



A - Training on Galaxy: Metabarcoding

June 2021 - webinar

GALAXY Practice

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

Objectives of the 1st part

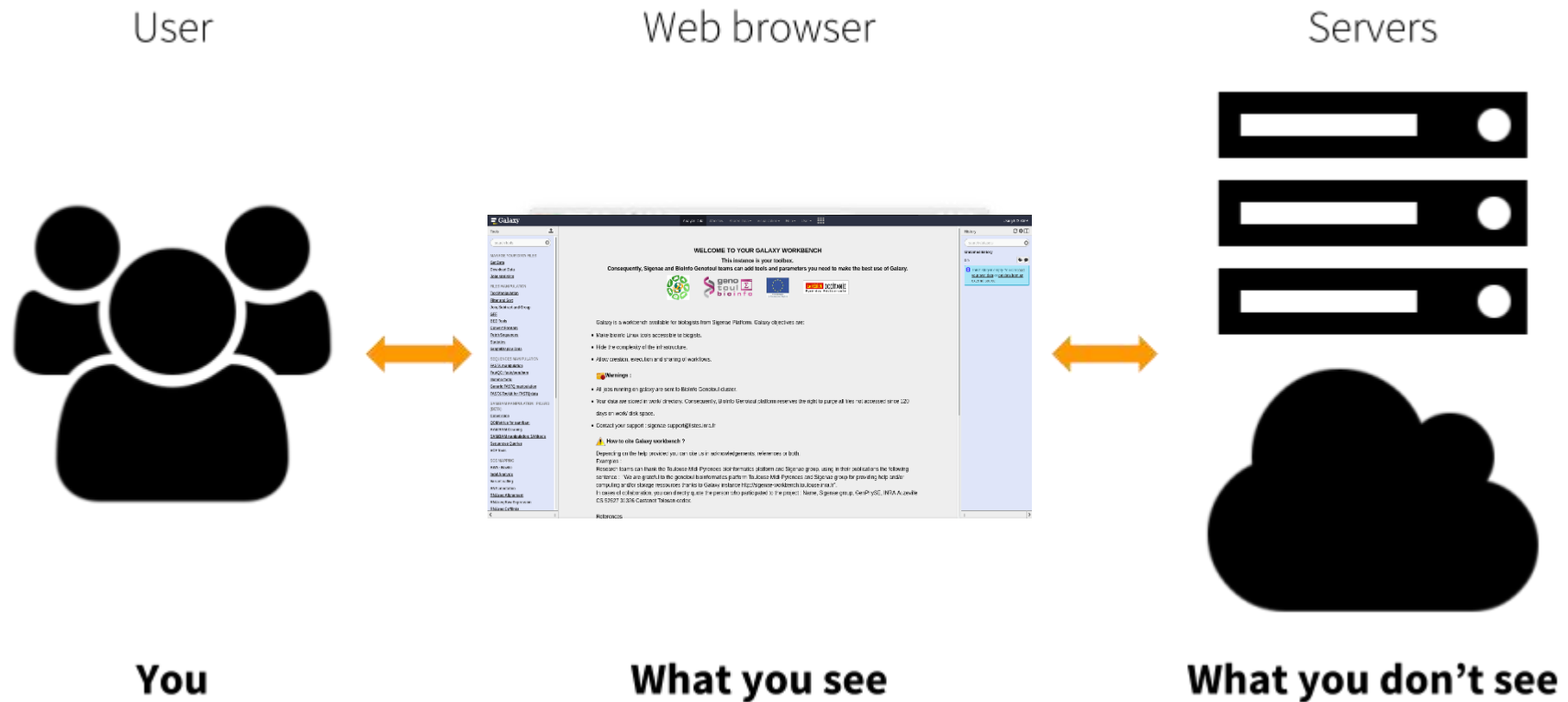
- Learn the basics of Galaxy
- Prepare the 2 datasets for FROGS formation

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



<https://galaxyproject.org/tutorials/g101/>

Where to use Galaxy?

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

- Genotoul Bioinfo



- Cluster: Many computers tightly connected that work together
- High performance computer:
 - More than 5000 cores
 - 34 TB of RAM
 - More than 1 Peta Byte (1024 TB) of hard drive



Where to use Galaxy?

- The Toulouse Galaxy platform is <https://vm-galaxy-prod.toulouse.inra.fr/galaxy/>

⚠ The different platforms of Galaxy are not connected together

- Galaxy is installed on many clusters across the world.
- Some tools are in our Galaxy platform but not in other platforms.
- Your data is not shared with other Galaxy platforms than ours.

Exemple of 2 INRAE Galaxy platforms

Galaxy Migale Analyse de données Workflow Visualize Données partagées Aide Authentification

Tools search tools

Get Data
Collection Operations
BASIC TOOLS
Text Manipulation
Convert Formats
Filter and Sort
Join, Subtract and Group
Statistics
Multiple Alignments
ncbi_blast
NGS TOOLS
Quality control
FASTQ manipulation
Mapping
RNAseq
Variant calling
Variant analyses
Migale Tools
SEQUENCE ANALYSIS TOOLS
GENOME ANALYSIS TOOLS
Genome annotation
METAGENOMICS TOOLS
Metabarcoding
METAPROTEOMICS TOOLS
Send Data
Lift-Over

migale
Welcome to the Migale Galaxy instance!

Global load of the nodes dedicated to Galaxy jobs: 5%

What's new

- 28th January 2021: Added [Parnsnp](#) in Section *Phylogeny*.
- 28th January 2021: Added [ECTyper](#) in Section *Genome Annotation*.
- 14th January 2021: Added [Filter SPAdes Output](#) in Section *Assembly*.
- 27th November 2020: Added [Seqtk](#) in Section *FAST/FASTQ Manipulation*.
- 27th November 2020: Added [SeqSero](#) in Section *Genome Annotation*.
- 23d November 2020: Added [staram](#) in Section *Genome Annotation*.
- 2nd October 2020: Added [Shovill](#) in Section *Assembly*.
- 27th July 2020: Added [Progressive Mauve](#) in Section *Multiple Alignements*.
- 27th July 2020: Added [Mummer4](#) in Section *Multiple Alignements*.
- 24th July 2020: Added [NCBI accession Download](#) in Section *Get Data*.

Exemple of 2 INRAE Galaxy platforms

Galaxy | Analyze Data | Workflow | Shared Data | Visualization | Help | User | Using 0%

Tools | search tools

MANAGE YOUR DATA FILES
[Get Data](#)
[Download Data](#)
[Jobs statistics](#)

FILES MANIPULATION
[Text Manipulation](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[GFF](#)
[BED Tools](#)
[Convert Formats](#)
[Fetch Sequences](#)
[Statistics](#)
[Graph/Display Data](#)

SEQUENCES MANIPULATION
[FASTA manipulation](#)
[FastQC: fastq/sam/bam](#)
[Illumina fastq](#)
[Generic FASTQ manipulation](#)
[FASTX-Toolkit for FASTQ data](#)

SAMBAM MANIPULATION : PICARD (BETA)
[Conversion](#)
[QC/Metrics for sam/bam](#)
[BAM/SAM Cleaning](#)
[SAMBAM manipulation: SAMtools](#)
[Sequences Queries](#)
[VCF Tools](#)

SGS MAPPING
[BWA - Bowtie](#)
[Indel Analysis](#)
[Variant calling](#)
[SNP annotation](#)
[RNAseq Alignment](#)
[RNAseq Raw Expression](#)
[RNAseq Cufflinks](#)

Genotoul Bioinfo | European Union | LA REGION OCCITANIE Pyrénées-Méditerranée

Welcome on Galaxy Sigenae / BioInfo Genotoul.
Contact your support : support.sigenae@inra.fr

Ready to use Workflows

- 1/ [FROGS - Find, Rapidly, Otus with Galaxy Solution](#)
- 2/ [Tax4Fun workflow](#)
- 3/ [Function Table for Tax4Fun matrix workflow](#)

Galaxy News

- 1/ How to use [datasets collections](#) ?
- 2/ How to [upload several files in Galaxy](#) ?
- 3/ New tools : [Tax4Fun \(manual\)](#), [VIP \(manual\)](#), [Salmon](#), [multiQC](#).
- 4/ Tools availables for [virologie analyses](#)
- 5/ [Virology tools](#)
- 6/ How to use [Function Table tool](#) ?
- 7/ How to [export](#) and [import](#) your histories ?

E-learning

<https://inra.classilio.com/Login>

2 e-learning availables : Galaxy initiation and Quality sequences analysis

Training resources : [BioInfo Genotoul / Sigenae](#)

History | search datasets

Data Collection
3 shown
33 b

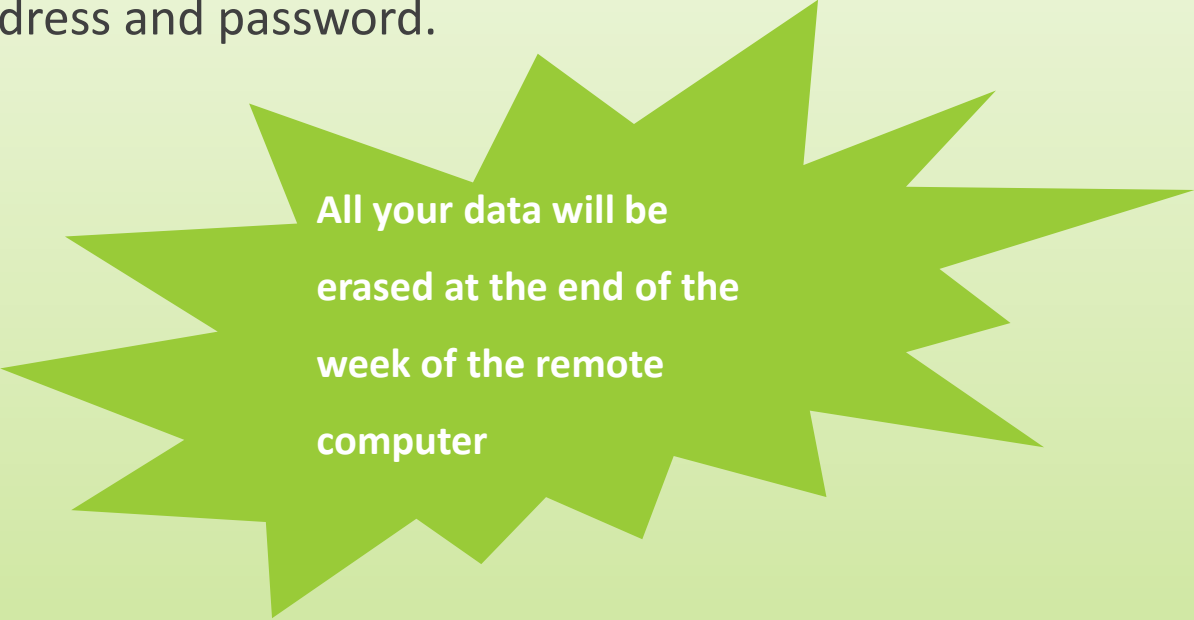
- 3: [seq1.fasta](#)
- 2: [seq2.fasta](#)
- 1: [seq3.fasta](#)

Practice:

CONNECT TO OUR GALAXY WORKBENCH

Exercise

Ask trainers for the address and password.




**All your data will be
erased at the end of the
week of the remote
computer**

Exercise

During this training, even you have a personal account, don't use it !
We use training accounts for more facilities


If you don't have a personal account, you can request a personal account later on :
<http://bioinfo.genotoul.fr/index.php/ask-for/create-an-account/>



All your data will be
erased at the end of the
week on the training
account.

Exercise

Our Galaxy platform is: <https://vm-galaxy-prod.toulouse.inra.fr/galaxy>




Before starting, check if your browser is not already connected to a Galaxy account. If this is the case, log out or use a private browser window!
Risk of data duplication

Exercise

Our Galaxy platform is: <https://vm-galaxy-prod.toulouse.inra.fr/galaxy/>

Be careful, to fully login you must enter your credentials twice:

- The first time in this pop-up window:



Authentication requise

 Le site <http://galaxy-workbench.toulouse.inra.fr> demande un nom d'utilisateur et un mot de passe. Le site indique : « Please enter your Genotoul LDAP password »

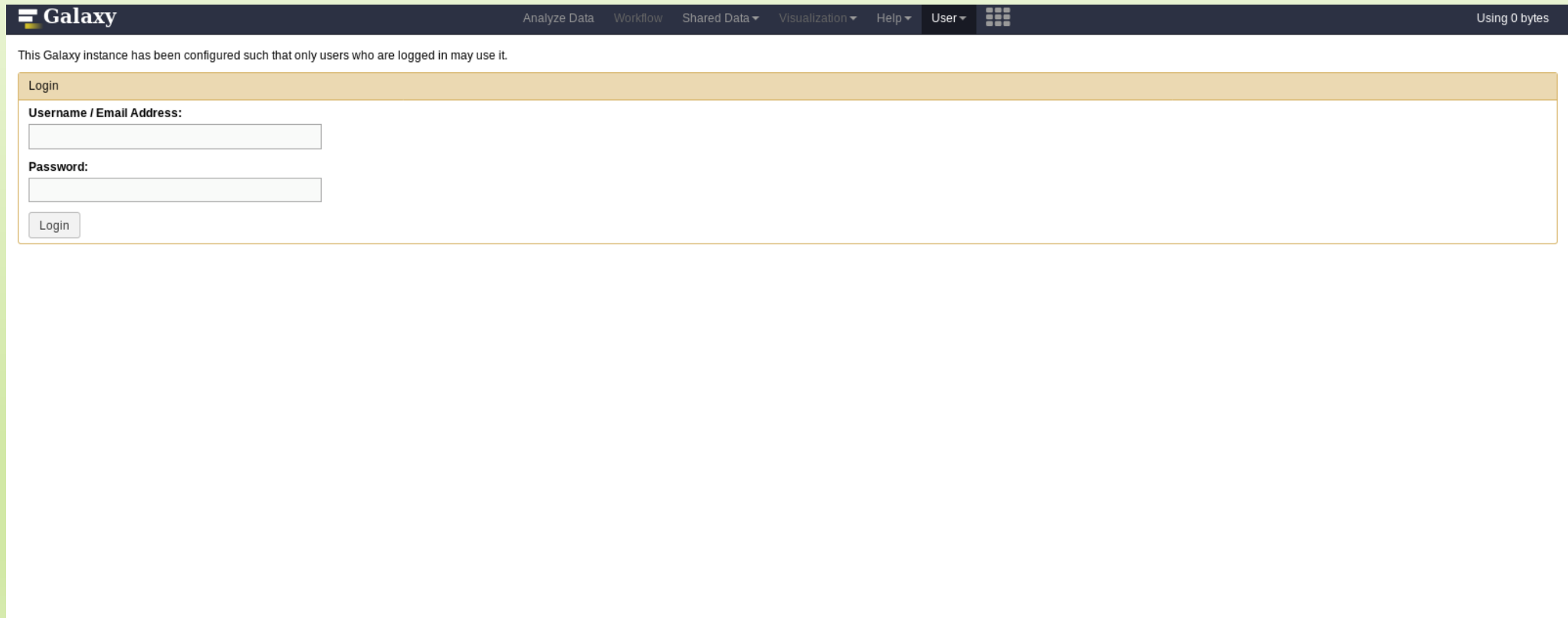
Utilisateur :

Mot de passe :

Annuler OK

Exercise

- And a second time, in the browser:



The screenshot shows the Galaxy web interface. At the top, there is a dark navigation bar with the Galaxy logo on the left and several menu items: 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. On the far right of the navigation bar, it says 'Using 0 bytes'. Below the navigation bar, a message states: 'This Galaxy instance has been configured such that only users who are logged in may use it.' Below this message is a 'Login' section with a light orange header. Inside the login section, there are two input fields: the first is labeled 'Username / Email Address:' and the second is labeled 'Password:'. Below the password field is a 'Login' button.

Galaxy
Analyze Data Workflow Shared Data Visualization Help User
Using 0%

Tools

search tools

MANAGE YOUR DATA FILES

[Get Data](#)

[Download Data](#)

[Jobs statistics](#)

FILES MANIPULATION

[Text Manipulation](#)

[Filter and Sort](#)

[Join, Subtract and Group](#)

[GFF](#)

[BED Tools](#)

[Convert Formats](#)

[Fetch Sequences](#)

[Statistics](#)

[Graphical Analysis](#)

SEQUENCES MANIPULATION

[FASTA manipulation](#)

[FastQC: fastq/sam/bam](#)

[Illumina fastq](#)

[Generic FASTQ manipulation](#)

[FASTX-Toolkit for FASTQ data](#)

SAMBAM MANIPULATION : PICARD (BETA)

[Conversion](#)

[QC/Metrics for sam/bam](#)

[BAM/SAM Cleaning](#)

[SAMBAM manipulation: SAMtools](#)

[Sequences Queries](#)

[VCF Tools](#)

SGS MAPPING

[BWA - Bowtie](#)

[Indel Analysis](#)


[Variant calling](#)

[SNP annotation](#)

[RNAseq Alignment](#)

[RNAseq Raw Expression](#)

[RNAseq Cufflinks](#)



How to cite Galaxy workbench ?

Depending on the help provided you can cite us in acknowledgements, references or both.


Examples :

Research teams can thank the Toulouse Midi-Pyrenees bioinformatics platform and Sigeneae group, using in their publications the following sentence : "We are grateful to the genotoul bioinformatics platform Toulouse Midi-Pyrenees and Sigeneae group for providing help and/or computing and/or storage resources thanks to Galaxy Instance <http://sigeneae-workbench.toulouse.inra.fr>".

In cases of collaboration, you can directly quote the person who participated to the project : Name, Sigeneae group, GenPhySE, INRA Auzeville CS 52627 31326 Castanet Tolosan cedex.


References
X. SIGENAE [<http://www.sigeneae.org/>]

RESULTS VISUALISATION —AND— TOOL PARAMETER WINDOW




Ready to use Workflows

- 1/ FROGS - Find, Rapidly, Otus with Galaxy Solution
- 2/ Tax4Fun workflow
- 3/ Function Table for Tax4Fun matrix workflow



Galaxy News

- 1/ How to use [datasets collections](#) ?
- 2/ How to [upload several files in Galaxy](#) ?
- 3/ New tools : Tax4Fun ([manual](#)), VIP ([manual](#)), Salmon, multiQC.
- 4/ Tools availables for [virologie analyses](#)
- 5/ Virology tools
- 6/ How to use [Function Table tool](#) ?



E-learning

<https://inra.classilio.com/Login>

2 e-learning availables : Galaxy initiation and Quality sequences analysis

History

search datasets

Data Collection
3 shown

33 b

3: seq1.fasta

2: seq2.fasta

1: seq3.fasta

DATASETS HISTORY

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:

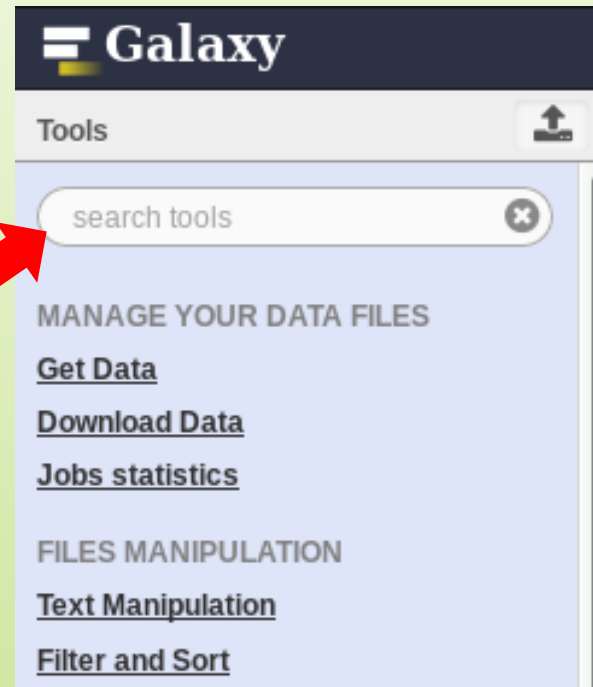
DISCOVER GALAXY

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

Search a tool by name.



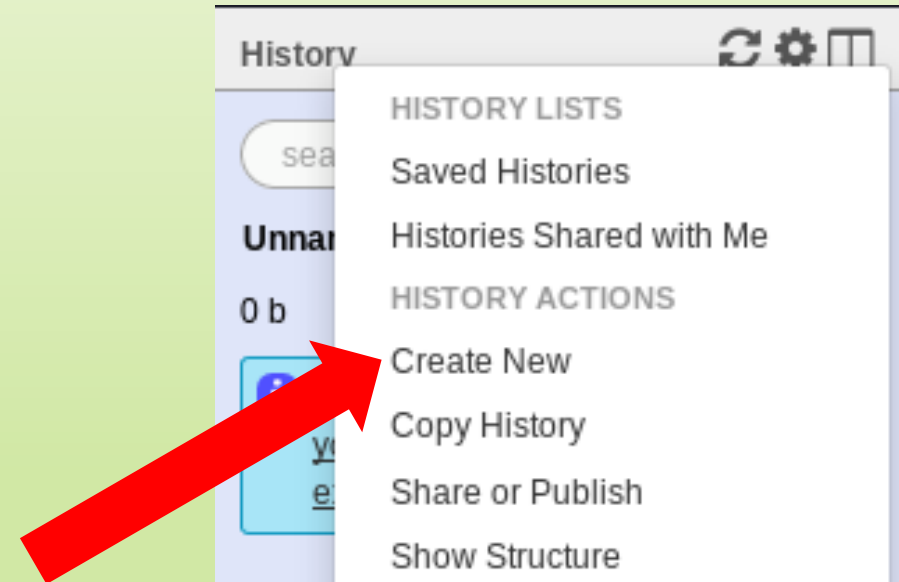
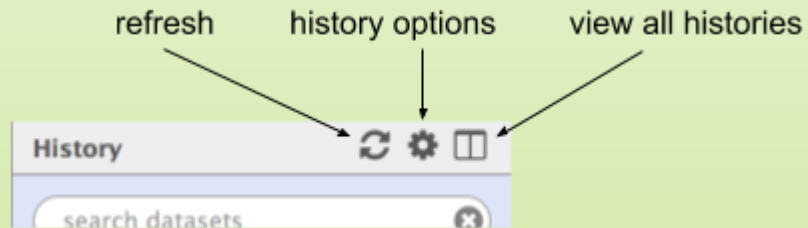
Manipulate Histories

Practice:

CREATE THE 2 HISTORIES NEEDED FOR THE FROGS FORMATION

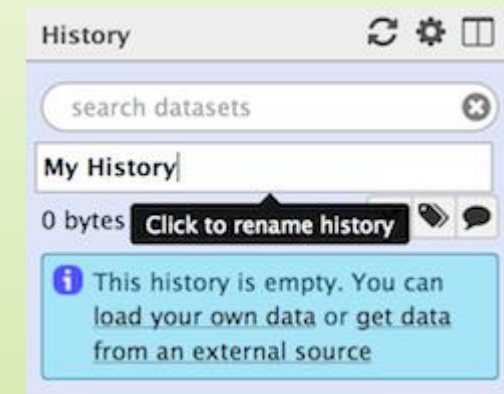
To create a new history:

- Click on the wheel.
- Click on « create new ».



To rename a history:

- Click on the history name (at the top).
- Enter « 16S ».
- Hit « **Enter** » to validate.



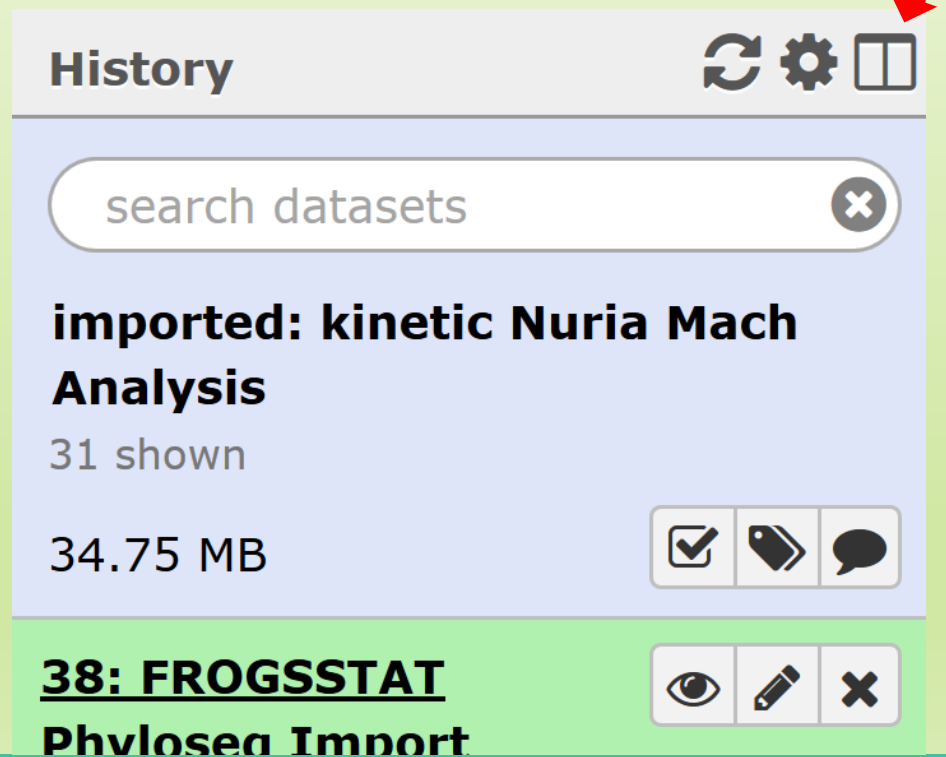
Don't use special
characters or
accents!

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?

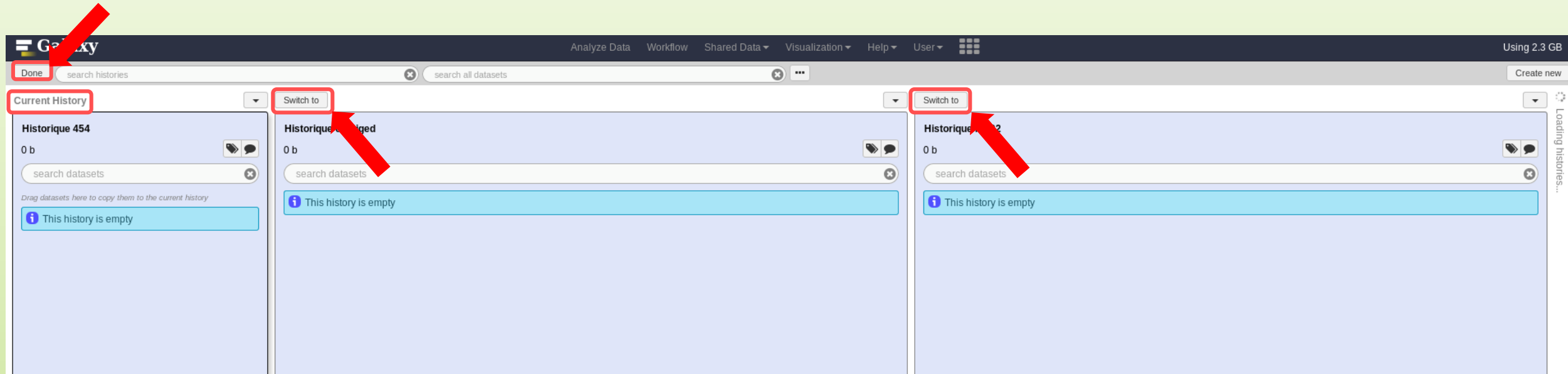
- To view all histories, click on this icon.



Explore the « View all histories » section

The screenshot displays the Galaxy web interface. At the top, the navigation bar includes 'Galaxy' and menu items: 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. Below the navigation bar, there are search bars for 'search histories' and 'search all datasets', and a 'Create new' button. The main content area is titled 'Current History' and shows three history panels. The first panel is 'imported: kinetic Nuria Mach Analysis' (31 shown, 34.75 MB). The second panel is 'Test Ta Thi's tools' (73 shown, 233.76 MB), which has a dropdown menu open showing 'Copy', 'Delete', and 'Purge' options. The third panel is 'Copy of 'full p... maria.bernard...' (14 shown, 11.32 MB). Below the panels, there are search bars for each history and a note: 'Drag datasets here to copy them to the current history'. The bottom of the screenshot shows the start of two other history panels: '38: FROGSSTAT Phylogen...' and '14: FROGS Phv...'. A red box highlights the dropdown arrow on the 'Test Ta Thi's tools' panel and the menu options.

Switch current history



- Switch to the history named « ITS » as current history.
- Click on “Done” 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/

Index of /~formation/15_FROGS/Webinar_data

<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
 Parent Directory		-	
 ITS1.tar.gz	04-Mar-2021 15:40	157M	
 chaillou withprimers 64renamedsamples V1V3 10000seq R1R2.tar.gz	04-Mar-2021 15:34	35M	
 metadata ITS.tsv	04-Mar-2021 15:37	394	
 metadata chaillou.tsv	04-Mar-2021 15:29	2.2K	

All samples are in archive .tar.gz

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



- By URL



We will learn how to import data with these 2 modalities.

- 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:

- From your computer



- By URL



- From Genotoul Bioinfo clusters



- Shared by other users of Galaxy



Upload file from URL

1. In your current history “16S”
2. Go to **Get Data > Upload File from your computer**
3. Click on **Paste/Fetch Data**
4. Copy the address of the file:
http://genoweb.toulouse.inra.fr/~formation/15_FROGS/Webinar_data/chaillou_withprimers_64renamedsamples_V1V3_10000seq_R1R2.tar.gz
5. Change the type!
6. Click on Start
7. You can put one address per line for multiple uploads

Download from web or upload from disk

Regular Composite

Name	Size	Type	Genome	Settings	Status
New File	128 b	tar	unspecified (?)		100% ✓

You can tell Galaxy to download data from web by entering URL in this box (one per line). You can also directly paste the contents of a file.

```
http://genoweb.toulouse.inra.fr/~formation/15_FROGS/Webinar_data/chaillou_withprimers_64renamedsamples_V1V3_10000seq_R1R2.tar.gz
```

Type (set all): Auto-detect Genome (set all): unspecified (?)

Choose local file Paste/Fetch data Pause Reset Start Close

Upload file from URL

The screenshot shows a web interface with a 'History' header. Below the header is a search bar containing the text 'search datasets'. Underneath the search bar, there is a section titled '16S' with the text '1 shown' and '34.85 MB'. To the right of this section are three icons: a checkmark, a folder, and a speech bubble. Below this, a specific file entry is highlighted in green. The entry is labeled '1:' and contains the following URL: http://genoweb.toulouse.inra.fr/~formation/15_FROGS/Webinar_data/chaillou_withprimers_64r_enamedsamples_V1V3_10000seq_R1R2.tar.gz. To the right of the URL are three icons: an eye, a pencil, and an 'X'.

Practice:

UPLOAD FILE FROM YOUR COMPUTER

How to import your data to Galaxy ?

- Several ways to upload your data to Galaxy:

- From your computer



- By URL



- From Genotoul Bioinfo clusters

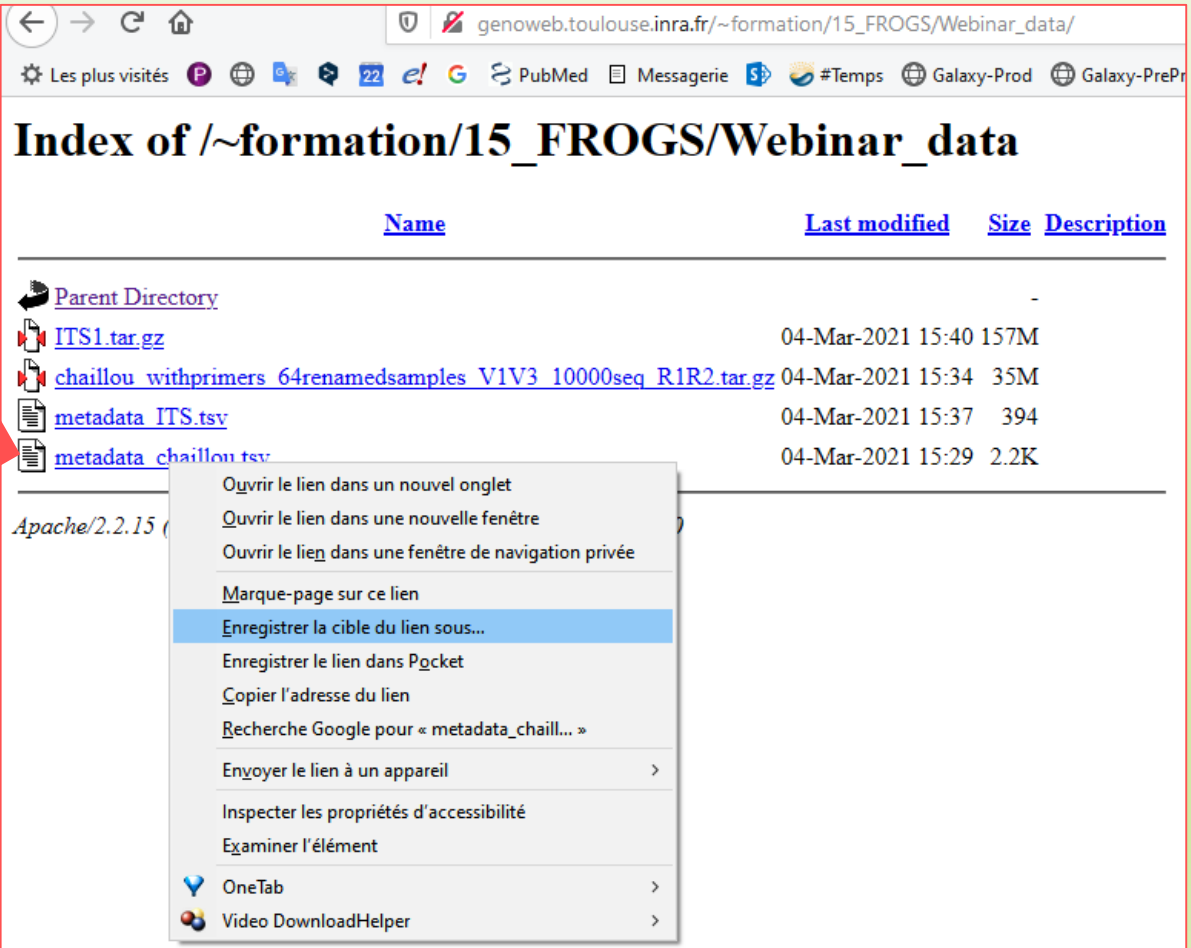


- Shared by other users of Galaxy



Upload a local file

- To simulate that you have already the file on your own computer, download the file “metadata_Chailou.tsv” as:



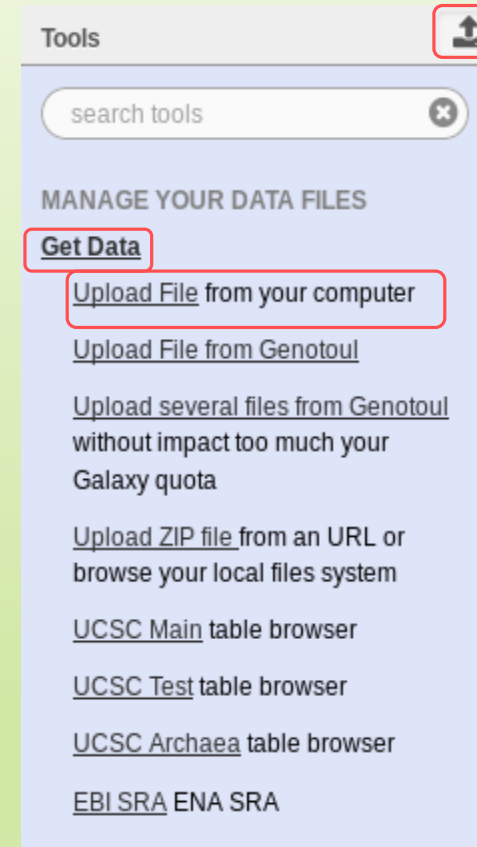
The screenshot shows a web browser window with the address bar displaying `genoweb.toulouse.inra.fr/~formation/15_FROGS/Webinar_data/`. The page title is "Index of /~formation/15_FROGS/Webinar_data". Below the title is a table with columns for "Name", "Last modified", "Size", and "Description". The table lists several files, including "metadata_chailou.tsv". A red arrow points to this file, and a context menu is open over it, with the option "Enregistrer la cible du lien sous..." selected.

<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
Parent Directory	-	-	-
ITS1.tar.gz	04-Mar-2021 15:40	157M	
chailou_withprimers_64renamedsamples_V1V3_10000seq_R1R2.tar.gz	04-Mar-2021 15:34	35M	
metadata ITS.tsv	04-Mar-2021 15:37	394	
metadata_chailou.tsv	04-Mar-2021 15:29	2.2K	

Upload a local file

- In Galaxy, your current history must be 16S
- Click on the « Get Data » tool.
- Upload file from your computer.

⚠ For files smaller than 2 GB for old generations of galaxy platform



Upload a local file

Download from web or upload from disk

Regular Composite

Drop files here

Click here to choose a file on your hard drive.

Type (set all): Genome (set all):

The image shows a software interface for uploading files. At the top, there's a title "Download from web or upload from disk" and two tabs: "Regular" (selected) and "Composite". Below the tabs is a large dashed-line box containing the text "Drop files here". A blue callout box with white text says "Click here to choose a file on your hard drive." with a red arrow pointing to the "Choose local file" button in the bottom toolbar. The toolbar also includes a search icon, a "Genome" dropdown menu, and buttons for "Pause", "Reset", "Start", and "Close".

Upload a local file

The datatype of a file indicates the content of it.

.tsv = tabular

.fastq = fastq

.tar or .tar.gz = tar

Select the type corresponding to your file.

Do not trust blindly
the auto-detect!




1. Select the type of file (Do not leave on Auto-Detect!).


2. Select your other files the same way.

3. Begin upload.



Download from web or upload from disk

[Regular](#) [Composite](#)

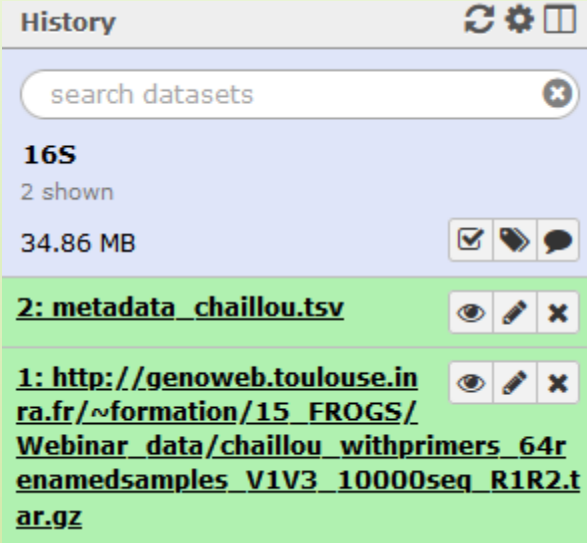
Name	Size	Type	Genome	Settings	Status
 metadata_chaillou.tsv	2.2 KB	tabular	unspecified (?)		100% 

Type (set all): 

Genome (set all):

 Choose local file  Paste/Fetch data

Upload a local file



History

search datasets

16S
2 shown
34.86 MB

2: metadata_chaillou.tsv

1: http://genoweb.toulouse.inra.fr/~formation/15_FROGS/Webinar_data/chaillou_withprimers_64r_enamedsamples_V1V3_10000seq_R1R2.tar.gz

Create and fill yourself the ITS history

Get data form here:

http://genoweb.toulouse.inra.fr/~formation/15_FROGS/Webinar_data/

Index of /~formation/15_FROGS/Webinar_data

	Name	Last modified	Size
	Parent Directory	-	-
	How create an archive.pdf	11-Mar-2021 15:55	781K
	How to put your data on genotoul server.pdf	11-Mar-2021 16:01	1.5M
	ITS1.tar.gz	04-Mar-2021 15:40	157M
	chailou withprimers 64renamedsamples V1V3 10000seq R1R2.tar.gz	04-Mar-2021 15:34	35M
	metadata ITS.tsv	04-Mar-2021 15:37	394
	metadata chailou.tsv	04-Mar-2021 15:29	2.2K



Download from web or upload from disk

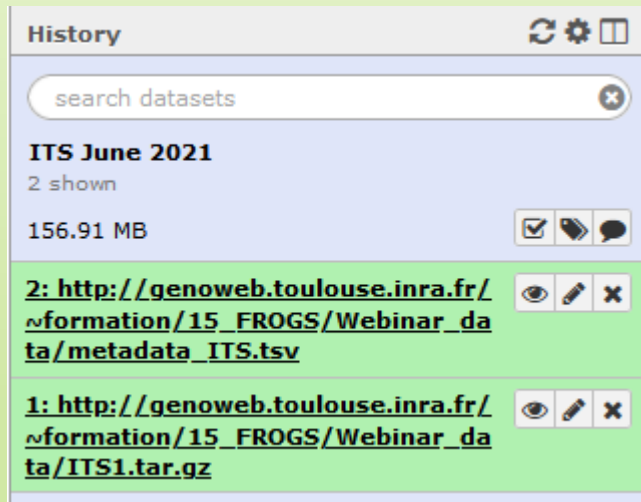
Regular Composite

Name	Size	Type	Genome	Settings	Status
New File	76 b	tar	unspecified (?)		100%
You can tell Galaxy to download data from web by entering URL in this box (one per line). You can also directly paste the contents of a file.					
<input type="text" value="http://genoweb.toulouse.inra.fr/~formation/15_FROGS/Webinar_data/ITS1.tar.gz"/>					
New File	81 b	tabular	unspecified (?)		100%
You can tell Galaxy to download data from web by entering URL in this box (one per line). You can also directly paste the contents of a file.					
<input type="text" value="http://genoweb.toulouse.inra.fr/~formation/15_FROGS/Webinar_data/metadata_ITS.tsv"/>					

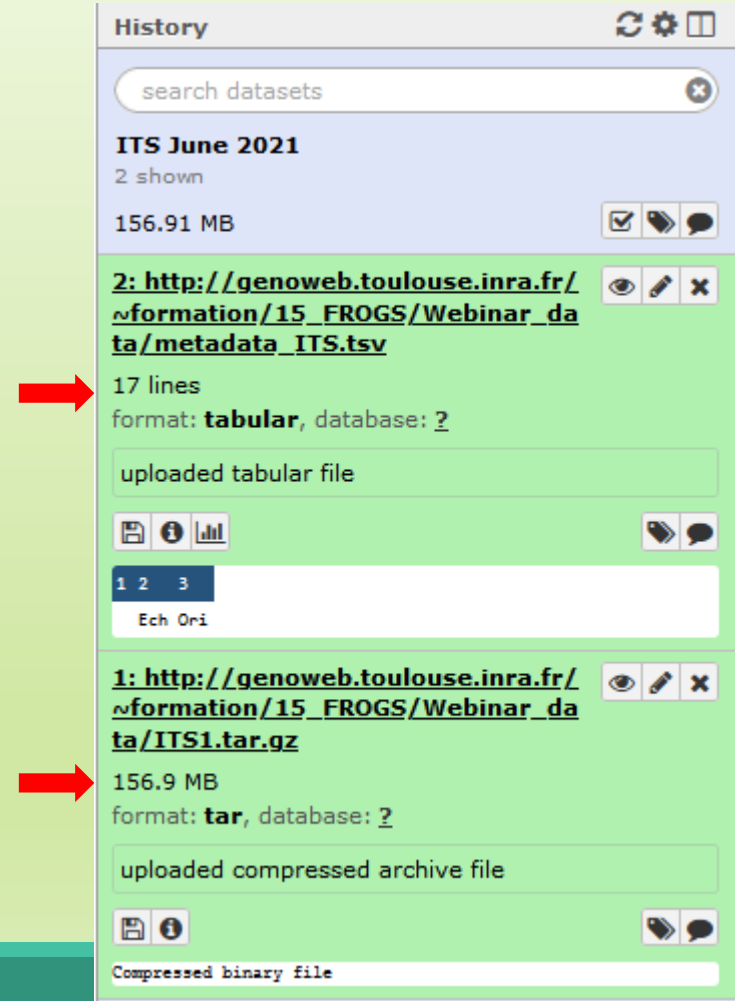
Type (set all): Genome (set all):

Create and fill yourself the ITS history

You have to obtain your second history:



The screenshot shows the 'History' panel with a search bar and two entries. The top entry is 'ITS June 2021' with a size of 156.91 MB. The second entry is a tabular file with 17 lines, and the first entry is a compressed archive file of 156.9 MB.



This screenshot provides a detailed view of the 'History' panel. A red arrow points to the second entry, which is a tabular file with 17 lines, format 'tabular', and database '?'. Below it, a text box contains 'uploaded tabular file'. Another red arrow points to the first entry, which is a compressed archive file of 156.9 MB, format 'tar', and database '?'. Below it, a text box contains 'uploaded compressed archive file'. The panel also includes a search bar, a refresh button, and a list of icons for each entry.

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



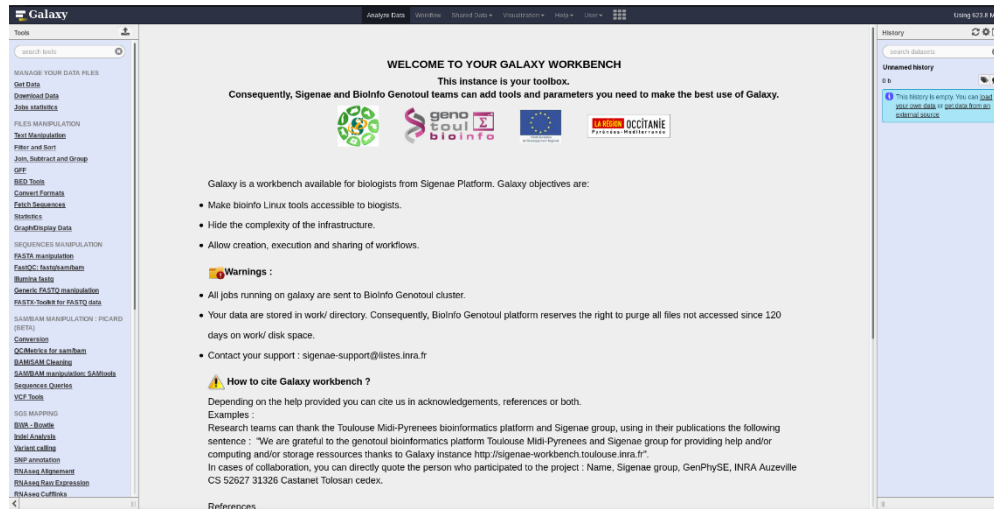
Upload your data from Genotoul

To download data from the server « Genotoul of Bioinformatics of Toulouse », you must already have your data on this server.

To learn how put your data on genotoul

[http://genoweb.toulouse.inra.fr/~formation/15_FROGS/Webinar_data/How to put your data on genotoul server.pdf](http://genoweb.toulouse.inra.fr/~formation/15_FROGS/Webinar_data/How_to_put_your_data_on_genotoul_server.pdf)

Advantage to put your data on Genotoul:



1TB of space



/work: 1TB (1024GB) of space
/save: 250GB of space

- This method allows you to have more space disk and to upload bigger files.
- Files not used in the last 180 days can be purged at any time.

Tools

search tools

MANAGE YOUR DATA FILES

Get Data

[Upload File from your computer](#)[Upload File from Genotoul](#)[Upload several files from Genotoul](#)[Upload ZIP file from an URL or browse your local files system](#)[UCSC Main table browser](#)[UCSC Test table browser](#)[UCSC Archaea table browser](#)[EBI SRA ENA SRA](#)

Download Data

Jobs statistics

FILES MANIPULATION

[Text Manipulation](#)[Filter and Sort](#)[Join, Subtract and Group](#)[GFF](#)[BED Tools](#)[Convert Formats](#)[Fetch Sequences](#)[Statistics](#)[Graph/Display Data](#)

SEQUENCES MANIPULATION

FASTA manipulation

Upload File from Genotoul (Galaxy Version 1.0.0)

Options

Path to file

Path must be like : /work/USERNAME/somewhere/afile

File type

Bam

Execute

Help

This program does

Path to file

This must be an absolute

valid path : /work/LinuxUserName/galaxy/file.extension
invalid path : /home/LinuxUserName/work/galaxy/file.extension

⚠ To use this tool and to maintain the confidentiality of yours directories:

1. Create a "galaxy" directory in your work : `mkdir galaxy`
2. `chmod a+x /work/LinuxUserName`

Example : `drwxr-x--x 4 smaman sigenae 16384 mar 9 14:15 /work/smaman`

3. `chmod a+r /work/LinuxUserName/dataGalaxy.fasta`

📘 Thanks to the fact that this tool requires you to enter your filepath (without "browse" button), you can manage your data privacy.

For example, if your data to download in Galaxy are: /work/LinuxUserName/galaxy/data.fasta:

1. Add "x" rights to "others" on /work/LinuxUserName/ and on galaxy/

It is not useful that "others" have "r" rights of these directories.

- Switch to a new history.
- Next go to Get Data > Upload File from Genotoul.

History

search datasets

ITS

0 b

📘 This history is empty. You can [load your own data](#) or [get data from an external source](#)

Galaxy Analyze Data Workflow Shared Data Visualization Help User Using 42%

Tools search tools

MANAGE YOUR DATA FILES

Get Data

- Upload File from your computer
- Upload File from Genotoul
- Upload several files from Genotoul
- Upload ZIP file from an URL or browse your local files system
- UCSC Main table browser
- UCSC Test table browser
- UCSC Archaea table browser
- EBI SRA ENA SRA

Download Data

Jobs statistics

FILES MANIPULATION

- Text Manipulation
- Filter and Sort
- Join, Subtract and Group
- GFF
- BED Tools
- Convert Formats
- Fetch Sequences
- Statistics
- Graph/Display Data

SEQUENCES MANIPULATION

FASTA manipulation

Upload File from Genotoul (Galaxy Version 1.0.0) Options

Path to file

Path must be like : /work/USERNAME/somewhere/afile

File type

What it does

This program allows you to use a file stored in your genotoul work directory and optimize Galaxy work space by creating symlinks.

Path to file

This must be an absolute path to a file located in your genotoul work directory. The path must start with **/work/YOUR_USER_NAME/blablabla.extension**

valid path : /work/LinuxUserName/galaxy/file.extension
 invalid path : /home/LinuxUserName/work/galaxy/file.extension

⚠ To use this tool and to maintain the confidentiality of yours directories:

1. Create a "galaxy" directory in your work : `mkdir galaxy`
2. `chmod a+x /work/LinuxUserName`

Example : `drwxr-x--x 4 smaman sigenae 16384 mar 9 14:15 /work/smaman`

3. `chmod a+r /work/LinuxUserName/dataGalaxy.fasta`

📄 Thanks to the fact that this tool requires you to enter your filepath (without "browse" button), you can manage your data privacy.

For example, if your data to download in Galaxy are: /work/LinuxUserName/galaxy/data.fasta:

1. Add "x" rights to "others" on /work/LinuxUserName/ and on galaxy/

It is not useful that "others" have "r" rights of these directories.

History search datasets

ITS

1 shown

1.5 MB

1: /work/formation/FROGS/ITS.tar.gz

5,613 lines

format: tar.gz, database: ?

Epilog : job finished at Fri Mar 5 11:37:55 CET 2021

complete-ADN-1_R1.fastq.gz

▪ write your file address

▪ Don't forget to change the Datatype!

Tools



search tools

MANAGE YOUR DATA FILES

Get Data

[Upload File from your computer](#)[Upload File from Genotoul](#)[Upload several files from Genotoul](#)[Upload ZIP file from an URL or browse your local files system](#)[UCSC Main table browser](#)[UCSC Test table browser](#)[UCSC Archaea table browser](#)[EBI SRA ENA SRA](#)

Download Data

Jobs statistics

FILES MANIPULATION

[Text Manipulation](#)[Filter and Sort](#)[Join, Subtract and Group](#)[GFF](#)[BED Tools](#)[Convert Formats](#)[Fetch Sequences](#)[Statistics](#)[Graph/Display Data](#)

SEQUENCES MANIPULATION

[FASTA manipulation](#)

Upload File from Genotoul (Galaxy Version 1.0.0)

Options

Path to file

Path must be like : /work/USERNAME/somewhere/afile

File type

Execute

What it does

This program allows you to use a file stored in your genotoul work directory and optimize Galaxy work space by creating symlinks.

Path to file

This must be an absolute path to a file located in your genotoul work directory. The path must start with **/work/YOUR_USER_NAME/blablabla.extension**

valid path : /work/LinuxUserName/galaxy/file.extension
invalid path : /home/LinuxUserName/work/galaxy/file.extension

⚠ To use this tool and to make

1. Create a "galaxy" directory
2. chmod a+x /work/LinuxUserName/galaxy/

```
Example : drwxr-x--x 4 smaman sigenae 16384 mar 9 14:15 /work/smaman
```

3. chmod a+r /work/LinuxUserName/dataGalaxy.fasta

ⓘ Thanks to the fact that this tool requires you to enter your filepath (without "browse" button), you can manage your data privacy.

For example, if your data to download in Galaxy are: /work/LinuxUserName/galaxy/data.fasta:

1. Add "x" rights to "others" on /work/LinuxUserName/ and on galaxy/

It is not useful that "others" have "r" rights of these directories.

After executing, the file is imported into Galaxy.

History

search datasets

ITS

1 shown

1.5 MB

1: /work/gpascal/FROGS_Formation/ITS.tar.gz

5,613 lines

format: tar.gz, database: ?

Epilog : job finished at Thu Mar 4 17:12:09 CET 2021

complexe-ADN-1_R1.fastq.gz

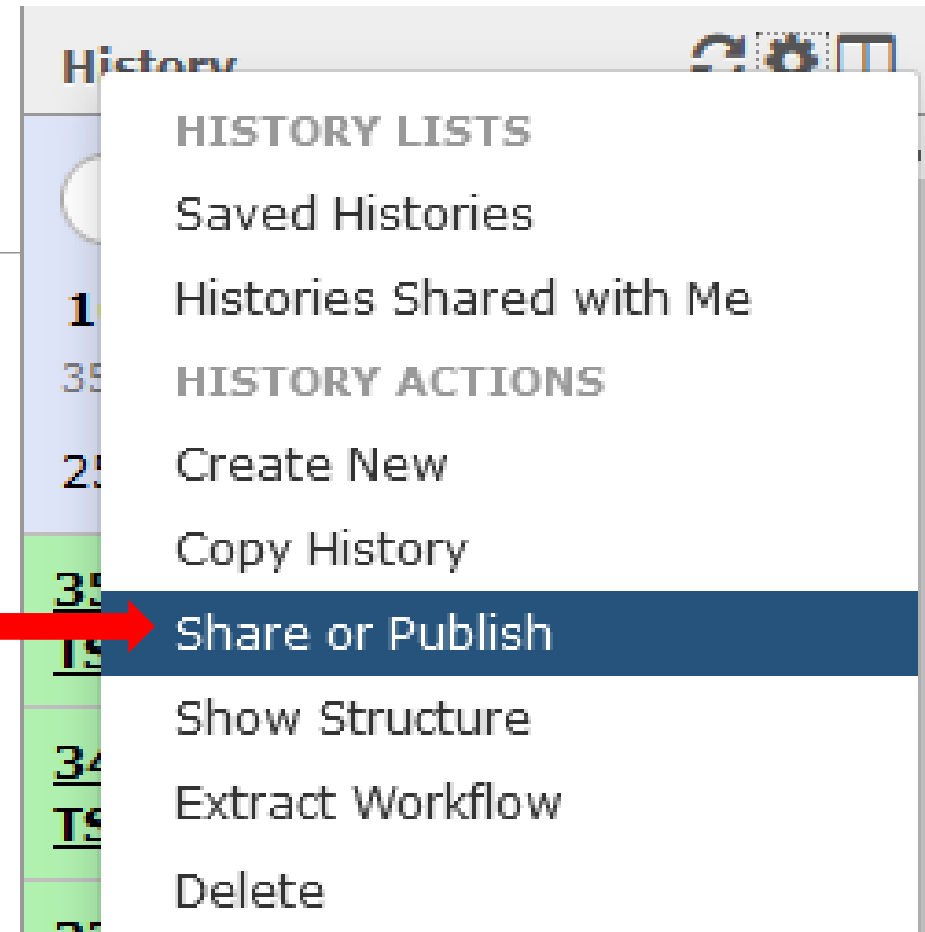
Share a History

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

Click on the wheel and click here.



Share an history

Share or Publish History '16S'

Make History Accessible via Link and Publish It

This history is currently restricted so that only you and the users listed below can access it. You can:

Make History Accessible via Link

Generates a web link that you can share with other people so that they can view and import the history.

Make History Accessible and Publish

Makes the history accessible via link (see above) and publishes the history to Galaxy's [Published Histories](#) section, where it is publicly listed and searchable.

Share History with Individual Users

You have not shared this history with any users.

Share with a user



[Back to Histories List](#)

Tools



search tools



MANAGE YOUR DATA FILES

[Get Data](#)[Download Data](#)[Jobs statistics](#)

FILES MANIPULATION

[Text Manipulation](#)[Filter and Sort](#)[Join, Subtract and Group](#)[GFF](#)[BED Tools](#)[Convert Formats](#)[Fetch Sequences](#)[Statistics](#)[Graph/Display Data](#)

SEQUENCES MANIPULATION

[FASTA manipulation](#)[FastQC: fastq/sam/bam](#)[Illumina fastq](#)[Generic FASTQ manipulation](#)[FASTX-Toolkit for FASTQ data](#)

SAM/BAM MANIPULATION : PICARD (BETA)

[Conversion](#)[QC/Metrics for sam/bam](#)[BAM/SAM Cleaning](#)[SAM/BAM manipulation: SAMtools](#)[Sequences Queries](#)[VCF Tools](#)

Share 1 histories

Histories to be shared:

History Name	Number of Datasets
Historique R1R2	3

Galaxy user emails with which to share histories

Select a user

Enter a Galaxy user email address or a comma-separated list of addresses if sharing with multiple users

Submit

- Enter an email address from a Galaxy user.
- Auto-completion helps you to find your co-worker

History



search datasets



Historique R1R2

3 shown

126.19 MB

**3: Upload several files from Genotoul (sampleA_R2)****2: Upload several files from Genotoul (sampleA_R1)****1: Upload several files from Genotoul**

empty

format: **txt**, database: ?

EpiLog : job finished at Tue Jun 27 22:53:52 CEST 2017



Share 1 histories

Histories to be shared:

History Name

ITS June 2021

Number of Datasets

26

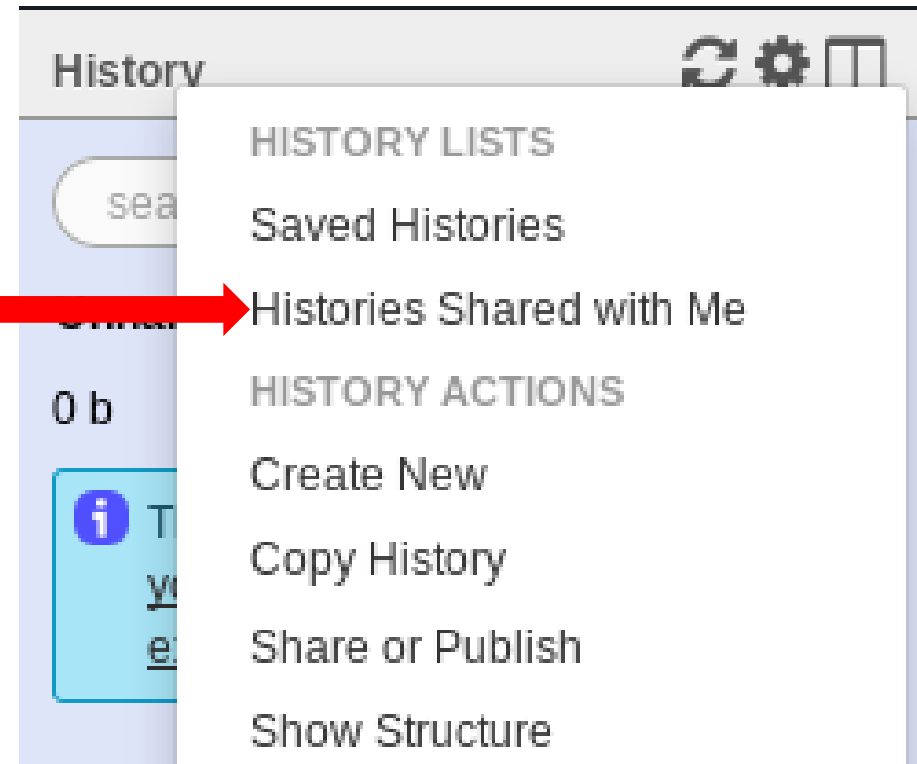
Galaxy user emails with which to share histories

comma-separated list of addresses if sharing with multiple users

- You must click on the address in the dropdown menu.
- Click on submit.
- The history will appear in your co-worker's history.

Import a shared history

Click on the wheel and click here.



Tools

search tools

MANAGE YOUR DATA FILES

Get Data
Download Data
Jobs statistics

FILES MANIPULATION

Text Manipulation
Filter and Sort
Join, Subtract and Group

GFF
BED Tools
Convert Formats

Fetch Sequences
Statistics
Graph/Display Data

SEQUENCES MANIPULATION

FASTA manipulation
FastQC: fastq/sam/bam
Illumina fastq
Generic FASTQ manipulation
FASTX-Toolkit for FASTQ data

SAM/BAM MANIPULATION : PICARD (BETA)

Conversion
QC/Metrics for sam/bam
BAM/SAM Cleaning
SAM/BAM manipulation: SAMtools
Sequences Queries
VCF Tools

Histories shared with you by others

<input type="checkbox"/>	Name	Datasets	Created	Last Updated	Shared by
<input type="checkbox"/>	chaillou_2021	53 4	Feb 17, 2021	Feb 19, 2021	laurent.cauquil@inra.fr
<input type="checkbox"/>	ForestSoil iTags	54 3	Jul 09, 2020	Jul 10, 2020	auer@insa-toulouse.fr

For 0 selected histories:

Check the box and click on « Copy » to import an history into your Galaxy account.

Name of the history.

The person who shared the history.

History

search datasets

Historique R1R2
3 shown
126.19 MB

3: Upload several files from Genotoul (sampleA_R2)

2: Upload several files from Genotoul (sampleA_R1)


1: Upload several files from Genotoul

empty
format: txt, database: ?
Epilog : job finished at Tue Jun 27 22:53:52 CEST 2017

How manipulate datasets

To rename a dataset

- Switch to ITS history



The screenshot shows a web interface titled "History". At the top, there is a search bar labeled "search datasets" and a refresh icon. Below this, the section "Historique ITS" indicates "1 shown" and "156.91 MB". A single dataset entry is listed with the ID "1:". The entry details include the URL http://genoweb.toulouse.inra.fr/~formation/15_FROGS/FROGS_ini/DATA/ITS.tar.qz, a size of "156.9 MB", and the format "tar, database: ?". A text box below the entry contains "uploaded compressed archive file". At the bottom, there is a label "Compressed binary file". A red arrow points from a blue callout box to the edit icon (a pencil) next to the dataset entry.

Click here to display attributes and change the name.

To rename a dataset

Change the name here
and call it « ITS1.tar.gz »

Attributes Convert Format Datatype Permissions

Edit Attributes

Name:

Info:

Annotation / Notes:

Add an annotation or notes to a dataset; annotations are available when editing a dataset.

Database/Build:

Save

Auto-detect

This will inspect the dataset and attempt to correct the above column values.

To change the datatype

Attributes Convert Format Datatype Permissions

Edit Attributes

Name:

Info:

Annotation / Notes:

Add an annotation or notes to a dataset; annotations are available when a history is viewed.

Database/Build:

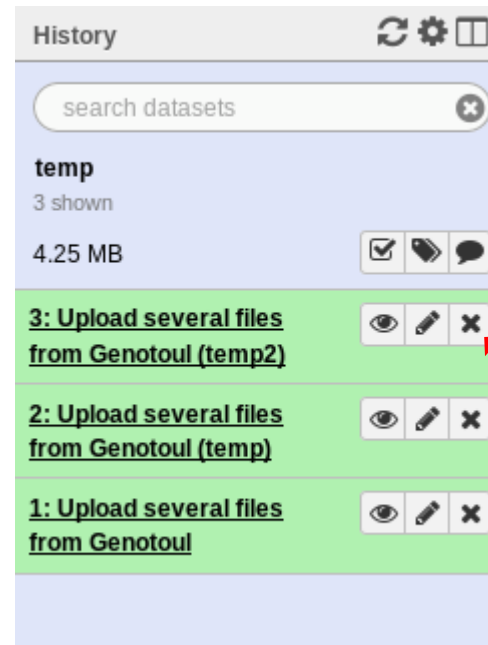
Save

Auto-detect

This will inspect the dataset and attempt to correct the above column values if they are not accurate.

If you put the wrong datatype in during the upload, you can change it here.

To delete a dataset



Click here to delete a dataset.

To delete a dataset

Click here.

History

search datasets

temp

2 shown, [1 deleted](#)

4.25 MB

2: Upload several files from Genotoul (temp)

1: Upload several files from Genotoul

History

search datasets

temp

3 shown, [hide deleted](#)

4.25 MB

This dataset has been deleted
Undelete it
[Permanently remove it from disk](#)

3: Upload several files from Genotoul (temp2)

2: Upload several files from Genotoul (temp)

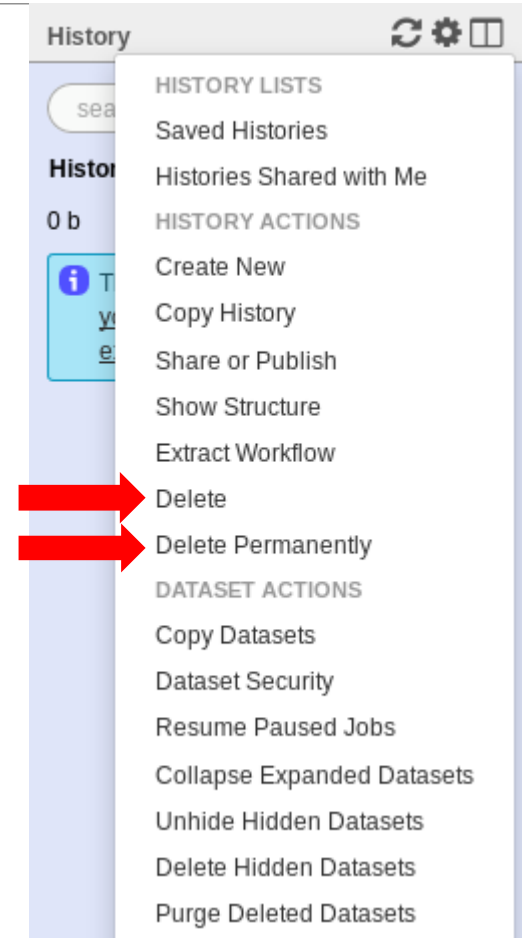
1: Upload several files from Genotoul

Delete this dataset permanently.

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

How to delete a history?

- Stay in the temp history.
- Click on the wheel.
- Click on delete.
- A deleted item on Galaxy is recoverable.
- To definitively delete it:
click on « Delete Permanently ».



Other tools available on Galaxy

- Text Manipulation
- Filter and Sort data
- FASTA manipulation

Galaxy support

- Mail: support.sigenae@inrae.fr
- If you need more training about bioinformatics and Galaxy, please connect to Sigenae e-learning platform:
<http://sig-learning.toulouse.inra.fr/>

How to cite Galaxy?

- Goecks, J, Nekrutenko, A, Taylor, J and The Galaxy Team. "*Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences.*" Genome Biol. 2010 Aug 25;11(8):R86.
- Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. "*Galaxy: a web-based genome analysis tool for experimentalists*". Current Protocols in Molecular Biology. 2010 Jan; Chapter 19:Unit 19.10.1-21.
- Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J, Miller W, Kent WJ, Nekrutenko A. "*Galaxy: a platform for interactive large-scale genome analysis.*" Genome Research. 2005 Oct; 15(10):1451-5.

How to cite Genotoul Galaxy workbench?

- 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.