



A - Training on Galaxy: Metabarcoding

October 2022 - webinar

GALAXY Practice

LUCAS AUER, MARIA BERNARD, LAURENT CAUQUIL, VINCENT DARBOT, MAHENDRA MARIADASSOU, GÉRALDINE PASCAL & OLIVIER RUÉ















Objectives of the 1st part

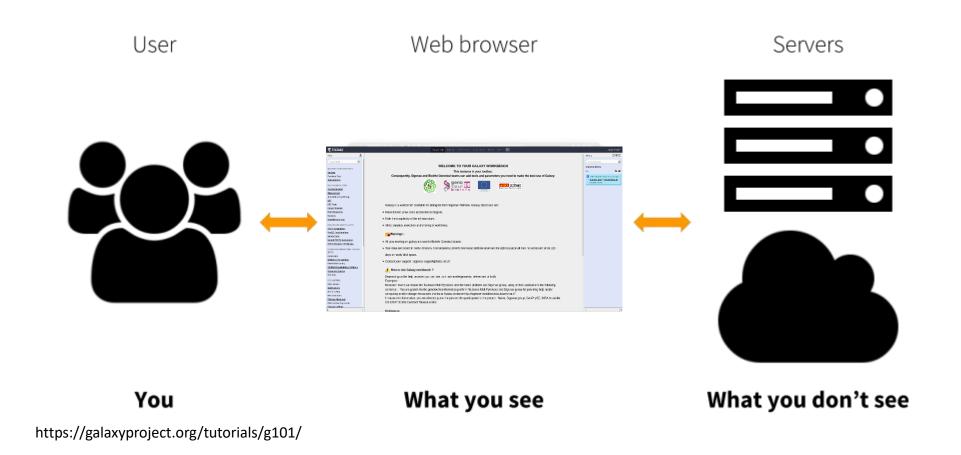
- Learn the basics of Galaxy
- Prepare the 2 datasets for FROGS formation

What is Galaxy?

Galaxy software framework

- Galaxy is an overlay that allows researchers without computer expertise to perform computational analyses online, automate them and share the results easily.
- Developed by the Galaxy Team Project.
- Relies on a computational infrastructure (Server).
- Open source

Galaxy software framework



Where to use Galaxy?

 Galaxy software can be installed and run on powerful server farms (Cluster) or on your individual PC.





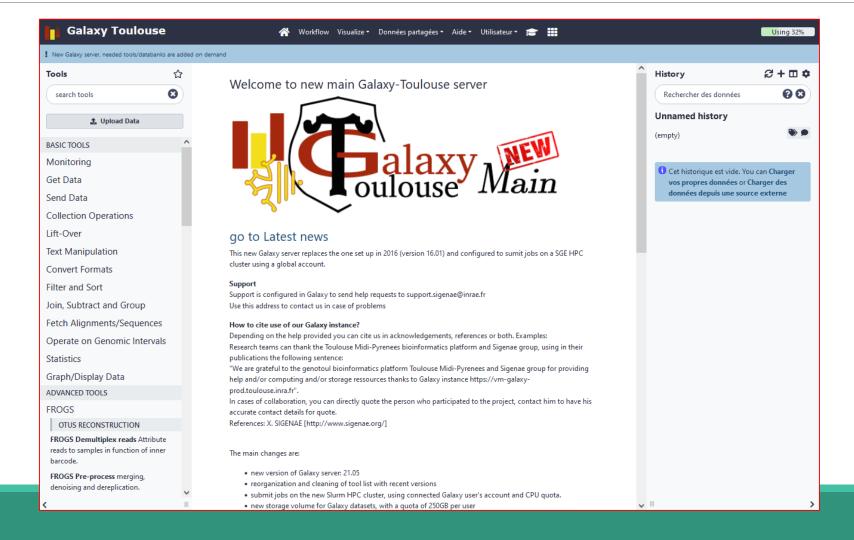
Where to use Galaxy?

- The Toulouse Galaxy platform: https://vm-galaxy-prod.toulouse.inrae.fr
- Institut Français Bioinformatique IFB (national server): https://metabarcoding.usegalaxy.fr
- Migale Plateform: galaxy.migale.inra.fr
- Others...
- ⚠ The different platforms of Galaxy are not connected together
- Galaxy is installed on many clusters across the world.
- Some tools are in <u>our</u> Galaxy platform but not in <u>other</u> platforms.
- Your data is not shared with other Galaxy platforms than ours.

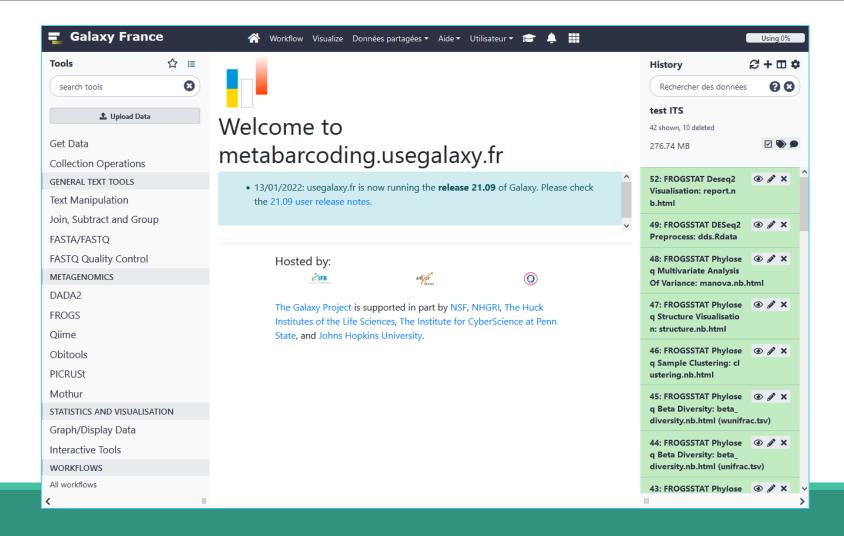
Exemple of Galaxy platforms



Exemple of Galaxy platforms



Exemple of Galaxy platforms



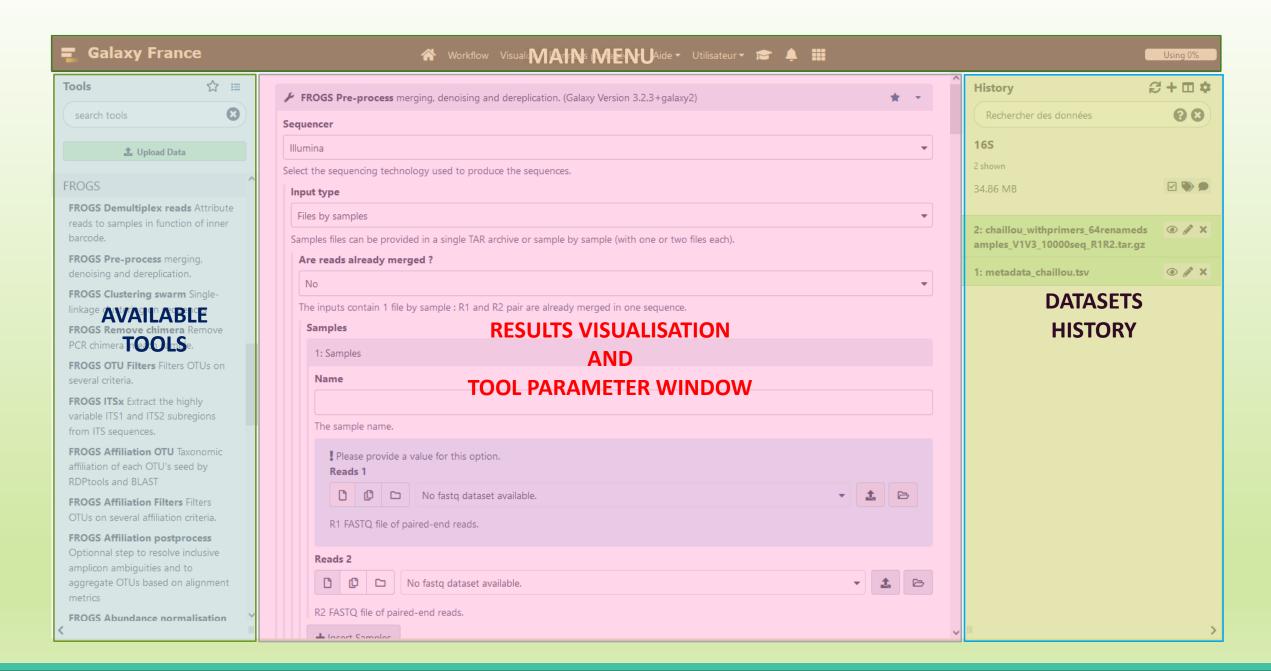
Practice:

CONNECT TO GALAXY WORKBENCH

Exercise

Enter your login and your password

Welcome to Galaxy, please log in
Public Name or Email Address
Password
Forgot password? Click here to reset your password. Login
Don't have an account? Registration for this Galaxy instance is disabled. Please contact an administrator for assistance.



Vocabulary of Galaxy

Tools:

- A tool has a function which is explained when you click on it.
- Each Galaxy platform has its own tools.

Datasets:

- A dataset is a file, uploaded to Galaxy by you or produced by a tool.
- Be careful: a dataset has a datatype.

Histories:

- A tool generates datasets and these datasets are stored in the current history.
- Everything is permanently saved.
- If you log off your computer or browser, it's ok, everything will keep running and be saved!

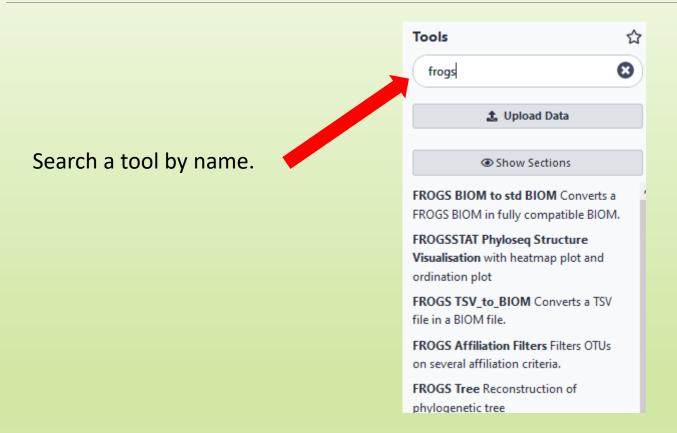
Pratice:

DISCOVER GALAXY

Exercise

- 1. Visit the Galaxy Platform.
- 2. Look at the tool list.
- 3. Display only FROGS tools.
- 4. Display all tools concerning fastq files.

Exercise





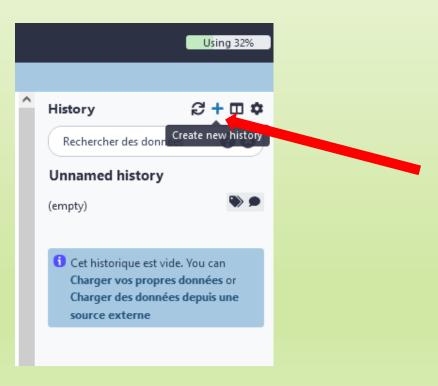
Manipulate Histories

Practice:

CREATE THE 2 HISTORIES NEEDED FOR THE FROGS FORMATION

To create a new history:

Click on the plus.



To rename a history:

Click on the "Unnamed history" (at the top).

Enter « 16S ».

Hit « Enter » to validate.

Don't use special characters or accents!

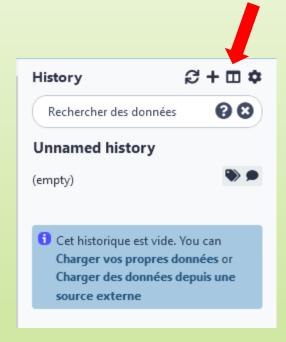


Exercise

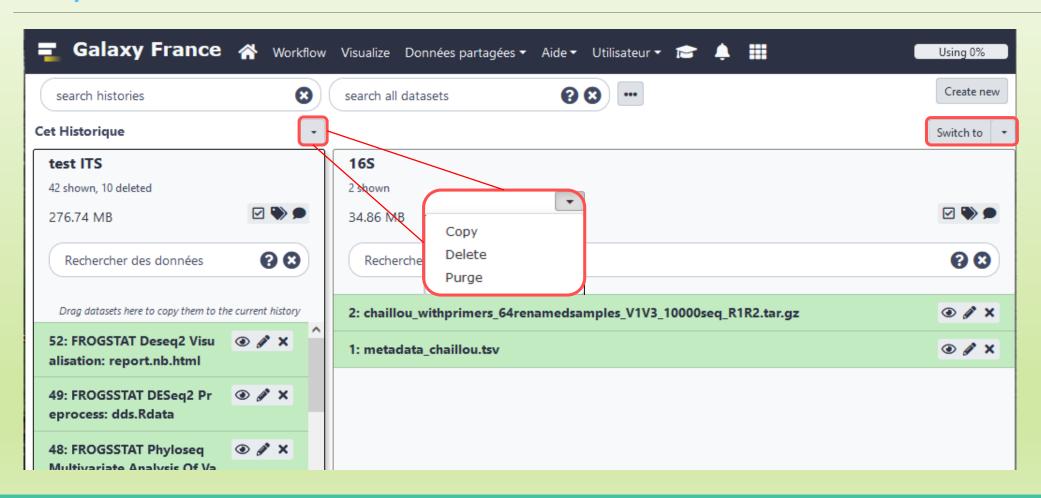
- Create another history named ITS
- Switch to the history named « 16S » as current history.
- Go back to the main interface.

How to list all histories?

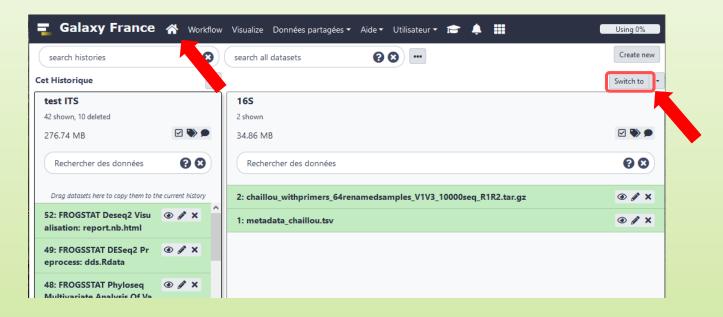
To view all histories, click on this icon.



Explore the « View all histories » section



Switch current history



- Switch to the history named « 16S » as current history.
- Click on "Home" to go back to the main interface.

Data import

How to import your data to Galaxy?

- Several ways to upload your data to Galaxy:
 - From your computer



By URL



From Genotoul Bioinfo clusters



Shared by other users of Galaxy



Formation data

All the training data are online on:
 http://genoweb.toulouse.inra.fr/~formation/15 FROGS/Webinar data/



Above all: how create an archive?

To learn how create an archive

http://genoweb.toulouse.inra.fr/~formation/15 FROGS/Webinar data /How create an archive.pdf

Create an archive is always possible!

It is our recommendation.

How to import your data to Galaxy?

- Several ways to upload your data to Galaxy:
 - From your computer



drag-and-drop files

By URL



Paste an url

From Genotoul Bioinfo clusters



Shared by other users of Galaxy



Practice

UPLOAD FILE FROM AN URL

How to import your data to Galaxy?

Several ways to upload your data to Galaxy:





By URL



From Genotoul Bioinfo clusters

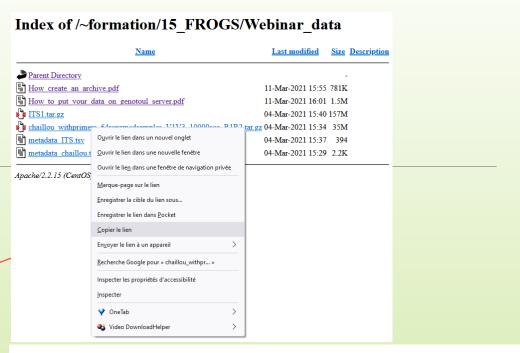


Shared by other users of Galaxy

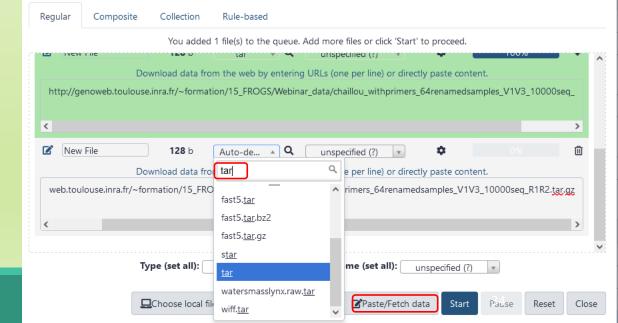


16S history creation

- In your current history "16S"
- 2. Go to **Get Data** > **Upload File from your computer**
- 3. Click on Paste/Fetch Data
- 4. Click right on the name file
- 5. Copy the address of the link
- 6. Paste the address
- 7. Change the data type!
- 8. Click on Start
- 9. You can put one address per line for multiple uploads
- 10. Start again with the second file



Download from web or upload from disk



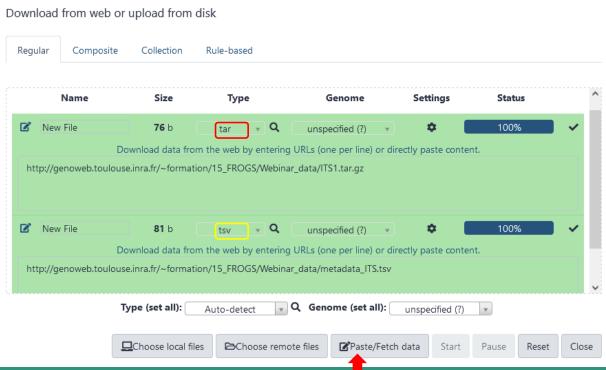
Create and fill yourself the ITS history

Get data form here:

http://genoweb.toulouse.inra.fr/~formation/15 FROGS/Webinar data/

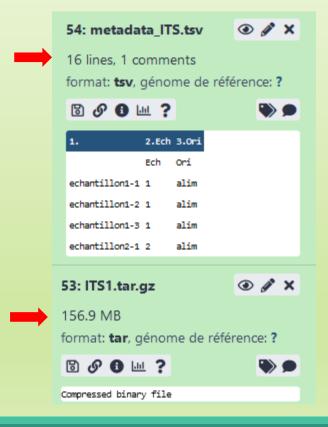






Create and fill yourself the ITS history

You have to obtain your second history:



Share a History

Share a history

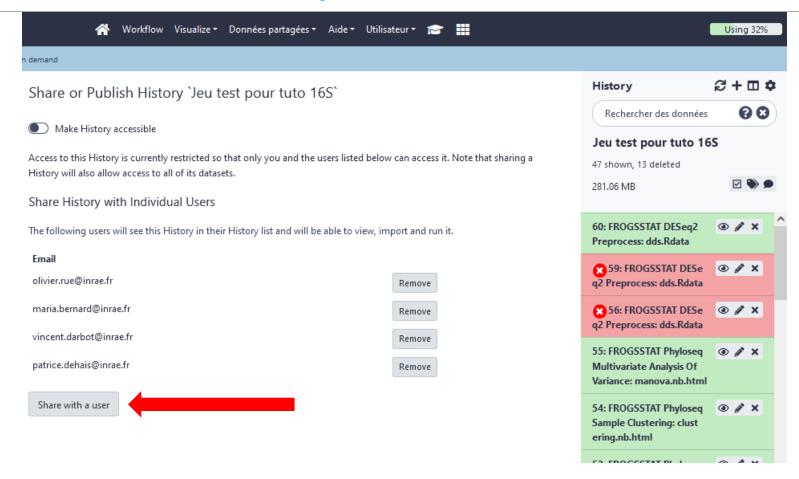
- You can share a history with another Galaxy user:
 - For working with your colleague.
 - For support, so we can help you better and faster.
- You can import shared history to your account too.

Share an history

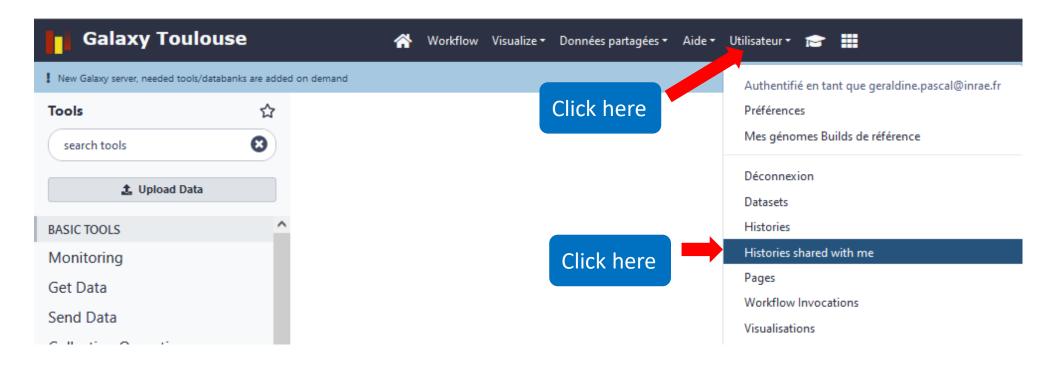


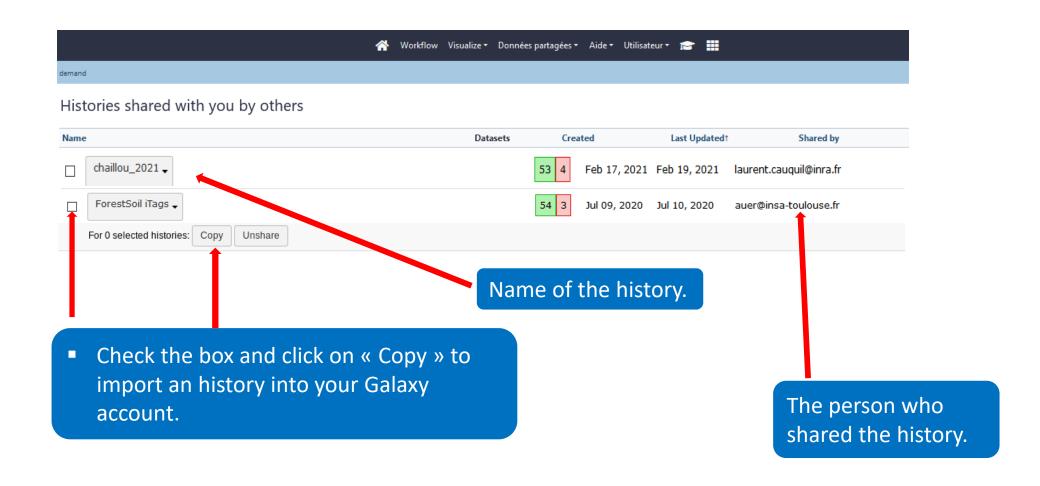


Share an history



Import a shared history





How manipulate datasets

To rename a dataset

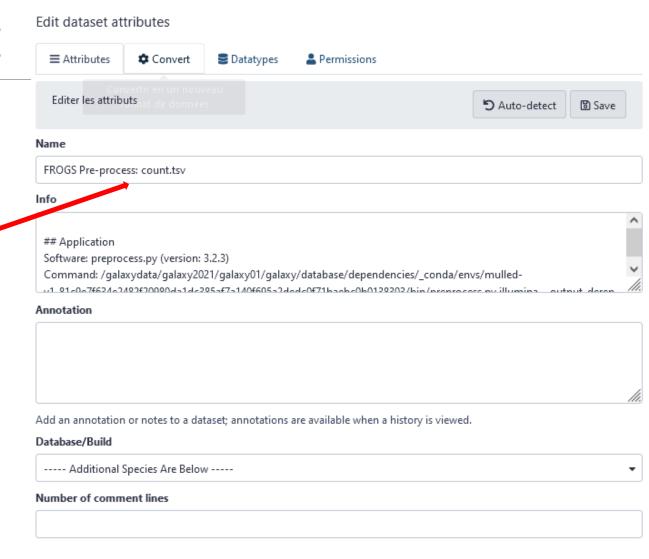
Switch to ITS history



Click here to display attributes and change the name.

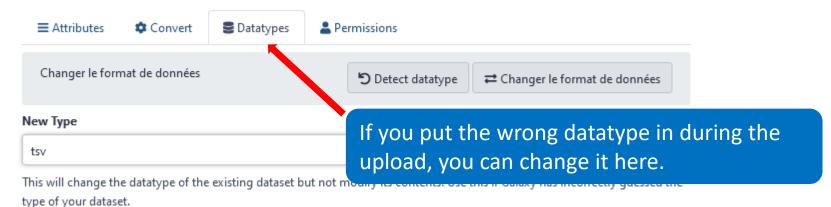
To rename a dataset

Change the name here

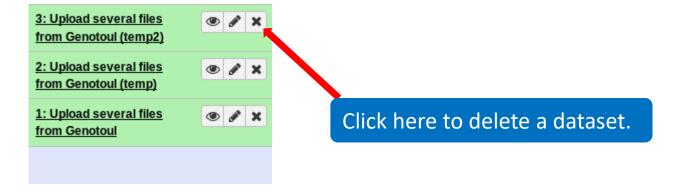


To change the datatype

Edit dataset attributes



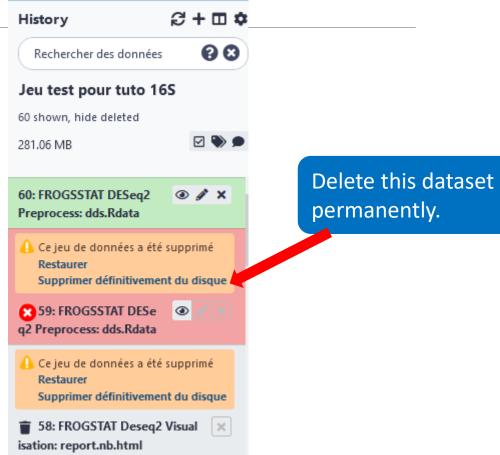
To delete a dataset



To delete a dataset



Nothing is deleted in Galaxy until you delete it permanently (=purge).



How to cite Galaxy workbench?

For Genotoul Bioinfo:

- Research teams can thank the Toulouse Occitanie bioinformatics platform and Sigenae group, using in their publications the following sentence: "We are grateful to the genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul, doi: 10.15454/1.5572369328961167E12) and Sigenae group for providing help and/or computing and/or storage ressources thanks to Galaxy instance http://sigenae-workbench.toulouse.inra.fr".
- In cases of collaboration, you can directly quote the person who participated to the project: Name, Sigenae group, GenPhySE, INRAE Auzeville CS 52627 31326 Castanet Tolosan cedex.

For IFB:

We are grateful to the Institut Français de Bioinformatique (ANR-11-INBS-0013) for providing help and/or computing and/or storage resources.

For Migale:

- Research teams that have used Migale services are expected to thank the Migale bioinformatics facility in the publications of their analyses with the following sentence: "We are grateful to the INRAE MIGALE bioinformatics facility (MIGALE, INRAE, 2020. Migale bioinformatics Facility, doi: 10.15454/1.5572390655343293E12) for providing help and/or computing and/or storage resources"
- In cases of collaboration, you can directly quote the person who participated to the project with the two following affiliations: Université Paris-Saclay, INRAE, MalAGE, 78350, Jouy-en-Josas, France Université Paris-Saclay, INRAE, BioinfOmics, MIGALE bioinformatics facility, 78350, Jouy-en-Josas, France