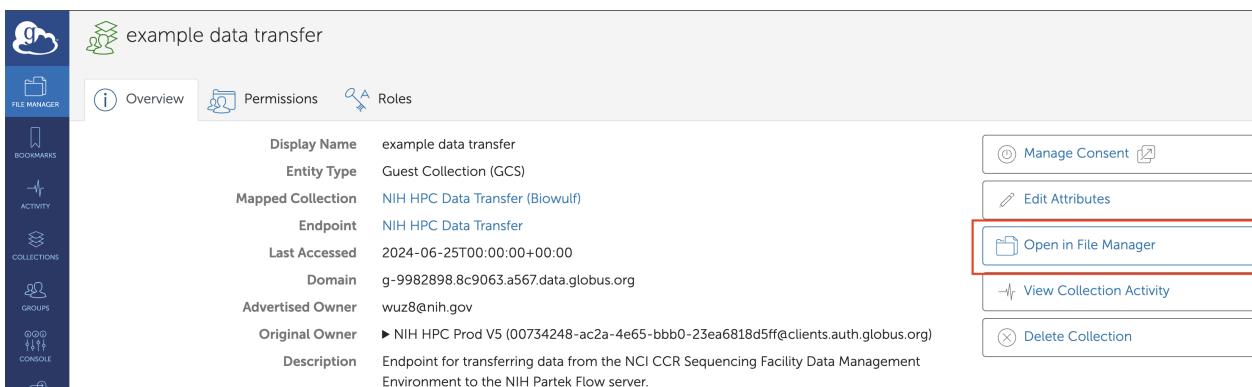
This screenshot shows the NIH Data Management Environment (DME) interface. The top navigation bar includes links for Dashboard, Register, Search, Browse, and Manage. The Manage link is highlighted with a red box. Below the navigation is a 'Download' section. A message box at the top of this section says: 'This page allows you to download the selected collection to a Globus endpoint, an AWS S3 bucket, Google Drive, Google Cloud or dbGaP.' A 'Selected Collection:' field shows the path: '/FNL_SF_Archive/PI_Lab_SF_Lab/Project_SFRnD_CS026880_4scRNAseq_030620/Flowcell_HYH2JBGXC/Sample_N_1395BL_NextGEM'. Under 'Destination Hierarchy:', the radio button 'Download to Specified Folder' is selected. A message box at the bottom of the download section says: 'Asynchronous download request is submitted successfully! Task Id: 899e5011-06ef-46ef-a3e4-6b644e0f3c3e'.

Click on "Manage" and then "Download Tasks" to check download progress. Each download is assigned a task ID.



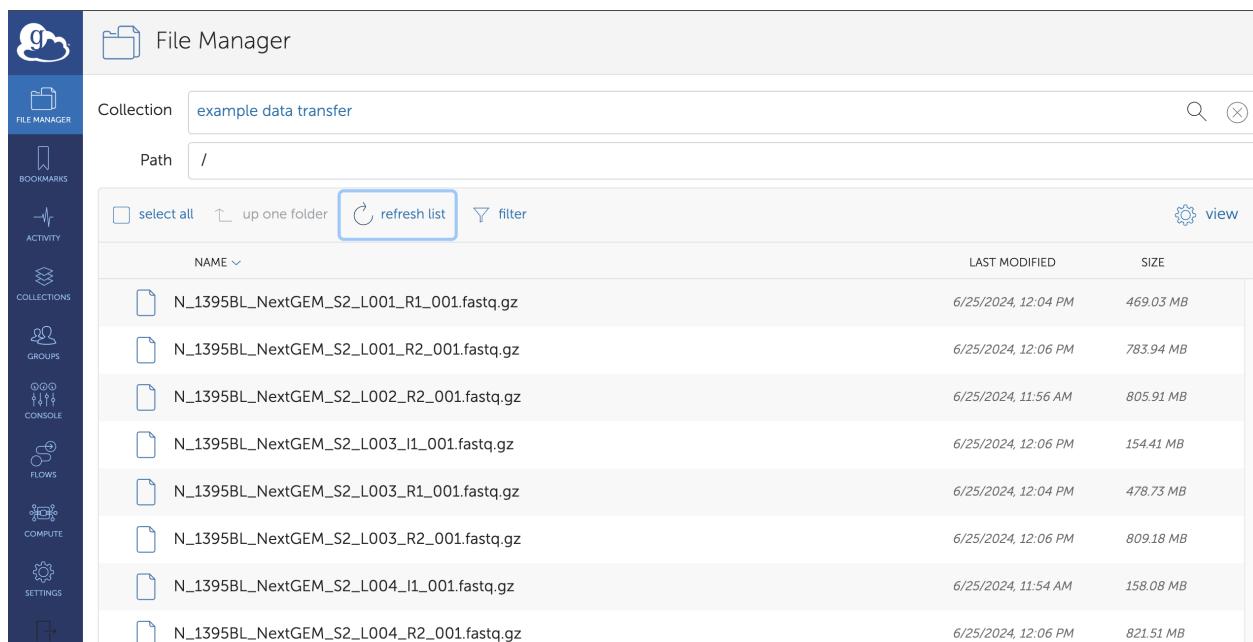
This screenshot shows the 'Download Tasks' page in the NIH DME. The 'Manage' link in the top navigation bar is highlighted with a red box. A dropdown menu for 'Download Tasks' is open, also highlighted with a red box. The table below shows a single task with the ID '899e5011-06ef-46ef-a3e4-6b644e0f3c3e'. The table has columns for Task ID, Path, Type, Destination Type, Created On, Completed On, Transfer Result, Retry User ID, and Error. The 'Task ID' column shows the ID '899e5011-06ef-46ef-a3e4-6b644e0f3c3e' with a red box around it. The 'Path' column shows the path '/FNL_SF_Archive/PI_Lab_SF_Lab/Project_SFRnD_CS026880_4scRNAseq_030620/Flowcell_HYH2JBGXC/Sample_N_1395BL_NextGEM'. The 'Type' column shows 'COLLECTION'. The 'Destination Type' column shows 'GLOBUS'. The 'Created On' and 'Completed On' columns both show the date '06/25/2024 11:42'. The 'Transfer Result' column shows 'In Process'.

Go back to the "example data transfer" endpoint on Globus and click on "Open in File Manager".



This screenshot shows the Globus File Manager interface for the 'example data transfer' endpoint. The left sidebar has icons for FILE MANAGER, BOOKMARKS, ACTIVITY, COLLECTIONS, GROUPS, and CONSOLE. The main area shows the endpoint details: Display Name 'example data transfer', Entity Type 'Guest Collection (GCS)', Mapped Collection 'NIH HPC Data Transfer (Biowulf)', Endpoint 'NIH HPC Data Transfer', Last Accessed '2024-06-25T00:00:00+00:00', Domain 'g-9982898.8c9063.a567.data.globus.org', Advertised Owner 'wuz8@nih.gov', Original Owner 'NIH HPC Prod V5 (00734248-ac2a-4e65-bbb0-23ea6818d5ff@clients.auth.globus.org)', and Description 'Endpoint for transferring data from the NCI CCR Sequencing Facility Data Management Environment to the NIH Partek Flow server'. On the right, there are several buttons: 'Manage Consent', 'Edit Attributes', 'Open in File Manager' (which is highlighted with a red box), 'View Collection Activity', and 'Delete Collection'.

The data will populate in the `data/username/PartekFlow/globus/example_data_transfer` folder (ie. the folder pointed to by the "example data transfer" endpoint) as the download proceeds.



The screenshot shows the Globus File Manager interface. On the left is a sidebar with icons for FILE MANAGER, BOOKMARKS, ACTIVITY, COLLECTIONS, GROUPS, CONSOLE, FLOWS, COMPUTE, and SETTINGS. The main area is titled 'File Manager' and shows a collection named 'example data transfer' at the path '/'. A list of files is displayed, each with a preview icon, name, last modified date, and size. The files are fastq.gz files from a NextGEM sequencing run, with names like 'N_1395BL_NextGEM_S2_L001_R1_001.fastq.gz' and sizes ranging from 154.41 MB to 821.51 MB. The 'refresh list' button is highlighted with a blue box.

NAME	LAST MODIFIED	SIZE
N_1395BL_NextGEM_S2_L001_R1_001.fastq.gz	6/25/2024, 12:04 PM	469.03 MB
N_1395BL_NextGEM_S2_L001_R2_001.fastq.gz	6/25/2024, 12:06 PM	783.94 MB
N_1395BL_NextGEM_S2_L002_R2_001.fastq.gz	6/25/2024, 11:56 AM	805.91 MB
N_1395BL_NextGEM_S2_L003_I1_001.fastq.gz	6/25/2024, 12:06 PM	154.41 MB
N_1395BL_NextGEM_S2_L003_R1_001.fastq.gz	6/25/2024, 12:04 PM	478.73 MB
N_1395BL_NextGEM_S2_L003_R2_001.fastq.gz	6/25/2024, 12:06 PM	809.18 MB
N_1395BL_NextGEM_S2_L004_I1_001.fastq.gz	6/25/2024, 11:54 AM	158.08 MB
N_1395BL_NextGEM_S2_L004_R2_001.fastq.gz	6/25/2024, 12:06 PM	821.51 MB

These changes are also reflected on Biowulf. Again, replace `username` with the user's Biowulf user name.

```
ls /data/username/PartekFlow/globus/example_data_transfer
```

```
N_1395BL_NextGEM_S2_L001_I1_001.fastq.gz
N_1395BL_NextGEM_S2_L001_R1_001.fastq.gz
N_1395BL_NextGEM_S2_L001_R2_001.fastq.gz
N_1395BL_NextGEM_S2_L002_I1_001.fastq.gz
N_1395BL_NextGEM_S2_L002_R1_001.fastq.gz
N_1395BL_NextGEM_S2_L002_R2_001.fastq.gz
N_1395BL_NextGEM_S2_L003_I1_001.fastq.gz
N_1395BL_NextGEM_S2_L003_R1_001.fastq.gz
N_1395BL_NextGEM_S2_L003_R2_001.fastq.gz
N_1395BL_NextGEM_S2_L004_I1_001.fastq.gz
N_1395BL_NextGEM_S2_L004_R1_001.fastq.gz
N_1395BL_NextGEM_S2_L004_R2_001.fastq.gz
```

Transferring Data to the NIH Partek Flow Server Using Command Line

Note

Command line uploads can be used when the files are already present on Biowulf. -- Biowulf (https://partekflow.cit.nih.gov/#upload_commandline)

Copy from User's Biowulf data Folder to Partek Flow uploads Folder

Copying from user's Biowulf /data directory to the uploads subfolder in PartekFlow is allowed. Be sure to be in an interactive session for this by using the `sinteractive` command. For instance:

To transfer the folder `hbr_uhr_fastq_download` in `/data/username` to the `uploads` subfolder in PartekFlow, do the following. Replace `username` with the user's Biowulf user name.

```
cp -r /data/username/hbr_uhr_fastq_download /data/username/PartekFlow/uploads/
```

Note that the Biowulf user owns the `hbr_uhr_fastq_download` PartekFlow. This enables users to delete the folder. The `ls` command lists directory content and the `-l` option list the content in long or detailed format.

```
[wuz8@cn4304 username]$ ls -l /data/username/PartekFlow/uploads/
drwxr-s---+ 2 username      partekfl 4096 May 28 11:55 hbr_uhr_fastq_download/
```

It is possible to change into the upload subfolder in `/data/username/PartekFlow` and make a directory using the `mkdir` command.

```
cd /data/username/PartekFlow/uploads
```

```
[username@cn4304 uploads]$ mkdir hbr_uhr_fastq_download/
```

Again the user owns the directory, and thus can delete it. Creating folders is a good way to keep data organized for analyses.

```
[wuz8@cn4304 uploads]$ ls -l
drwxrws---+ 2 username      partekfl 4096 May 28 13:12 hbr_uhr_fastq
```

Tip

File transfer using command line:

- Copy from data folder to Partek Flow uploads folder is **OK**
- Copy from data folder to Partek Flow directory is **OK**
- Copy from data folder to Partek Flow project folder is **not OK**
- Copy from Partek Flow uploads folder to project folder is **not OK**

Caution

While the above command line methods for data transfer are valid, users should remember that: "Virtually all file transfer activities should be run from within the web interface, rather than from the command line. We have implemented a permissions policy on users' PartekFlow directories to prevent inadvertent file removal mistakes that breaks the old way of moving files around." -- Biowulf staff

scp from Personal Computer to the Partek Flow uploads Folder.

The `scp` command enables users to securely transfer data to and from personal computer to a remote computer such as a high performance computing system. To transfer data from personal computer to the Partek Flow `uploads` folder do the following.

Note

Helix is the node on the Biowulf system that is intended for data transfer.

```
scp local_file username@helix.nih.gov:/data/username/PartekFlow/uploads
```

Enter user's Biowulf password.

For example, the following command construct will transfer the folder `hbr_uhr_fastq` on local computer that contains FASTQ files to the Partek Flow `uploads` folder on Biowulf. The `-r` option in `scp` will copy a folder and its content.

```
scp -r hbr_uhr_fastq_files username@helix.nih.gov:/data/username/PartekFlow/uploads
```

```
ls -l /data/username/PartekFlow/uploads
```

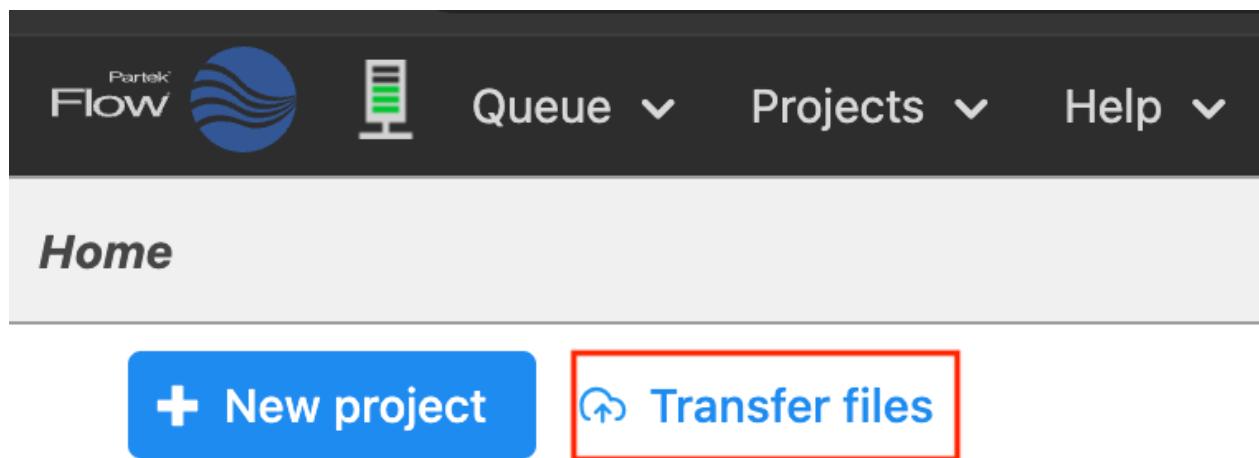
Again, the Biowulf account user owns the directory and can modify and delete it.

```
drwxr-s---+ 2 username  partekfl 4096 Jun 25 21:13 hbr_uhr_fastq_file
```

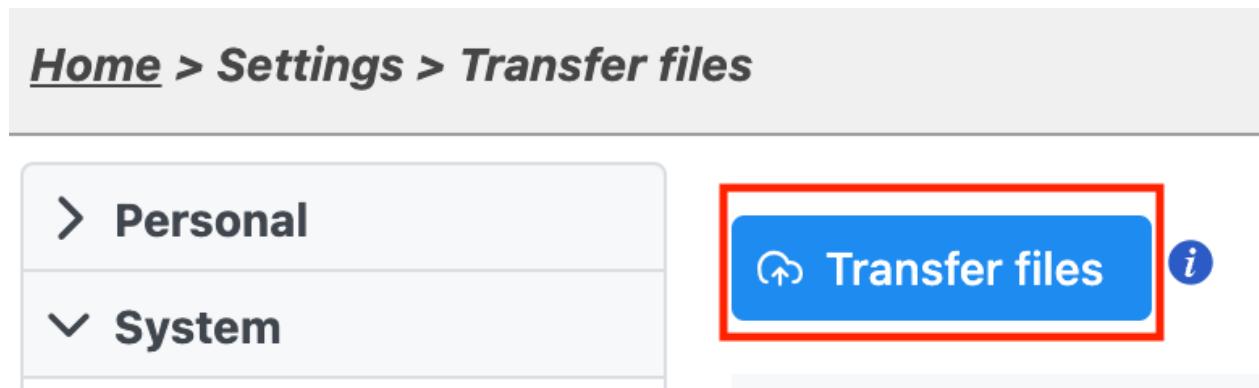
Transferring Data to the NIH Partek Flow Server Using the Web Tool

Using the Partek Flow web tool to transfer data

If users have data stored on a personal computer, the Partek Flow web tool can be used to transfer data onto the server. Upon signing in to Partek Flow, a button labeled "Transfer files" is available. Click on this.



This takes the users to the "Transfer files" section of the settings page. Click on the blue "Transfer files" tab at the top.

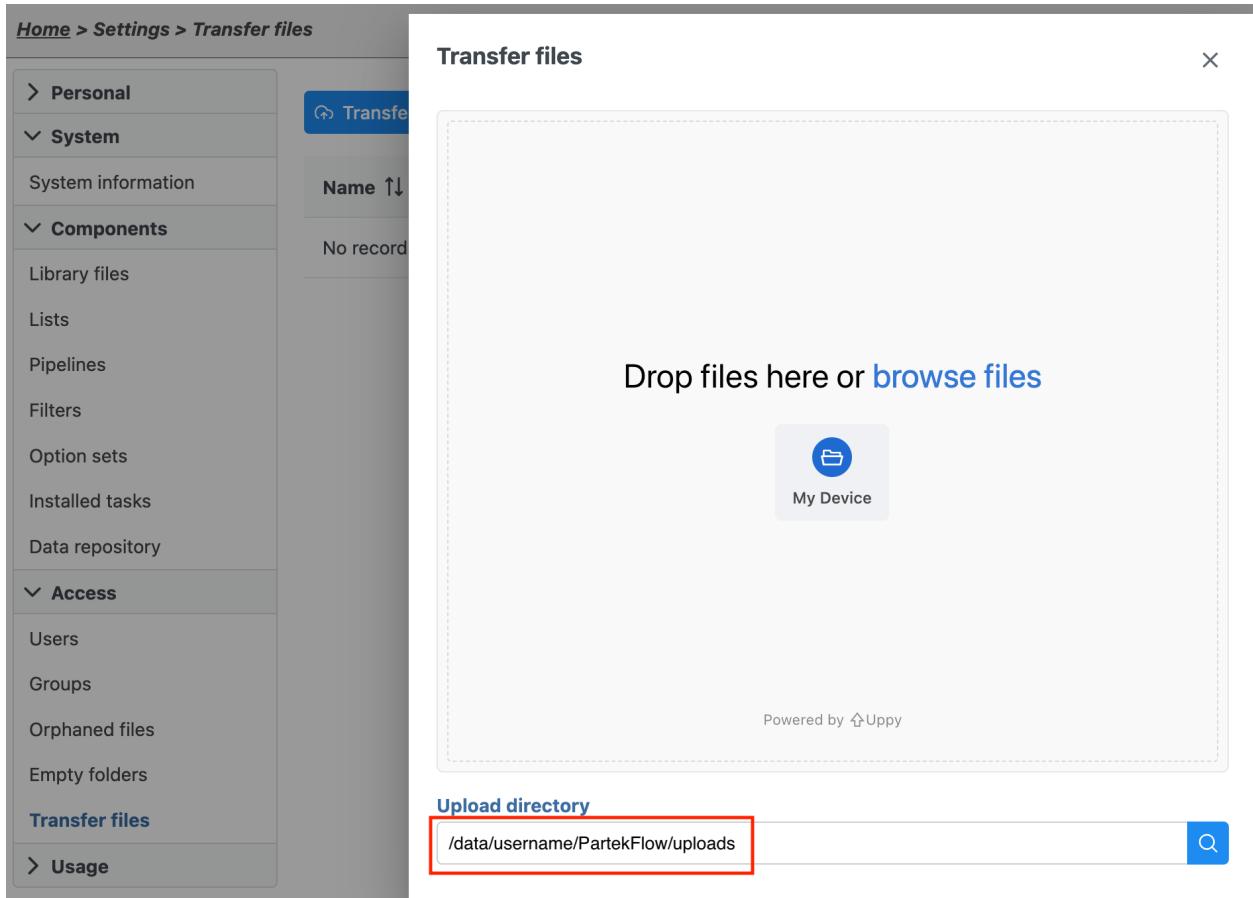


A dialogue box that assists the user with data transfer appears.

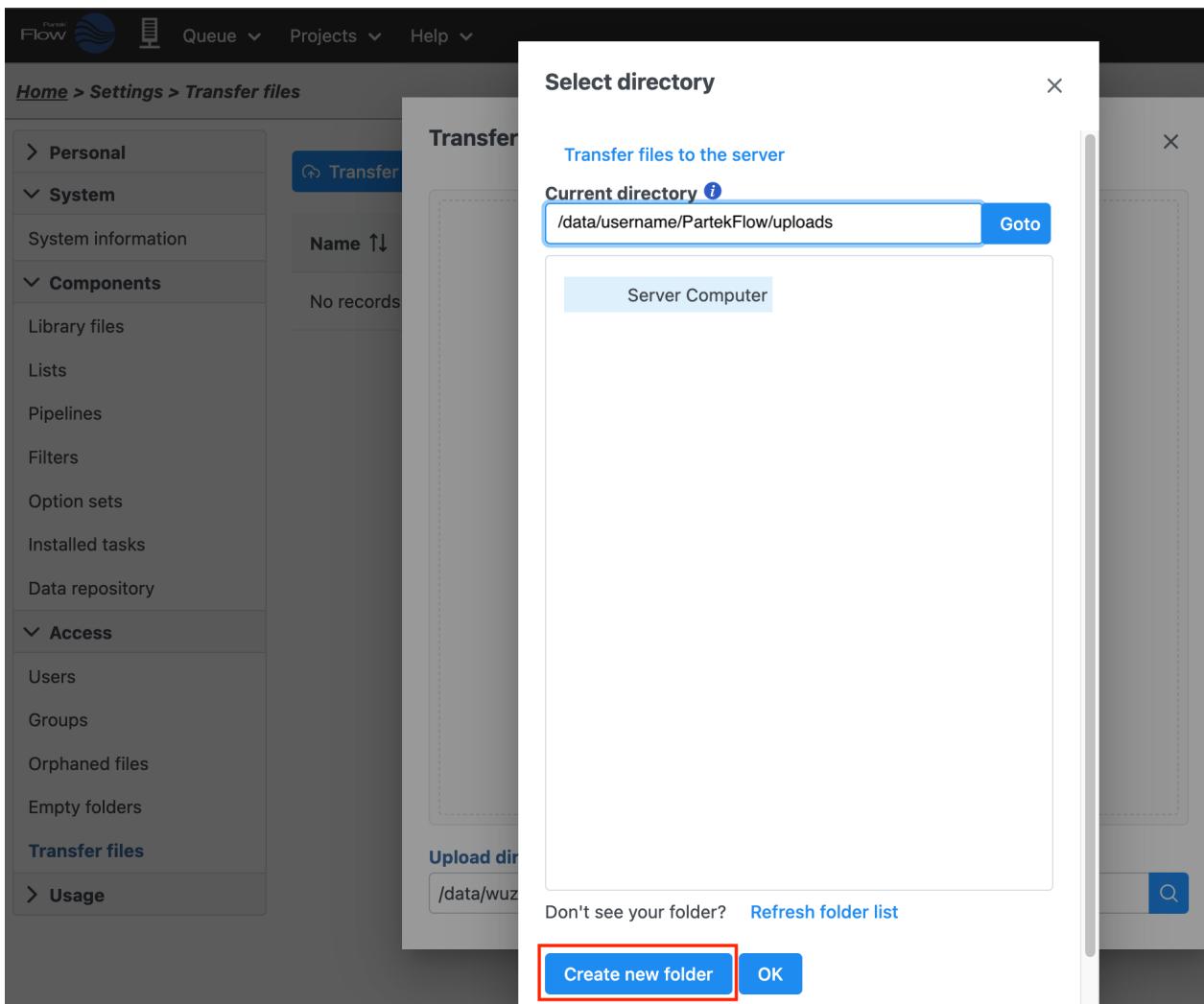
Note

Users can only transfer to the Partek Flow uploads folder or subfolders within uploads.

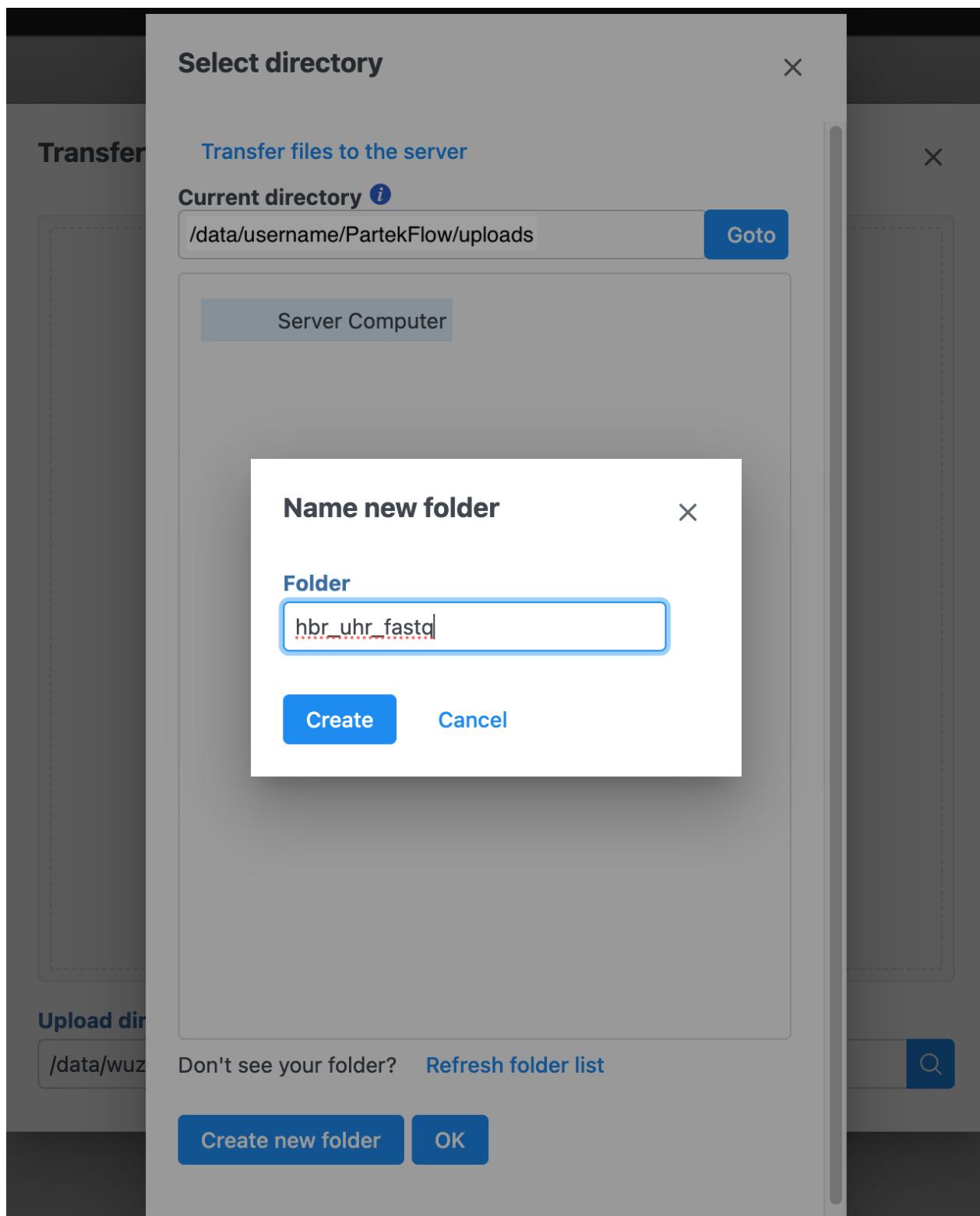
The current "Upload directory" as indicated at the bottom of the file transfer dialogue box is /data/username/PartekFlow/uploads. Click on the magnifying glass next to it to create new folders within Partek Flow uploads to store data uploaded for specific projects. It is a good idea to create a folder to store data for each project.



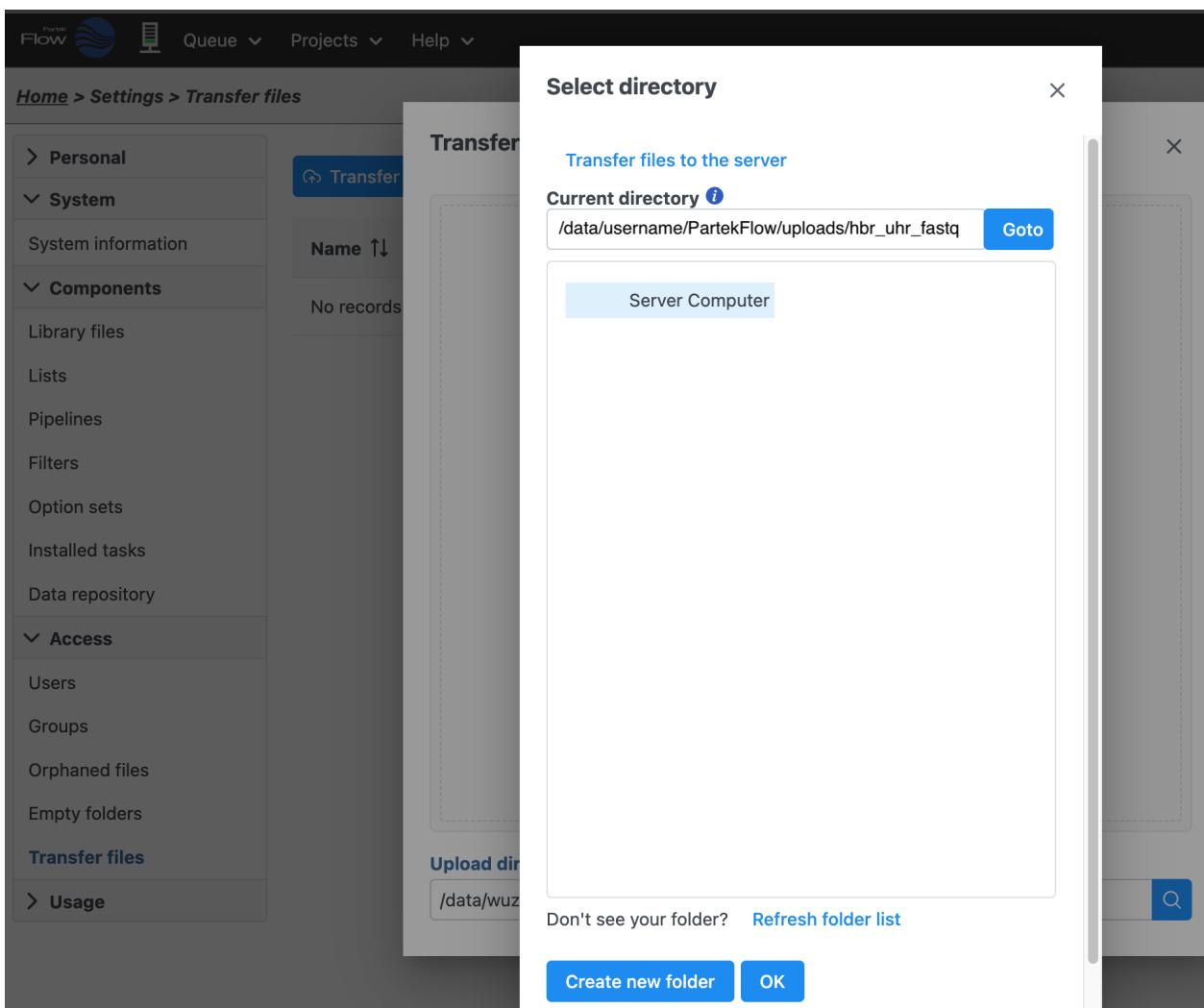
To create a new folder, click on the blue "Create new folder" tab.



Enter the name of the folder (ie. hbr_uhr_fastq) and click the blue "Create" tab when ready.



The "Current directory" is updated to reflect that the user is now in /data/username/PartekFlow/uploads/hbr_uhr_fastq.



Click "Ok" to return to the file transfer dialogue box. Note that the "Upload directory" has been updated to `/data/username/PartekFlow/uploads/hbr_uhr_fastq` as well.

Twelve FASTQ files will be transferred to the folder `hbr_uhr_fastq` in `uploads`. Users can either drag and drop or browse to select files for upload. Click the green "Upload" button when ready (the number shown inside this button depends on the number of items to be transferred).

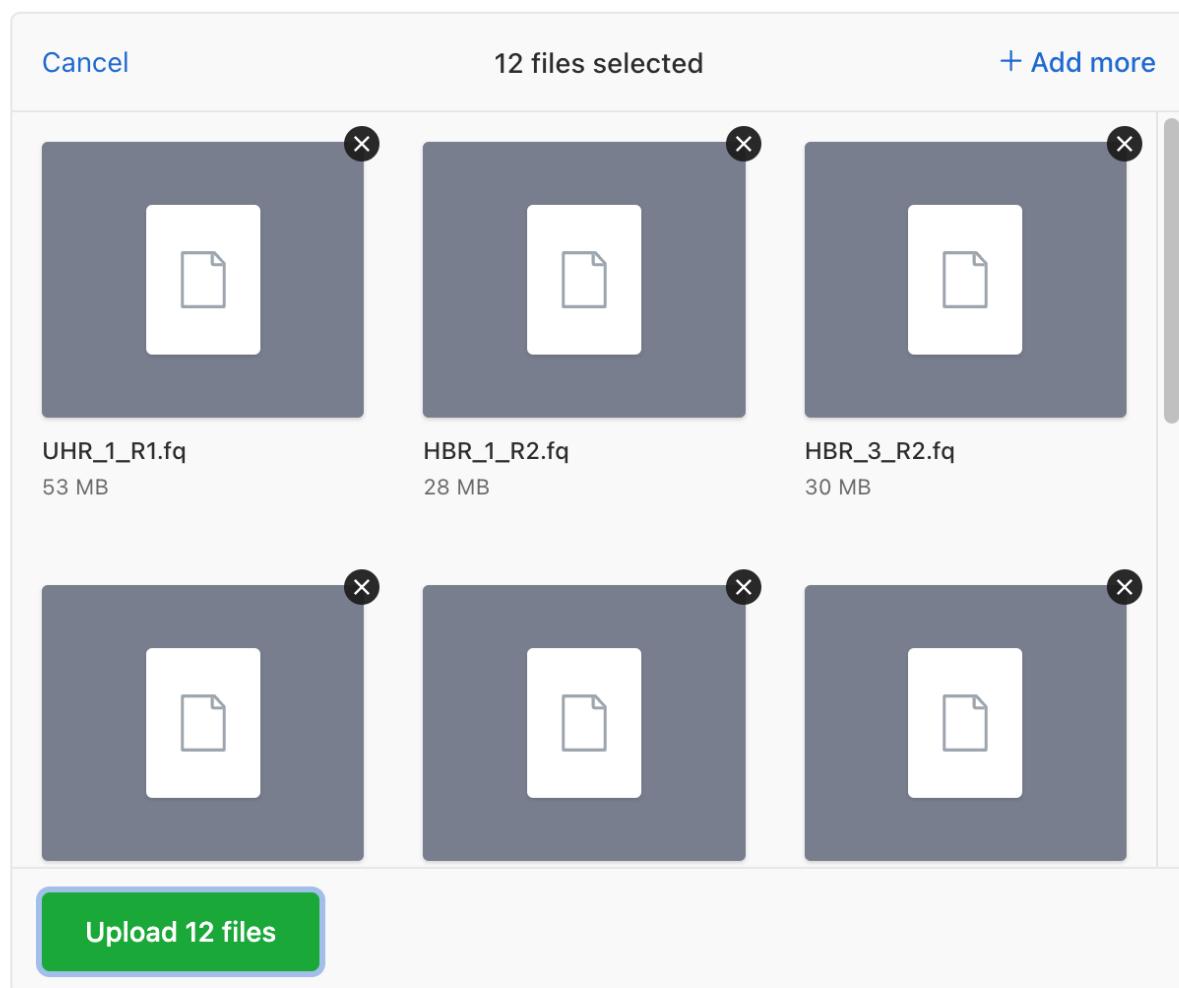
Transfer files



Cancel 12 files selected + Add more

		
UHR_1_R1.fq 53 MB	HBR_1_R2.fq 28 MB	HBR_3_R2.fq 30 MB
		

Upload 12 files



Upload directory

/data/username/PartekFlow/uploads/hbr_uhr_fastq



The warning message below will appear once the data transfer has commenced. Click "I understand".

Transfer files

Cancel Uploading 12 files + Add more

Transfer files

The transfer has started. Please do not close this browser tab or let the computer go to sleep or shut down until the transfer is complete.

I understand

UHR_1_R1.fq
53 MB

UHR_3_R1.fq

HBR_2_R2.fq

UHR_2_R1.fq

Uploading: 0%
0 of 12 files uploaded · 0 B of 453 MB · 0s left

Upload directory

/data/username/PartekFlow/uploads/hbr_uhr_fastq

A bar indicating the transfer status will also appear.

Transfer files



Cancel Uploading 10 files + Add more

UHR_1_R1.fq 53 MB	HBR_1_R2.fq 28 MB	HBR_3_R2.fq 30 MB
UHR_3_R1.fq	HBR_2_R2.fq	UHR_2_R1.fq
<div style="border: 2px solid red; padding: 5px;"><p>Uploading: 28% 2 of 12 files uploaded · 126 MB of 453 MB · 11m 40s left</p>[undo] [cancel]</div>		
Upload directory <input type="text" value="/data/wuz8/PartekFlow/uploads/hbr_uhr_fastq"/> [search]		

After the transfer is complete, users can click on the arrow next to the /data/username/PartekFlow/upload folder to view its subfolders and contents of subfolders.

Note

The user will not own contents transferred via the web tool. Thus, the only way to delete is using the "trash can" under the "Action" column. If user attempts to delete via command line, the message "Permission denied" will appear.

Name	Creator	Size	Last modified	Status	Actions
/data/username/PartekFlow/uploads/hbr_uhr_fastq		453.43 MB	-	-	Delete folder
HBR_3_R1.fq	username	30.38 MB	26 Jun 2024, 02:24 PM EDT	Complete	Delete individual files
UHR_2_R1.fq	username	38.01 MB	26 Jun 2024, 02:24 PM EDT	Complete	Delete individual files
UHR_1_R2.fq	username	53.24 MB	26 Jun 2024, 02:24 PM EDT	Complete	Delete individual files

Listing the contents of the `/data/username/PartekFlow/uploads` folder on Biowulf will reveal the `hbr_uhr_fastq` subdirectory.

```
[wuz8@cn4304 username]$ ls -l /data/username/PartekFlow/uploads/
drwxrws---+ 2 partekfl partekfl 4096 May 28 14:26 hbr_uhr_fastq
```

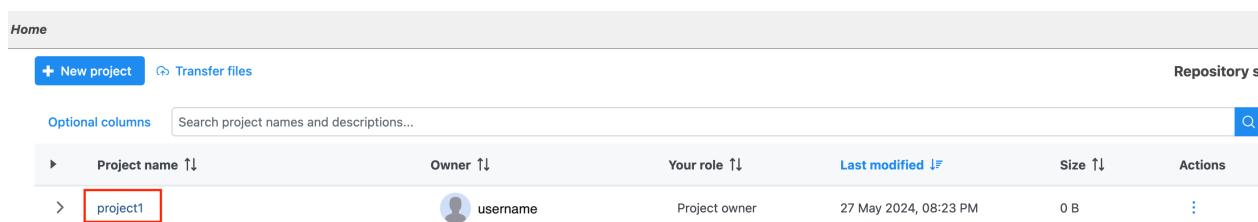
An attempt to remove the `hbr_uhr_folder` will result in a message saying "Permission denied". In the `rm` command below, `-r` tells `rm` to remove a folder and `f` forces `rm` to delete without asking for confirmation.

```
[wuz8@cn4304 username]$ rm -rf /data/username/PartekFlow/uploads/hbr_
rm: cannot remove '/data/username/PartekFlow/uploads/hbr_uhr_fastq/hb
```

Importing Data to the Partek Flow Projects

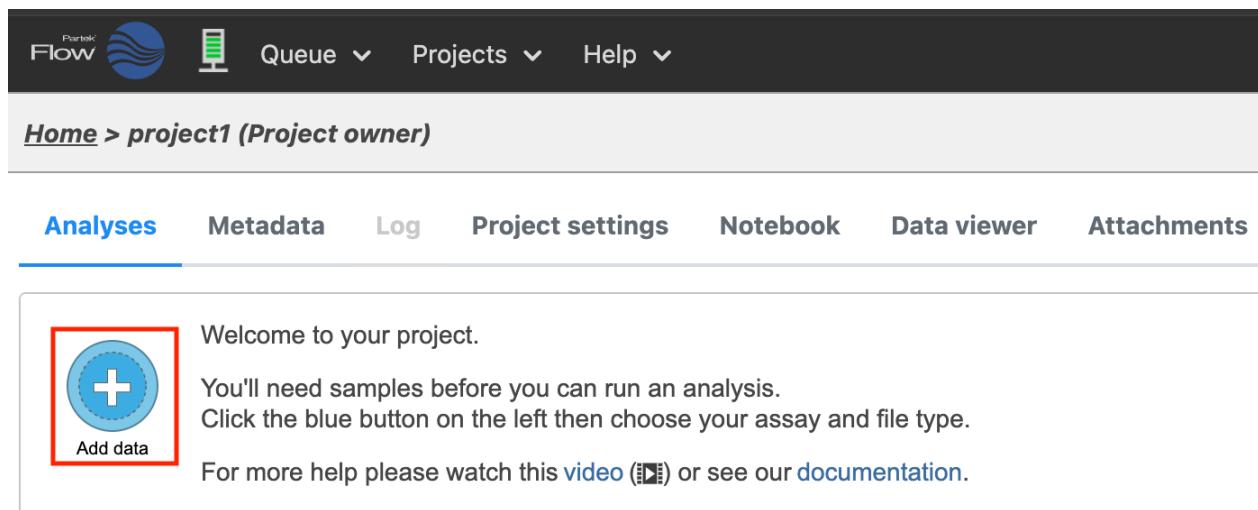
Data import into Partek Flow projects

After data has been transferred to the NIH Partek Flow server, the next step is to import them into a project. Sign onto Partek Flow and hit the "New project" button click to create a new project called project1. The project will then appear in the project table. Click on project1 to enter the analysis interface.



Project name ↑↓	Owner ↑↓	Your role ↑↓	Last modified ↴	Size ↑↓	Actions
> project1	username	Project owner	27 May 2024, 08:23 PM	0 B	⋮

Click on the round button labeled "Add data" to begin adding data to the project.



Welcome to your project.

You'll need samples before you can run an analysis.
Click the blue button on the left then choose your assay and file type.

For more help please watch this [video](#) or see our [documentation](#).

Users will be greeted with a range of options for data import. In this example, click "Bulk" and then select "RNA-Seq". FASTQ files will be imported. Click "Next" when ready.

Home > project1 > Initial import

Single cell Bulk Microarray Other

RNA-Seq Chip/ATAC-Seq DNA-Seq Metagenomics Proteomics

Select the format

fastq
Import unaligned reads. Acceptable file types are fastq, fastq.gz, fastq,bz2, fq, fq.gz, fq,bz2

bam
Import aligned reads. Acceptable file types are bam, sam, and ubam

Generic Count matrix
Import quantified data (e.g. gene counts per sample). Acceptable file types are txt, csv, tsv, txt.gz, csv.gz, tsv.gz

Back Next

In the subsequent page, the current directory is displayed in the file menu on the left.

Note

Users will only be able to access content in the `/data/username/PartekFlow` folder once signed on. So make sure to put the data there using one the methods described in this tutorial.

This example imports the FASTQ file in `uploads/hbr_uhr_fastq`. Be sure to check the box (labeled "Name") to select all files.

Home > project1 > Initial import > Import sample files > fastq

File select

Partek Flow Server URL GEO / ENA

Transfer files to the server

Current directory [?](#)
/data/username/PartekFlow/uploads/hbr_uhr_fastq [Goto](#)

12 files selected

<input checked="" type="checkbox"/> Name	Size
<input checked="" type="checkbox"/> HBR_1_R1.fq	27.76 MB
<input checked="" type="checkbox"/> HBR_1_R2.fq	27.76 MB
<input checked="" type="checkbox"/> HBR_2_R1.fq	33.91 MB
<input checked="" type="checkbox"/> HBR_2_R2.fq	33.91 MB
<input checked="" type="checkbox"/> HBR_3_R1.fq	30.38 MB
<input checked="" type="checkbox"/> HBR_3_R2.fq	30.38 MB
<input checked="" type="checkbox"/> UHR_1_R1.fq	53.24 MB
<input checked="" type="checkbox"/> UHR_1_R2.fq	53.24 MB
<input checked="" type="checkbox"/> UHR_2_R1.fq	38.01 MB
<input checked="" type="checkbox"/> UHR_2_R2.fq	38.01 MB

Don't see your folder? [Refresh folder list](#)

Valid files are: fastq

Scroll to the bottom of the data import page and select the data type (ie. mRNA) and click "Finish".

File select

Partek Flow Server URL GEO / ENA

Transfer files to the server

Current directory [?](#)
/data/username/PartekFlow/uploads/hbr_uhr_fastq [Goto](#)

12 files selected

<input checked="" type="checkbox"/>	Name	Size
<input checked="" type="checkbox"/>	HBR_1_R1.fq	27.76 MB
<input checked="" type="checkbox"/>	HBR_1_R2.fq	27.76 MB
<input checked="" type="checkbox"/>	HBR_2_R1.fq	33.91 MB
<input checked="" type="checkbox"/>	HBR_2_R2.fq	33.91 MB
<input checked="" type="checkbox"/>	HBR_3_R1.fq	30.38 MB
<input checked="" type="checkbox"/>	HBR_3_R2.fq	30.38 MB
<input checked="" type="checkbox"/>	UHR_1_R1.fq	53.24 MB
<input checked="" type="checkbox"/>	UHR_1_R2.fq	53.24 MB
<input checked="" type="checkbox"/>	UHR_2_R1.fq	38.01 MB
<input checked="" type="checkbox"/>	UHR_2_R2.fq	38.01 MB

Don't see your folder? [Refresh folder list](#) Valid files are: fastq

Select data type: (only required for multi assay project)

If none of the preset options matches your data, choose custom and add a custom type.

mRNA [▼](#)

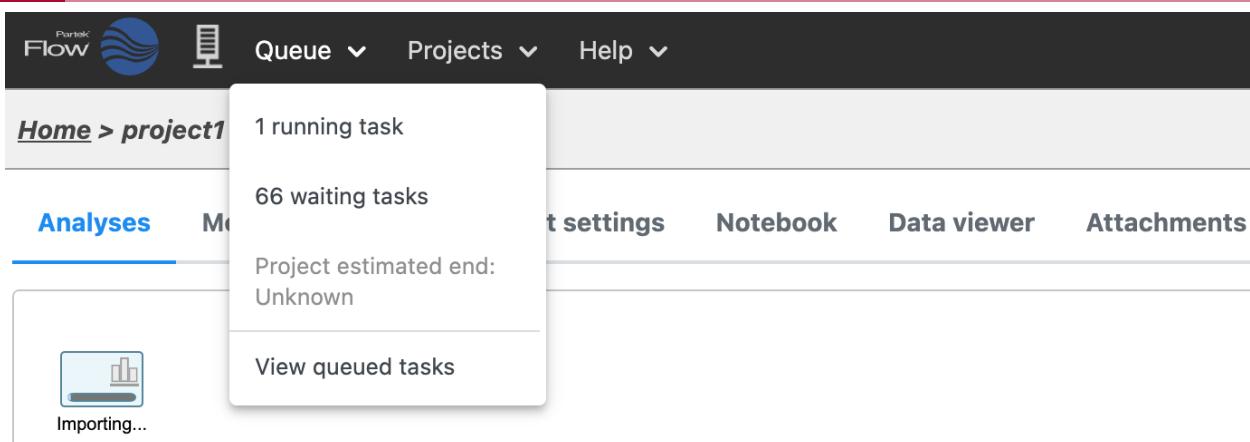
[Back](#) [Finish](#)

A rectangle labeled "Importing" will appear as data is being imported into a project.

Analyses Metadata Log Project settings Notebook Data viewer Attachments

Tip

The time it takes to import or to complete a task in Partek Flow depends on the queue. If the queue is long, then it's time for a coffee break.



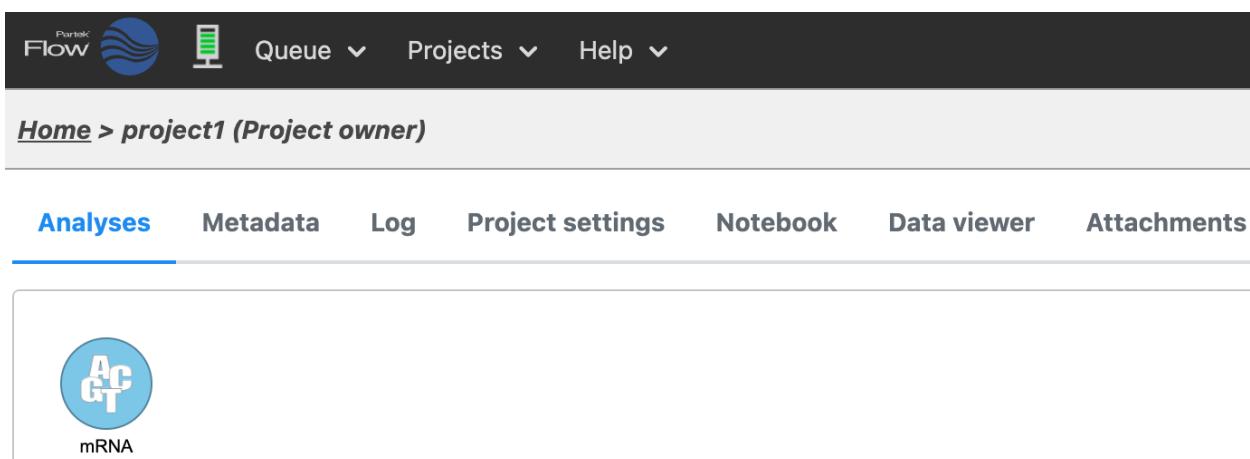
1 running task

66 waiting tasks

Project estimated end:
Unknown

[View queued tasks](#)

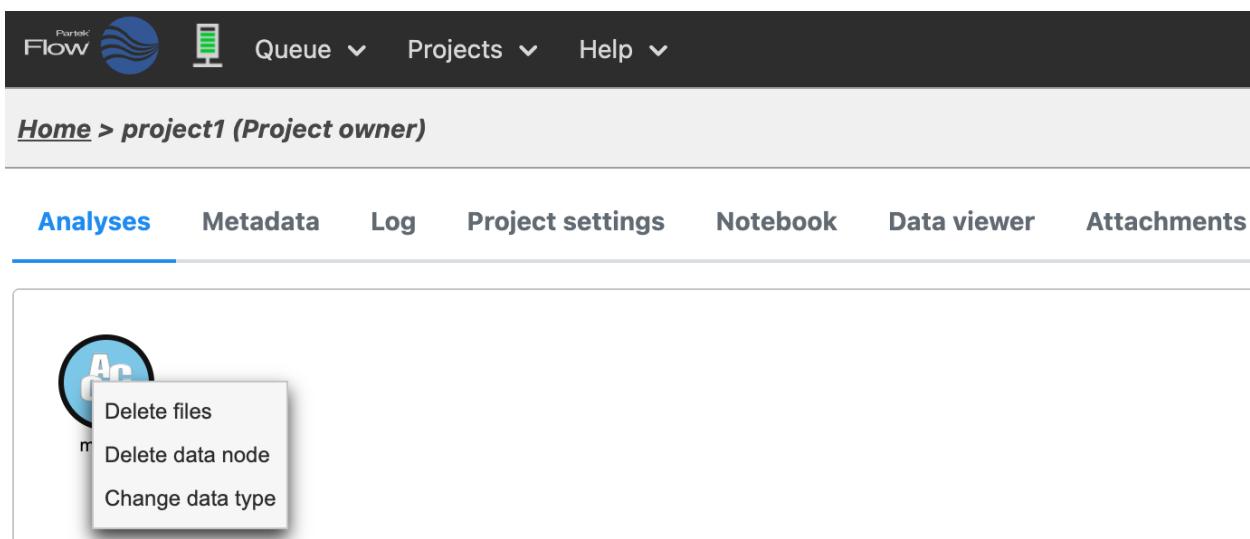
Upon successfully importing data, there will be a circular data node in the project.



[Project settings](#)

mRNA

Right click on the data node to either delete this node, files, or change data type.



[Delete files](#)

[Delete data node](#)

[Change data type](#)

Next, click on the "Metadata" tab to add metadata for the samples. Under the "Sample attributes" sections, users can manually add metadata or assign it from a file. This example will demonstrate manually adding metadata by clicking on "Manage".

Analyses Metadata Log Project settings Notebook Data viewer Attachments

Import

- Add data
- Sample attributes**
- Manage ←
- Assign values from file ←
- Add system-wide attribute

	Sample name	⋮
1	HBR_1	⋮
2	HBR_2	⋮
3	HBR_3	⋮
4	UHR_1	⋮
5	UHR_2	⋮
6	UHR_3	⋮

>Show data files Download

In the next page, select "Add new attribute".

Flow Queue Projects Help

Home > project1 > Metadata > Manage sample attributes

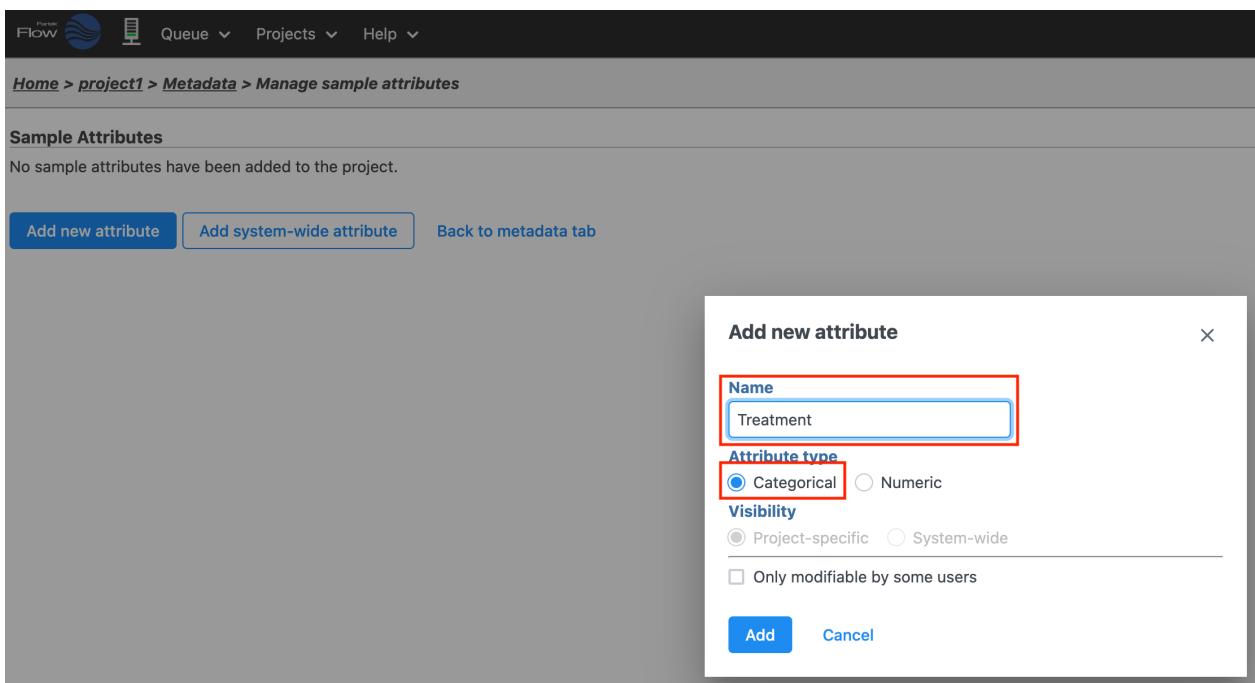
Sample Attributes

No sample attributes have been added to the project.

Add new attribute **Add system-wide attribute** **Back to metadata tab**

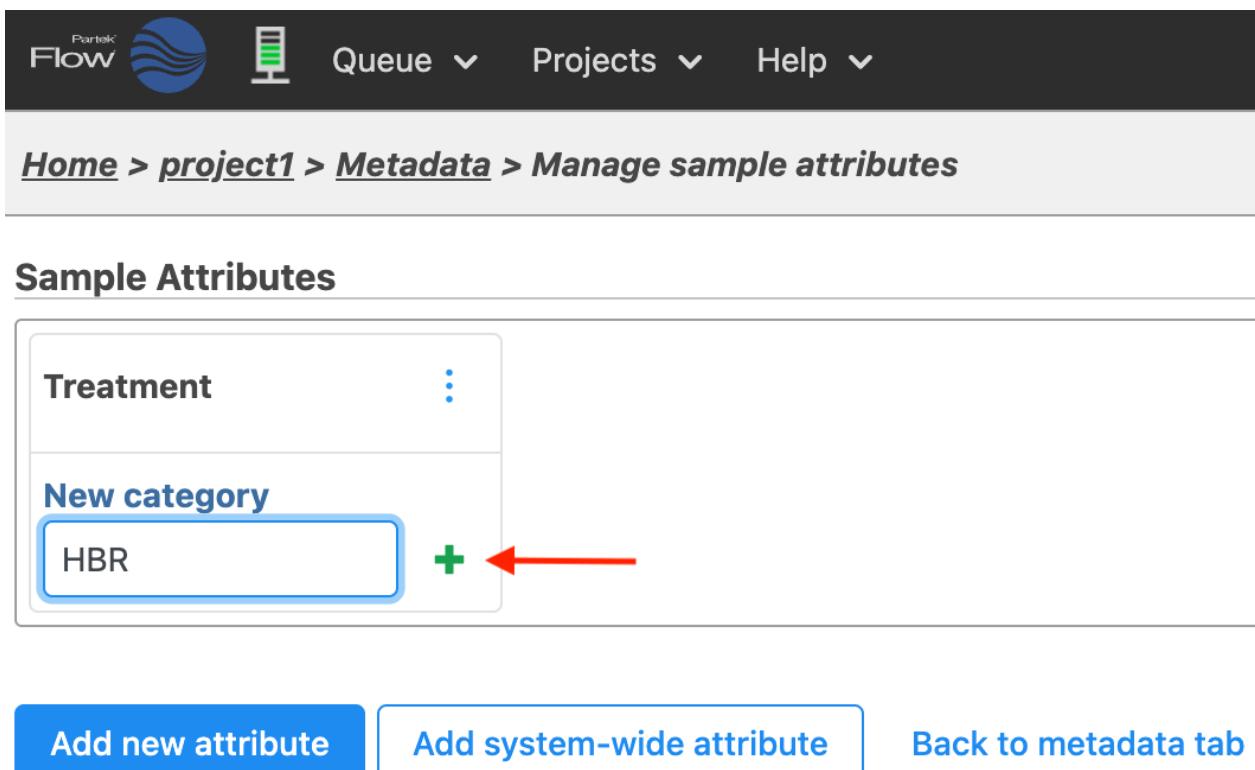
Create a new attribute to begin assigning

In the dialogue box that appears, name this attribute (think of an attribute as a variable) "Treatment" so that users can distinguish the treatment group in which each sample belongs. Be sure that "Categorical" is selected as this attribute type (ie. treatment group is categorical). Click "Add" when ready.



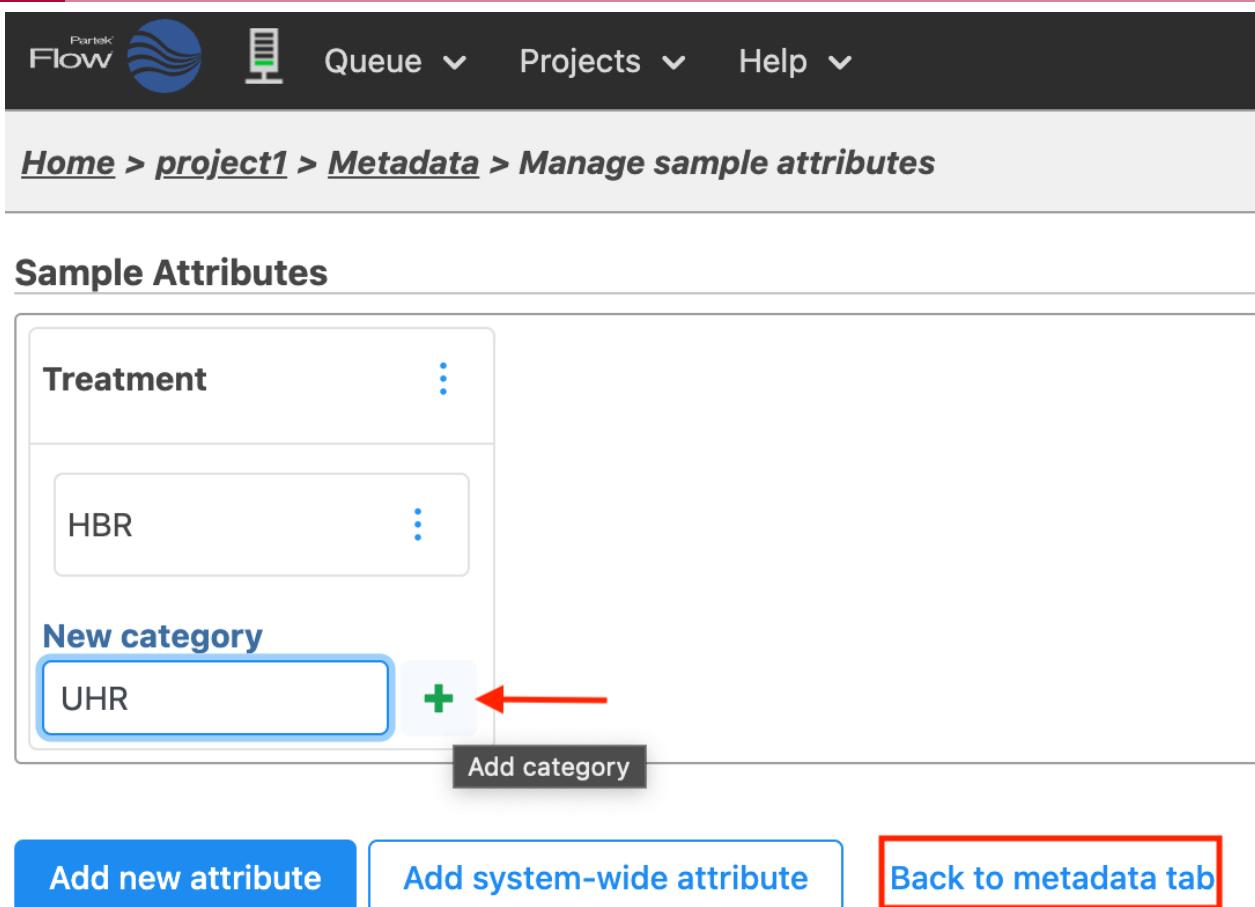
The screenshot shows the 'Manage sample attributes' page in Partek Flow. At the top, there are navigation links: 'Home', 'project1', 'Metadata', and 'Manage sample attributes'. Below this, a section titled 'Sample Attributes' states 'No sample attributes have been added to the project.' There are three buttons: 'Add new attribute' (highlighted with a red box), 'Add system-wide attribute', and 'Back to metadata tab'. A modal window titled 'Add new attribute' is open, containing fields for 'Name' (Treatment), 'Attribute type' (Categorical, selected), and 'Visibility' (Project-specific). The 'Add' and 'Cancel' buttons are at the bottom of the modal.

Then, under the "New category" box, enter the first treatment group (ie. HBR). Be sure to click the green "+" to add it.



The screenshot shows the 'Manage sample attributes' page. The 'Treatment' category is listed. A new category 'HBR' is added under 'New category', indicated by a red arrow pointing to the green '+' button next to the entry. The 'Add new attribute', 'Add system-wide attribute', and 'Back to metadata tab' buttons are at the bottom.

After that, add the second treatment group (ie. UHR) and click on the green "+" to add. Click the "Back to metadata tab" when finishing adding treatment groups.



Partek Flow Queue ▾ Projects ▾ Help ▾

Home > project1 > Metadata > Manage sample attributes

Sample Attributes

Treatment	⋮
HBR	⋮

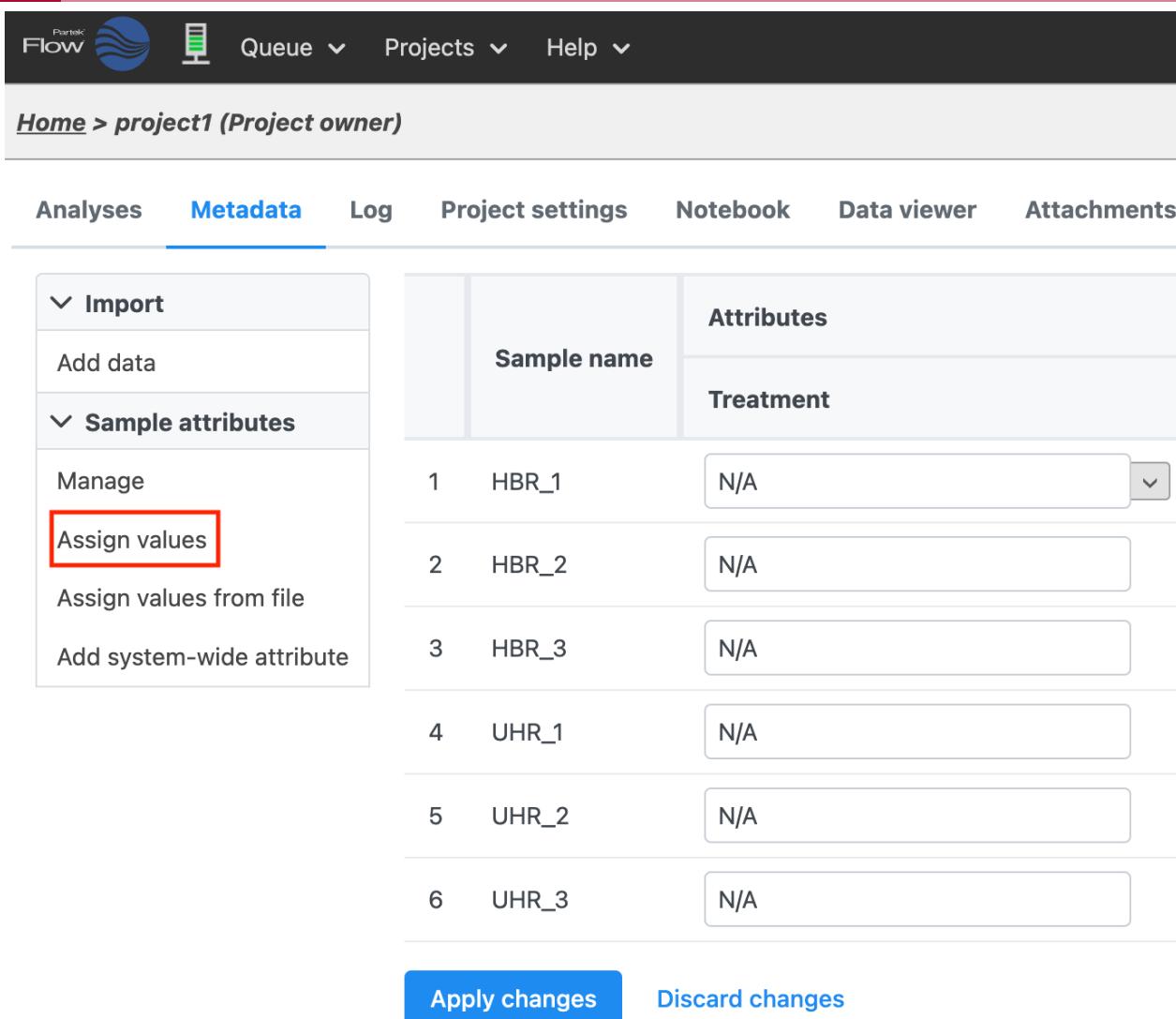
New category

UHR + ←

Add category

Add new attribute **Add system-wide attribute** **Back to metadata tab**

Hit the "Assign values" tab.

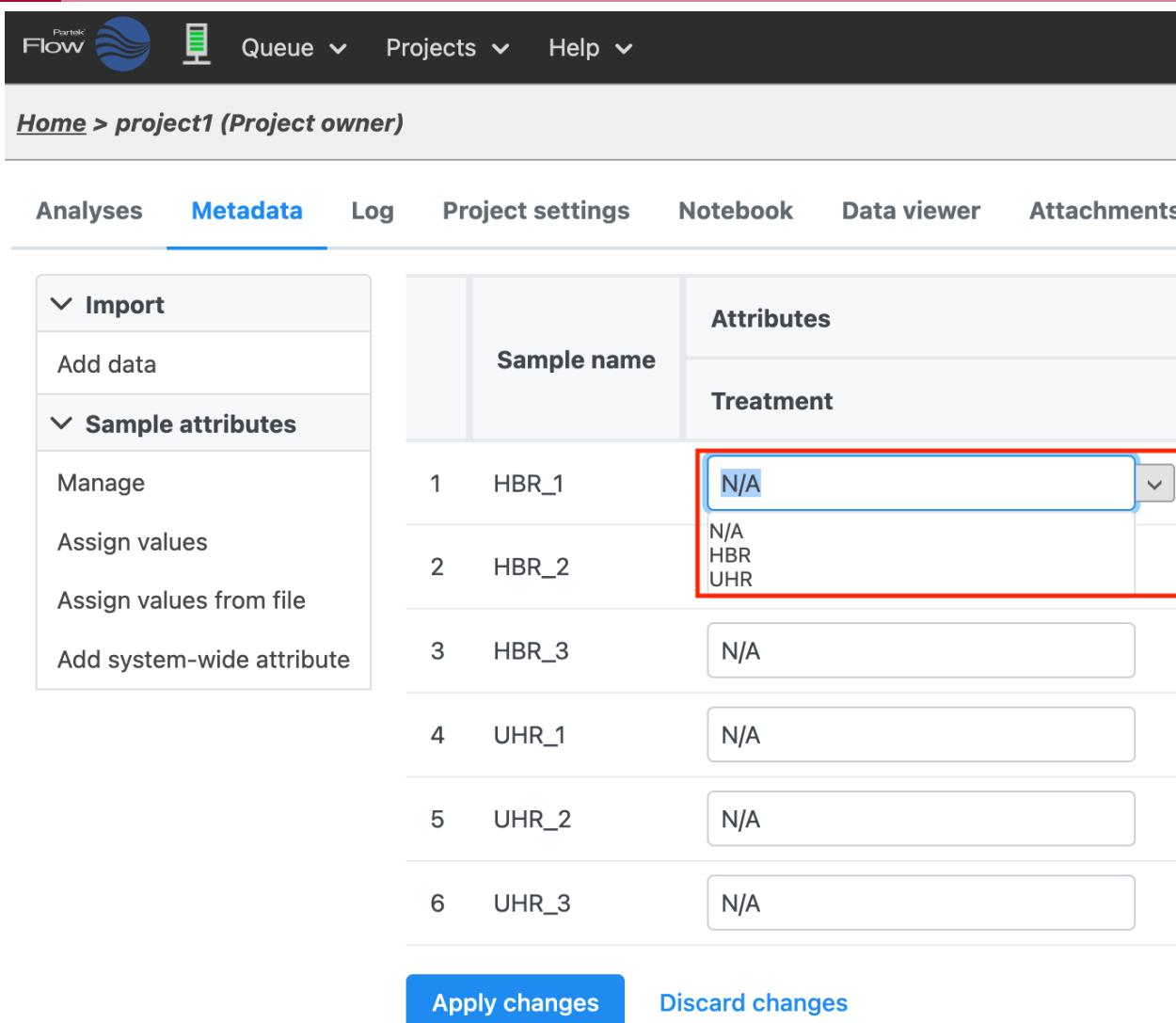


The screenshot shows the Partek Flow interface with the 'Metadata' tab selected. On the left, a sidebar contains 'Import' and 'Sample attributes' sections. The 'Assign values' option in the 'Sample attributes' section is highlighted with a red box. The main area displays a table of sample attributes. The table has columns for 'Sample name' and 'Attributes'. The 'Treatment' column for all samples shows 'N/A'.

	Sample name	Attributes
1	HBR_1	N/A
2	HBR_2	N/A
3	HBR_3	N/A
4	UHR_1	N/A
5	UHR_2	N/A
6	UHR_3	N/A

Apply changes **Discard changes**

Then, select the appropriate treatment for each sample using the drop down box. Click on "Apply changes" when finished and then the "Analyses" tab to go back to the analysis.



Analyses Metadata Log Project settings Notebook Data viewer Attachments

Import

- Add data

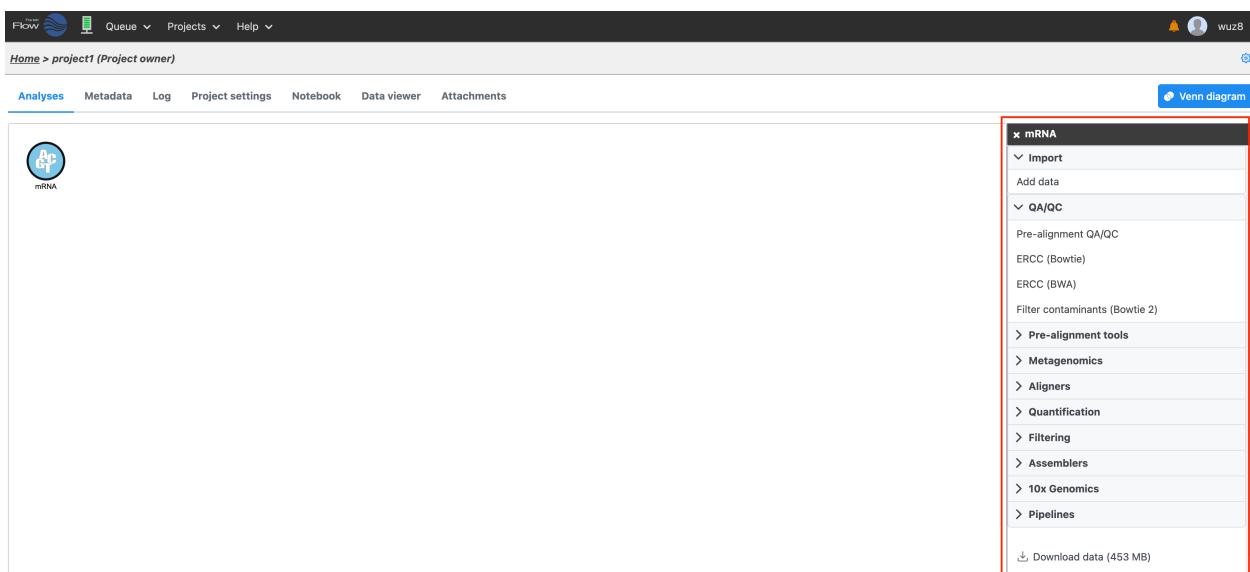
Sample attributes

- Manage
- Assign values
- Assign values from file
- Add system-wide attribute

	Sample name	Attributes
1	HBR_1	N/A N/A HBR UHR
2	HBR_2	N/A
3	HBR_3	N/A
4	UHR_1	N/A
5	UHR_2	N/A
6	UHR_3	N/A

Apply changes Discard changes

Click on a data node to see a menu of tasks that can be performed. For instance, the first step in analyzing high throughput sequencing data from FASTQ files is to perform pre-alignment QC.



Analyses Metadata Log Project settings Notebook Data viewer Attachments

mRNA

Import

Add data

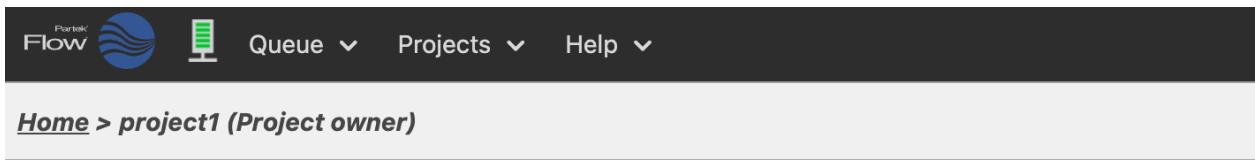
QA/QC

- Pre-alignment QA/QC
 - ERCC (Bowtie)
 - ERCC (BWA)
 - Filter contaminants (Bowtie 2)
- Pre-alignment tools
- Metagenomics
- Aligners
- Quantification
- Filtering
- Assemblers
- 10x Genomics
- Pipelines

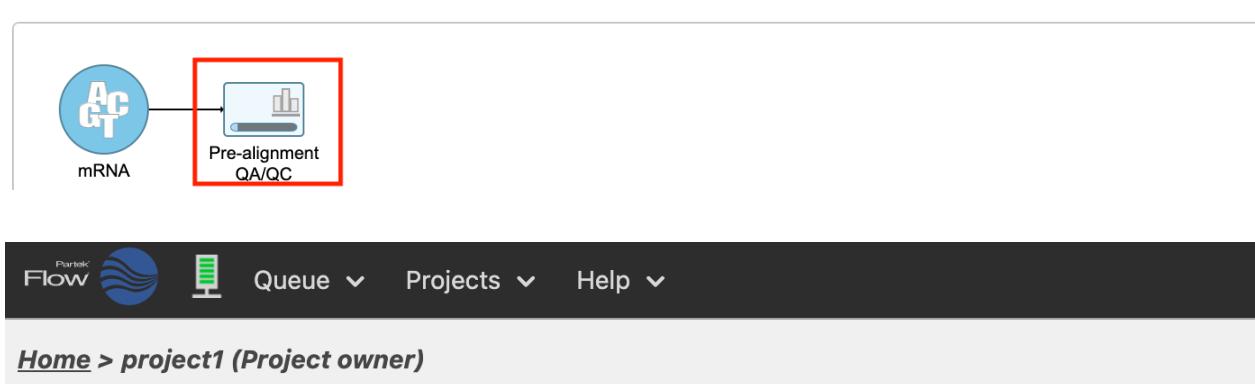
Download data (453 MB)

Note

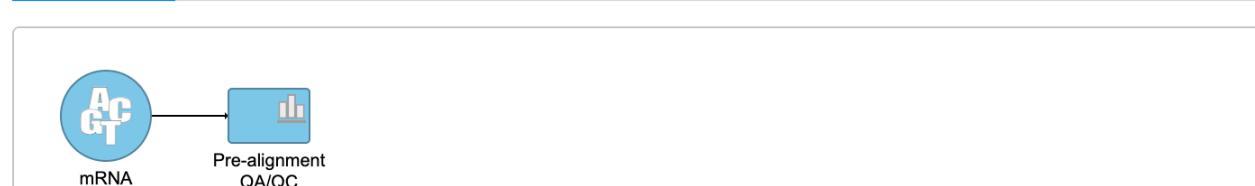
In Partek Flow, data nodes are round and task nodes are rectangular. Running tasks are light blue rectangles with a status bar inside. Once the task is complete, the rectangle turns into a darker shade of blue.



The screenshot shows the Partek Flow software interface. At the top, there is a dark header bar with the "Partek Flow" logo, a "Queue" icon, a "Projects" icon, and a "Help" icon. Below the header, the URL "Home > project1 (Project owner)" is displayed. The main content area has a light gray background and contains a "Analyses" tab (which is blue, indicating it is selected), and other tabs: "Metadata", "Log", "Project settings", "Notebook", "Data viewer", and "Attachments". Below the tabs, there is a diagram showing a circular "mRNA" node on the left connected by an arrow to a rectangular "Pre-alignment QA/QC" node on the right. The "Pre-alignment QA/QC" node is highlighted with a red border, and a status bar at the bottom of the rectangle shows the text "Pre-alignment QA/QC".



The screenshot shows the Partek Flow software interface, similar to the one above. The "Analyses" tab is selected. The "Pre-alignment QA/QC" task node is now a solid blue rectangle, indicating it has completed. The status bar at the bottom of the node is no longer visible.



The screenshot shows the Partek Flow software interface, similar to the ones above. The "Analyses" tab is selected. The "Pre-alignment QA/QC" task node is now a solid blue rectangle, indicating it has completed. The status bar at the bottom of the node is no longer visible.