



Advanced Molecular Detection Southeast Region Bioinformatics



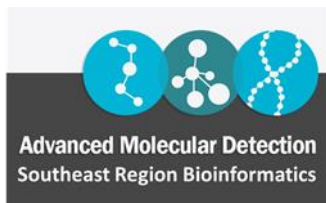
Galaxy Data Library Permissions

April 17, 2020

BPHL-SEbioinformatics@flhealth.gov

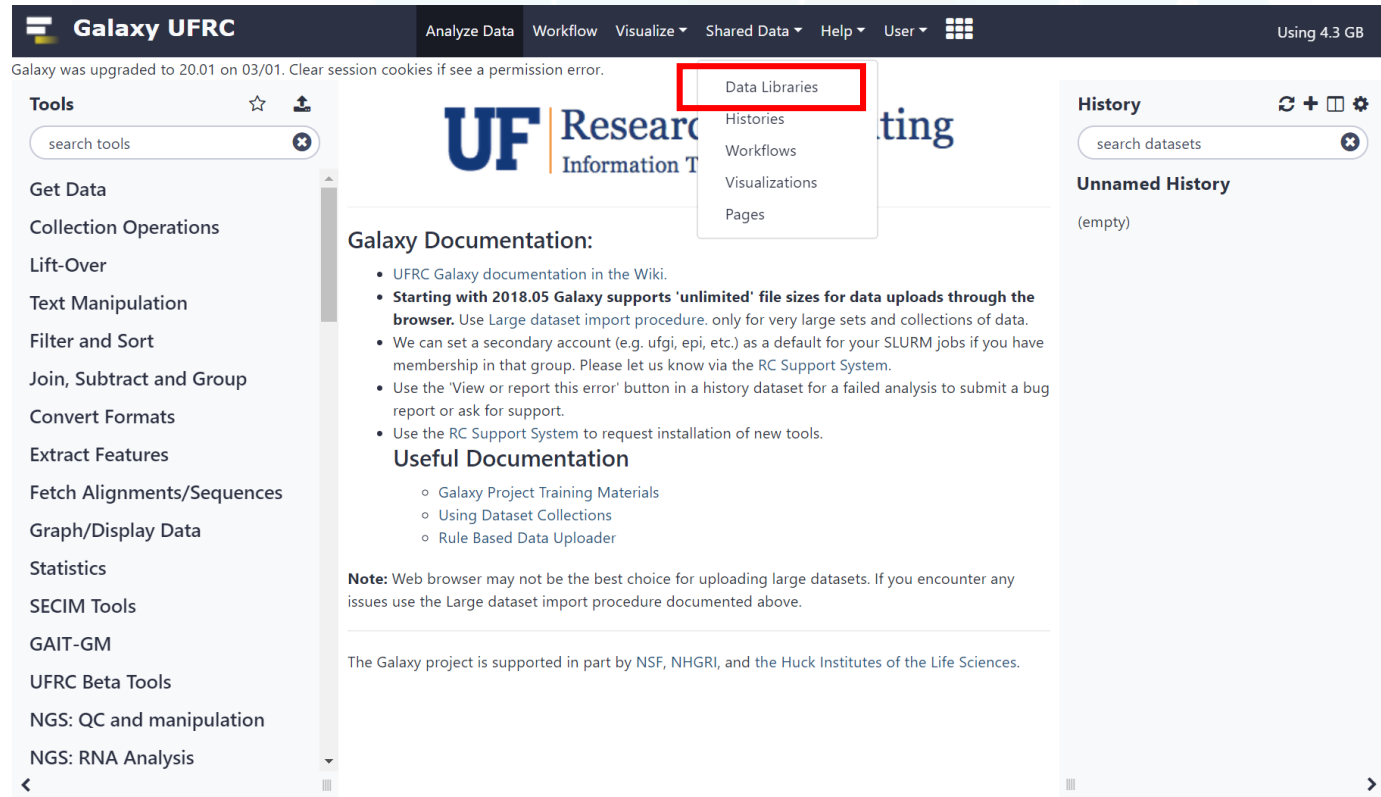
Galaxy Data Libraries

- Shared space to share data with users from your state group
 - Each user doesn't need to upload the same data if you are working on the same datasets.
 - One user can add the data and then it is in Galaxy for all to use!
 - You also can share results and output files with your group using libraries.
- Each state has one admin to manage your data library
- Permissions can be set to add new users
 - Allow users to access data in the data library
 - Allow users to add data to the data library



How to access your state's Galaxy data library

- Log in to Galaxy
- Click on “Shared Data” in the top center of your screen
- Click on “Data Libraries”



The screenshot displays the Galaxy UFRc web interface. At the top, a dark navigation bar contains the 'Galaxy UFRc' logo, a series of tabs (Analyze Data, Workflow, Visualize, Shared Data, Help, User), and a 'Using 4.3 GB' indicator. Below the navigation bar, a message states: 'Galaxy was upgraded to 20.01 on 03/01. Clear session cookies if see a permission error.' The main content area is divided into three sections. On the left is a 'Tools' sidebar with a search bar and a list of tool categories: Get Data, Collection Operations, Lift-Over, Text Manipulation, Filter and Sort, Join, Subtract and Group, Convert Formats, Extract Features, Fetch Alignments/Sequences, Graph/Display Data, Statistics, SECIM Tools, GAIT-GM, UFRc Beta Tools, NGS: QC and manipulation, and NGS: RNA Analysis. The center section features the 'UF Research Information Technology' logo and 'Galaxy Documentation' with a bulleted list of updates and a 'Useful Documentation' section. On the right is a 'History' sidebar with a search bar and an 'Unnamed History' section showing '(empty)'. A red rectangle highlights the 'Data Libraries' option in the top navigation bar's dropdown menu.

Galaxy UFRc

Analyze Data Workflow Visualize Shared Data Help User

Using 4.3 GB

Galaxy was upgraded to 20.01 on 03/01. Clear session cookies if see a permission error.

Tools

search tools

Get Data

Collection Operations

Lift-Over

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

Extract Features

Fetch Alignments/Sequences

Graph/Display Data

Statistics

SECIM Tools

GAIT-GM

UFRc Beta Tools

NGS: QC and manipulation

NGS: RNA Analysis

UF Research Information Technology

Galaxy Documentation:

- UFRc Galaxy documentation in the Wiki.
- **Starting with 2018.05 Galaxy supports 'unlimited' file sizes for data uploads through the browser.** Use Large dataset import procedure. only for very large sets and collections of data.
- We can set a secondary account (e.g. ufgi, epi, etc.) as a default for your SLURM jobs if you have membership in that group. Please let us know via the RC Support System.
- Use the 'View or report this error' button in a history dataset for a failed analysis to submit a bug report or ask for support.
- Use the RC Support System to request installation of new tools.

Useful Documentation

- Galaxy Project Training Materials
- Using Dataset Collections
- Rule Based Data Uploader

Note: Web browser may not be the best choice for uploading large datasets. If you encounter any issues use the Large dataset import procedure documented above.

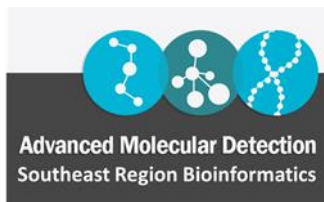
The Galaxy project is supported in part by NSF, NHGRI, and the Huck Institutes of the Life Sciences.

History

search datasets

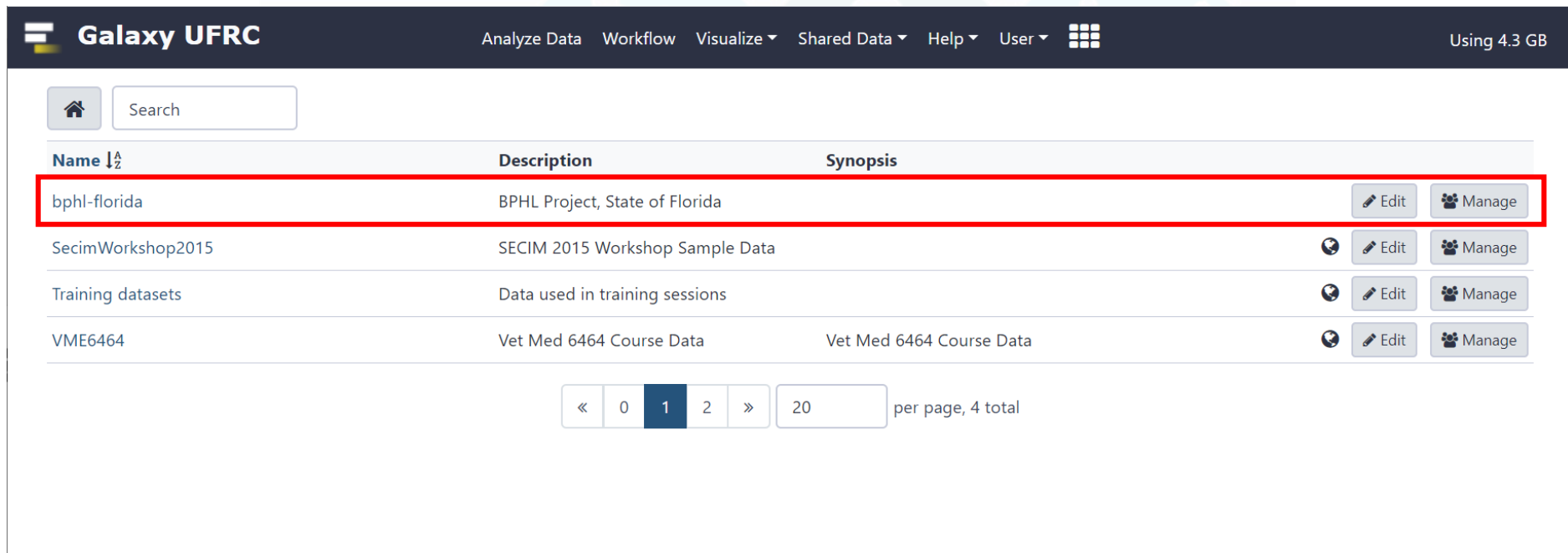
Unnamed History

(empty)



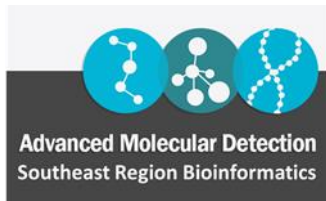
How to access your state's Galaxy data library

- You will see your state's private data library in the list
 - This library is private to users within your state group



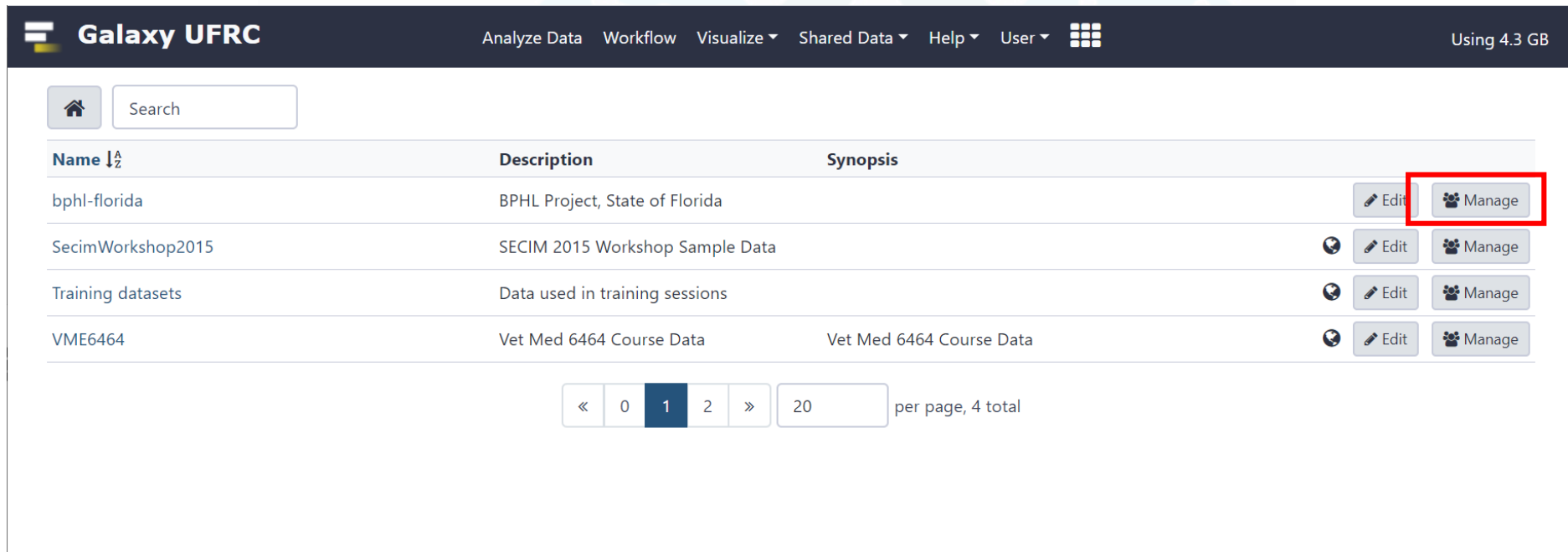
The screenshot shows the Galaxy UFRC interface. At the top, there is a navigation bar with links: Analyze Data, Workflow, Visualize, Shared Data, Help, User, and a grid icon. The text 'Using 4.3 GB' is on the right. Below the navigation bar is a search bar with a home icon and the text 'Search'. The main content area is a table with three columns: Name, Description, and Synopsis. The first row, 'bphl-florida', is highlighted with a red box. The second row is 'SecimWorkshop2015', the third is 'Training datasets', and the fourth is 'VME6464'. Each row has 'Edit' and 'Manage' buttons. At the bottom, there is a pagination control showing '1' of 4 pages, with a '20' per page, 4 total.

Name	Description	Synopsis	
bphl-florida	BPHL Project, State of Florida		Edit Manage
SecimWorkshop2015	SECIM 2015 Workshop Sample Data		Edit Manage
Training datasets	Data used in training sessions		Edit Manage
VME6464	Vet Med 6464 Course Data	Vet Med 6464 Course Data	Edit Manage



How to allow other users in your state to access the data library

- Click on “Manage”.

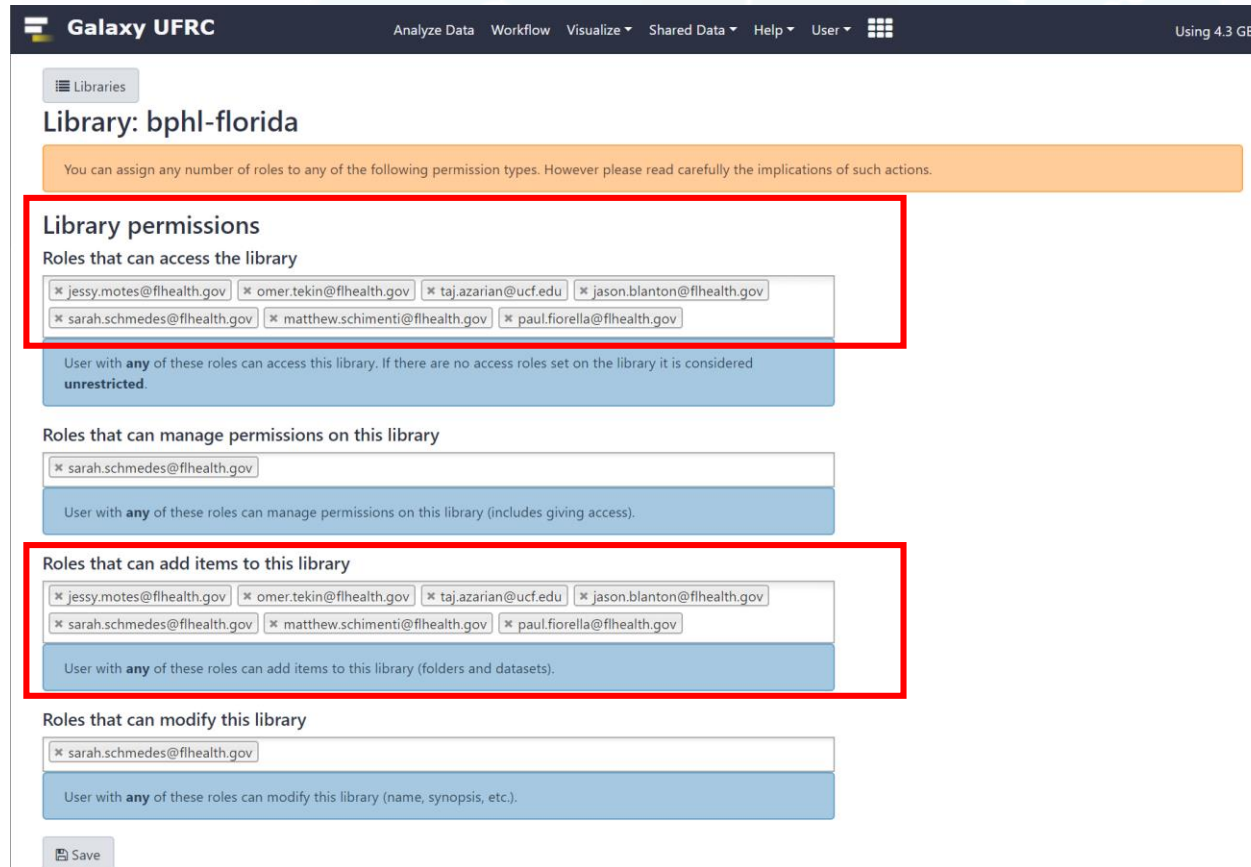


The screenshot shows the Galaxy UFRC interface. At the top, there is a navigation bar with links: Analyze Data, Workflow, Visualize, Shared Data, Help, User, and a grid icon. The text "Using 4.3 GB" is on the right. Below the navigation bar is a search bar with a home icon and the text "Search". The main content area displays a table with three columns: Name, Description, and Synopsis. The table lists four datasets: bphl-florida, SecimWorkshop2015, Training datasets, and VME6464. Each dataset row has two buttons: "Edit" and "Manage". The "Manage" button for the first dataset, "bphl-florida", is highlighted with a red rectangle. Below the table, there is a pagination control showing "20 per page, 4 total" and a set of buttons for navigating between pages (0, 1, 2, etc.).

Name	Description	Synopsis	Edit	Manage
bphl-florida	BPHL Project, State of Florida		Edit	Manage
SecimWorkshop2015	SECIM 2015 Workshop Sample Data		Edit	Manage
Training datasets	Data used in training sessions		Edit	Manage
VME6464	Vet Med 6464 Course Data	Vet Med 6464 Course Data	Edit	Manage

How to allow other users in your state to access the data library

- You will see your email address in each permission field since you are the admin for your group.
- Add your user email addresses to the “access” and “add items” permission fields.
 - You may already see a few email addresses for your users added. Galaxy Admins started this process for you, but you need to add any additional users that may not be added.
- Click “Save”.

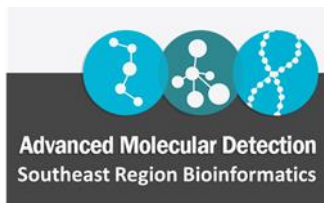


The screenshot shows the Galaxy UFRC interface for managing library permissions. The top navigation bar includes links for Analyze Data, Workflow, Visualize, Shared Data, Help, and User, along with a grid icon and a storage indicator 'Using 4.3 GB'. The main content area is titled 'Library: bphl-florida' and includes a warning message: 'You can assign any number of roles to any of the following permission types. However please read carefully the implications of such actions.'

The 'Library permissions' section is divided into four categories, each with a list of roles and a description of their permissions:

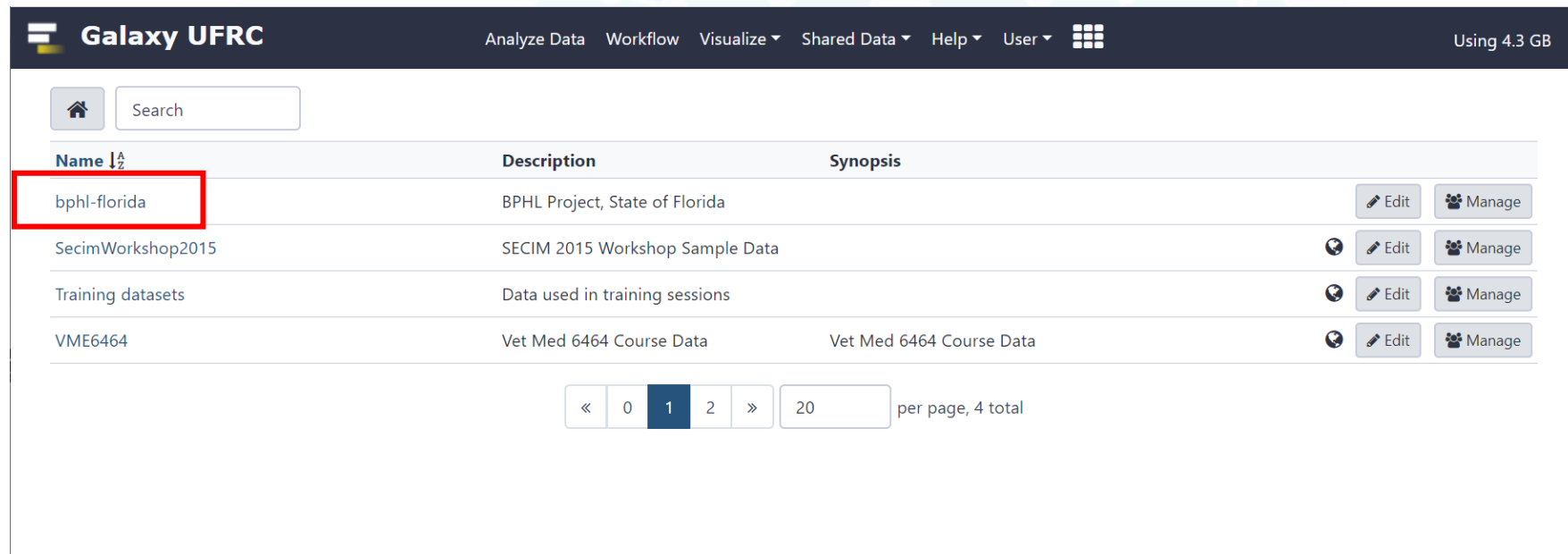
- Roles that can access the library:** This section is highlighted with a red box. It lists six email addresses: jessy.motes@flhealth.gov, omer.tekin@flhealth.gov, taj.azarian@ucf.edu, jason.blanton@flhealth.gov, sarah.schmedes@flhealth.gov, matthew.schimenti@flhealth.gov, and paul.fiorella@flhealth.gov. Below the list, a blue box states: 'User with **any** of these roles can access this library. If there are no access roles set on the library it is considered **unrestricted**.'
- Roles that can manage permissions on this library:** This section lists one email address: sarah.schmedes@flhealth.gov. Below the list, a blue box states: 'User with **any** of these roles can manage permissions on this library (includes giving access).'
- Roles that can add items to this library:** This section is also highlighted with a red box. It lists the same seven email addresses as the 'access' section. Below the list, a blue box states: 'User with **any** of these roles can add items to this library (folders and datasets).'
- Roles that can modify this library:** This section lists one email address: sarah.schmedes@flhealth.gov. Below the list, a blue box states: 'User with **any** of these roles can modify this library (name, synopsis, etc).'

At the bottom of the form, there is a 'Save' button.



How to add data to your data library

- Go into your data library by clicking on the name.

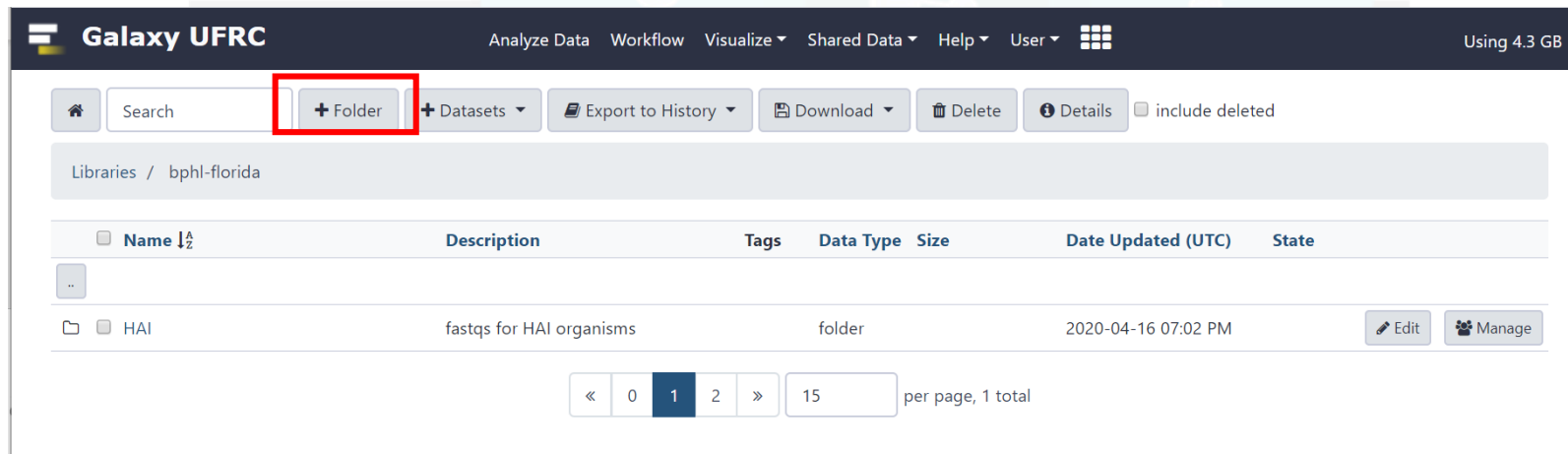


The screenshot shows the Galaxy UFRC data library interface. The top navigation bar includes links for Analyze Data, Workflow, Visualize, Shared Data, Help, and User, along with a grid icon and a status indicator 'Using 4.3 GB'. Below the navigation bar is a search bar with a home icon and a search input field. The main content area displays a table of data entries. The first entry, 'bphl-florida', is highlighted with a red box. The table has three columns: Name, Description, and Synopsis. The 'bphl-florida' entry has a description of 'BPHL Project, State of Florida' and a synopsis of 'BPHL Project, State of Florida'. The other entries are 'SecimWorkshop2015', 'Training datasets', and 'VME6464'. Each entry has 'Edit' and 'Manage' buttons. At the bottom of the table, there is a pagination control showing '0 1 2' and a '20 per page, 4 total' indicator.

Name	Description	Synopsis	Actions
bphl-florida	BPHL Project, State of Florida	BPHL Project, State of Florida	Edit Manage
SecimWorkshop2015	SECIM 2015 Workshop Sample Data	SECIM 2015 Workshop Sample Data	Edit Manage
Training datasets	Data used in training sessions	Data used in training sessions	Edit Manage
VME6464	Vet Med 6464 Course Data	Vet Med 6464 Course Data	Edit Manage

How to add data to your data library

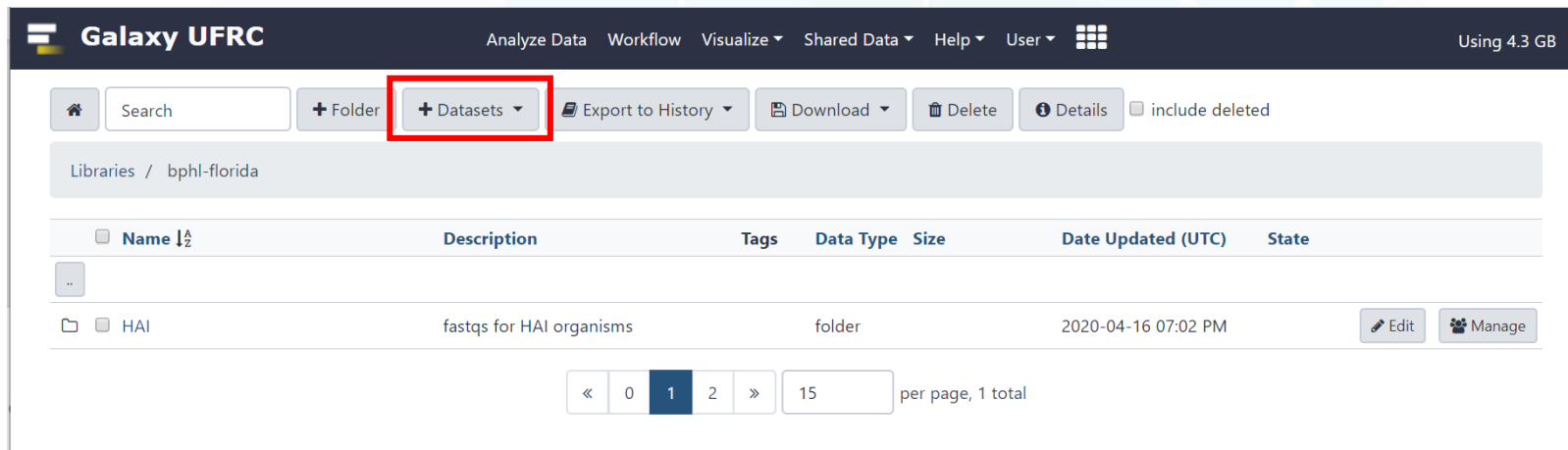
- You can add folders to organize your data or output files.
 - You do not need to manage the permissions for each folder. The permissions for your data library apply to all folders and data within the library.



The screenshot shows the Galaxy UFRC web interface. At the top, there's a navigation bar with links like 'Analyze Data', 'Workflow', 'Visualize', 'Shared Data', 'Help', 'User', and a 'Using 4.3 GB' status indicator. Below this is a toolbar with buttons for 'Search', '+ Folder' (highlighted with a red box), '+ Datasets', 'Export to History', 'Download', 'Delete', 'Details', and an 'include deleted' checkbox. The main content area shows a table of data items. The table has columns for 'Name', 'Description', 'Tags', 'Data Type', 'Size', 'Date Updated (UTC)', and 'State'. One item is visible: a folder named 'HAI' with the description 'fastqs for HAI organisms', data type 'folder', and updated on '2020-04-16 07:02 PM'. At the bottom, there's a pagination control showing '1' of 1 items per page.

How to add data to your data library

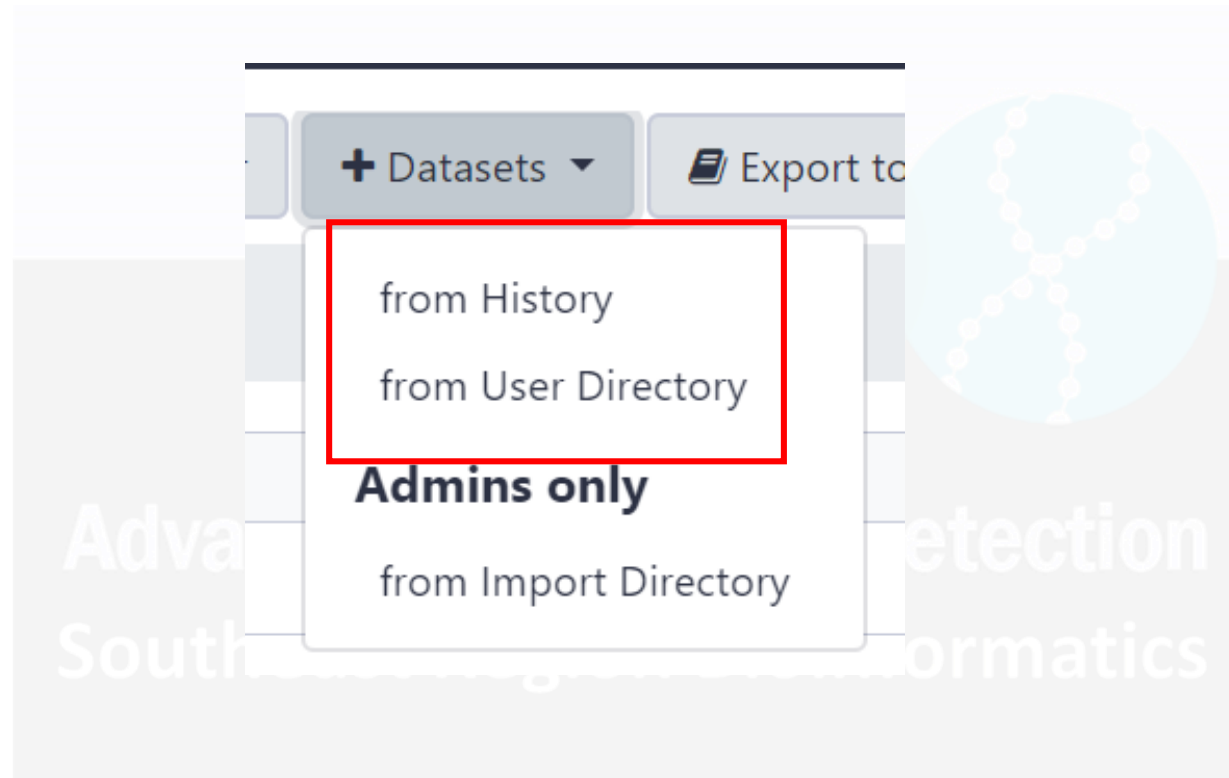
- To import or add data files to the data library, click “+ Datasets”.



The screenshot shows the Galaxy UFRC web interface. The top navigation bar includes links for 'Analyze Data', 'Workflow', 'Visualize', 'Shared Data', 'Help', and 'User'. The main content area has a search bar, a '+ Folder' button, and a '+ Datasets' button (highlighted with a red box). Below these are buttons for 'Export to History', 'Download', 'Delete', 'Details', and a checkbox for 'include deleted'. The breadcrumb trail shows 'Libraries / bphl-florida'. A table lists datasets with columns: Name, Description, Tags, Data Type, Size, Date Updated (UTC), and State. One dataset is listed: 'HAI' with description 'fastqs for HAI organisms', data type 'folder', and date '2020-04-16 07:02 PM'. At the bottom, there are pagination controls showing '1' of 1 page and '15' items per page.

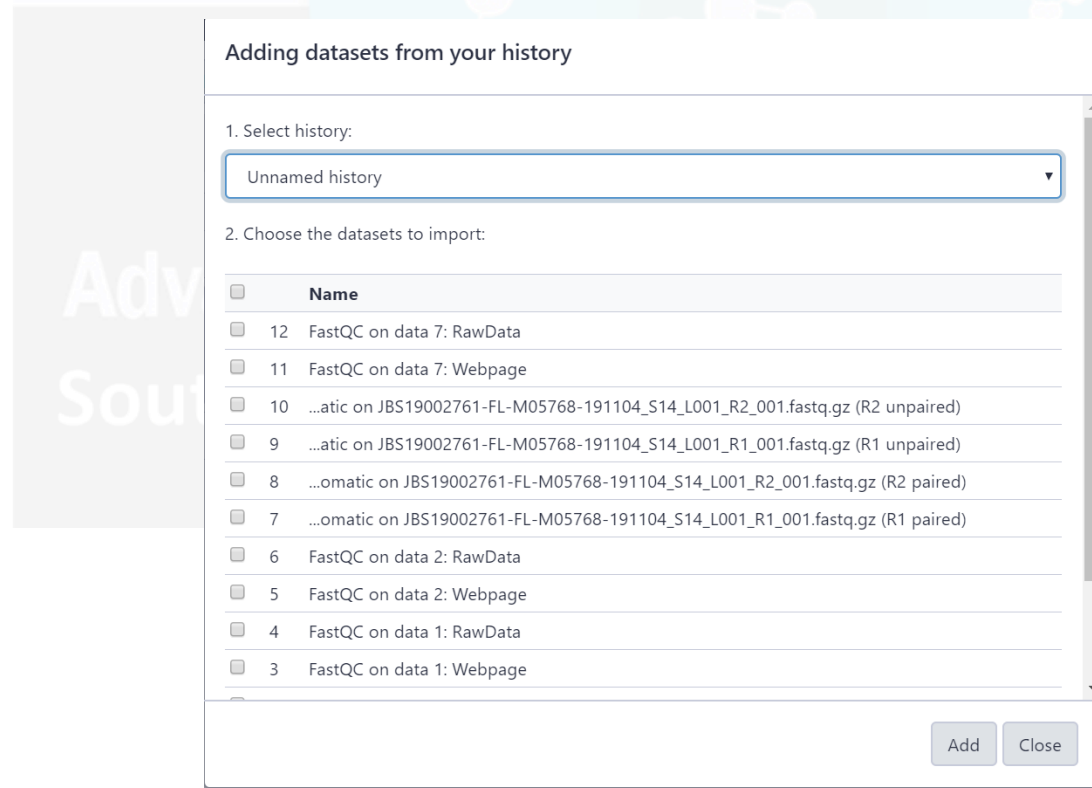
How to add data to your data library

- Select “from History” or “from User Directory.”



How to add data to your data library

- From History:
 - You can add data files from an existing History in Galaxy (i.e., data you have already imported into Galaxy or outputs from analyses within Galaxy).
 - Select the history that contains the file you want to add to your library in the drop down menu.
 - Then check the boxes to the left of the files you want to add.



Adding datasets from your history

1. Select history:

Unnamed history

2. Choose the datasets to import:

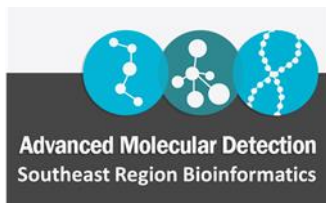
<input type="checkbox"/>	Name
<input type="checkbox"/>	12 FastQC on data 7: RawData
<input type="checkbox"/>	11 FastQC on data 7: Webpage
<input type="checkbox"/>	10 ...atic on JBS19002761-FL-M05768-191104_S14_L001_R2_001.fastq.gz (R2 unpaired)
<input type="checkbox"/>	9 ...atic on JBS19002761-FL-M05768-191104_S14_L001_R1_001.fastq.gz (R1 unpaired)
<input type="checkbox"/>	8 ...omatic on JBS19002761-FL-M05768-191104_S14_L001_R2_001.fastq.gz (R2 paired)
<input type="checkbox"/>	7 ...omatic on JBS19002761-FL-M05768-191104_S14_L001_R1_001.fastq.gz (R1 paired)
<input type="checkbox"/>	6 FastQC on data 2: RawData
<input type="checkbox"/>	5 FastQC on data 2: Webpage
<input type="checkbox"/>	4 FastQC on data 1: RawData
<input type="checkbox"/>	3 FastQC on data 1: Webpage

Add Close

How to add data to your data library

- From User Directory:
 - You can add data files from your user Galaxy data upload folder on HiPerGator (see previous hand-out, “Instructions to Upload Data into Galaxy via HiPerGator Server” to add data to this folder).
 - Select the files you want to add to your library.

Warning: There is currently a bug with the “User Directory” option right now. Galaxy admin staff is currently working to address this. In the meantime, only use the “from History” option. However, you can still use your Galaxy user upload directory on HiPerGator to upload data into a History (just not directly into a Data Library). I will let you know when this is fixed.



Please select folders or files

☒ Choose Files ☐ Choose Folders

All files you select will be imported into the current folder ignoring their folder structure.

☐ Link files instead of copying ☒ Convert line endings to POSIX ☐ Convert spaces to tabs

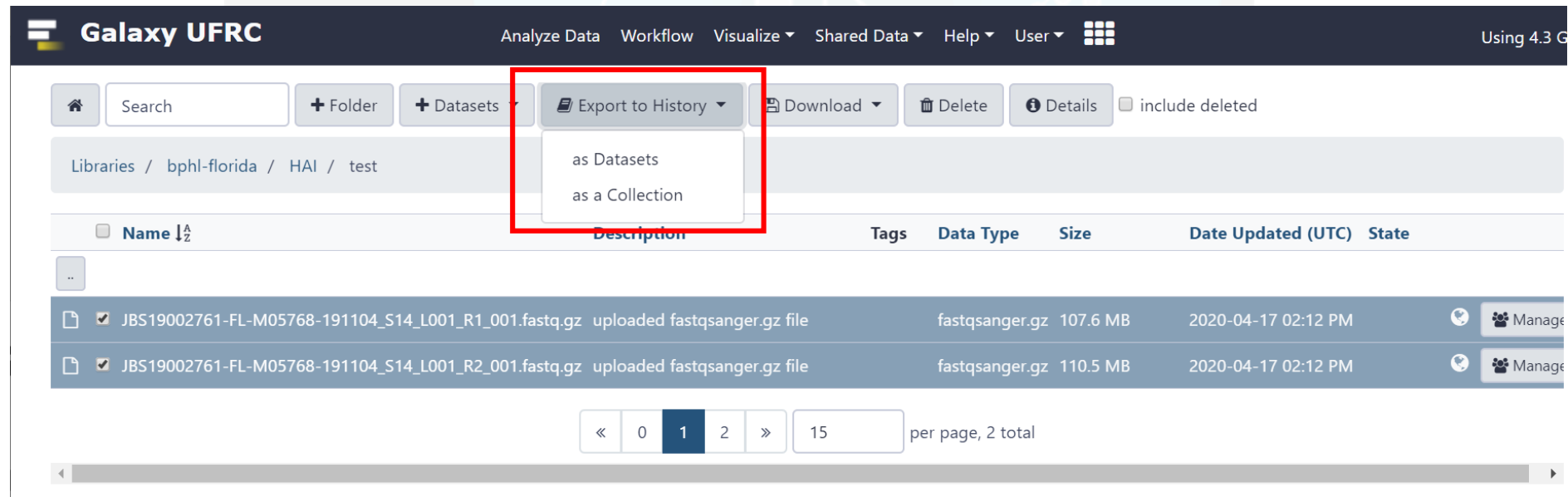
Select all Unselect all

- ☐ JBI20000342.gff
- ☐ JBI20000346.gff
- ☐ JBI20000354.gff
- ☐ JBI20000363.gff
- ☐ JBS19001187-FL-M06033-190703_S13_L001_R1_001.fastq.gz
- ☐ JBS19001187-FL-M06033-190703_S13_L001_R2_001.fastq.gz
- ☐ JBS19002761-FL-M05768-191104_S14_L001_R1_001.fastq.gz
- ☐ JBS19002761-FL-M05768-191104_S14_L001_R2_001.fastq.gz
- ☐ data

Import Close

How to use data stored your data library

- Once data is stored in your data library, you and other users from your state can use this data for any number of analyses by transferring a copy of the data into a History.
- Select the files you want to export to a History.
- Then click on “Export to History”. Select “as Datasets” or “as a Collection”.

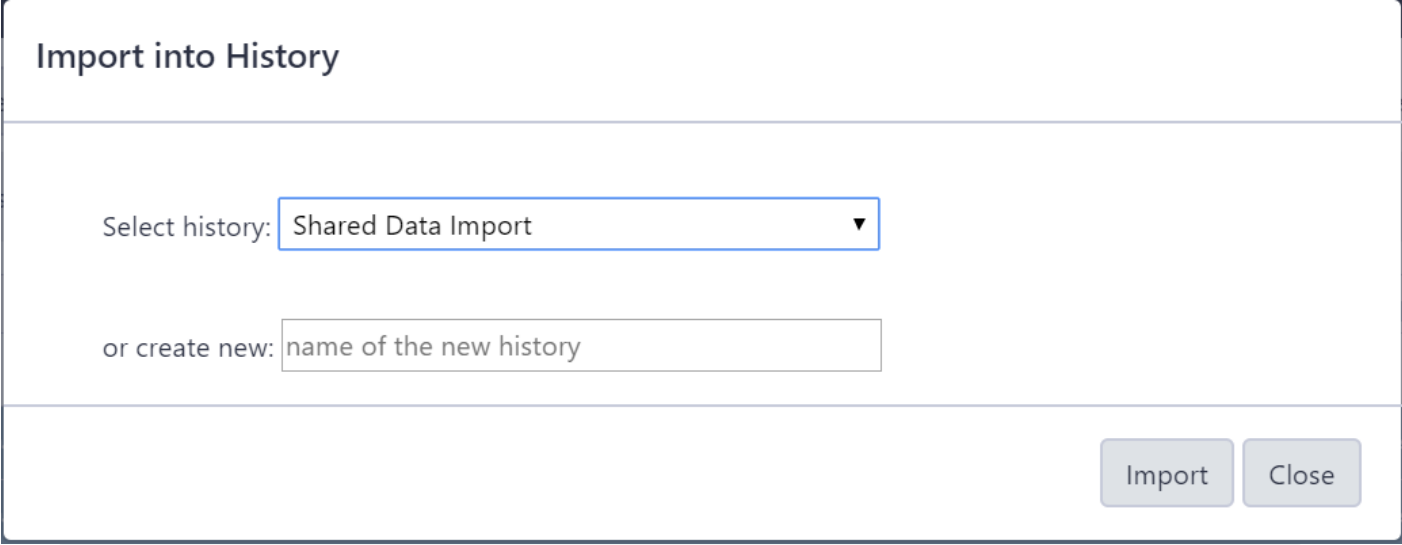


The screenshot displays the Galaxy UFRC web interface. At the top, the navigation bar includes links for 'Analyze Data', 'Workflow', 'Visualize', 'Shared Data', 'Help', and 'User', along with a grid icon and the text 'Using 4.3 G'. Below the navigation bar, there is a search bar and buttons for '+ Folder', '+ Datasets', and 'Export to History'. The 'Export to History' button is highlighted with a red rectangle, and its dropdown menu is open, showing two options: 'as Datasets' and 'as a Collection'. Below the dropdown, there is a table with columns for 'Name', 'Description', 'Tags', 'Data Type', 'Size', 'Date Updated (UTC)', and 'State'. The table contains two rows of data, both representing 'JBS19002761-FL-M05768-191104_S14_L001' files. At the bottom, there is a pagination bar showing '1' selected out of '2' items, with a '15' per page setting.

Name	Description	Tags	Data Type	Size	Date Updated (UTC)	State
JBS19002761-FL-M05768-191104_S14_L001_R1_001.fastq.gz	uploaded fastqsanger.gz file		fastqsanger.gz	107.6 MB	2020-04-17 02:12 PM	Manage
JBS19002761-FL-M05768-191104_S14_L001_R2_001.fastq.gz	uploaded fastqsanger.gz file		fastqsanger.gz	110.5 MB	2020-04-17 02:12 PM	Manage

How to use data stored your data library

- Either select an existing History you want to add your data to, or create a new History.
- Then click “Import”.



The dialog box is titled "Import into History". It contains two main sections. The first section is labeled "Select history:" and features a dropdown menu with the text "Shared Data Import" and a downward arrow. The second section is labeled "or create new:" and features a text input field with the placeholder text "name of the new history". At the bottom right of the dialog box, there are two buttons: "Import" and "Close".

Import into History

Select history: Shared Data Import ▼

or create new: name of the new history

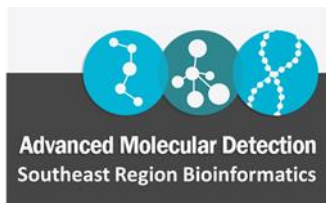
Import Close

How to use data stored your data library

- Once your data is added to a History, go back to your main screen by clicking “Analyze Data” at the top center of your screen.
- You will now see your data in your History you selected or created.

The screenshot displays the Galaxy UFRCC web interface. The top navigation bar includes links for 'Analyze Data', 'Workflow', 'Visualize', 'Shared Data', 'Help', and 'User', along with a 'Using 4.3 GB' indicator. The left sidebar contains a 'Tools' section with a search bar and a list of tool categories: Get Data, Collection Operations, Lift-Over, Text Manipulation, Filter and Sort, Join, Subtract and Group, Convert Formats, Extract Features, Fetch Alignments/Sequences, Graph/Display Data, Statistics, SECIM Tools, GAIT-GM, UFRCC Beta Tools, NGS: QC and manipulation, and NGS: RNA Analysis. The main content area features the 'UF Research Computing Information Technology' logo and 'Galaxy Documentation' with several bullet points. A 'History' panel on the right shows a search bar and a list of datasets. Two datasets are highlighted in a red box:

Dataset ID	File Name	Size
2: JBS19002761-FL-M0576 8-191104_S14_L001_R2_00	1.fastq.gz	218.08 MB
1: JBS19002761-FL-M0576 8-191104_S14_L001_R1_00	1.fastq.gz	



How to download files from your data library to your local computer

- You can also download data and output files from your data library to your local computer.
- Select the files you want to download.
- Click “Download”, and select the type of file compression for your download.

The screenshot shows the Galaxy UFRC web interface. At the top, there's a navigation bar with links like 'Analyze Data', 'Workflow', 'Visualize', 'Shared Data', 'Help', and 'User'. Below this, a search bar and several action buttons are visible. A red box highlights the 'Download' button, which has a dropdown menu open showing three options: '.tar.gz', '.tar.bz', and '.zip'. Below the buttons, there's a breadcrumb trail: 'Libraries / bphl-florida / HAI / test'. A table lists files with columns for Name, Description, Size, Date Updated (UTC), and State. Two files are listed, both starting with 'JBS19002761-FL-M05768-191104_S14_L001_R1_001.fastq.gz' and 'R2_001.fastq.gz'. At the bottom, there's a pagination control showing '1' of 2 pages, with a 'per page, 2 total' label.

Name	Description	Size	Date Updated (UTC)	State
JBS19002761-FL-M05768-191104_S14_L001_R1_001.fastq.gz	uploaded fastqsanger.gz file fastqsanger.gz	107.6 MB	2020-04-17 02:12 PM	
JBS19002761-FL-M05768-191104_S14_L001_R2_001.fastq.gz	uploaded fastqsanger.gz file fastqsanger.gz	110.5 MB	2020-04-17 02:12 PM	



Advanced Molecular Detection Southeast Region Bioinformatics

Questions???

BPHL-SEbioinformatics@flhealth.gov

Sarah Schmedes, PhD
BRR Lead, Southeast Region

Taj Azarian, PhD
Bioinformatician
Assistant Professor, UCF

Jason Blanton, PhD
Molecular Administrator
State Sequencing Coordinator