



Formation FROGS juillet 2019

Géraldine Pascal, Laurent Cauquil

► To cite this version:

Géraldine Pascal, Laurent Cauquil. Formation FROGS juillet 2019: Trainingon Galaxy: Metagenomics / ITS analysisin FROGS? / Trainingon Galaxy: Statistics to explore metagenomics. Master. (Trainingon Galaxy: Metagenomics / ITS analysisin FROGS? / Trainingon Galaxy: Statistics to explore metagenomics), 2019. hal-02790029

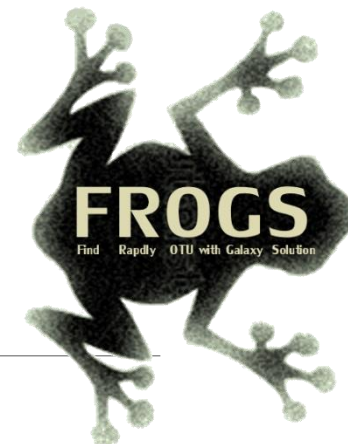
HAL Id: hal-02790029

<https://hal.inrae.fr/hal-02790029>

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Training on Galaxy: Metagenomics

July 2019

Find, Rapidly, OTUs with Galaxy Solution

FRÉDÉRIC ESCUDIÉ* and LUCAS AUER*, MARIA BERNARD, LAURENT CAUQUIL, SARAH MAMAN, MAHENDRA MARIADASSOU, SYLVIE COMBES, GUILLERMINA HERNANDEZ-RAQUET, GÉRALDINE PASCAL & MALO LE BOULCH

*THESE AUTHORS HAVE CONTRIBUTED EQUALLY TO THE PRESENT WORK.



Introduction to Galaxy platform and preparation of FROGS training

July 2019

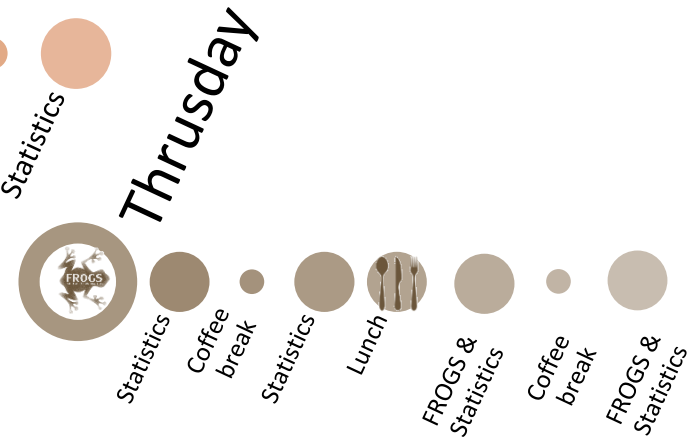
MALO LE BOULCH, SARAH MAMAN, GÉRALDINE PASCAL

Chitchat time!

- What is your computer skills level?
- Have you ever heard of or used Galaxy?

Feedback on metagenomics:

- What are your needs in “metagenomics”?
- 454 / MiSeq?
- 16S/ITS ?
- Your background ?



9 am to 5 pm



2 short coffee breaks
morning and afternoon



Lunch
12.30 to 2.00 pm

Objectives of the 1st part

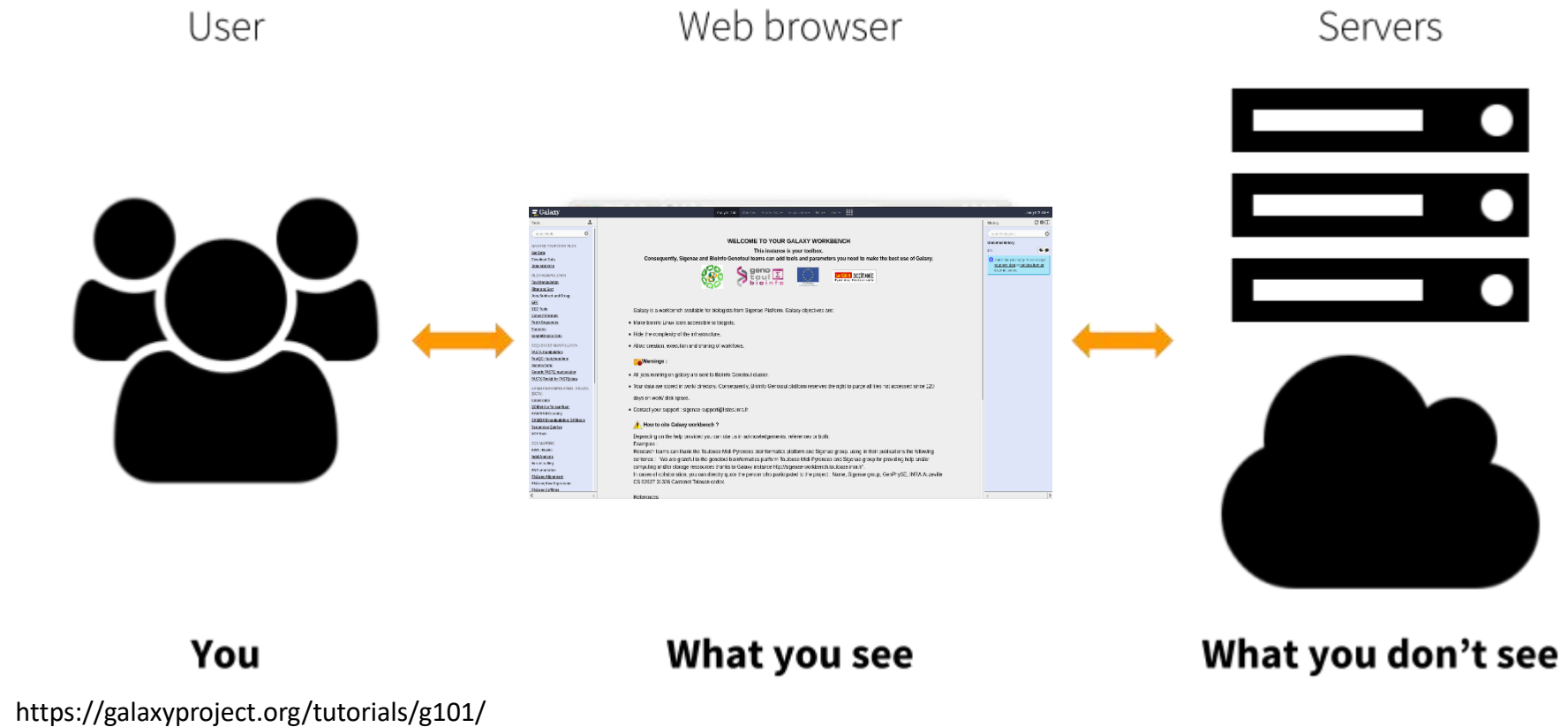
- Learn the basics of Galaxy
- Being independent when using it
- Prepare the datasets for FROGS formation

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.

- 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?

- Our Galaxy platform is <http://sigenae-workbench.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 INRA Galaxy platforms

Galaxy / Migale

Analyze Data Workflow Shared Data Visualization Help User

Tools

search tools

Migale Tools

Get Data

Send Data

Lift-Over

Gene Annotation

Gene Prediction

Text Manipulation

Filter and Sort

Join, Subtract and Group

GFFtools

Convert Formats

Extract Features

Fetch Sequences

Fetch Alignments

Get Genomic Scores

Statistics

Graph/Display Data

Phenotype Association

Sequence Alignment/Pairwise Alignment

Metagenomic analyses

Metagenomics FROGS

Metagenomics Qiime

Metagenomics Mothur

FASTA manipulation

NGS: QC and manipulation

NGS: Assembly

NGS: Mapping

NGS: RNA Analysis

NGS: SAM Tools

NGS: Peak Calling

SNP/WGA: Data; Filters

Variant Analysis

NGS: Picard

NGS: Variant analysis

Workflows

All workflows

Bienvenue sur le portail Galaxy de la plateforme Migale. Pour tous renseignements, demandes ou remarques, veuillez contacter galaxy-help@jouy.inra.fr



Plateforme de BioInformatique - INRA Jouy en Josas

Galaxy est une plateforme qui propose une « constellation » d'outils pour analyser, manipuler et visualiser des données génomiques, sans avoir besoin de connaissance en programmation. Elle est développée par The Center for Comparative Genomics and Bioinformatics. L'utilisateur peut réaliser quatre grands types d'opérations :

- **manipulation de fichiers** : ajout ou suppression de colonnes, trier les fichiers, concaténer plusieurs fichiers, comparaison de listes, ...
- **opérations sur les données** : sommer, moyenner, soustraire, calculer la couverture d'une région déterminée, ...
- **analyse de séquences** : calculer des corrélations, utiliser des outils d'EMBOSS, aligner les données de séquençage, ...
- **visualisation des données** : afficher des alignements multiples, générer des graphiques, ...

Contact : galaxy-help@jouy.inra.fr

Pour toutes demandes d'intégration de nouvel outil au sein du portail, veuillez remplir le [formulaire](#) mis à votre disposition sur le [site web de la plateforme Migale](#).

Une [Foire Aux Questions](#) autour de Galaxy a été mise en place sur le [site web de la plateforme Migale](#).

Si vous utilisez le portail Galaxy de la plateforme Migale pour effectuer vos analyses, merci de [citer Galaxy](#) ET de [remercier la plateforme dans les acknowledgements](#).

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



Exemple of 2 INRA Galaxy platforms

The screenshot displays the Galaxy Sigene / BioInfo Genotoul web interface. The top navigation bar includes links for 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar contains a 'Tools' section with a search bar and a list of categories: 'MANAGE YOUR DATA FILES', 'FILES MANIPULATION', 'SEQUENCES MANIPULATION', 'SAMBAM MANIPULATION : PICARD (BETA)', and 'SGS MAPPING'. The main content area features a large banner with the text 'Welcome on Galaxy Sigene / BioInfo Genotoul.' and 'Contact your support : support.sigene@inra.fr'. Below the banner, there are three columns of information: 'Ready to use Workflows' with a list of links, 'Galaxy News' with a list of updates, and 'E-learning' with a link to the training resources. The right sidebar shows a 'History' section with a search bar and a list of datasets.

Galaxy

Analyze Data Workflow Shared Data Visualization Help User

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

LA REGION OCCITANIE Pyrénées-Méditerranée

Using 0%

History

search datasets

Data Collection

3 shown

33 b

3: seq1.fasta

2: seq2.fasta

1: seq3.fasta

Welcome on Galaxy Sigene / BioInfo Genotoul.

Contact your support : support.sigene@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](#) / [Sigene](#)

Your Turn!

CONNECT TO OUR GALAXY WORKBENCH

Exercise

During this formation, we will use a remote computer.

Connect to a distant computer by clicking on the remote desktop connection icon in the task bar.



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 formation if you have a personal account, use it.

This account will save all the work done during this week and your future work.

If you don't have a personal account ask us for a temporary account.

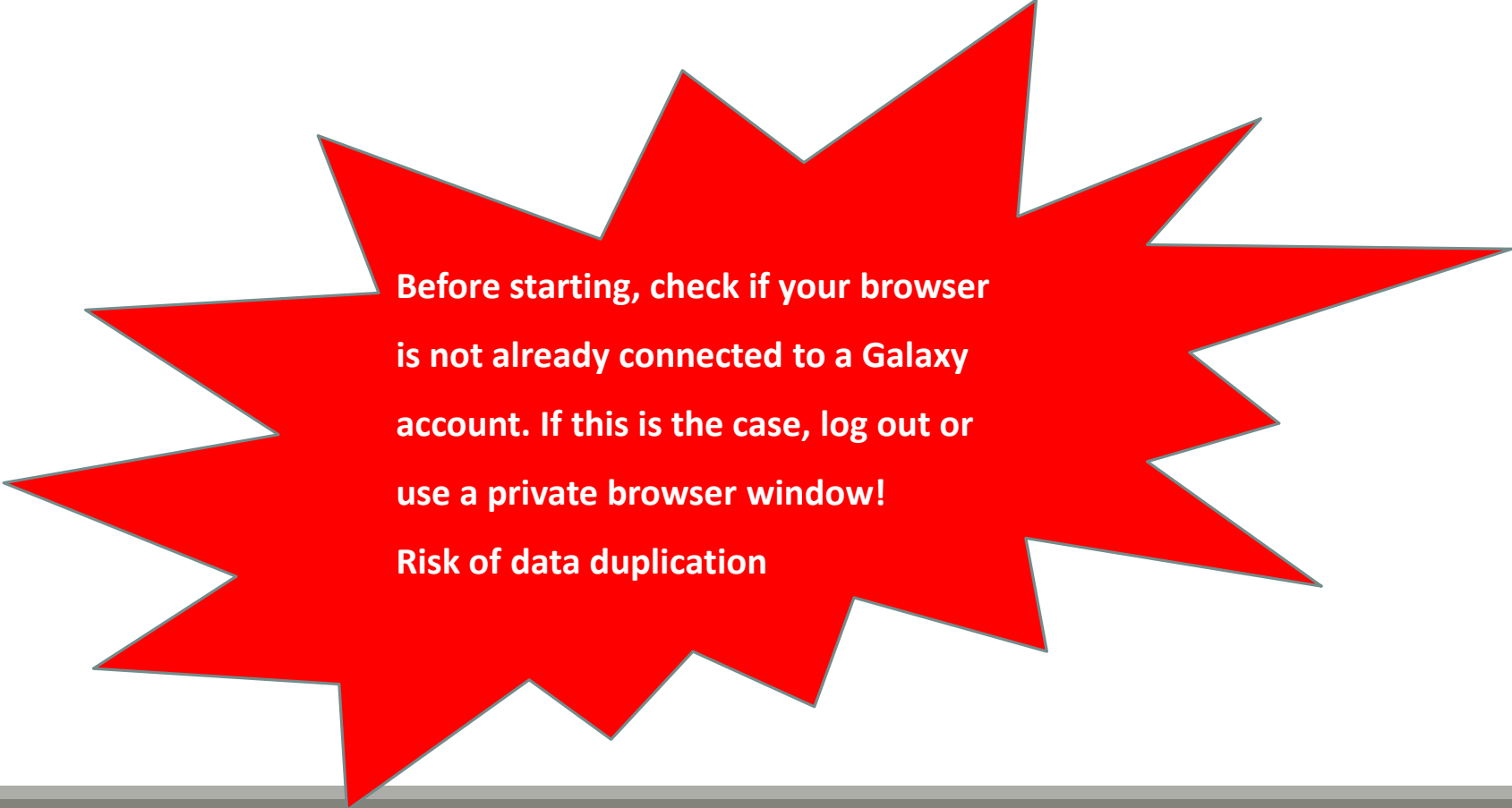
You can request a personal account later on : <http://bioinfo.genotoul.fr/index.php/ask-for/create-an-account/>

A red starburst graphic with multiple points, containing white text.

**All your data will be
erased at the end of the
week on a temporary
account**

Exercise

Our Galaxy platform is: <http://sigenae-workbench.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: <http://sigenae-workbench.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 »

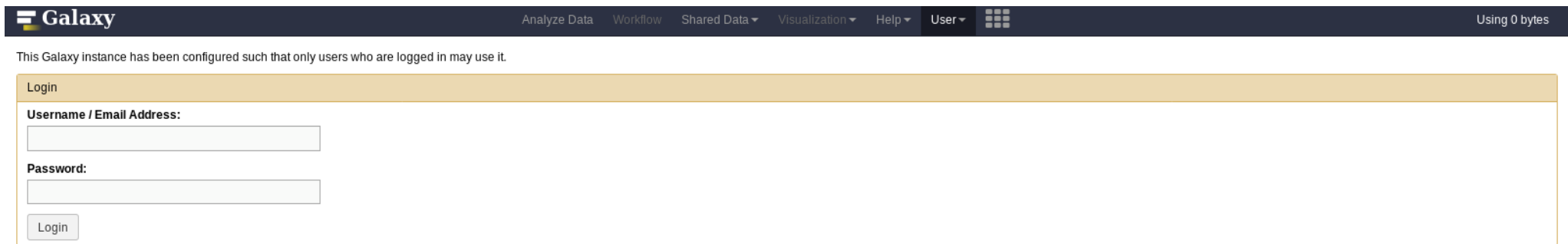
Utilisateur :

Mot de passe :

Annuler OK

Exercise

- And a second time, in the browser:



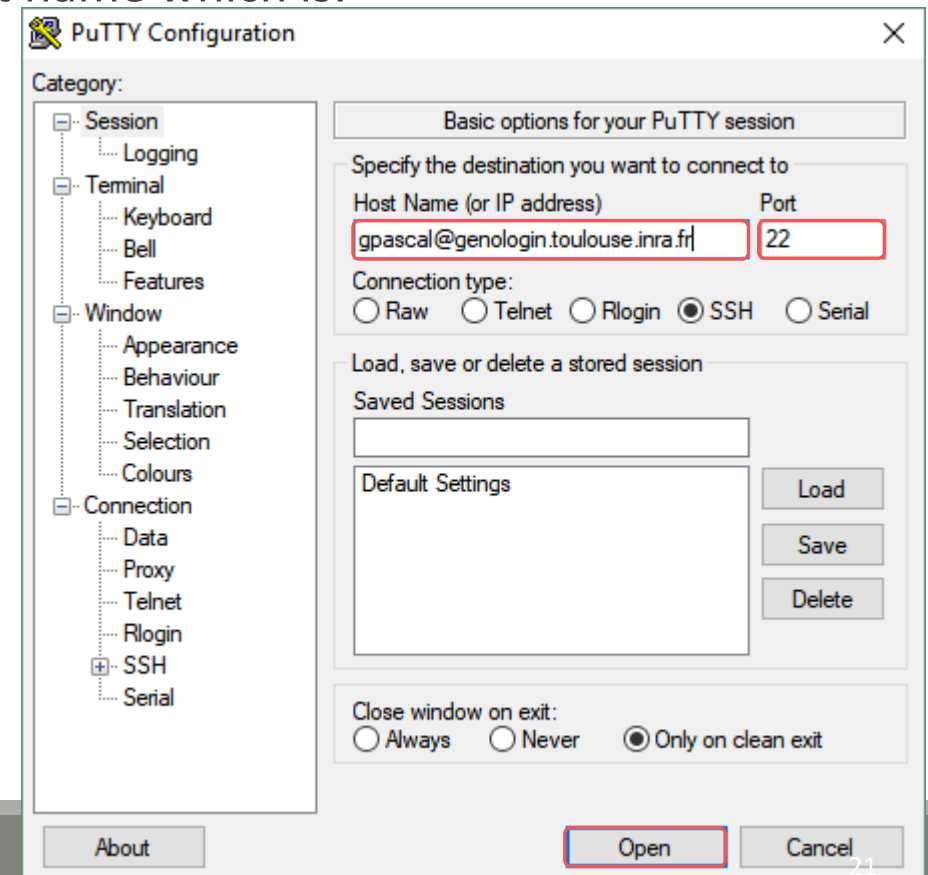
The screenshot shows the Galaxy web interface. At the top is a dark navigation bar with the Galaxy logo on the left and several menu items: 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. On the far right of the navigation bar, it says 'Using 0 bytes'. Below the navigation 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 light-colored box titled 'Login'. Inside the 'Login' box, 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.

Change your password

- In order to change your password, you need to download PuTTY: <http://www.putty.org/>
- PuTTY is a terminal emulator, it allows to connect directly to the server in command line.
- You can not change your password via the Galaxy's interface for the moment.

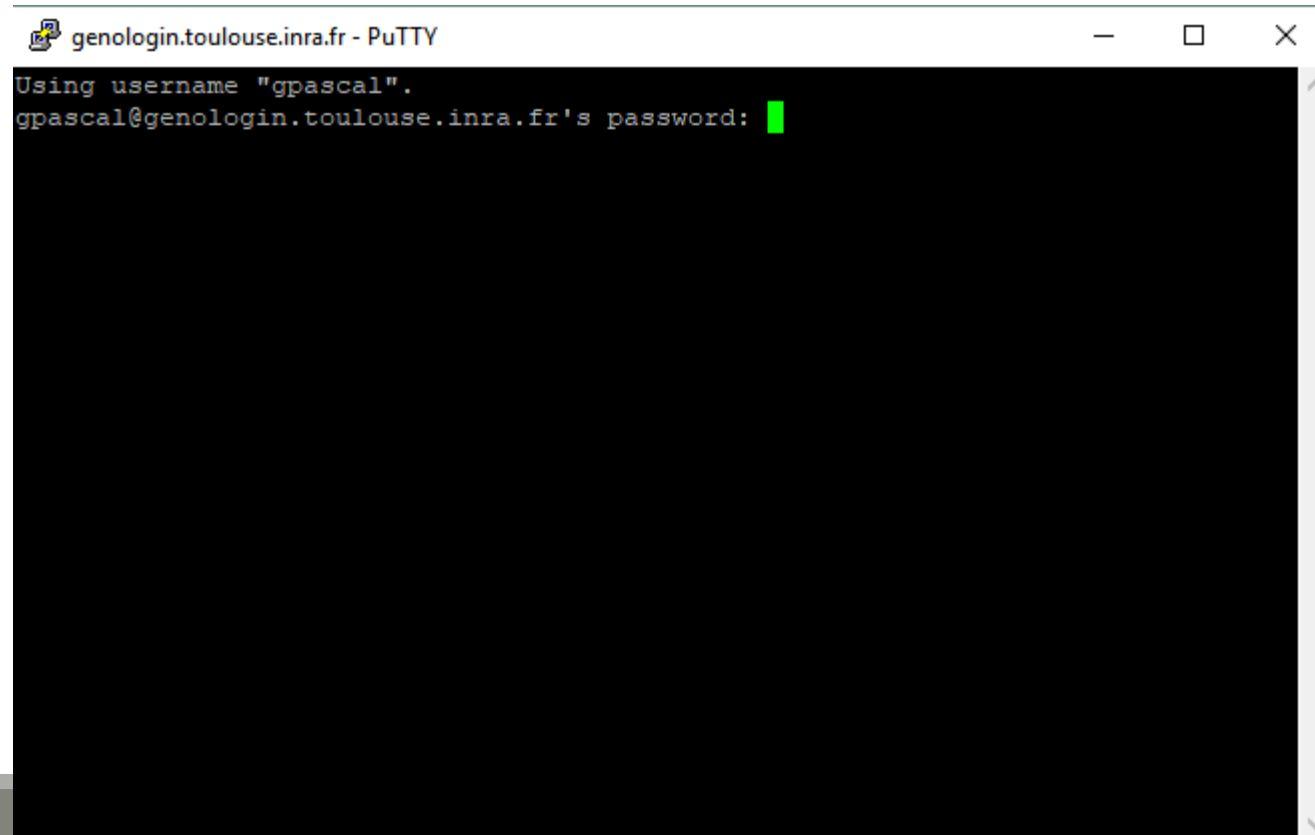
Change your password

- Launch PuTTY.
- In the following window, you must enter your host name which is:
YourGenotoulName@genologin.toulouse.inra.fr
- The connection type must be SSH.
- Click on « Open » .



Change your password

- A new window appear, click on « Yes ».
- In the following window, enter your password and hit « Enter ».



```
genologin.toulouse.inra.fr - PuTTY
Using username "gpascal".
gpascal@genologin.toulouse.inra.fr's password: █
```

Change your password

- Type « passwd » and hit « Enter ».



```
genologin.toulouse.inra.fr - PuTTY
e cluster).
Pour obtenir plus de memoire, veuillez consulter la FAQ de notre site web (cf +
bas)

=====
Informations concernant le quota de temps de calcul
=====

Il existe un quota de temps de calcul annuel de 100.000H pour les academiques (5
00H pour les entreprises privees).
Au dela il faut renseigner le formulaire de demande de ressources exceptionnelle
s sur notre site web.

Vous pouvez verifier votre quota de calcul avec la commande: squota_cpu

=====
Support
=====

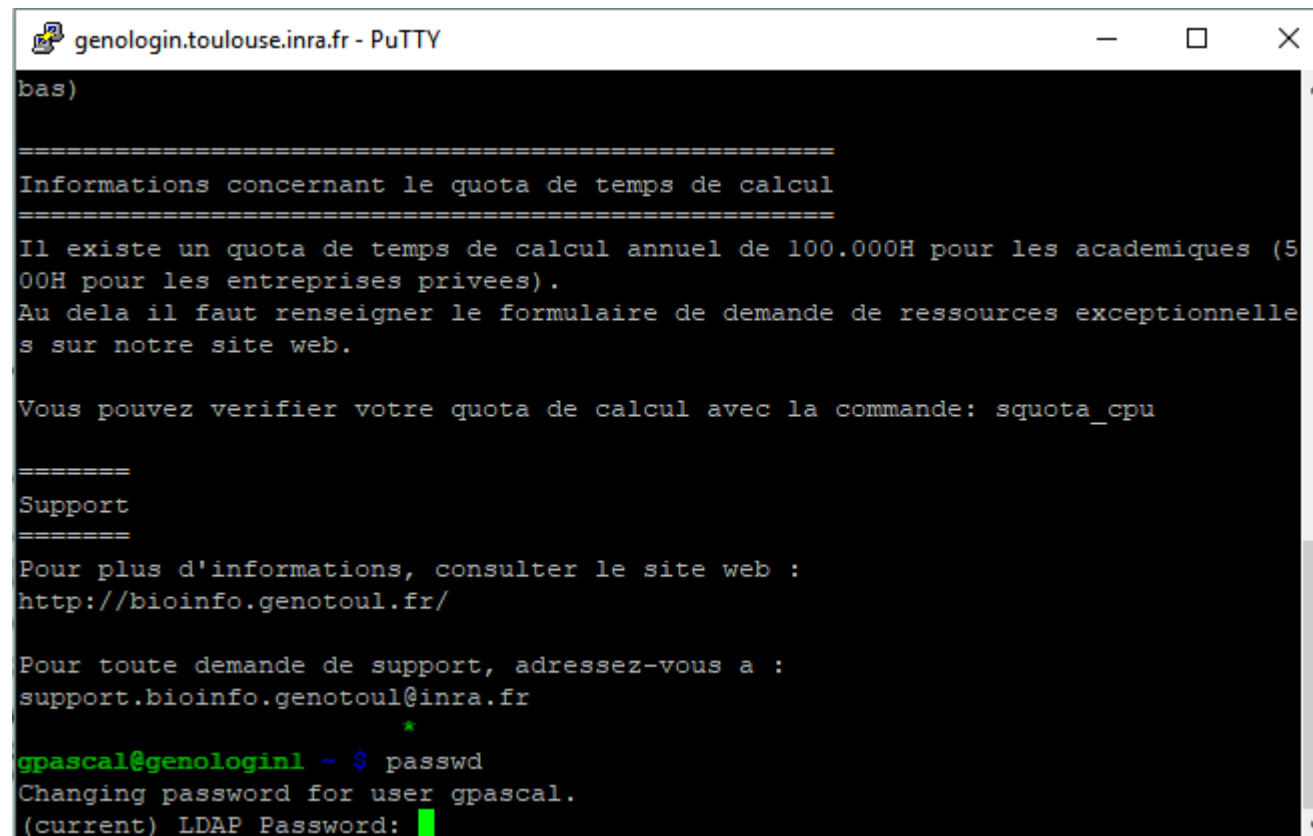
Pour plus d'informations, consulter le site web :
http://bioinfo.genotoul.fr/

Pour toute demande de support, adressez-vous a :
support.bioinfo.genotoul@inra.fr

gpascal@genologin1 ~ $ passwd
```


Change your password

- Enter your current password and hit « Enter ».



The screenshot shows a PuTTY terminal window titled "genologin.toulouse.inra.fr - PuTTY". The terminal output includes a welcome message in French about calculation quotas, followed by a support contact section. At the bottom, the user "gpascal" has entered the command "passwd", and the system prompts for the "current) LDAP Password:" with a redacted input field.

```
genologin.toulouse.inra.fr - PuTTY
bas)

=====
Informations concernant le quota de temps de calcul
=====

Il existe un quota de temps de calcul annuel de 100.000H pour les academiques (5
00H pour les entreprises privees).
Au dela il faut renseigner le formulaire de demande de ressources exceptionnelle
s sur notre site web.

Vous pouvez verifier votre quota de calcul avec la commande: squota_cpu

=====
Support
=====

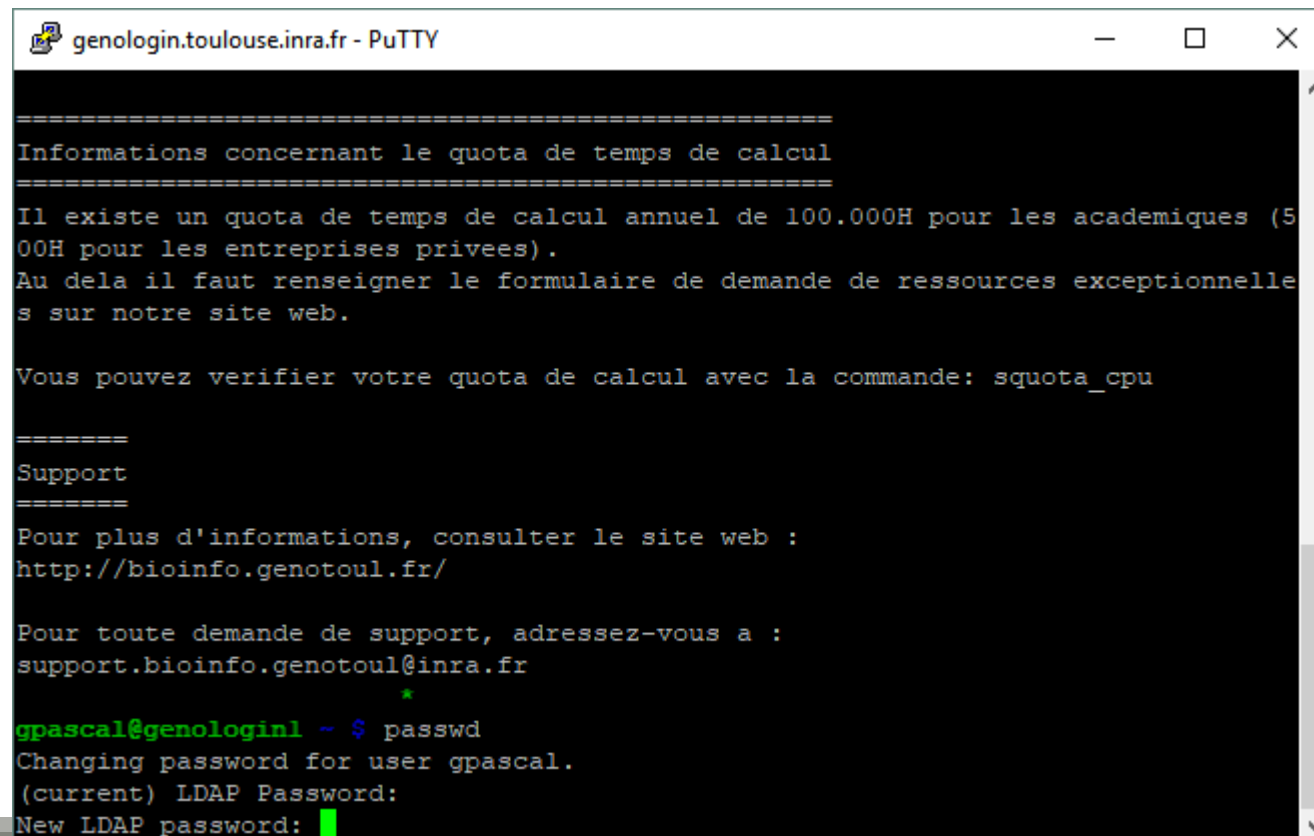
Pour plus d'informations, consulter le site web :
http://bioinfo.genotoul.fr/

Pour toute demande de support, adressez-vous a :
support.bioinfo.genotoul@inra.fr

gpascal@genologin1 ~ $ passwd
Changing password for user gpascal.
(current) LDAP Password: 
```

Change your password

- Enter your new password (with an upper case, a number and a special character in it) and hit « Enter ».



The screenshot shows a PuTTY terminal window titled "genologin.toulouse.inra.fr - PuTTY". The terminal displays a message about calculation time quotas, followed by instructions on how to verify the quota and contact support. At the bottom, the user "gpascal" runs the "passwd" command to change their password. The prompt "(current) LDAP Password:" is shown, followed by "New LDAP password:" with a green cursor.

```
genologin.toulouse.inra.fr - PuTTY

=====
Informations concernant le quota de temps de calcul
=====

Il existe un quota de temps de calcul annuel de 100.000H pour les academiques (5
00H pour les entreprises privees).
Au dela il faut renseigner le formulaire de demande de ressources exceptionnelle
s sur notre site web.

Vous pouvez verifier votre quota de calcul avec la commande: squota_cpu

=====
Support
=====

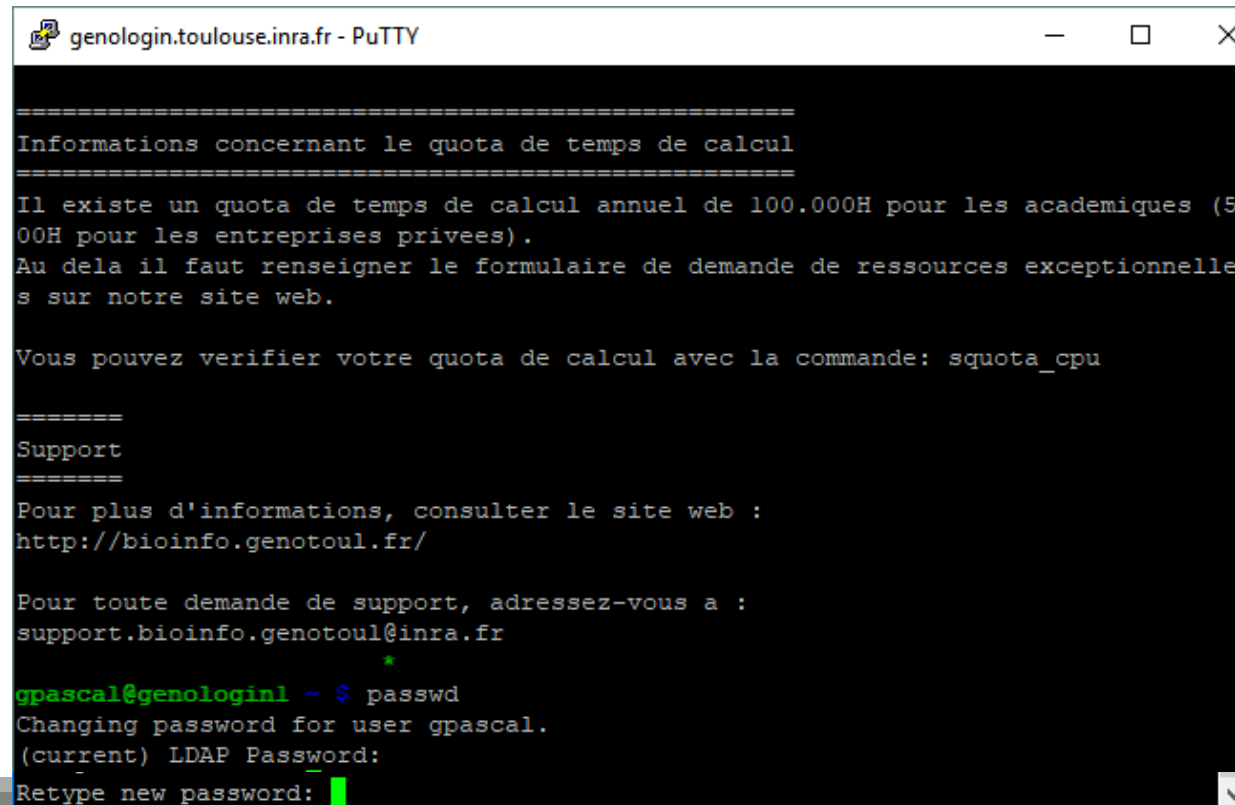
Pour plus d'informations, consulter le site web :
http://bioinfo.genotoul.fr/

Pour toute demande de support, adressez-vous a :
support.bioinfo.genotoul@inra.fr

gpascal@genologin1 ~ $ passwd
Changing password for user gpascal.
(current) LDAP Password:
New LDAP password: █
```

Change your password

- Enter a second time your new password and hit « Enter ». You have changed your Genotoul password, you can close PuTTY.



```
genologin.toulouse.inra.fr - PuTTY

=====
Informations concernant le quota de temps de calcul
=====
Il existe un quota de temps de calcul annuel de 100.000H pour les academiques (5
00H pour les entreprises privees).
Au dela il faut renseigner le formulaire de demande de ressources exceptionnelle
s sur notre site web.

Vous pouvez verifier votre quota de calcul avec la commande: squota_cpu

=====
Support
=====
Pour plus d'informations, consulter le site web :
http://bioinfo.genotoul.fr/

Pour toute demande de support, adressez-vous a :
support.bioinfo.genotoul@inra.fr

gpascal@genologin1 ~ $ passwd
Changing password for user gpascal.
(current) LDAP Password:
Retype new password:
```

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

SAM/BAM MANIPULATION : PICARD (BETA)

Conversion

QC/Metrics for sam/bam

BAM/SAM Cleaning

SAM/BAM 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 ressources 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/>]

History

search datasets

Data Collection

3 shown

33 b

3: seq1.fasta

2: seq2.fasta

1: seq3.fasta

DATASETS HISTORY

AVAILABLE TOOLS

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.classililo.com/Login>
2 e-learning availables : Galaxy initiation and Quality sequences analysis

Vocabulary of Galaxy

- **Tools:**

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

- **Dataset:**

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

- **History:**

- 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!

Your Turn!

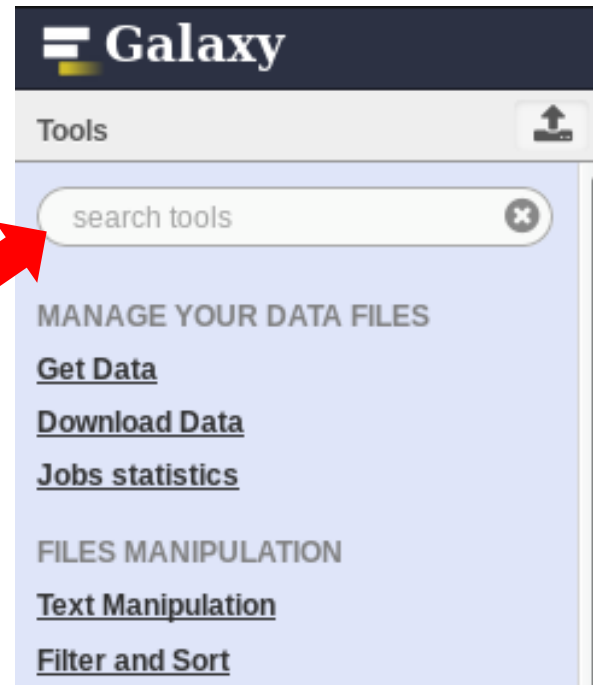
DISCOVER GALAXY

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.



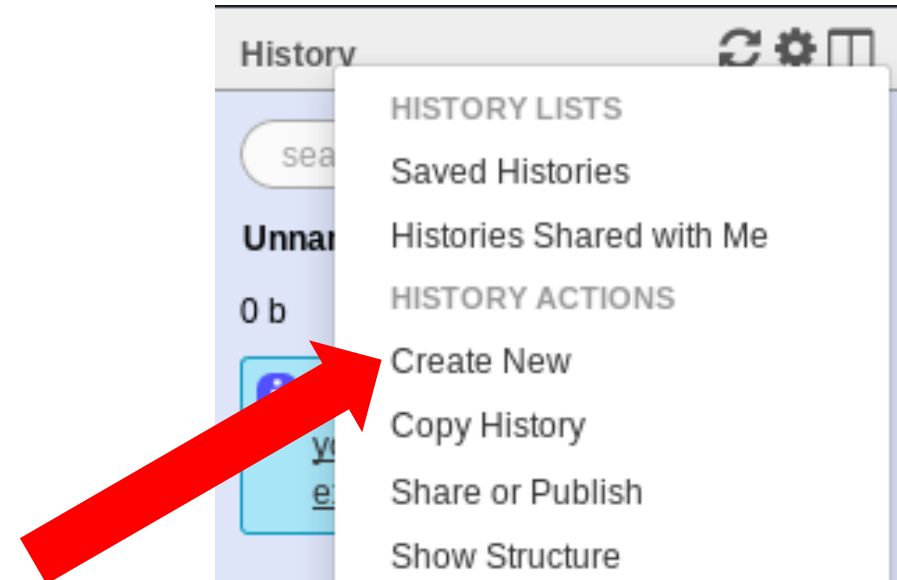
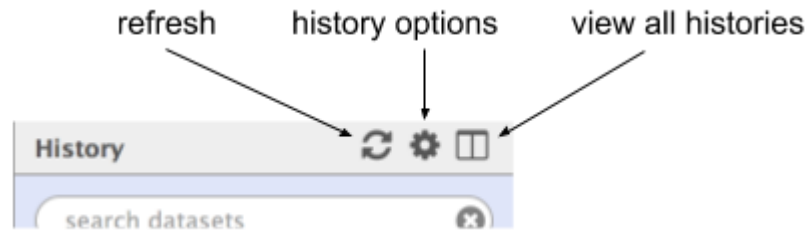
Manipulate Histories

Your Turn!

CREATE THE 4 HISTORIES NEEDED FOR THE FROGS FORMATION

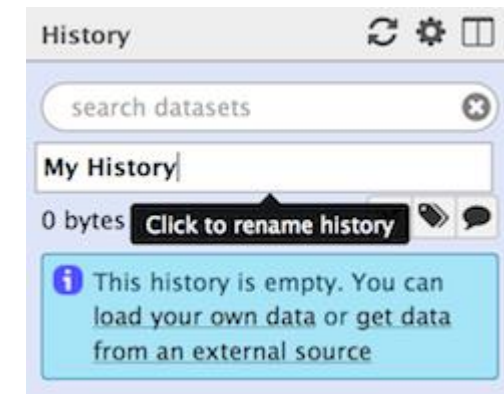
To create a new history:

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



To rename a history:

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



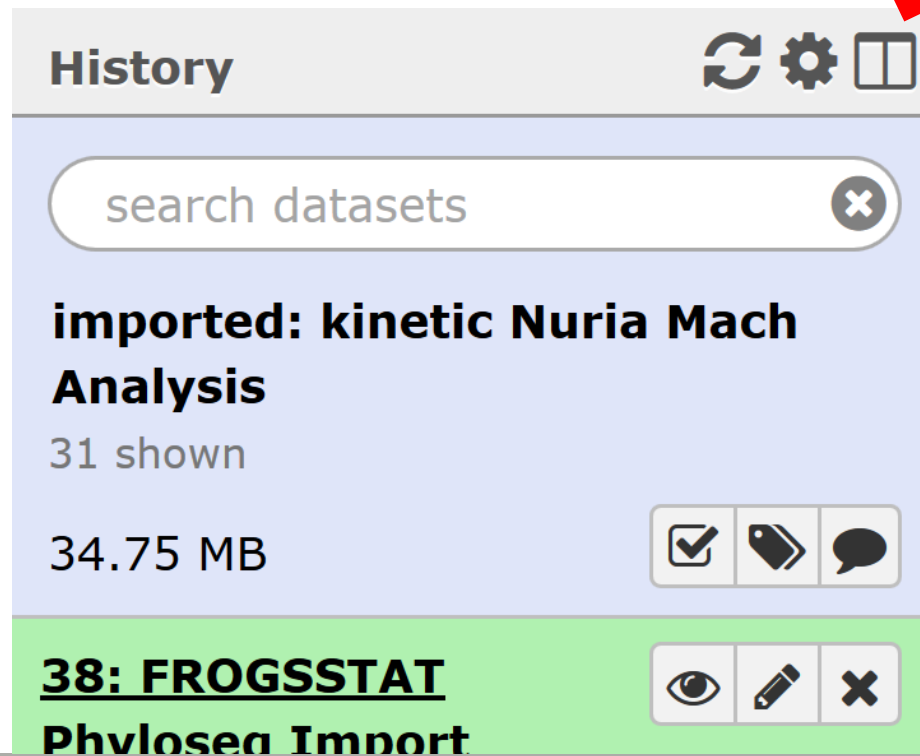
**Don't use special
characters or
accents!**

Exercise

- Create histories named:
 - 454 or ITS
 - merged
 - temp
- Switch to the history named « multiplex » 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, a dark navigation bar contains the Galaxy logo and links for 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. Below this is a light gray toolbar with a 'Done' button, two search bars ('search histories' and 'search all datasets'), a 'Create new' button (highlighted with a red box), and a menu icon. The main content area is titled 'Current History' and features three history panels. The first panel, 'imported: kinetic Nuria Mach Analysis', shows 31 datasets and 34.75 MB. The second panel, 'Test Ta Thi's tools', shows 73 datasets and 233.76 MB; its dropdown menu is open, showing 'Copy', 'Delete', and 'Purge' options (all highlighted with a red box). The third panel, 'Copy of \'full ph... maria.bernard...', shows 14 datasets and 11.32 MB. Each panel includes a 'Switch to' button and a 'search datasets' input. A red line connects the dropdown arrow of the first panel to the open menu of the second panel.

Galaxy Analyze Data Workflow Shared Data Visualization Help User

Done search histories search all datasets Create new

Current History

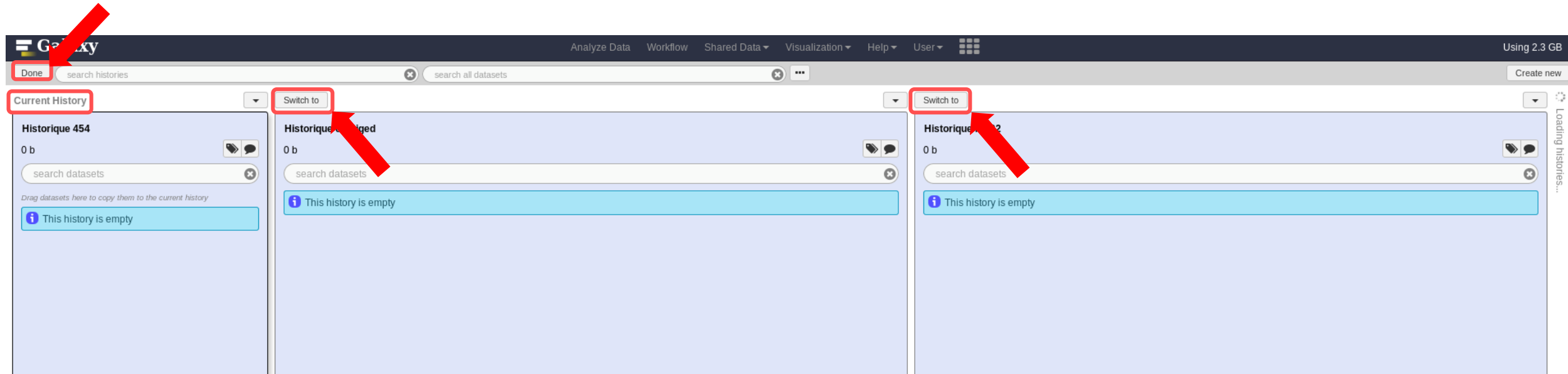
imported: kinetic Nuria Mach Analysis
31 shown
34.75 MB
search datasets
Drag datasets here to copy them to the current history

Test Ta Thi's tools
73 shown
233.76 MB
search datasets
76: FROGS Pre-process: report.html

Copy of 'full ph... maria.bernard'
14 shown
11.32 MB
search datasets
14: FROGS Phv

- Copy
- Delete
- Purge

Switch current history



- Switch to the history named « multiplex » as current history.
- Click on “Done” to go back to the main interface.

Data import

How to import your data to Galaxy

- 5 ways to upload your data to Galaxy:

- From your computer



- By URL



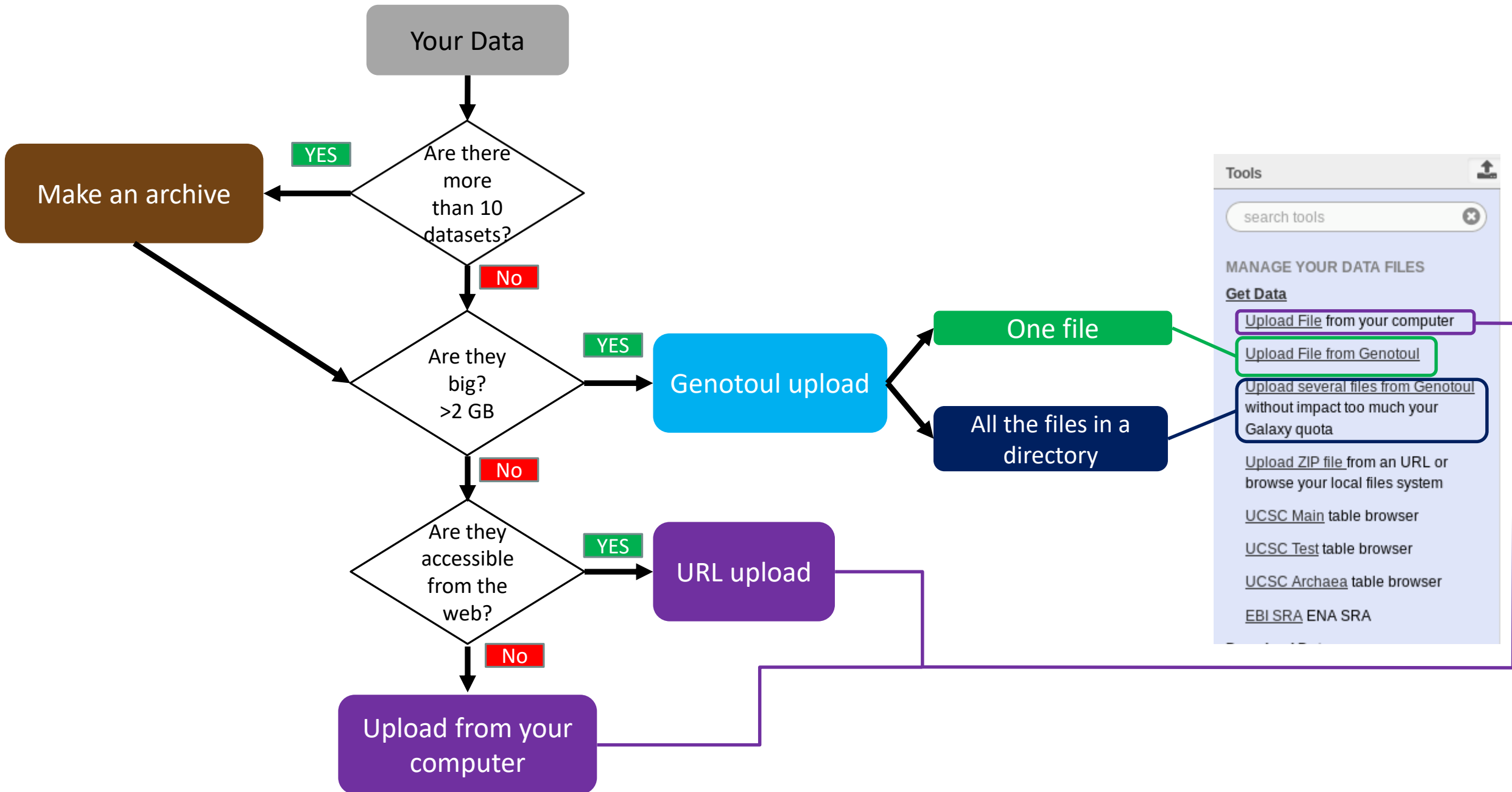
- From Genotoul Bioinfo clusters



- Shared by other users of Galaxy



How to choose your upload method?

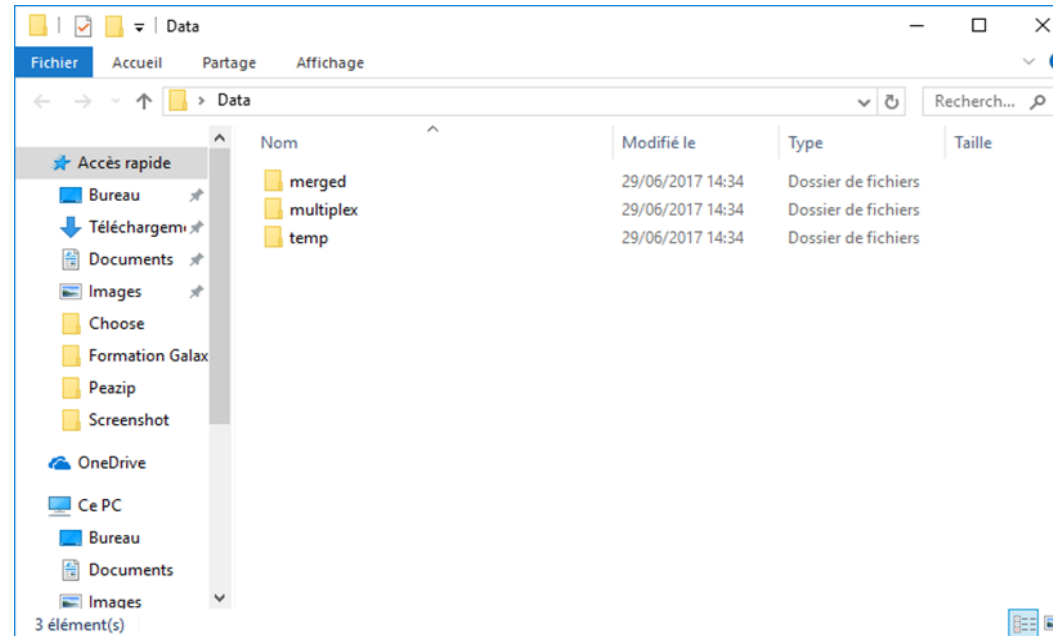


Your Turn!

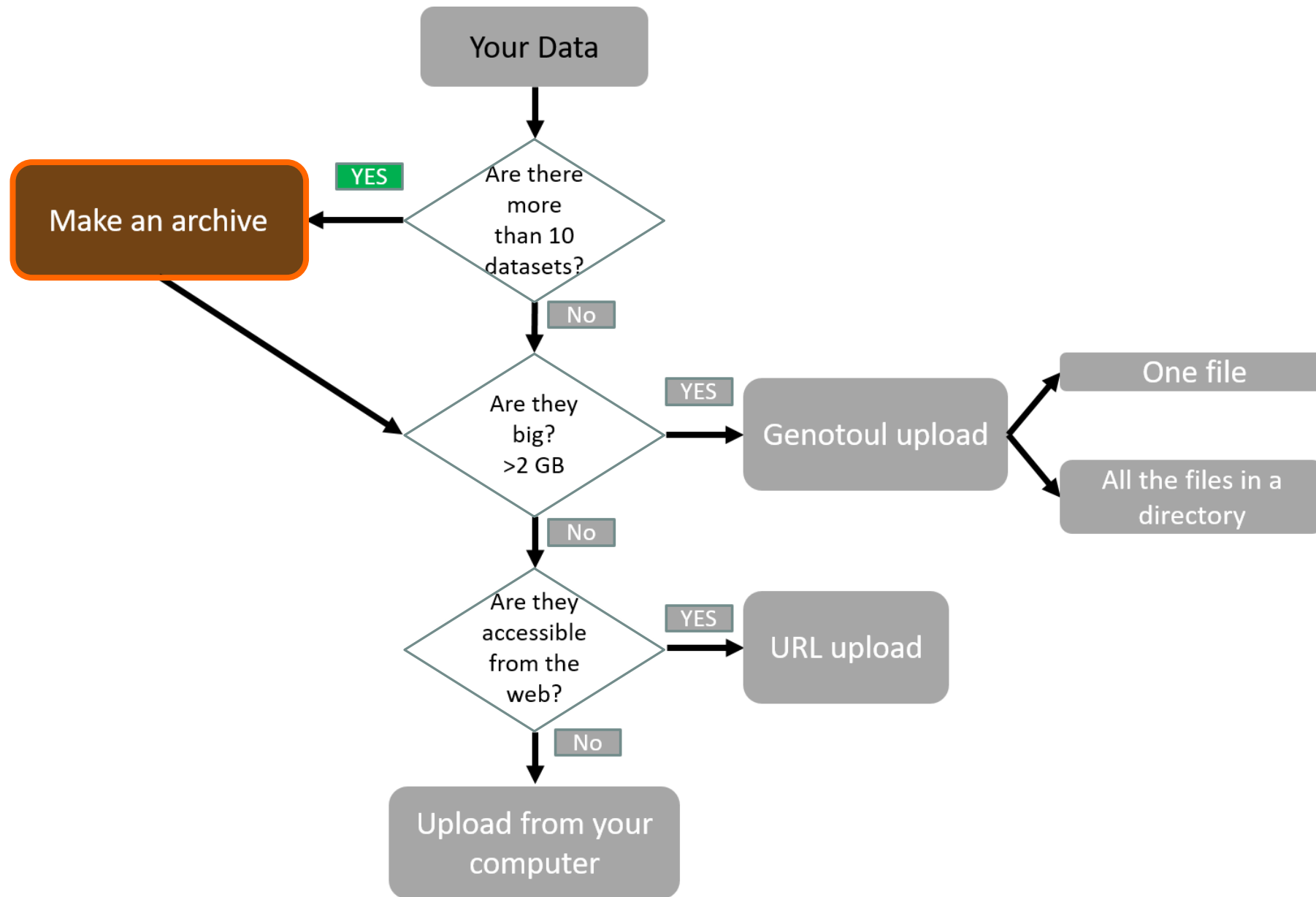
PREPARE FILES

Formation data

- All the formation data are in the Data folder on your desktop.

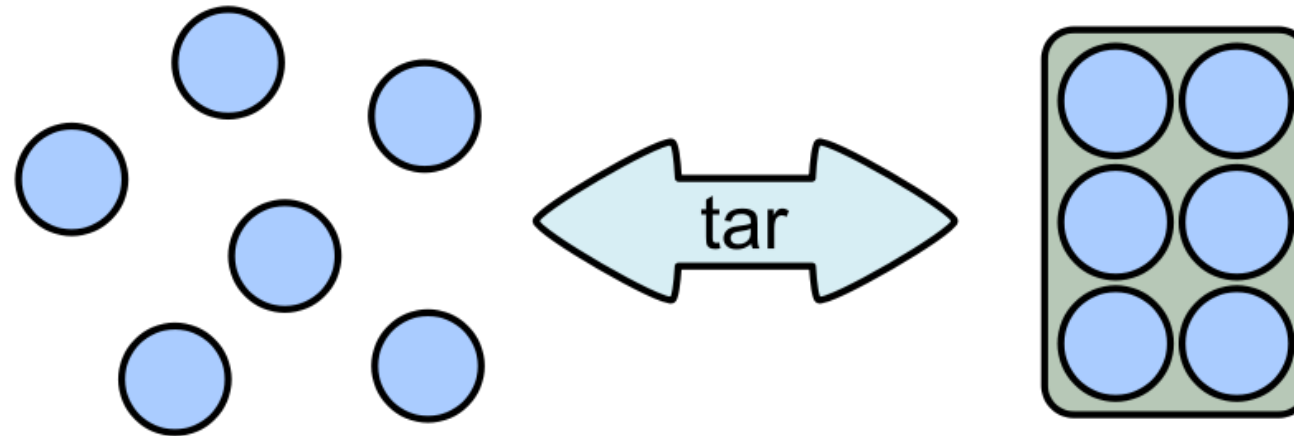


- If you don't have this directory on your desktop, ask us.
- You can find all the formation data online on:
http://genoweb.toulouse.inra.fr/~formation/15_FROGS/FROGS_ini/DATA/Dataset/



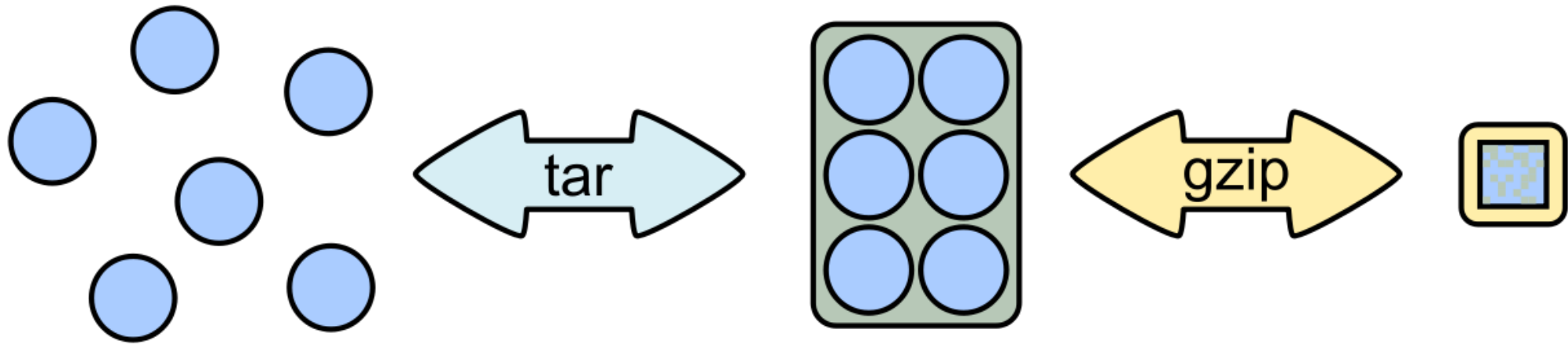
Create a Tar file

- Uploading multiple local files is time consuming.
- Solution: put all your files in an archive!
- What is a tar file?



Create a Tar.gz

- Moreover, we can compress the archive to free up space.



Your Turn!

CREATE AN ARCHIVE WITH THE FILES IN MERGED

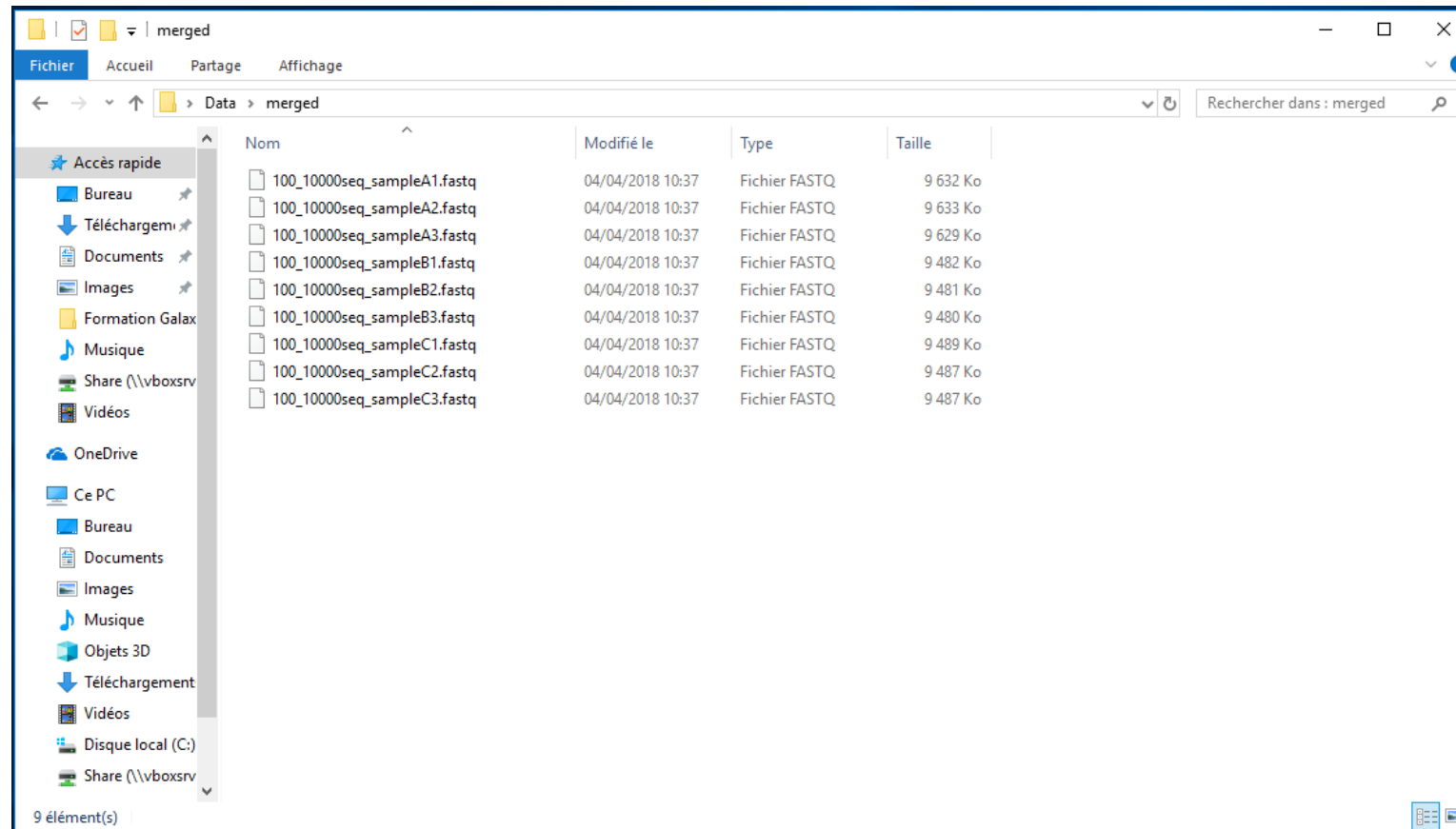
7-Zip

- 7-Zip is a software called a file archiver.
- Can archive and compress files.
- Open source
- At the lab, you could download it at: <http://www.7-zip.org/>



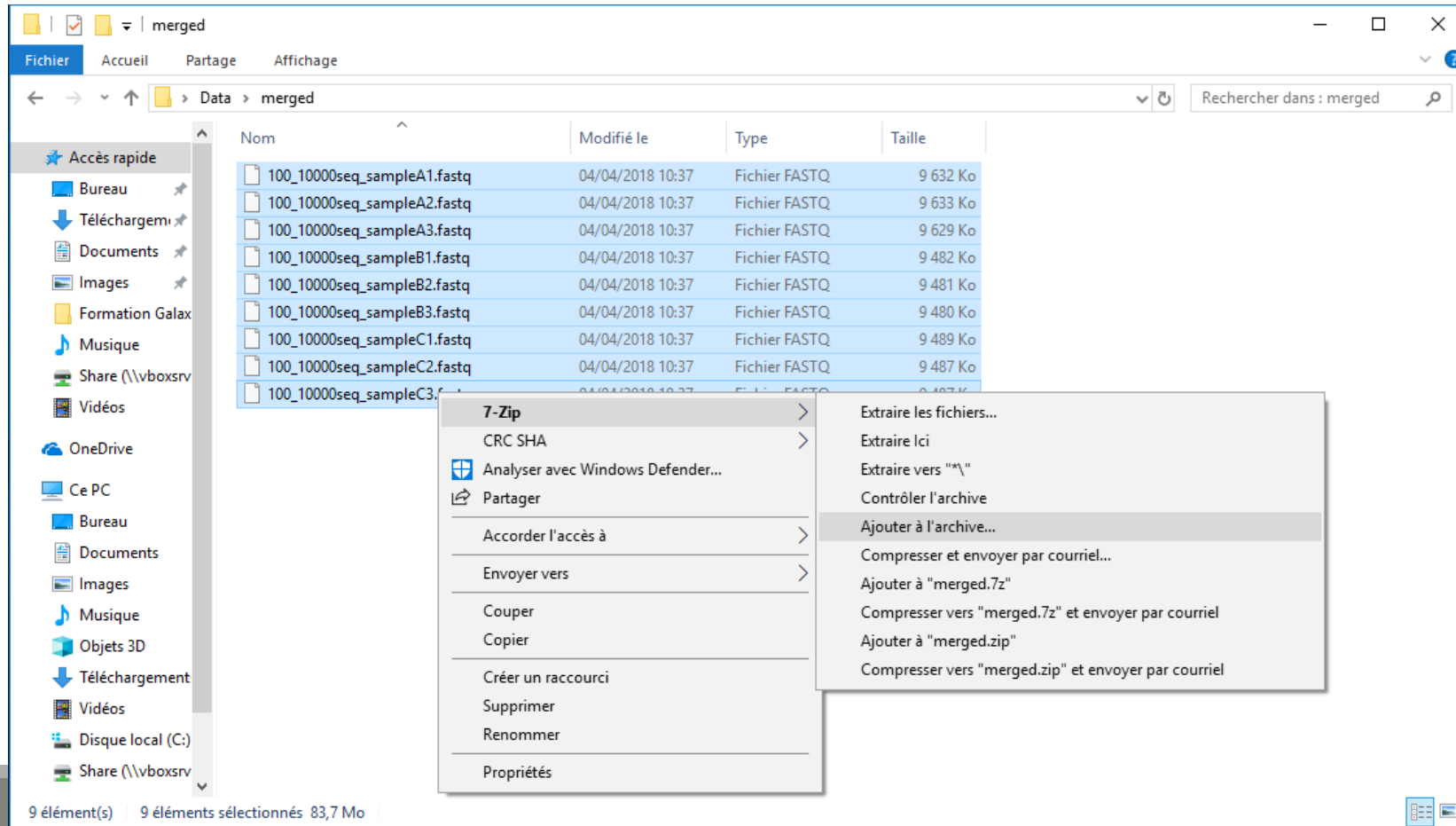
Create a Tar file with 7-Zip

- Browse into the Data directory and next into the merged directory.



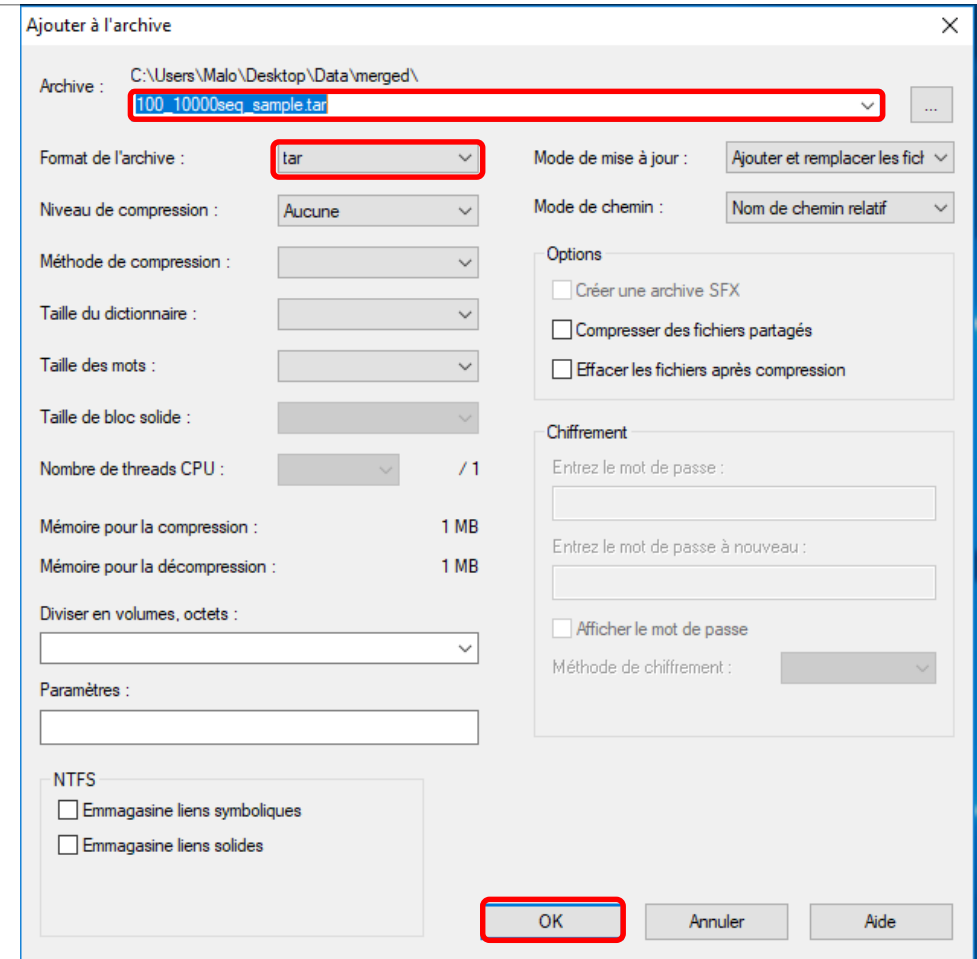
Create a Tar file with 7-Zip

- Select all files and right click on it > 7-Zip > Add to archive . . .



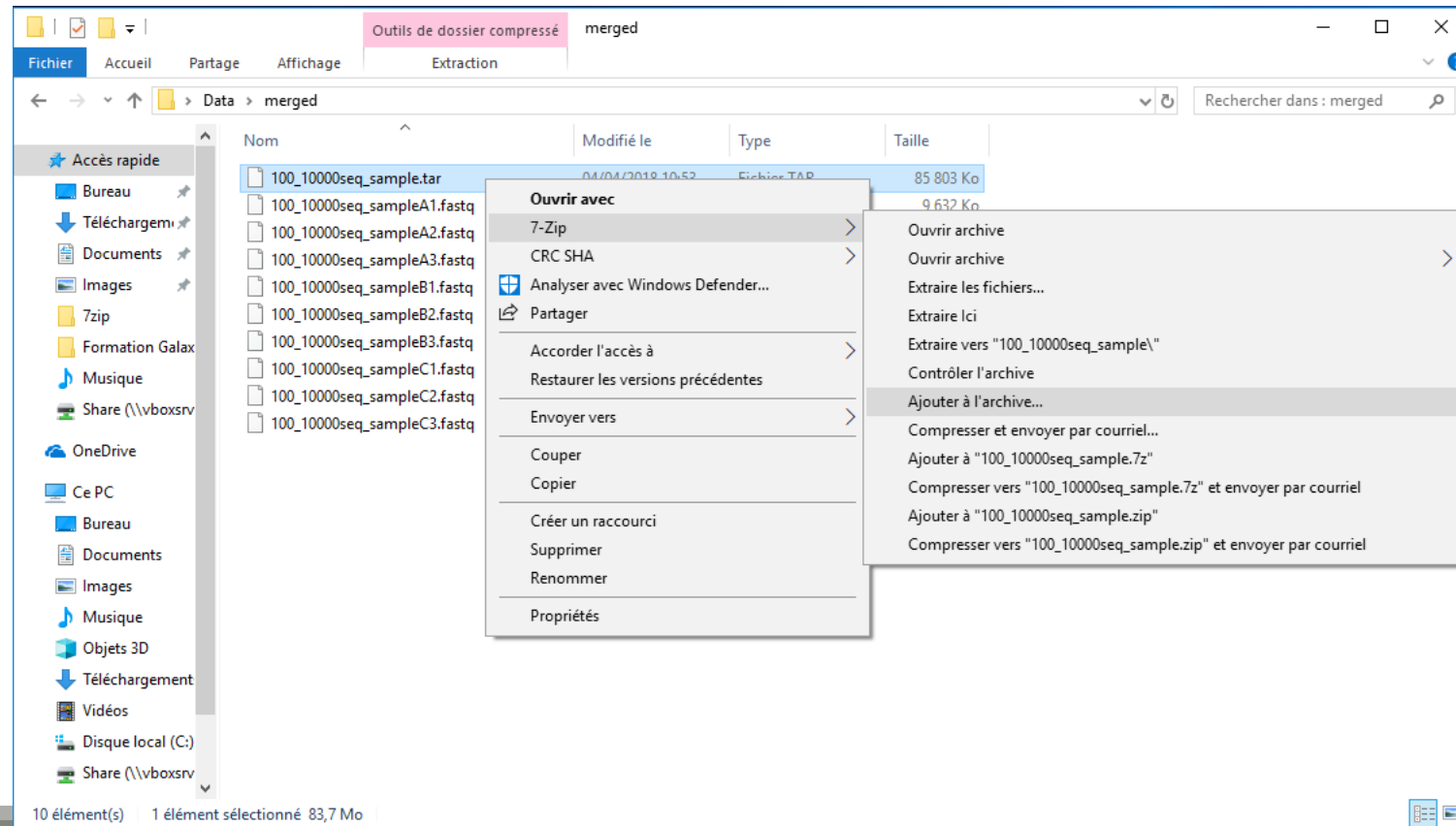
Create a Tar file with 7-Zip

- Change the name of the archive to:
« 100_10000seq_sample.tar »
- Choose « tar » as archive format .
- Click on OK .



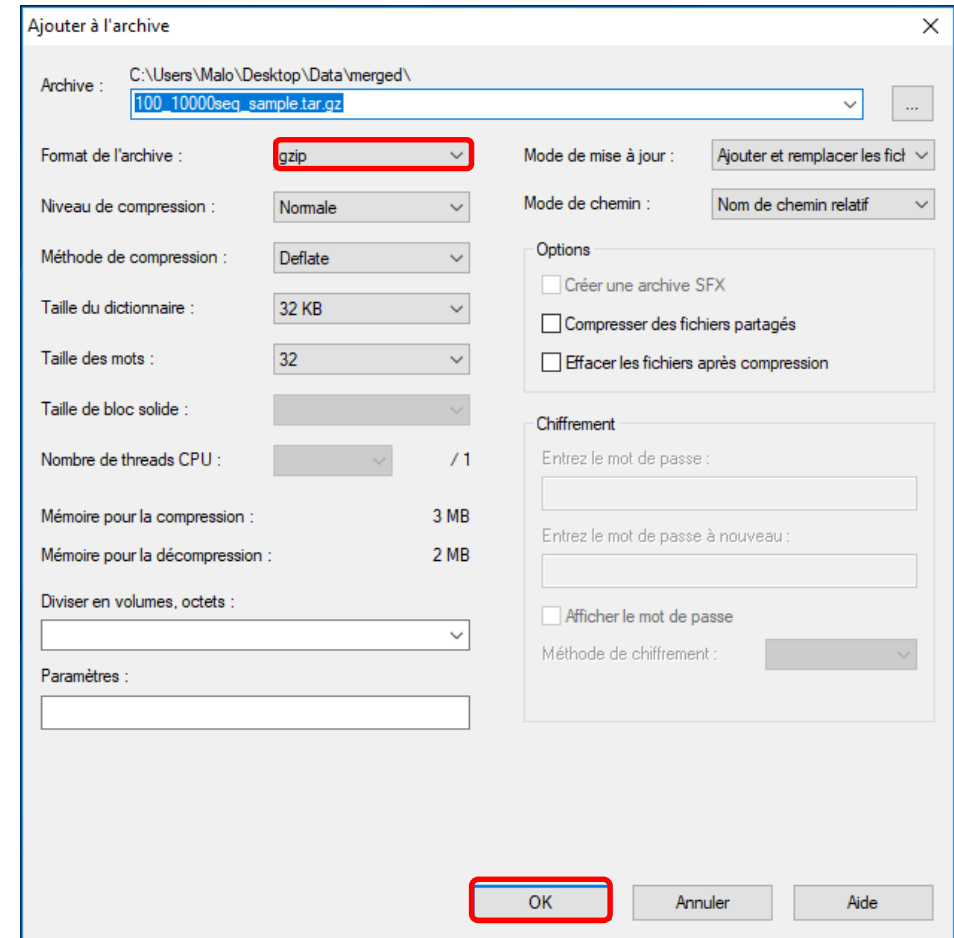
Create a Tar file with 7-Zip

- Select the newly created archive and right click on it > 7-Zip > Add to archive . . .

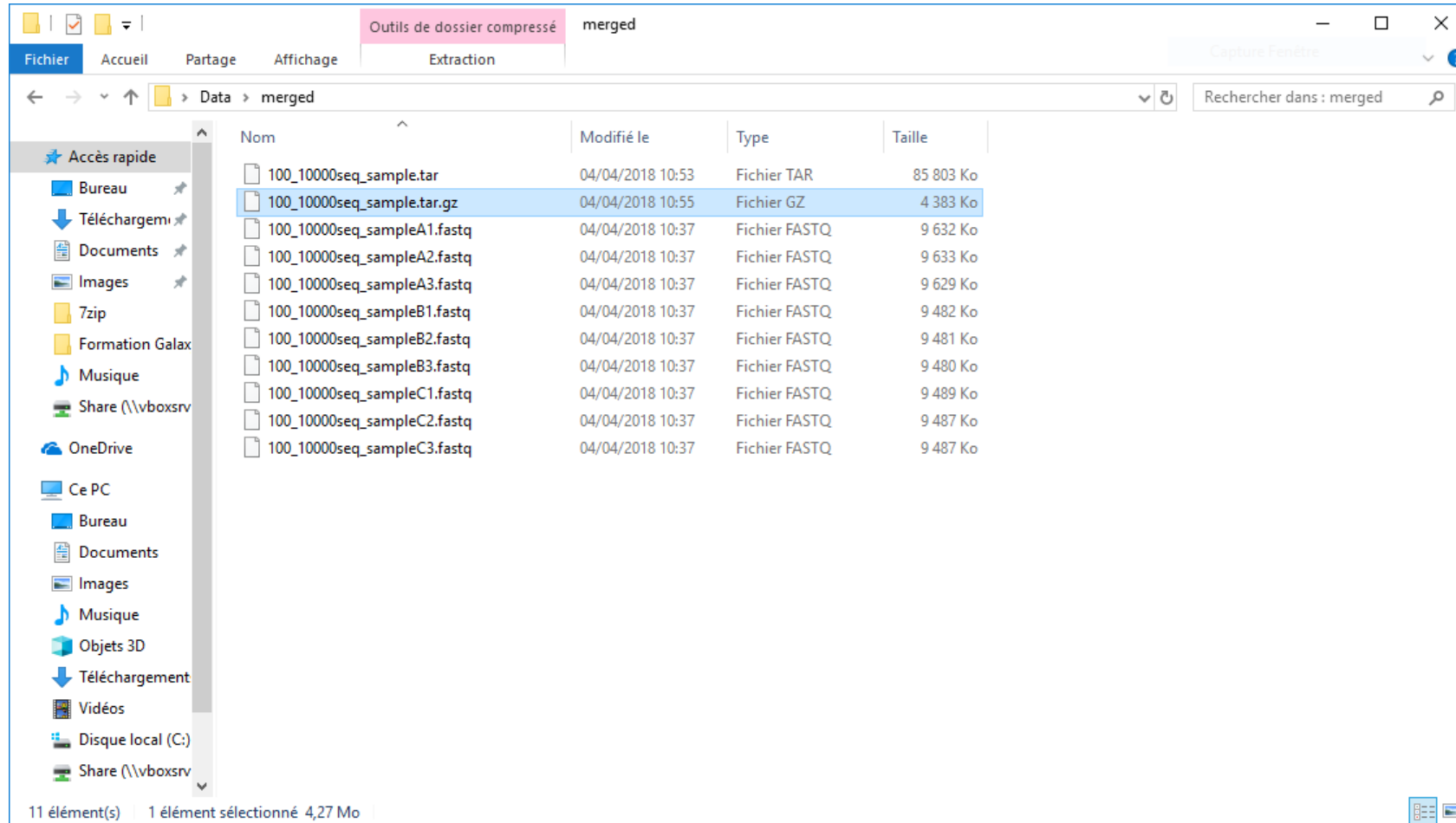


Create a Tar file with 7-Zip

- Choose « gzip » as archive format .
- Click on OK .

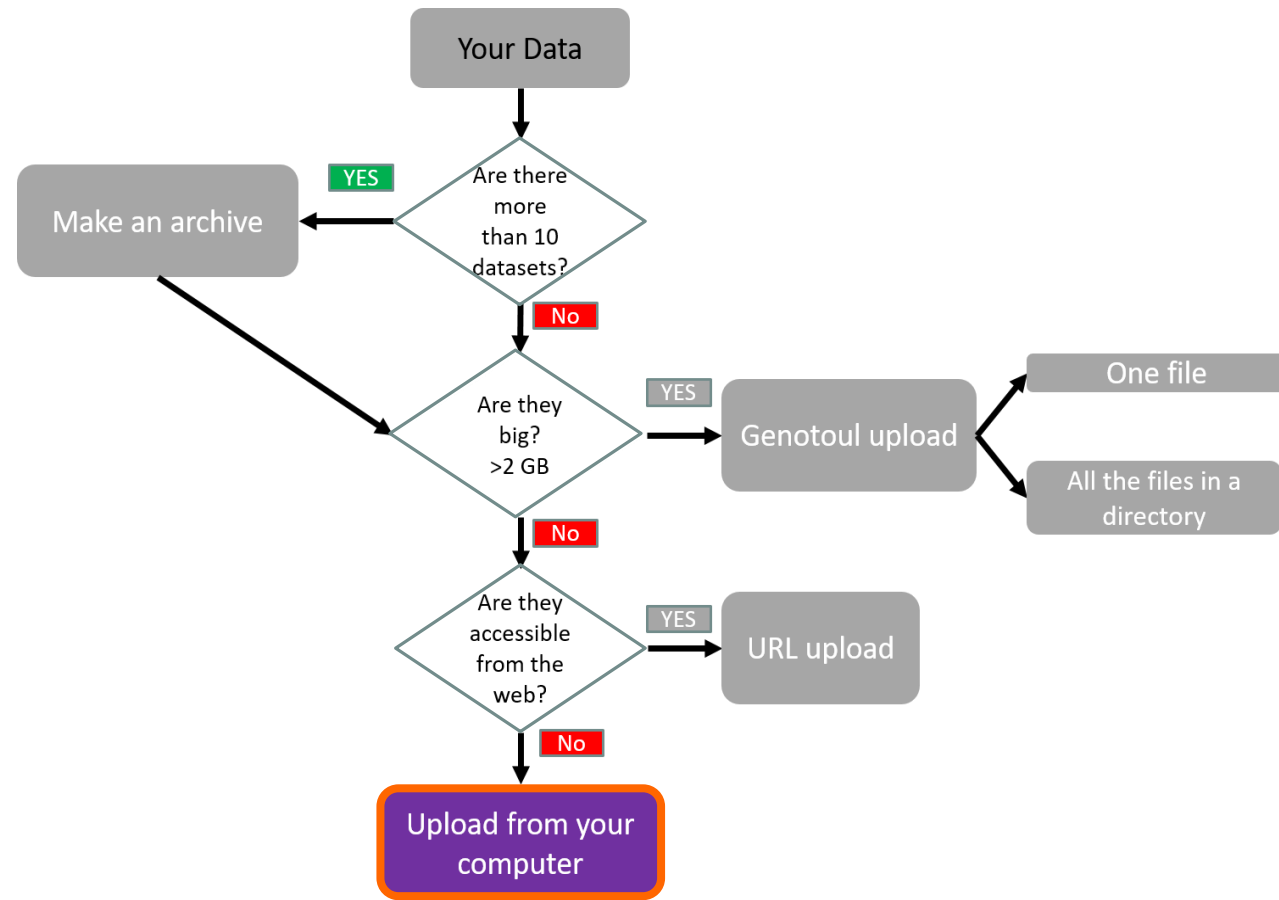


Create a Tar file with 7-Zip



Your Turn!

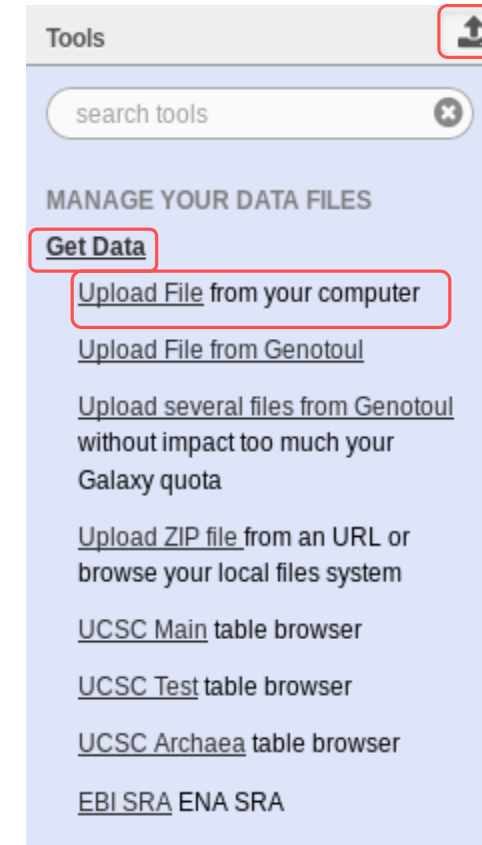
UPLOAD FILES FROM YOUR COMPUTER



Upload local files

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

⚠ For files smaller than 2 GB



Upload local files

Download from web or upload from disk

Regular

Composite

Drop files here

Click here to choose a file on your hard drive.

Upload one of the two files in the multiplex directory.

Type (set all):

Auto-detect



Genome (set all):

unspecified (?)

Choose local file

Paste/Fetch data

Pause

Reset

Start

Close

Upload local files

The datatype of a file indicates the content of it.

.fasta = fasta

.fastq = fastq

.tar = tar





Select the type corresponding to your file.

Do not trust blindly the auto-detect!

Download from web or upload from disk







Regular Composite

You added 2 file(s) to the queue. Add more files or click 'Start' to proceed.

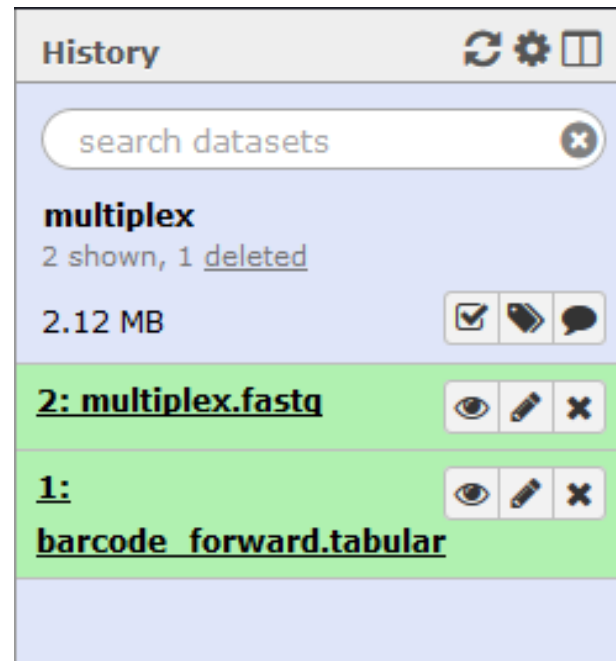
Name	Size	Type	Genome	Settings	Status
 barcode_forward.tabular	260 b	Auto-det...	unspecified (?)		0%
 multiplex.fastq	2.1 MB	Auto-det...	unspecified (?)		0%

1. Select the type of file (Do not leave on Auto-Detect!).
2. Select your other files the same way.
3. Begin upload.

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

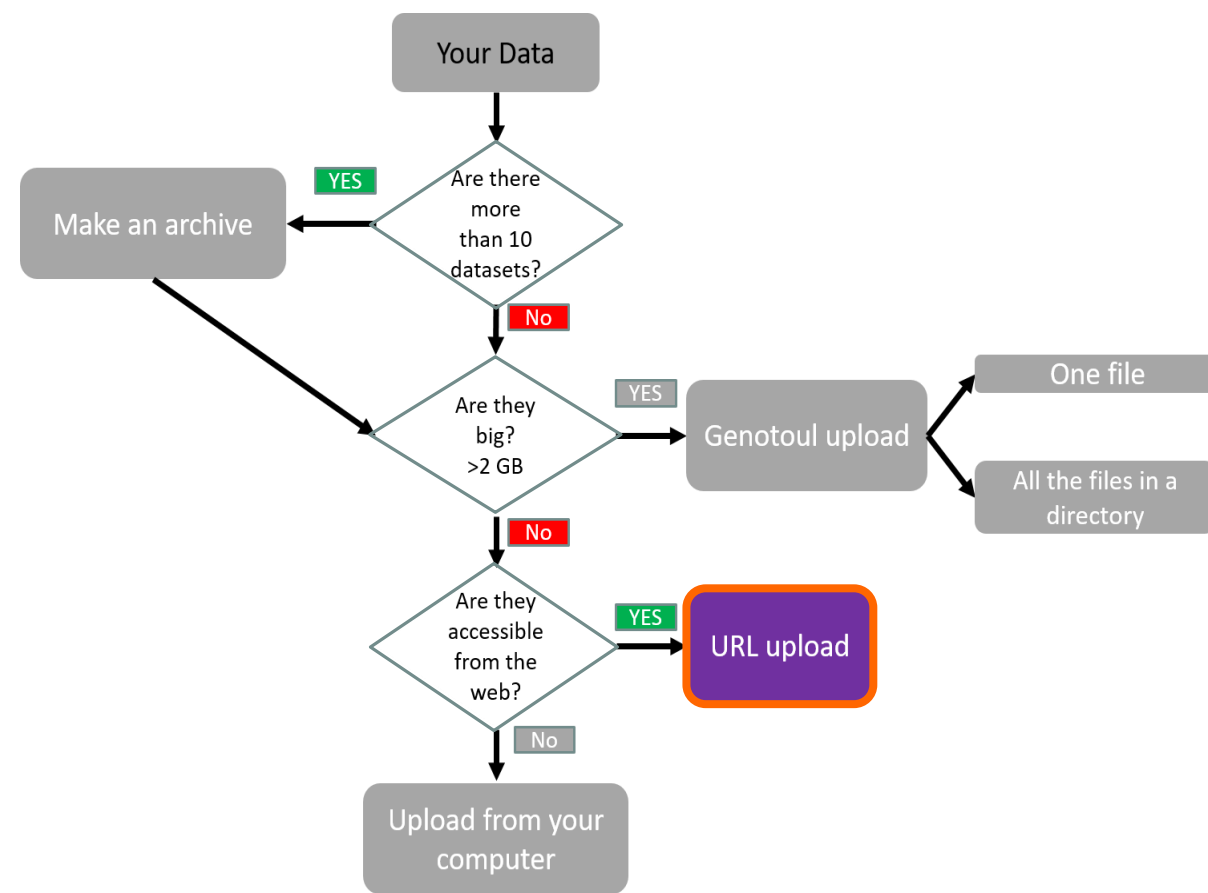
 Choose local file  Paste/Fetch data    

Upload local files



Your Turn!

UPLOAD FILE FROM AN URL



Upload file from URL

1. Switch to 454 **or ITS** history as current history.
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/FROGS_ini/DATA/454.fastq

Or

http://genoweb.toulouse.inra.fr/~formation/15_FROGS/FROGS_ini/DATA/ITS.tar.gz

1. Change the type!
2. Click on Start.
3. You can put one address per line for multiple uploads.

Download from web or upload from disk

Regular Composite

You added 1 file(s) to the queue. Add more files or click 'Start' to proceed.

Name	Size	Type	Genome	Settings	Status
New File	76 b	Auto-detect	unspecified (?)		

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.

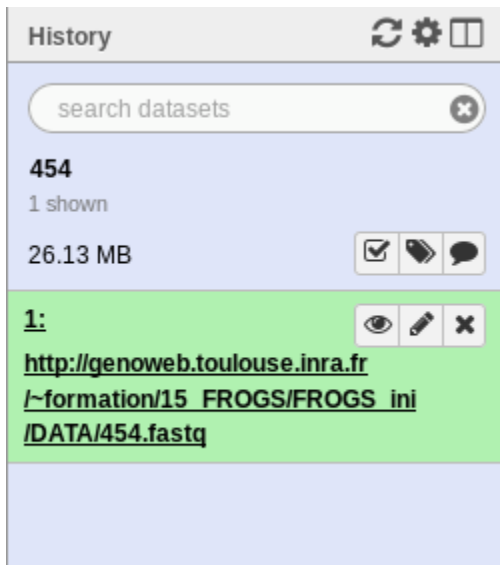
http://genoweb.toulouse.inra.fr/~formation/15_FROGS/FROGS_ini/DATA/454.fastq

http://genoweb.toulouse.inra.fr/~formation/15_FROGS/FROGS_ini/DATA/ITS.tar.gz

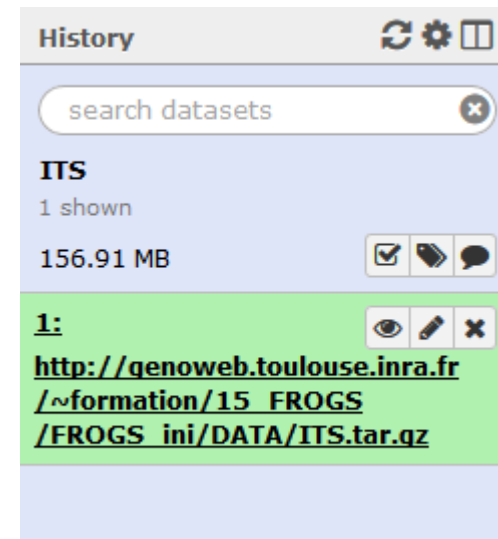
Type (set all): Auto-detect Genome (set all): unspecified (?)

Choose local file Paste/Fetch data Pause Reset **Start** Close

Upload file from URL

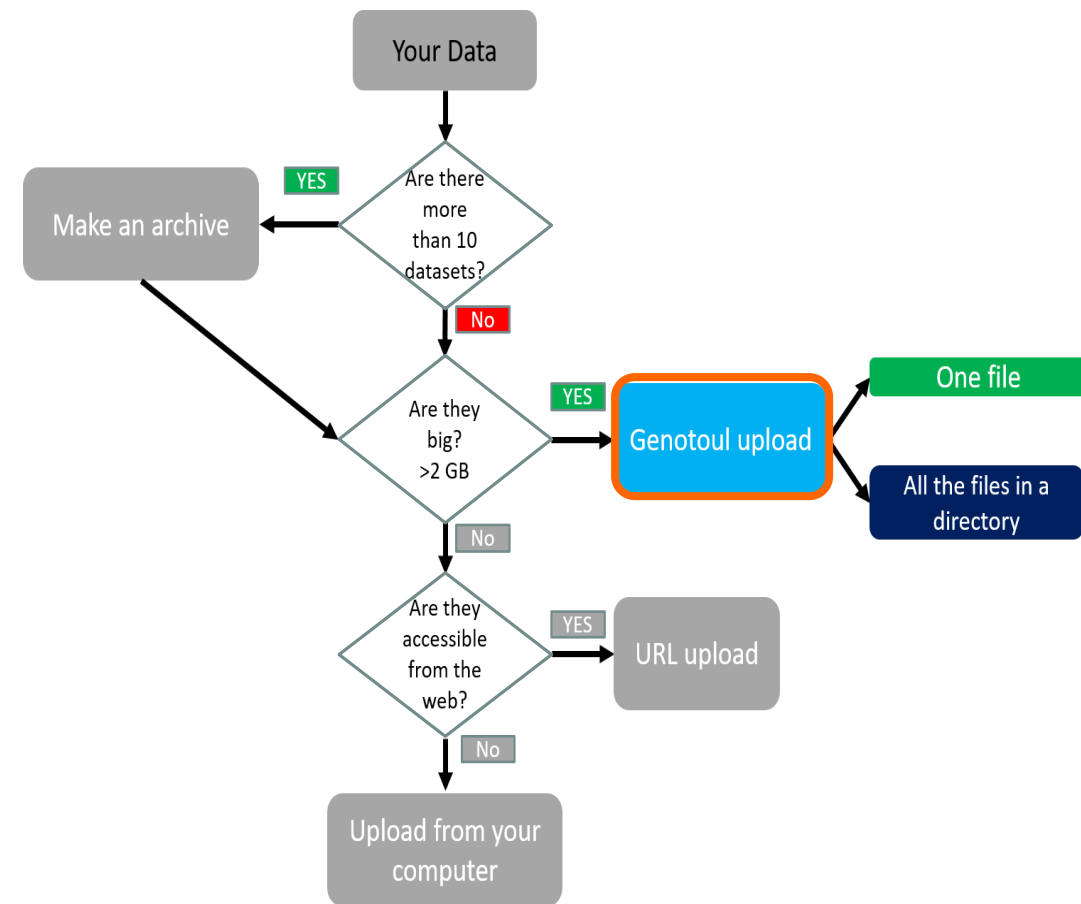


OR

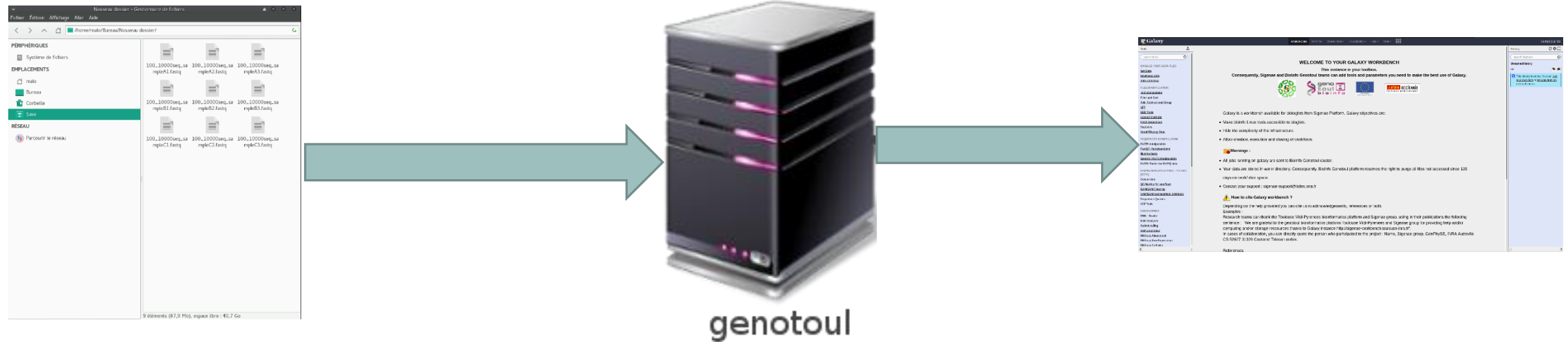


Your Turn!

UPLOAD FILES TO GENOTOUL AND LINK IT TO GALAXY



Objectives

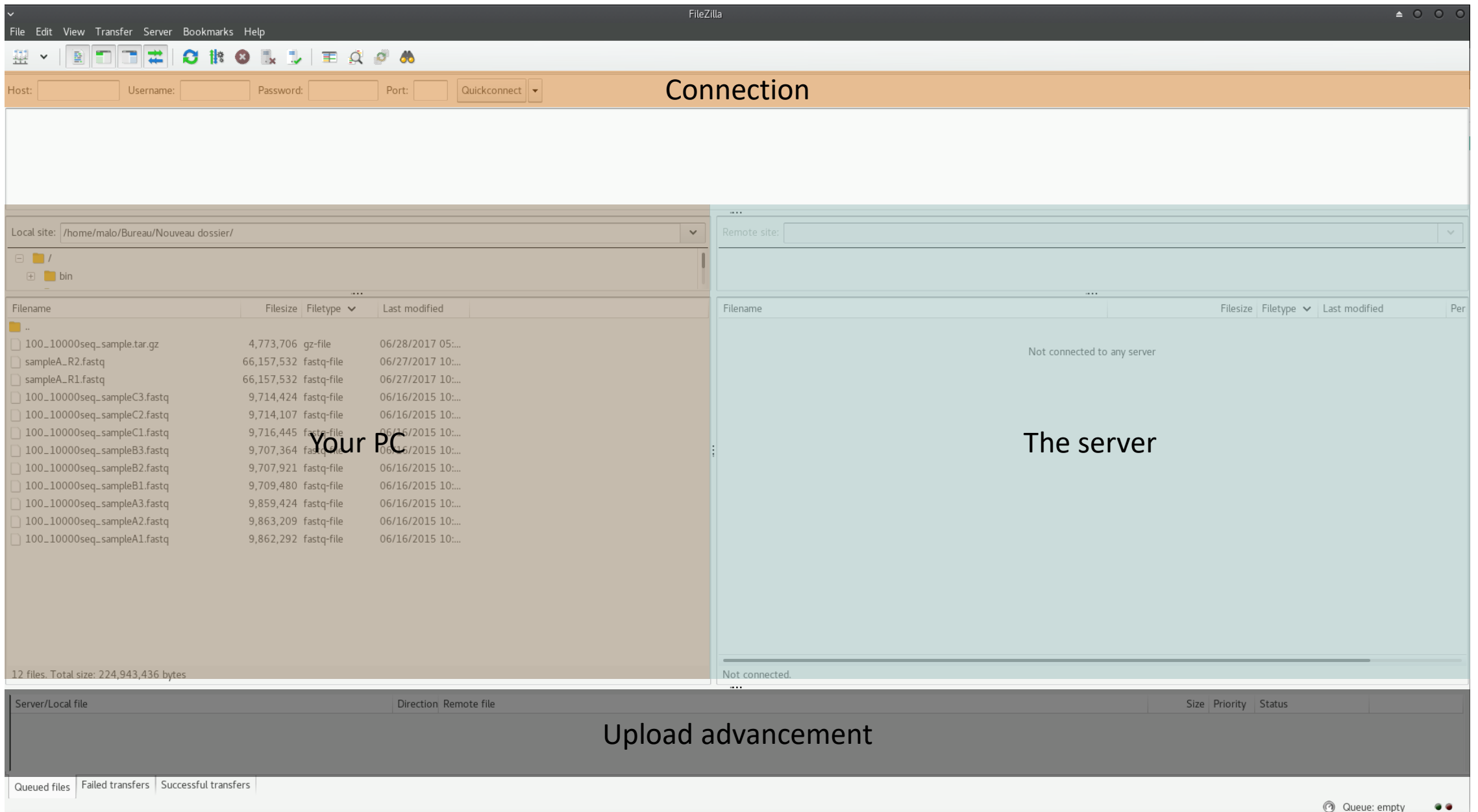


Transfer your files to your Genotoul account and link the file to Galaxy.

Preparation

- Open the data directory on your desktop.
- Launch Filezilla.
 - Filezilla is a FTP client *i.e.* can transfer files to a distant server.
 - You can download it on : <https://filezilla-project.org/>





Host:

Username:

Password:

Port:

Quickconnect

Status: Retrieving directory listing of /home/malo/Bureau/Data/

Status: Listing directory /work/mleboulch

Status: Directory listing of "/work/mleboulch" successful

Status: Retrieving directory listing of "/work/mleboulch/Formation"...

Status: Listing directory /work/mleboulch/Formation

Status: Directory listing of "/work/mleboulch/Formation" successful

Status: Deleting 2 files from "/work/mleboulch/Formation"

Status: Disconnected from server

Local site: /home/malo/Bureau/Data/

Data

merged

Filename	Filesize	Filetype	Last modified
..			
temp		Directory	06/29/2017 05:...
multiplex		Directory	06/29/2017 05:...
merged		Directory	06/29/2017 05:...

3 directories

Remote site: /work/mleboulch/Formation

mleboulch

Formation

Filename	Filesize	Filetype	Last modified	Per
..				

Empty directory listing

Empty directory.

Server/Local file

Direction

Remote file

Size

Priority

Status

Queued files

Failed transfers

Successful transfers (2)

Queue: empty

Browse to the data directory on your desktop.

Host: Username: Password: Port: Quickconnect

Status: Retrieving directory listing of /home/maleboulch/Data/ ...
Status: Listing directory /work/mleboluch
Status: Directory listing of "/work/mleboluch" successful
Status: Retrieving directory listing of "/work/mleboluch/Formation" ...
Status: Listing directory /work/mleboluch/Formation
Status: Directory listing of "/work/mleboluch/Formation" successful
Status: Deleting 2 files from "/work/mleboluch/Formation"
Status: Disconnected from server

Local site: /home/maleboulch/Data/

☐ Data
☐ merged

Filename	Filesize	Filetype <input type="button" value="v"/>	Last modified
..			
temp		Directory	06/29/2017 05:...
multiplex		Directory	06/29/2017 05:...
merged		Directory	06/29/2017 05:...

3 directories

Remote site: /work/mleboluch/Formation

☐ mleboluch
☐ Formation

Filename	Filesize	Filetype <input type="button" value="v"/>	Last modified	Per
..				

Empty directory listing

Empty directory.

Server/Local file	Direction	Remote file	Size	Priority	Status
-------------------	-----------	-------------	------	----------	--------

Queued files | Failed transfers | Successful transfers (2)

Queue: empty

To connect to Genotoul click on this icon.

- Host: genologin.toulouse.inra.fr
 - Port: 22
 - Protocol: SFTP
 - User: your Genotoul login
 - Logon Type: Ask for password
 - Click on connect.
-
- Password: your password

Site Manager

Select Entry:

- My Sites
 - Formation
 - Genotoul

Buttons: New Site, New Folder, New Bookmark, Rename, Delete, Duplicate

General Tab:

Protocol: SFTP - SSH File Transfer Protocol

Host: genologin.toulouse.inra.fr Port: 22

Logon Type: Ask for password

User: mleboulch

Password:

Background color: None

Comments:

Buttons: Connect, OK, Cancel

sftp://mleboulch@genotoul.toulouse.inra.fr - FileZilla

File Edit View Transfer Server Bookmarks Help

Host: Username: Password: Port: Quickconnect

Status: Connecting to genotoul.toulouse.inra.fr...
Status: Connected to genotoul.toulouse.inra.fr...
Status: Retrieving directory listing...
Status: Listing directory /home/mleboulch
Status: Directory listing of "/home/mleboulch" successful

Local site: /home/malo/Bureau/Data/ Remote site: /home/mleboulch

Local site directory listing:

Filename	Filesize	Filetype	Last modified
..			
temp		Directory	06/29/2017 05:...
multiplex		Directory	06/29/2017 05:...
merged		Directory	06/29/2017 05:...

Remote site directory listing:

Filename	Filesize	Filetype	Last modified
..			
.recently-used.xbel	4,162	xbel-file	03/30/2017 02:26:33 PM
.viminfo	615	File	02/14/2017 09:59:45 AM
.python_history	542	File	06/01/2017 11:08:10 AM
.bashrc	52	File	02/14/2017 09:59:45 AM
.bash_profile	226	File	05/15/2017 03:39:59 PM
.bash_history	17,308	File	06/23/2017 11:15:47 AM
.Xauthority	4,608	File	06/23/2017 11:15:15 AM
work		Directory	10/21/2016 04:37:06 PM
save		Directory	10/21/2016 04:37:06 PM
AIC-prefs		Directory	02/14/2017 10:04:44 AM
.ssh		Directory	01/10/2017 02:28:16 PM
.pki		Directory	03/30/2017 02:56:02 PM
.pathway-tools		Directory	06/01/2017 10:45:23 PM
.gnome2		Directory	02/01/2017 06:14:17 PM
.gconfd		Directory	03/30/2017 02:26:51 PM
.gconf		Directory	03/30/2017 09:08:21 AM
.dbus		Directory	02/01/2017 06:12:40 PM
.config		Directory	04/25/2017 04:13:46 PM

3 directories

Selected 1 directory.

Server/Local file Direction Remote file Size Priority Status

Queued files Failed transfers Successful transfers

Queue: empty

Double click on work directory to access it.

sftp://mleboulch@genotoul.toulouse.inra.fr - FileZilla

File Edit View Transfer Server Bookmarks Help

Host: Username: Password: Port: Quickconnect

Status: Deleting 2 files from /work/mleboulch/Formation
Status: Disconnected from server
Status: Retrieving directory listing of "/work/mleboulch"...
Status: Listing directory /work/mleboulch
Status: Directory listing of "/work/mleboulch" successful
Status: Retrieving directory listing of "/work/mleboulch/Formation"...
Status: Listing directory /work/mleboulch/Formation
Status: Directory listing of "/work/mleboulch/Formation" successful

Local site: /home/malo/Bureau/Data/ Remote site: /work/mleboulch/

Local site contents: Data, merged

Remote site contents: mleboulch, Formation

Local site file list:

Filename	Filesize	Filetype	Last modified
..		Directory	06/29/2017 05:...
temp			
multiplex			
merged			

3 directories

Remote site file list:

Filename	Filesize	Filetype	Last modified	Per
..				

Empty directory listing

Empty directory.

Server/Local file | Direction | Remote file | Size | Priority | Status

Queued files | Failed transfers | Successful transfers (2)

Queue: empty

- Create new directory with right click > Create directory.
- Name it « Formation ».
- Go inside this directory next.

sftp://mleboulch@genotoul.toulouse.inra.fr - FileZilla

File Edit View Transfer Server Bookmarks Help

Host: Username: Password: Port: Quickconnect

Status: Listing directory /home/mleboulch
Status: Directory listing of "/home/mleboulch" successful
Status: Retrieving directory listing of "/home/mleboulch/work"...
Status: Listing directory /work/mleboulch
Status: Directory listing of "/work/mleboulch" successful
Status: Retrieving directory listing of "/work/mleboulch/Formation"...
Status: Listing directory /work/mleboulch/Formation
Status: Directory listing of "/work/mleboulch/Formation" successful

Local site: /home/malo/Bureau/Data/ Remote site: /work/mleboulch/Formation

Local site: /
bin

Remote site: mleboulch
Formation

Filename	Filesize	Filetype	Last modified
..			
temp		Directory	06/29/2017 05:...
multiplex		Directory	06/29/2017 05:...
merged		Directory	06/29/2017 05:...

Empty directory listing

- Select the temp directory and drag and drop it following the arrow.
- Double click on the merge directory.

Selected 1 directory.

Empty directory.

Server/Local file	Direction	Remote file	Size	Priority	Status
-------------------	-----------	-------------	------	----------	--------

Queued files Failed transfers Successful transfers

Queue: empty

sftp://mleboulch@genotoul.toulouse.inra.fr - FileZilla

File Edit View Transfer Server Bookmarks Help

Host: Username: Password: Port: Quickconnect

Status: Listing directory /work/mleboulch/Formation/temp
Command: put "/home/malo/Bureau/Data/temp/temp2.fastq" "temp2.fastq"
Command: local:/home/malo/Bureau/Data/temp/temp2.fastq => remote:/work/mleboulch/Formation/temp/temp2.fastq
Status: File transfer successful, transferred 2,226,378 bytes in 1 second
Status: File transfer successful, transferred 2,226,378 bytes in 1 second
Status: Retrieving directory listing of "/work/mleboulch/Formation"...
Status: Listing directory /work/mleboulch/Formation
Status: Directory listing of "/work/mleboulch/Formation" successful

Local site: /home/malo/Bureau/Data/merged/

Remote site: /work/mleboulch/Formation

Local site directory listing:

Filename	Filesize	Filetype	Last modified
..			
100_10000seq_sample.tar.gz	4,773,706	gz-file	06/28/2017 05:...
100_10000seq_sampleC3.fastq	9,714,424	fastq-file	06/16/2015 10:...
100_10000seq_sampleC2.fastq	9,714,107	fastq-file	06/16/2015 10:...
100_10000seq_sampleC1.fastq	9,716,445	fastq-file	06/16/2015 10:...
100_10000seq_sampleB3.fastq	9,707,364	fastq-file	06/16/2015 10:...
100_10000seq_sampleB2.fastq	9,707,921	fastq-file	06/16/2015 10:...
100_10000seq_sampleB1.fastq	9,709,480	fastq-file	06/16/2015 10:...
100_10000seq_sampleA3.fastq	9,859,424	fastq-file	06/16/2015 10:...
100_10000seq_sampleA2.fastq	9,863,209	fastq-file	06/16/2015 10:...
100_10000seq_sampleA1.fastq	9,862,292	fastq-file	06/16/2015 10:...

Remote site directory listing:


Filename	Filesize	Filetype	Last modified	Per
..				
temp		Directory	06/29/2017 06:01:32 PM	drw...

Selected 1 file. Total size: 4,773,706 bytes

Server/Local file

Queued files Failed transfers Successful transfers (2)

Queue: empty



■ Select the 100_1000seq_sample.tar.gz file and drag and drop it following the arrow.

■ Double click on the merge directory.

sftp://mleboulch@genotoul.toulouse.inra.fr - FileZilla

File Edit View Transfer Server Bookmarks Help

Host: Username: Password: Port: Quickconnect

Status: Directory listing of "/work/.vle" successful
Status: Retrieving directory listing of "/work/BPMP" ...
Status: Listing directory /work/BPMP
Status: Directory listing of "/work/BPMP" successful
Status: Retrieving directory listing of "/work/mlebeurier2" ...
Command: cd "/work/mlebeurier2"
Error: Directory /work/mlebeurier2: permission denied
Error: Failed to retrieve directory listing

Local site: /home/malo/Bureau/Data/merged/ Remote site: /work/mlebeurier2

Data
merged

Filename Filesize Filetype Last modified

100_10000seq_sampleB1.fastq	9,709,480	fastq-file	04/04/2018 10:...
100_10000seq_sampleC2.fastq	9,714,107	fastq-file	04/04/2018 10:...
100_10000seq_sampleC3.fastq	9,714,424	fastq-file	04/04/2018 10:...
100_10000seq_sampleC1.fastq	9,716,445	fastq-file	04/04/2018 10:...
100_10000seq_sampleA3.fastq	9,859,424	fastq-file	04/04/2018 10:...
100_10000seq_sampleA1.fastq	9,862,292	fastq-file	04/04/2018 10:...

Select your work directory.

Right click on it and click on File Attributes.

mlebeulier2

Database

Filename Filesize Filetype Last modified Permissions

Directory		Directory	06/08/2018 11:19:08 AM	drwxr-xr-x
Directory		Directory	03/14/2018 02:21:50 PM	drwxr-xr-x
Directory		Directory	03/13/2018 11:50:00 AM	drwxr-xr-x
Directory		Directory	03/09/2018 04:09:45 PM	drwxr-xr-x
Directory		Directory	03/06/2018 04:25:02 PM	drwxr-xr-x
Directory		Directory	03/06/2018 01:24:27 PM	drwxr-xr-x
Directory		Directory	02/28/2018 03:48:40 PM	drwxr-xr-x
Directory		Directory	02/28/2018 03:47:01 PM	drwxr-xr-x
Directory		Directory	02/28/2018 09:29:46 AM	drwxr-xr-x
Directory		Directory	10/30/2017 05:07:06 PM	drwxr-xr-x
Directory		Directory	06/30/2017 04:26:59 PM	drwxr-xr-x

11 directories

Queued files Failed transfers Successful transfers

Queue: empty

Change file attributes

Please select the new attributes for the directory "Formation".

Owner permissions

<input checked="" type="checkbox"/> Read	<input checked="" type="checkbox"/> Write	<input checked="" type="checkbox"/> Execute
--	---	---

Group permissions

<input checked="" type="checkbox"/> Read	<input type="checkbox"/> Write	<input checked="" type="checkbox"/> Execute
--	--------------------------------	---

Public permissions

<input checked="" type="checkbox"/> Read	<input type="checkbox"/> Write	<input checked="" type="checkbox"/> Execute
--	--------------------------------	---

Numeric value:

You can use an x at any position to keep the permission the original files have.

☒ Recurse into subdirectories

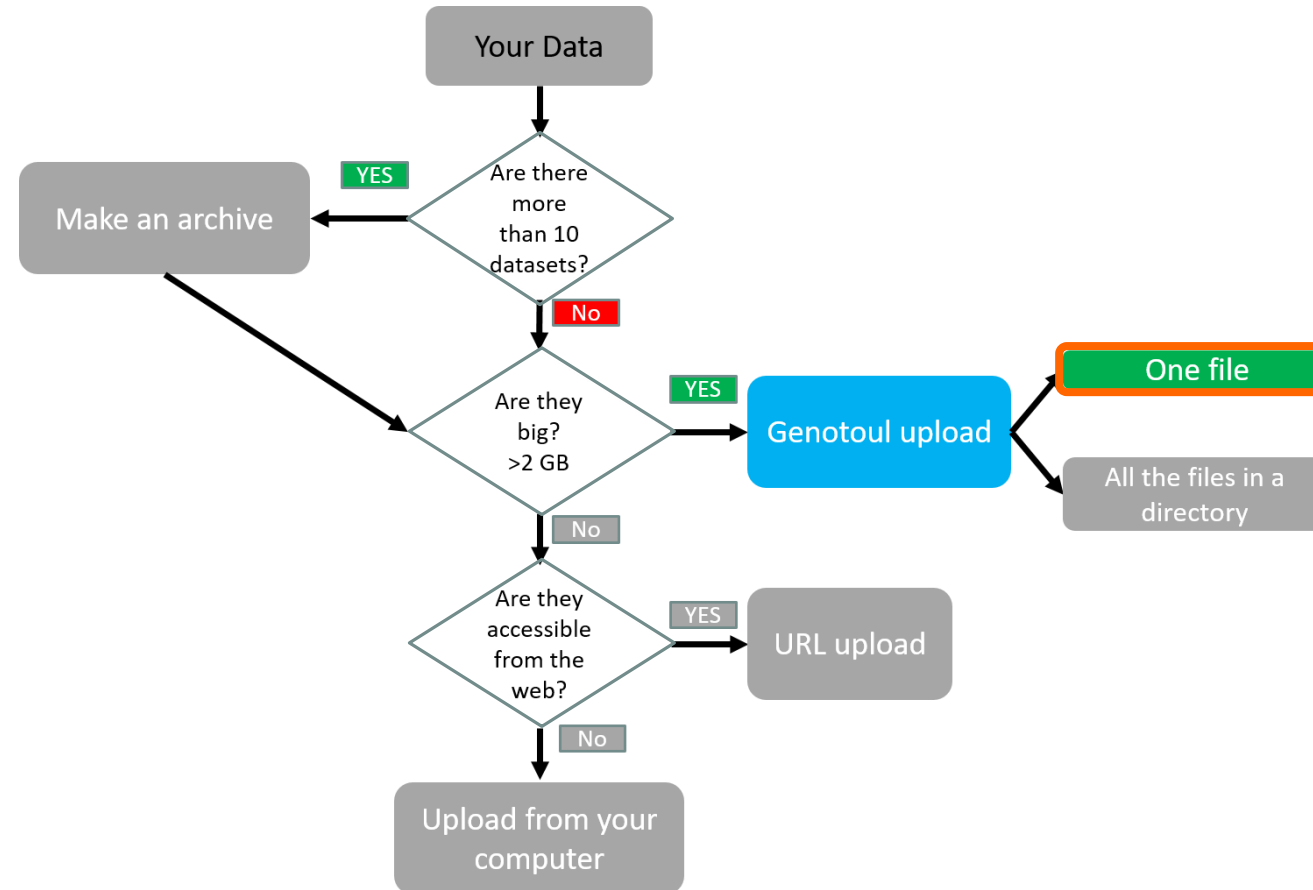
☒ Apply to all files and directories

☐ Apply to files only

☐ Apply to directories only

- Check that all boxes for execute and read are checked.
- Recurse this action to all files and subdirectories.
- That allows Galaxy to access your files on Genotoul.
 - Click on Ok.

Upload files from Genotoul



sftp://mleboulch@genotoul.toulouse.inra.fr - FileZilla

File Edit View Transfer Server Bookmarks Help

Host: Username: Password: Port: Quickconnect

Status: Setting permissions of /work/mleboulch/Formation/temp/temp2.fastq to 755
Status: /work/mleboulch/Formation/temp/temp2.fastq: 0644 -> 0755
Status: Retrieving directory listing of "/work/mleboulch/Formation"...
Status: Listing directory /work/mleboulch/Formation
Status: Directory listing of "/work/mleboulch/Formation" successful
Status: Retrieving directory listing of "/work/mleboulch/Formation/temp"...
Status: Listing directory /work/mleboulch/Formation/temp
Status: Directory listing of "/work/mleboulch/Formation/temp" successful

Local site: /home/malo/Bureau/Data/merged/

Data
merged

Filename	Filesize	Filetype	Last modified
100_10000seq_sampleB2.fastq	9,707,921	fastq-file	06/16/2015 10:...
100_10000seq_sampleB1.fastq	9,709,480	fastq-file	06/16/2015 10:...
100_10000seq_sampleA3.fastq	9,859,424	fastq-file	06/16/2015 10:...
100_10000seq_sampleA2.fastq	9,863,209	fastq-file	06/16/2015 10:...
100_10000seq_sampleA1.fastq	9,862,292	fastq-file	06/16/2015 10:...

10 files. Total size: 92,628,372 bytes

Remote: /work/mleboulch/Formation

Formation
temp

Filename	Filesize	Filetype	Last modified	Per
100_10000seq_sample.tar.gz	4,773,706	gz-file	06/29/2017 06:04:35 PM	-rw
temp		Directory	06/29/2017 06:01:32 PM	drw

1 file and 1 directory. Total size: 4,773,706 bytes

Server/Local file Direction Remote file

Queued files Failed transfers Successful transfers (3)

Queue: empty

Select this text and copy it by pressing the keyboard keys Ctrl+C.

Galaxy

Analyze DataWorkflowShared DataVisualizationHelpUser

Using 808.6 MB

Tools

search tools

MANAGE YOUR DATA FILES

Get Data

[Upload File from your computer](#)
[Upload File from Genotoul](#)
[Upload several files from Genotoul without impact too much your Galaxy quota](#)
[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](#)
[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](#)

Upload File from Genotoul (Galaxy Version 1.0.0)

Options

Path to file

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

File type

tar.gz

Execute

What it does

This procedure 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

valid path : /work/LinuxUser/invalid path : /home/L

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

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

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

- `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:

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

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

- Add "r" rights (only) to "others" on data.fasta file.

Thus, Galaxy can access and read data.fasta but all other files in those directories will not be accessible or readable.

fastq files have to be uploaded in a correct format (for instance, fastqsanger) in order to be used by Galaxy tools. If this is not the case, your fastq files uploaded will not be listed among available datasets in Galaxy tools.

Version Galaxy Tool : V1.0

Versions of bioinformatics tools used : No bioinformatique tool used.

History

search datasets

merged

0 b

This history is empty. You can load your own data or get data from an external source

Switch to merged history.

Next go to Get Data > Upload File from Genotoul.

Galaxy

Analyze DataWorkflowShared DataVisualizationHelpUser

Using 808.6 MB

Tools

search tools

MANAGE YOUR DATA FILES

Get Data

Upload File from your computer

Upload File from Genotoul

Upload several files from Genotoul without impact too much your Galaxy quota

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

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

Upload File from Genotoul (Galaxy Version 1.0.0)

Options

Path to file

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

File type

tar.gz

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 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.

2. Add "r" rights (only) to "others" on `data.fasta` file.

Thus, Galaxy can access and read `data.fasta` but all other files in those directories will not be accessible or readable.

fastq files have to be uploaded in a correct format (for instance, `fatsqsanger`) in order to be used by Galaxy tools. If this is not the case, your `fatsq` files uploaded will not be listed among available datasets in Galaxy tools.

Version Galaxy Tool : V1.0

Versions of bioinformatics tools used : No bioinformatique tool used.

History

search datasets

merged

0 b

This history is empty. You can load your own data or get data from an external source

Copy the text here by pressing Ctrl+V.

Don't forget to change the Datatype!

80

Path to file

/work/mleboulch/Formation

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



Path to file

/work/mleboulch/Formation/100_10000seq_sample.tar.gz

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

You must add « / » and the name of the file at the end of the text.

Galaxy

Analyze DataWorkflowShared DataVisualizationHelpUser

Using 813.2 MB

Tools

search tools

MANAGE YOUR DATA FILES

Get Data

Upload File from your computer

Upload File from Genotoul

Upload several files from Genotoul without impact too much your Galaxy quota

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

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

1 job has been successfully added to the queue - resulting in the following datasets:

1: /work/mleiboulch/Formation/100_10000seq_sample.tar.gz

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

History

search datasets

merged

1 shown

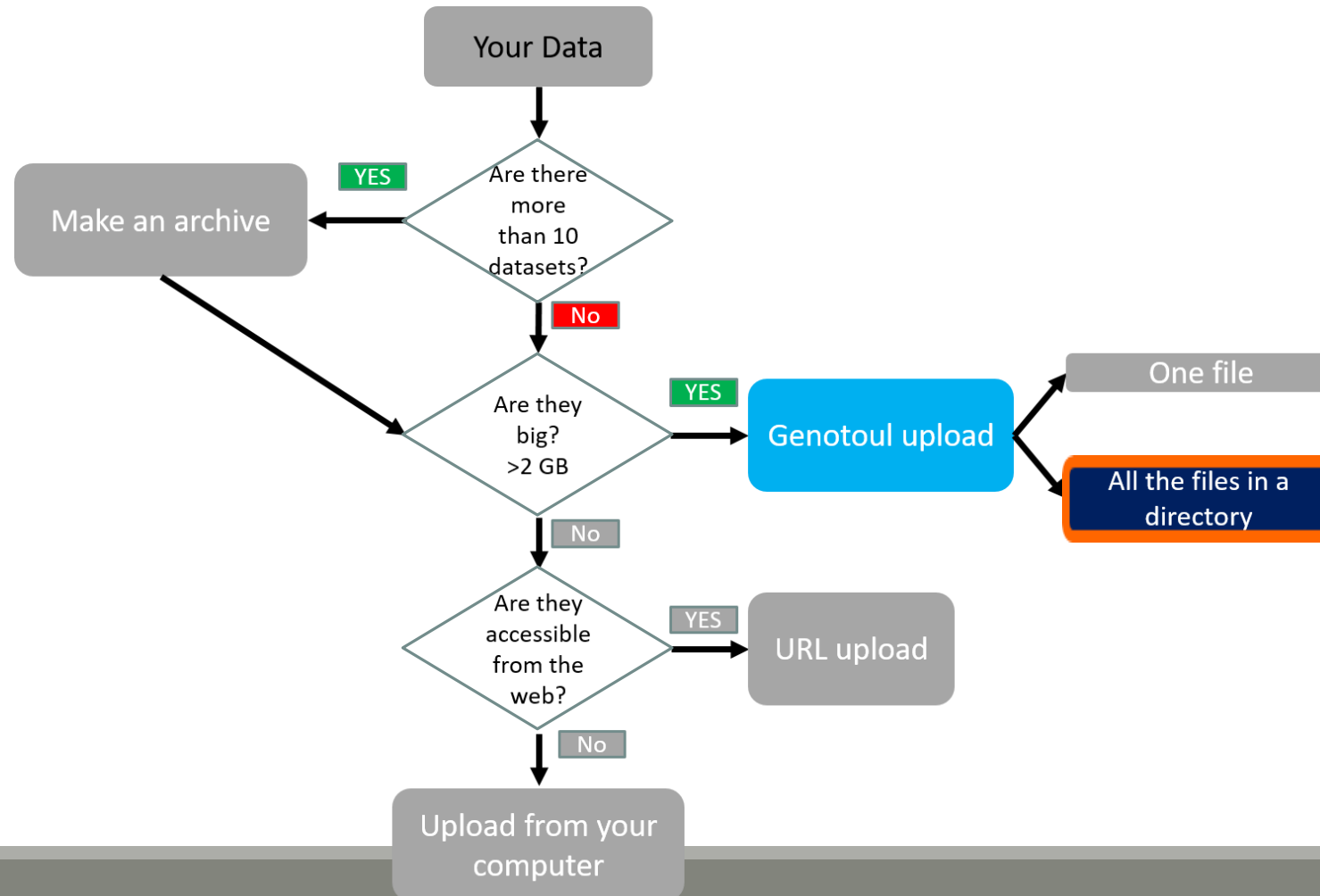
4.55 MB

1: /work/mleiboulch /Formation /100_10000seq_sample.tar.gz

After executing, the file is imported into Galaxy.

82

Upload files from Genotoul



sftp://mleboulch@genotoul.toulouse.inra.fr - FileZilla

File Edit View Transfer Server Bookmarks Help

Host: Username: Password: Port: Quickconnect

Status: Setting permissions of /work/mleboulch/Formation/temp/temp2.fastq to 755
Status: /work/mleboulch/Formation/temp/temp2.fastq: 0644 -> 0755
Status: Retrieving directory listing of "/work/mleboulch/Formation"...
Status: Listing directory /work/mleboulch/Formation
Status: Directory listing of "/work/mleboulch/Formation" successful
Status: Retrieving directory listing of "/work/mleboulch/Formation/temp"...
Status: Listing directory /work/mleboulch/Formation/temp
Status: Directory listing of "/work/mleboulch/Formation/temp" successful

Local site: /home/malo/Bureau/Data/merged/

Remote site: /work/mleboulch/Formation

Local site directory listing:

Filename	Filesize	Filetype	Last modified
..			
100_10000seq_sample.tar.gz	4,773,706	gz-file	06/28/2017 05:...
100_10000seq_sampleC3.fastq	9,714,424	fastq-file	06/16/2015 10:...
100_10000seq_sampleC2.fastq	9,714,107	fastq-file	06/16/2015 10:...
100_10000seq_sampleC1.fastq	9,716,445	fastq-file	06/16/2015 10:...
100_10000seq_sampleB3.fastq	9,707,364	fastq-file	06/16/2015 10:...
100_10000seq_sampleB2.fastq	9,707,921	fastq-file	06/16/2015 10:...

10 files. Total size: 92,628,372 bytes

Remote site directory listing:

Filename	Filesize	Filetype	Last modified	Per
..				
100_10000seq_sample.tar.gz	4,773,706	gz-file	06/29/2017 06:04:35 PM	-rw-rw-r--
temp		Directory	06/29/2017 06:01:32 PM	drwxr-xr-x

1 file and 1 directory. Total size: 4,773,706 bytes

Server/Local file | Direction | Remote file | Size | Priority | Status

Queued files | Failed transfers | Successful transfers (3)

Queue: empty

Double click on the temp directory.

sftp://mleboulch@genotoul.toulouse.inra.fr - FileZilla

File Edit View Transfer Server Bookmarks Help

Host: Username: Password: Port: Quickconnect

Status: Listing directory /work/mleboulch/Formation
Status: Directory listing of "/work/mleboulch/Formation" successful
Status: Retrieving directory listing of "/work/mleboulch/Formation/temp"...
Status: Listing directory /work/mleboulch/Formation/temp
Status: Directory listing of "/work/mleboulch/Formation/temp" successful
Status: Retrieving directory listing of "/work/mleboulch/Formation/temp"...
Status: Listing directory /work/mleboulch/Formation/temp
Status: Directory listing of "/work/mleboulch/Formation/temp" successful

Local site: /home/malo/Bureau/Data/merged/ Remote site: /work/mleboulch/Formation/temp

Local site tree: Data merged

Remote site tree: Formation temp

Local site files:

Filename	Filesize	Filetype	Last modified
100_10000seq_sample.tar.gz	4,773,706	gz-file	
100_10000seq_sampleC3.fastq	9,714,424	fastq-file	
100_10000seq_sampleC2.fastq	9,714,107	fastq-file	
100_10000seq_sampleC1.fastq	9,716,445	fastq-file	06/16/2015 10:...
100_10000seq_sampleB3.fastq	9,707,364	fastq-file	06/16/2015 10:...
100_10000seq_sampleB2.fastq	9,707,921	fastq-file	06/16/2015 10:...
100_10000seq_sampleB1.fastq	9,709,480	fastq-file	06/16/2015 10:...
100_10000seq_sampleA3.fastq	9,859,424	fastq-file	06/16/2015 10:...
100_10000seq_sampleA2.fastq	9,863,209	fastq-file	06/16/2015 10:...
100_10000seq_sampleA1.fastq	9,862,292	fastq-file	06/16/2015 10:...

10 files. Total size: 92,628,372 bytes

Remote site files:

Filename	Filesize	Filetype	Last modified	Per
temp2.fastq	2,226,378	fastq-file	06/29/2017 06:01:33 PM	-rw
temp.fastq	2,226,378	fastq-file	06/29/2017 06:01:33 PM	-rw

2 files. Total size: 4,452,756 bytes

Server/Local file Direction Remote file Size Priority Status

Queued files Failed transfers Successful transfers (3)

Queue: empty

Copy this address.

Galaxy

Analyze DataWorkflowShared DataVisualizationHelpUser

Using 813.2 MB

Tools

search tools

MANAGE YOUR DATA FILES

Get Data

Upload File from your computer

Upload File from Genotoul

Upload several files from Genotoul without impact too much your Galaxy quota

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

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

Upload several files from Genotoul without impact too much your Galaxy quota (Galaxy Version 1.0.1)

Options

Path to your directory which contains several files

/work/mleiboulch/Formation/temp

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

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/directory

valid path : /work/LinuxUserName/directory

invalid path : /home/LinuxUserName/work/directory

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/LinuxUserNam

Example : drwxr-x--x 4

Thanks to the fact that this tool

For example, if your data to down

Add "x" rights to "others" on /work

It is not useful that "others" ha

Thus, Galaxy can access and rea

Version Galaxy Tool : V1.0

Versions of bioinformatics tools us

Contacts (noms et emails) : sigen

E-learning available : Yes.

Please cite :

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

Examples : Acknowledgements We wish to thank the SIGENAE group for

References : SIGENAE [http://www.sig

History

search datasets

temp

0 b

This history is empty. You can load your own data or get data from an external source

- Go back to Galaxy and switch to temp history.
- Go to Get Data > Upload Several files from Genotoul
 - Paste the address into field.
- Be careful all the files from the directory will be uploaded!

Galaxy

Analyze DataWorkflowShared DataVisualizationHelpUser

Using 819.5 MB

Tools

search tools

MANAGE YOUR DATA FILES

Get Data

Upload File from your computer

Upload File from Genotoul

Upload several files from Genotoul without impact too much your Galaxy quota

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

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

1 job has been successfully added to the queue - resulting in the following datasets:

1: Upload several files from Genotoul

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

History

search datasets

temp

3 shown

4.25 MB

3: Upload several files from Genotoul (temp2)

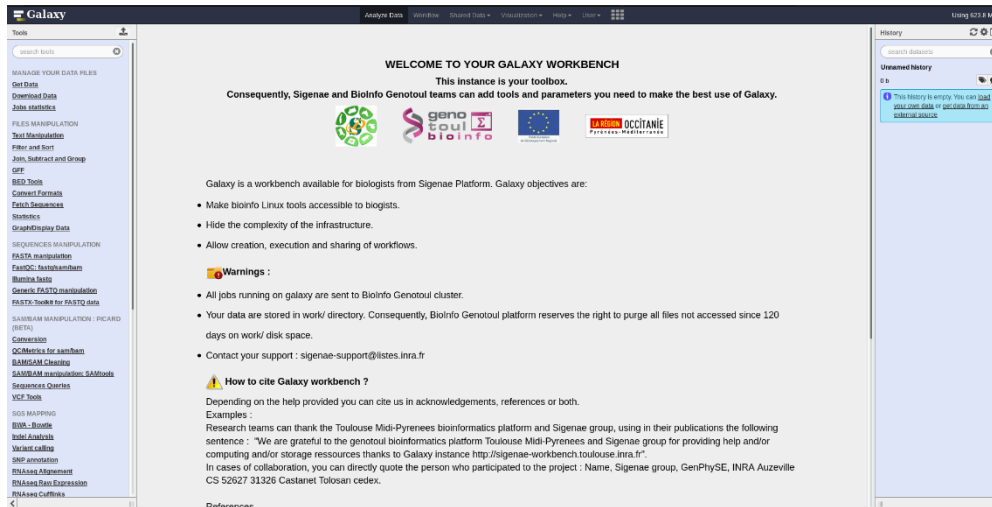
2: Upload several files from Genotoul (temp)

1: Upload several files from Genotoul

Click on execute.

All the files from the directory are uploaded.

Upload by Genotoul



1TB of space

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

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

Share a History

Share a history

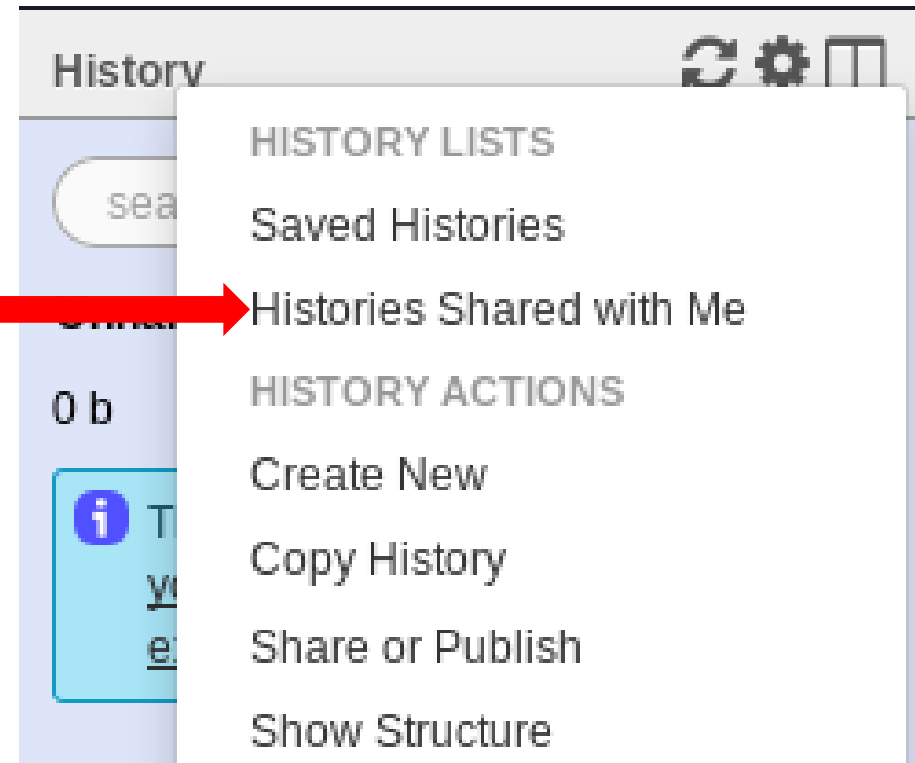
- You can share a history with another Galaxy user:
 - For working with your colleague.
 - For support, so we can help you better and faster.
- You can import shared history to your account too.

Your Turn!

IMPORT A SHARED HISTORY TO YOUR ACCOUNT

Import a shared history

Click on the wheel and click here.



Galaxy

Analyze DataWorkflowShared DataVisualizationHelpUser

Using 841.3 MB

Tools

search tools

MANAGE YOUR DATA FILES

Get DataDownload DataJobs statistics

FILES MANIPULATION

Text ManipulationFilter and SortJoin, Subtract and GroupGFFBED ToolsConvert FormatsFetch SequencesStatisticsGraph/Display Data

SEQUENCES MANIPULATION

FASTA manipulationFastQC: fastq/sam/bamIllumina fastqGeneric FASTQ manipulationFASTX-Toolkit for FASTQ data

SAM/BAM MANIPULATION : PICARD (BETA)

ConversionQC/Metrics for sam/bamBAM/SAM CleaningSAM/BAM manipulation: SAMtoolsSequences QueriesVCF Tools

Histories shared with you by others

<input type="checkbox"/>	Name	Datasets	Created	Last Updated	Shared by
<input type="checkbox"/>	ASMA2016	27	Nov 18, 2016	Mar 09, 2017	geraldine.pascal@inra.fr
<input type="checkbox"/>	Copy of 'ASMA_2016_FE0802' shared by laurent.cauquil@toulouse.inra.fr (active items only)	24	Dec 09, 2016	Jan 03, 2017	geraldine.pascal@inra.fr

for 0 selected histories: Copy Unshare

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

Check the box and click on copy to import R1R2 history into your Galaxy account.

You must import the « R1R2 » history.

Name of the history.

The person who shared the history.

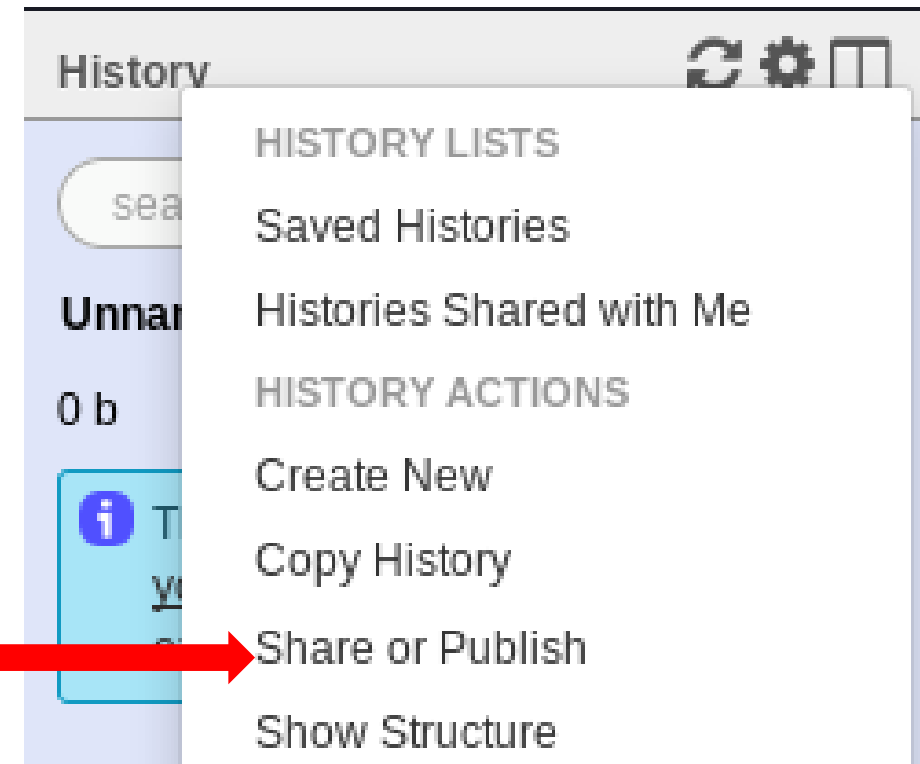
93

Your Turn!

SHARE A HISTORY WITH YOUR NEIGHBOUR

Share a history

- Switch to the history you want to share.
- 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](#)

Share or Publish History 'Historique R1R2'

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](#)

Click on « share with a user ».

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



Galaxy

Analyze DataWorkflowShared DataVisualizationHelpUser

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

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

Enter an email address from a Galaxy user.

Auto-completion helps you to find your co-worker

97

Galaxy user emails with which to share histories

malo.leboulch@inra.fr

malo.leboulch@inra.fr

Submit

a-separat



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

Manipulate datasets

Your Turn!

RENAME A DATASET

Rename a dataset

- Switch to 454 history or ITS history

The image shows two side-by-side screenshots of a web interface titled 'History'. The left screenshot shows 'Historique 454' with a dataset named '2: 454.fastq'. The right screenshot shows 'Historique ITS' with a dataset named '1:'. Both datasets have an edit icon (pencil) next to their names. A blue callout box on the right contains the text 'Click here to display attributes and change the name.' with two red arrows pointing to the edit icons in both screenshots.

Historique 454
1 shown, 1 deleted
26.13 MB

2: 454.fastq
26.1 MB
format: **fastq**, database: ?
uploaded fastq file

Historique ITS
1 shown
156.91 MB

1:
http://genoweb.toulouse.inra.fr/~formation/15_FROGS/FROGS_ini/DATA/ITS.tar.gz
156.9 MB
format: **tar**, database: ?
uploaded compressed archive file

Compressed binary file

Click here to display attributes and change the name.

Rename a dataset

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:

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

OR

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

Rename a dataset

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.

Rename a dataset

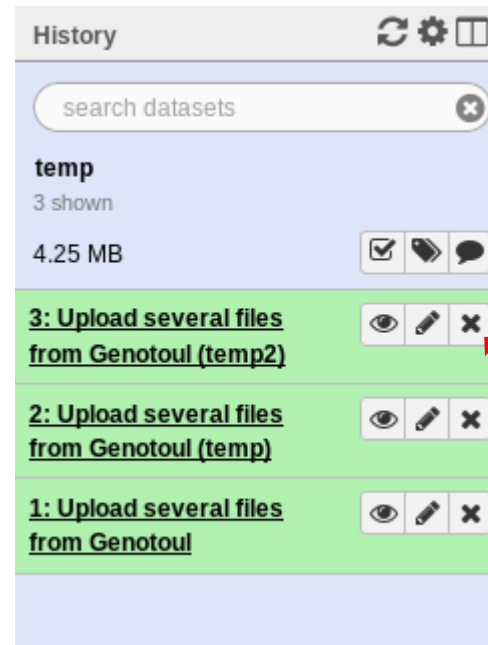
- Do the same with the merged history:
 - Switch to the merged history.
 - Change the name of the file to «100_10000seq_sample.tar.gz ».

Your Turn!

DELETE A DATASET

Delete a dataset

- Switch to temp history.



Click here to delete a dataset.

Delete a dataset

The diagram illustrates the process of deleting a dataset in Galaxy through two sequential screenshots of the 'History' panel, connected by a large blue arrow.

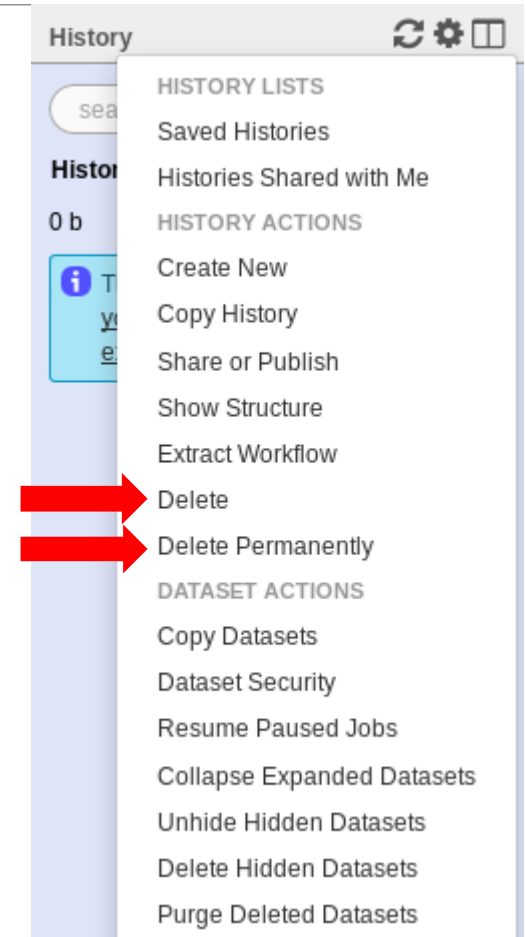
Left Screenshot: The 'History' panel shows a dataset named 'temp' with a size of 4.25 MB. A red box highlights the text '1 deleted' next to the dataset name. A blue callout box with the text 'Click here.' has a red arrow pointing to this text.

Right Screenshot: The 'History' panel shows the same dataset 'temp' with a size of 4.25 MB. A yellow warning box appears, stating 'This dataset has been deleted'. Inside this box, there are two options: 'Undelete it' and 'Permanently remove it from disk'. A red box highlights the 'Permanently remove it from disk' option. A blue callout box with the text 'Delete this dataset permanently.' has a red arrow pointing to this option.

Below the screenshots, a text box states: '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 ».



Current History

2 shown, 1 [deleted](#)

2.12 MB

search datasets

Drag datasets here to copy them to the current history

2: multiplex.fastq

1: barcode_forward.tabular

merged

1 shown

4.55 MB

search datasets

1: 100_10000seq_sample.tar.gz

454

1 shown

26.13 MB

search datasets

1: 454.fastq

Historique R1R2

2 shown, 1 [deleted](#)

126.19 MB

search datasets

3: sampleA_R2

2: sampleA_R1

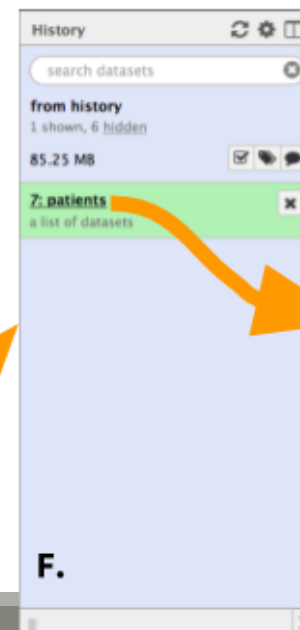
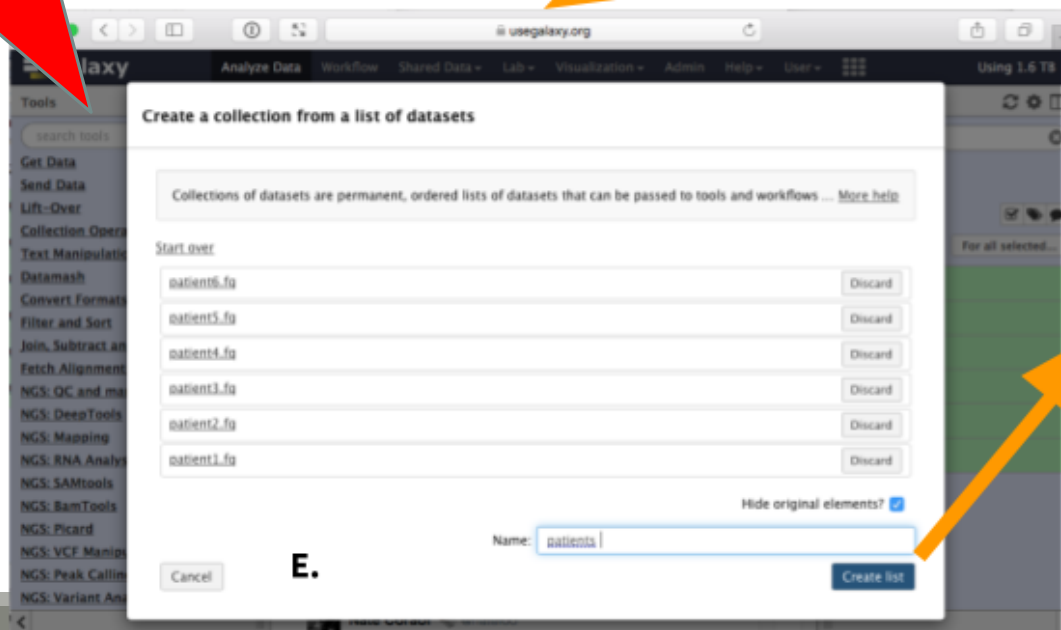
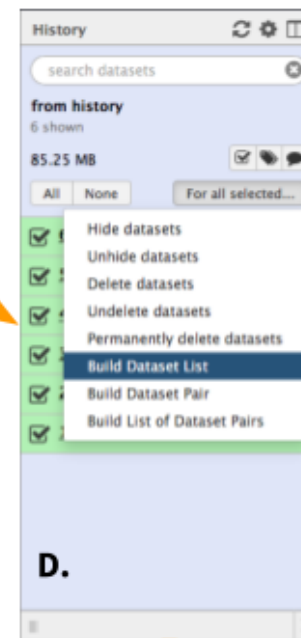
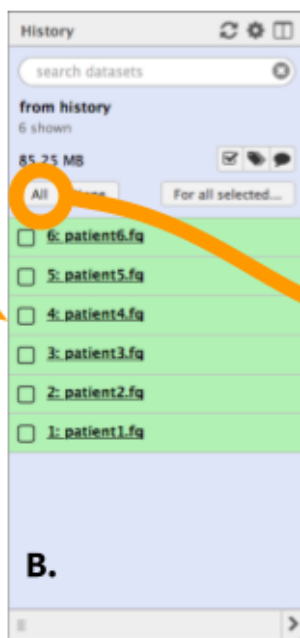
109

One word about data collections

- Used to perform the same analysis on multiple files.
- Gather multiple datasets in one collection.



Don't work with FROGS!



Others tools available on galaxy

- Text Manipulation
- Filter and Sort
- FASTA manipulation

Galaxy support

- Mail: support.sigenae@inra.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 Midi-Pyrenees bioinformatics platform and Sigenae group, using in their publications the following sentence : "We are grateful to the genotoul bioinformatics platform Toulouse Midi-Pyrenees and Sigenae group for providing help and/or computing and/or storage resources 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, INRA Auzeville CS 52627 31326 Castanet Tolosan cedex.