

Introduction to Galaxy

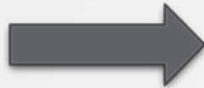
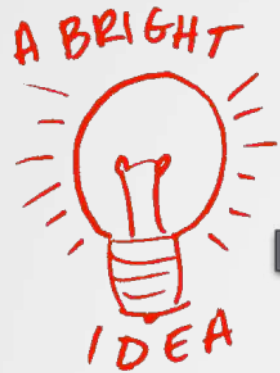
Stéphanie Le Gras
(slegras@igbmc.fr)

Guidelines

- Analyzing biological data with informatics tools
- Presentation of the Galaxy project
- Description of the main features of the Galaxy platform

Analyzing biological data with informatics tools

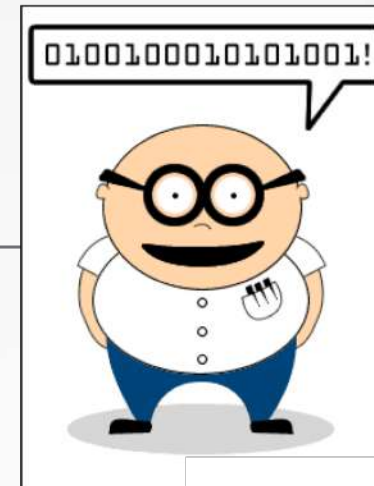
Bioinformatics analyses



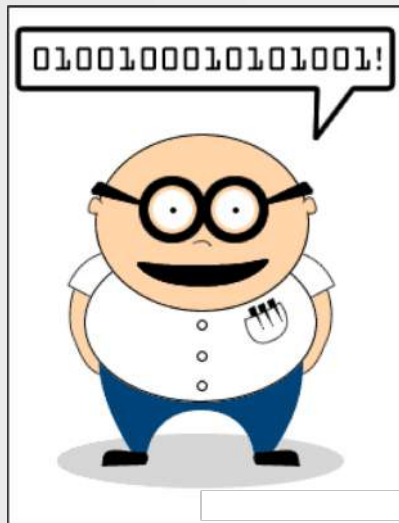
Informatics
data



nature



Bioinformatics analyses



Scripts, softwares

```
#!/usr/bin/perl

use strict;
use warnings;
use Getopt::Long;

## Date : 22 fev 2011
## Author : Stephanie Le Gras

## Objectives :

my $num_arg = scalar @ARGV;
my $programe = "ExtractID.pl";
my $input;
my $out;
my $id;

my $result = GetOptions(
    "id=s" => \$id,
    "out=s" => \$out,
    "input=s" => \$input,
);

my $usage = <<END;

Usage: $programe --id=FILENAME --out=FILENAME --input=FILENAME

END

die $usage unless ($result);
my @files = @ARGV;
die "Enter at least two files\n$usage" if ( $num_arg < 2 );
die $usage if ( $num_arg == 0 );

my %ids;
$out = ( defined $out ) ? $out : "results.txt";

## first, every lines of each files are put in the hash table ids. Variant ids are used as keys of the
## hash table and it contains a table.
```

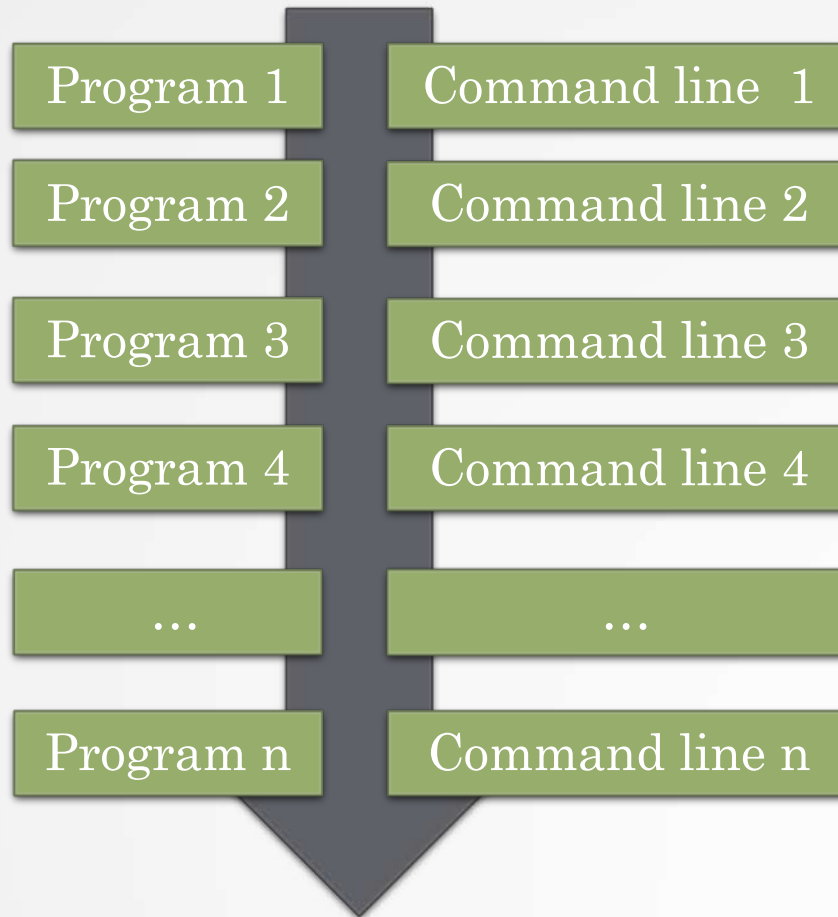
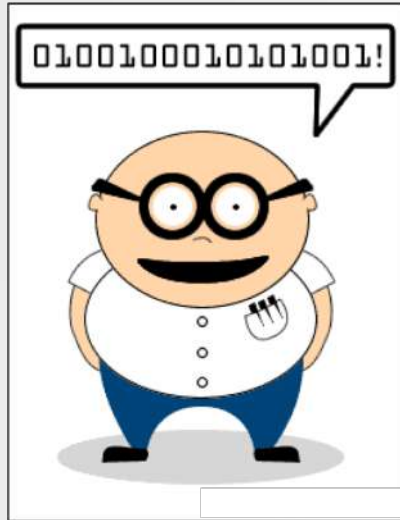
Command line

```
macs14 -t treatment.sort.bed -c control.bed -f BED -g mm --name=name1 --llocal=50000
--slocal=5000 > macs1.nohup 2>&1 &
```

Bioinformatics analyses



Bioinformatics analyses



**PIPELINE/
WORKFLOW**

Galaxy ?





Galaxy

PROJECT

Galaxy project

What is Galaxy ?

Galaxy is a **computing platform** that enables people to **run complex bioinformatics tools** on a **compute cluster** through a **simple web interface**.

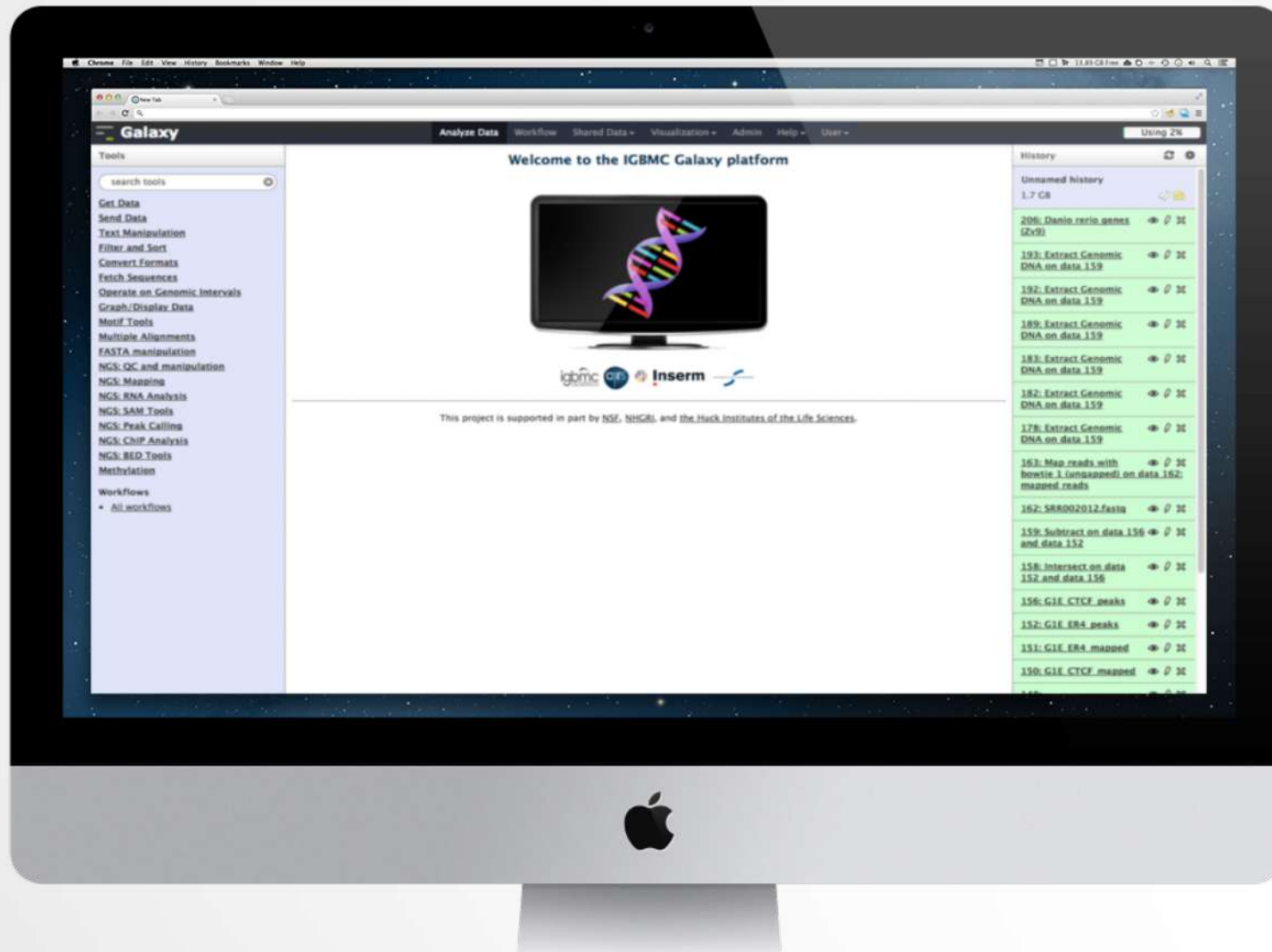


EMORY
UNIVERSITY

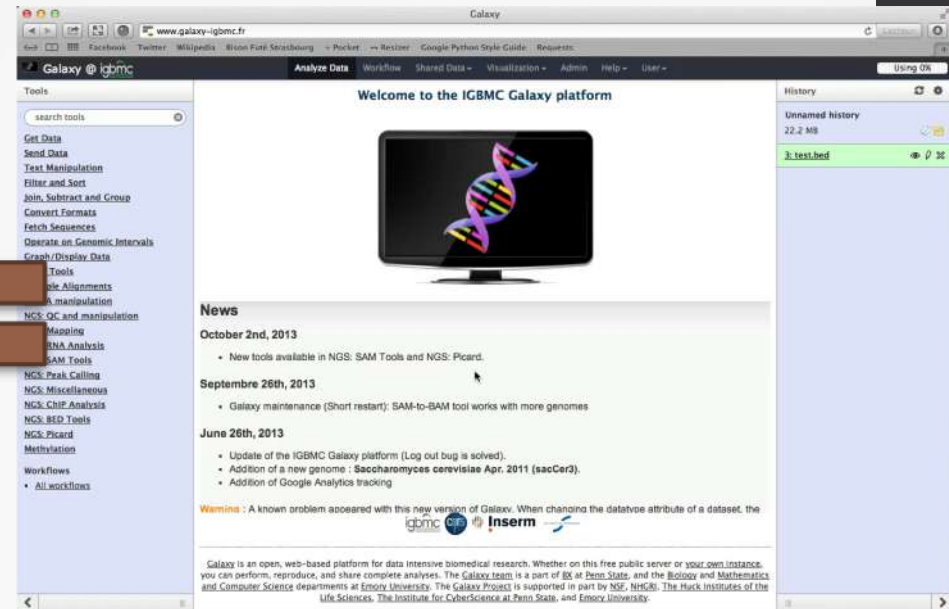


National
Human Genome
Research Institute

This is Galaxy



Running analyses with tools



Galaxy philosophy

- **Perform, and share** complete analysis
- **No programming skills** required
- **Open source** and **free** solution
- **Very large** and **active** community
- **Reproducibility/Usability/Transparency**

How to use Galaxy

Use Galaxy

- **Public servers**
- **Local servers**
- Clouds (Public, Commercial or Academic)
- Docker
- Virtual Machines

Galaxy public servers

- Galaxy Project's public server (<https://usegalaxy.org/>)
- There are several public remote Galaxy instances worldwide (156)
 - Genomics Servers
 - Domain Servers
 - Tool Publishing Servers

Public Galaxy Servers list :
<https://galaxyproject.org/use/>
Last Update on: 2019, November 14th

Galaxy public servers

- All analyses are run on remote computing infrastructures
- No need to have a Supercomputer to use Galaxy
- Web browser



Galaxy public servers

Your research institute



Connect to Galaxy web site
through a web browser
(<https://usegalaxy.org/>)

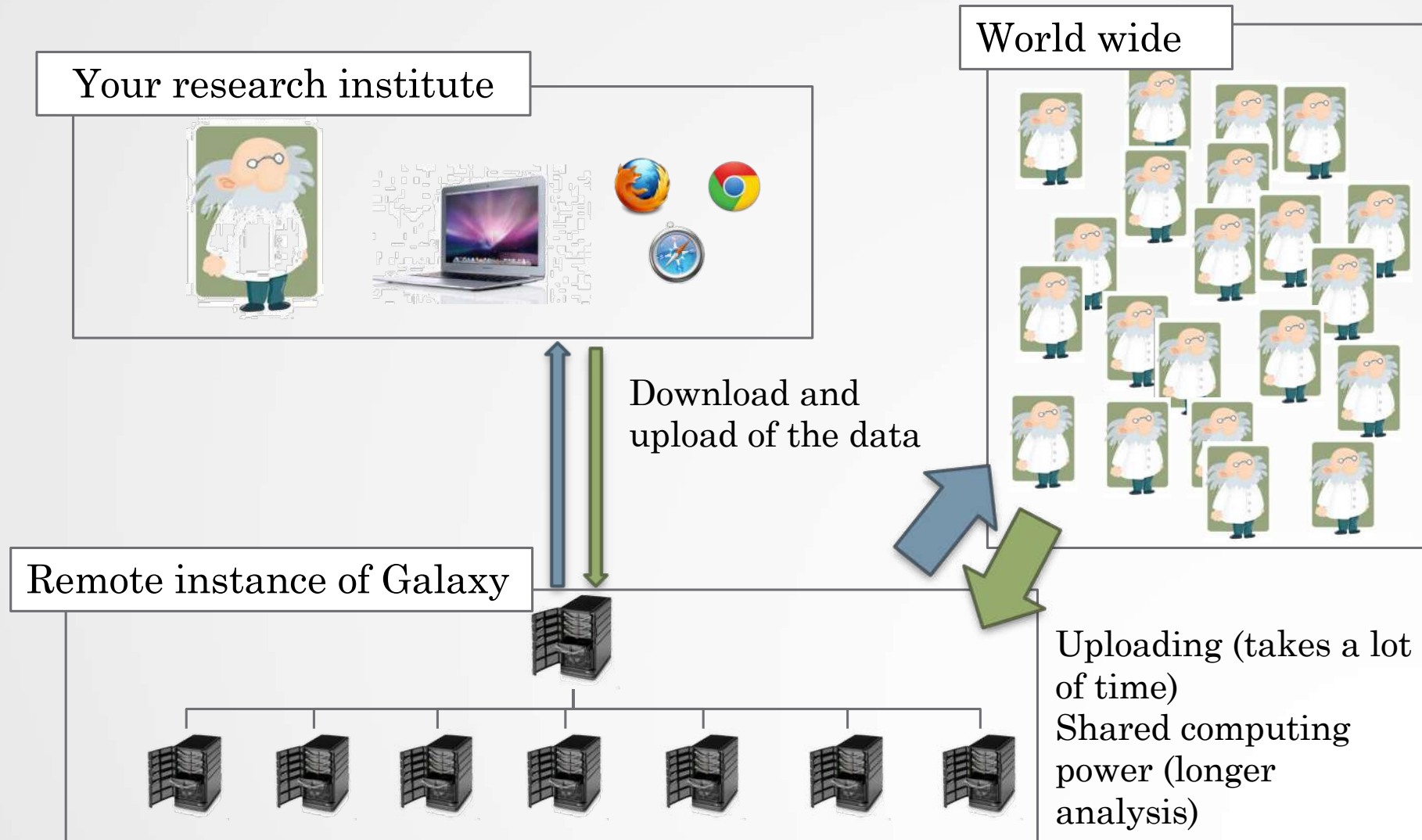
Download and
upload of the data

Remote instance of Galaxy



Run analyses

Galaxy public servers



Galaxy local server

- Run a local production Galaxy because you want to
 - install and use tools unavailable on public Galaxies
 - use sensitive data (e.g. clinical)
 - process large datasets that are too big for public Galaxies
 - plug-in new datasources
 - Develop Galaxy tools
 - Develop Galaxy itself



Description of the main features of Galaxy

Galaxy web interface

Top menu

The screenshot shows the Galaxy web interface with a dark blue top navigation bar. The main content area is divided into three vertical panels. The left panel is the 'Tools' panel, containing a search bar and a list of tool categories. The middle panel is the 'Data display and tools dialog window', showing a welcome message and a 'Galaxy Help' banner. The right panel is the 'History' panel, showing a search bar and a message that the history is empty. Three callout boxes with arrows point to these panels: 'Tool panel' on the left, 'Data display and tools dialog window' in the middle, and 'History panel' on the right.

Tools

search tools

Get Data

Collection Operations

GENERAL TEXT TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Datamash

GENOMIC FILE MANIPULATION

FASTA/FASTQ

FASTQ Quality Control

SAM/BAM

BED

VCF/BCF

Nanopore

Convert Formats

Lift-Over

COMMON GENOMICS TOOLS

O

F

G

Galaxy

Analyze Data Workflow Visualize Shared Data Help Login or Register Using 0%

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start here or consult our help resources. You can install your own Galaxy by following the tutorial and choose from thousands of tools from the Tool Shed.

Galaxy Help

Got Questions?
Get Answers.

help.galaxyproject.org

Tweets by @galaxyproject

Galaxy Project
@galaxyproject

Group tags for complex experimental designs, new GTN tutorial by @mariusvdbEEK
[training.galaxyproject.org/training-mater...](http://training.galaxyproject.org/training-material...)
#usegalaxy

History

search datasets

Unnamed history

(empty)

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

Tool panel

Data display and tools dialog window

History panel

Top menu

The image shows a screenshot of the Galaxy web interface. The top navigation bar is dark blue and contains the following items: **Analyze Data**, **Workflow**, **Visualize**, **Shared Data**, **Help**, **Login or Register**, and a grid icon. A callout box labeled "Run workflows" points to the "Analyze Data" menu item. A callout box labeled "Get Help" points to the "Help" menu item. On the left side, there is a "Tools" sidebar with a search bar and a list of tool categories: "Get Data", "Collection Operations", "GENERAL TEXT TOOLS", "Text Manipulation", "Filter and Sort", "Join, Subtract and Group", "Datamash", "GENOMIC FILE MANIPULATION", "FASTA/FASTQ", "FASTQ Quality Control", "SAM/BAM", "BED", and "VCF/BCF". A callout box labeled "Run analyses" points to the "Tools" sidebar. In the center, there is a large banner for "Try Galaxy on the Cloud" with the text "Now you can have a personal Galaxy within the infinite Universe". A callout box labeled "Access public data" points to the banner. On the right side, there is a "History" sidebar with a search bar and a list of "Unnamed" items. A callout box labeled "Log in/out, manage your account" points to the "Login or Register" menu item.

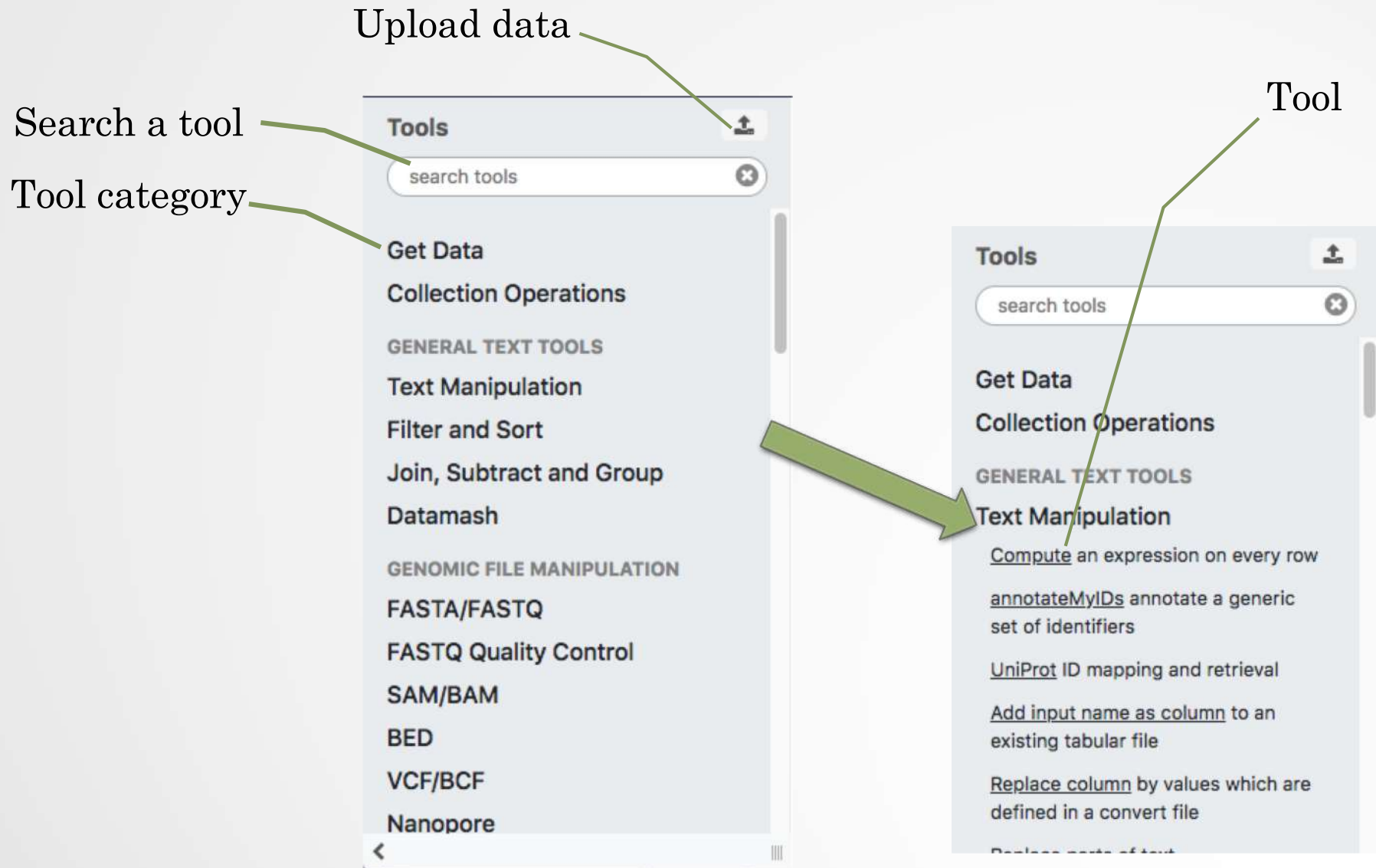
Exercise 1 : Log in

- Go to <http://use.galaxeast.fr>
- Log in with your training account (look to the post-it pasted into your booklet).

Tool Panel / Run analyses

The screenshot displays the Galaxy web interface. On the left is the **Tools** panel, which includes a search bar and a list of tool categories: **Get Data**, **Collection Operations**, **GENERAL TEXT TOOLS** (Text Manipulation, Filter and Sort, Join, Subtract and Group, Datamash), **GENOMIC FILE MANIPULATION** (FASTA/FASTQ, FASTQ Quality Control, SAM/BAM, BED, VCF/BCF, Nanopore, Convert Formats, Lift-Over), and **COMMON GENOMICS TOOLS**. A callout box labeled "Tool panel" with a downward arrow points to this section. The main content area features a header with navigation links (Analyze Data, Workflow, Visualize, Shared Data, Help, Login or Register) and a "Using 0%" indicator. Below the header is a introductory text about Galaxy and a "Galaxy Help" banner with the text "Got Questions? Get Answers." and the URL "help.galaxyproject.org". A "Tweets" section follows, showing a tweet from @galaxyproject about group tags for complex experimental designs. On the right is the **History** panel, which includes a search bar and a message stating "This history is empty. You can load your own data or get data from an external source."

Tool Panel / Run analyses



Tools dialog window

Tools

search tools

Get Data

Collection Operations

GENERAL TEXT TOOLS

Text Manipulation

[Compute an expression on every row](#)

[annotateMyIDs](#) annotate a generic set of identifiers

[UniProt ID mapping and retrieval](#)

[Add input name as column](#) to an existing tabular file

[Replace column](#) by values which are defined in a convert file

[Replace](#) parts of text

[Text transformation](#) with sed

[Unfold](#) columns from a table

[Unique lines](#) assuming sorted input file

[Replace Text](#) in entire line

[Replace Text](#) in a specific column

[Multi-Join](#) (combine multiple files)

[Select last](#) lines from a dataset (tail)

[Cut](#) columns from a table (cut)

[Create text file with recurring lines](#)

Analyze Data Workflow Visualize Shared Data Help Login or Register Using 0%

Compute an expression on every row (Galaxy Version 1.2.0) Versions Options

Add expression

c3-c2

as a new column to

No tabular dataset available. Dataset missing? See TIP below

Round result?

NO

Skip a header line

no

characters are already considered as comments and kept

Execute

TIP: If y

What it do

This tool co

- Colu
- c3-c

Example

If this is yo

chr1 1510

chr1 1510

History

search datasets

Unnamed history

(empty)

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

Data display and tools dialog window

- Set parameters
- Run tools
- Get help on tools
- Display content of dataset

History

The screenshot displays the Galaxy web interface. At the top, the navigation bar includes the Galaxy logo, 'Analyze Data', 'Workflow', 'Visualize', 'Shared Data', 'Help', 'Login or Register', and a grid icon. The main content area is divided into three sections:

- Tools Panel (Left):** A sidebar with a search bar and a list of tool categories: 'Get Data', 'Collection Operations', and 'GENERAL TEXT TOOLS'. Under 'Text Manipulation', several tools are listed, including 'Compute an expression on every row', 'annotateMyIDs', 'UniProt ID mapping and retrieval', 'Add input name as column', 'Replace column', 'Replace parts of text', 'Text transformation with sed', 'Unfold columns from a table', 'Unique lines', 'Replace Text in entire line', 'Replace Text in a specific column', 'Multi-Join', 'Select last lines from a dataset (tail)', 'Cut columns from a table (cut)', and 'Create text file with recurring lines'.
- Tool Configuration Panel (Center):** The 'Compute an expression on every row (Galaxy Version 1.2.0)' tool is selected. It features a search bar, 'Versions' and 'Options' buttons, and input fields for 'Add expression' (containing 'c3-c2'), 'as a new column to' (with a dropdown menu showing 'No tabular dataset available.'), 'Round result?' (set to 'NO'), and 'Skip a header line' (set to 'no'). A blue 'Execute' button is visible. A tip states: 'TIP: If your data is not TAB delimited, use Text Manipulation->Convert'. Below, the 'What it does' section explains that the tool computes an expression for every row and appends the result as a new column. An 'Example' section shows a snippet of genomic data: 'chr1 151077881 151077918 2 200 -'.
- History Panel (Right):** A panel titled 'History' with a search bar and a 'Using 0%' indicator. It shows 'Unnamed history (empty)'. A blue information box contains the text: 'This history is empty. You can load your own data or get data from an external source'.

History panel

Keep track of each job run

History

View all histories

History options

Refresh History

Search datasets

History name

History

search datasets

Unnamed history

(empty)

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

HISTORY LISTS

Saved Histories

Histories Shared with Me

CURRENT HISTORY

Create New

Copy History

Share or Publish

Show Structure

Extract Workflow

Delete

Delete Permanently

Make Data Private

DATASET ACTIONS

Copy Datasets

Dataset Security

Resume Paused Jobs

Collapse Expanded Datasets

Unhide Hidden Datasets

Delete Hidden Datasets

Purge Deleted Datasets

DOWNLOADS

Export Tool Citations

Export History to File

OTHER ACTIONS

Import from File

View all histories

The screenshot displays the Galaxy web interface with a dark blue header. The header includes the Galaxy logo, navigation links for 'Analyze Data', 'Workflow', 'Visualize', 'Shared Data', 'Help', and 'User', and a 'Using 0%' indicator. Below the header, there are search bars for 'search histories' and 'search all datasets', along with a 'Create new' button. The main content area is divided into several vertical panels, each representing a history. The first panel on the left is titled 'Current History' and is currently empty, with a message 'This history is empty'. The other panels are titled 'Unnamed history' and contain lists of datasets. Each dataset entry includes a number, a filename, and icons for viewing, editing, and deleting. The datasets listed include: '6: L1spa_ORF1.1.fasta', '5: TALs.fasta', '4: Count on data 3', '3: Galaxy14-(Intersect on data 13 and data 1).bed', '2: shuffleseq on data 1', '1: lkpeaks_notRBPJ_+-150_random80_adjacentSeq.fasta', '5: Correspondance_JASPAR_CORE.txt', '4: fimo.txt', and '2: Brn2_Day2_rtt_rep2.sort.bed'. The interface also shows a 'Switch to' dropdown for each history panel.

Exercise 2 : History

- Create a new history
- Change the name of the history to “RNA-seq data analysis”

Import data into Galaxy

- Your own data (from your computer)
- Shared data
- Data from external sources

Import your own data to Galaxy



Display the drag and drop utility used to upload local files

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
Mouse_ChIP-Seq_example_Control_Data_chr19_mm9.fastq	84.1 MB	Auto-det...	unspecified (?)		

Type (set all): Auto-detect

Genome (set all): unspecified (?)

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

Name of the dataset

Size of the dataset

File format

Genome

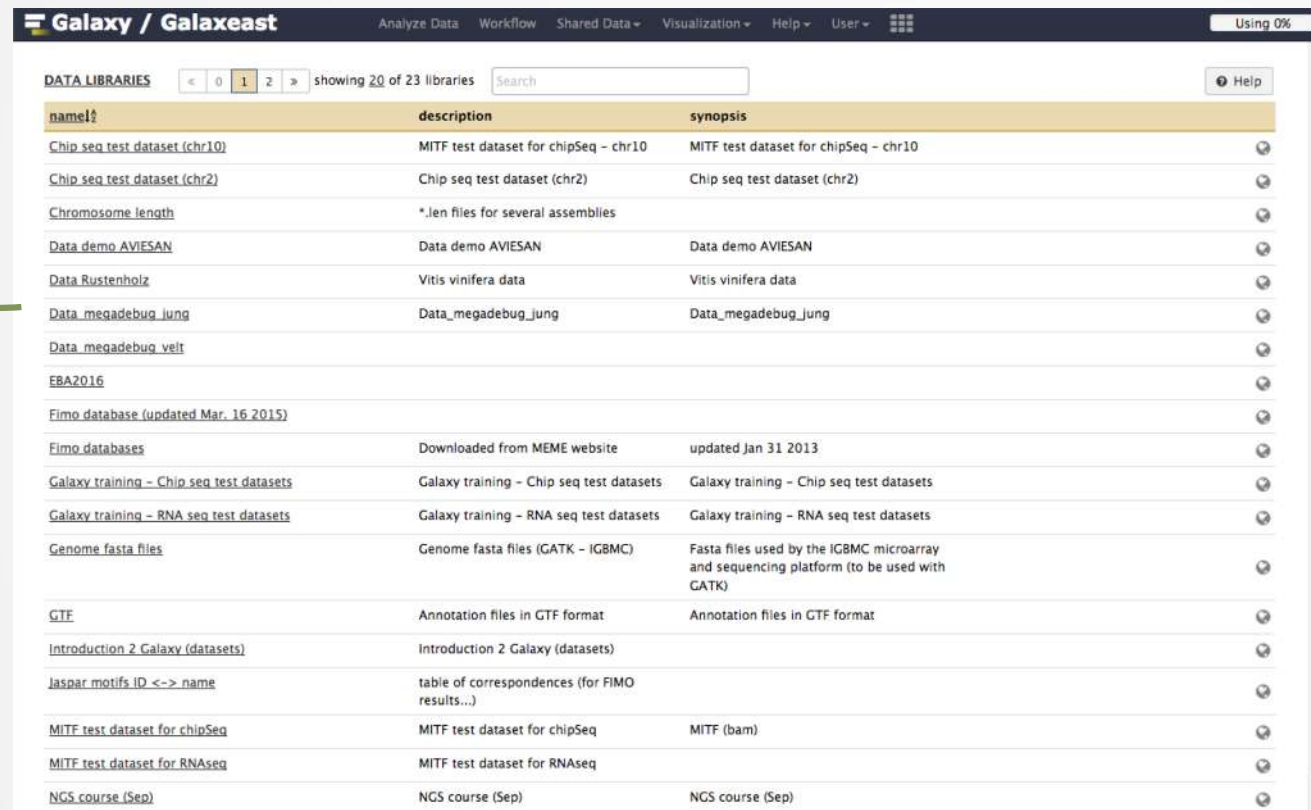
Import shared data (data libraries)

Shared Data ▾ Visualiza

(Top menu)

- Data Libraries
- Histories
- Workflows
- Visualizations
- Pages

Data Libraries



Galaxy / Galaxeast Analyze Data Workflow Shared Data ▾ Visualization ▾ Help ▾ User ▾ Using 0%

DATA LIBRARIES < 0 1 2 > showing 20 of 23 libraries Search Help

name	description	synopsis
Chip seq test dataset (chr10)	MITF test dataset for chipSeq - chr10	MITF test dataset for chipSeq - chr10
Chip seq test dataset (chr2)	Chip seq test dataset (chr2)	Chip seq test dataset (chr2)
Chromosome length	*.len files for several assemblies	
Data demo AVIESAN	Data demo AVIESAN	Data demo AVIESAN
Data Rustenholz	Vitis vinifera data	Vitis vinifera data
Data_megadebug_lung	Data_megadebug_lung	Data_megadebug_lung
Data_megadebug_veit		
EBA2016		
Fimo database (updated Mar. 16 2015)		
Fimo databases	Downloaded from MEME website	updated Jan 31 2013
Galaxy training - Chip seq test datasets	Galaxy training - Chip seq test datasets	Galaxy training - Chip seq test datasets
Galaxy training - RNA seq test datasets	Galaxy training - RNA seq test datasets	Galaxy training - RNA seq test datasets
Genome fasta files	Genome fasta files (GATK - IGBMC)	Fasta files used by the IGBMC microarray and sequencing platform (to be used with GATK)
GTF	Annotation files in GTF format	Annotation files in GTF format
Introduction 2 Galaxy (datasets)	Introduction 2 Galaxy (datasets)	
Jaspar motifs ID <-> name	table of correspondences (for FIMO results...)	
MITF test dataset for chipSeq	MITF test dataset for chipSeq	MITF (bam)
MITF test dataset for RNAseq	MITF test dataset for RNAseq	
NGS course (Sep)	NGS course (Sep)	NGS course (Sep)

Import shared data (data libraries)

2. Import selected dataset to history

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

DATA LIBRARIES << 0 1 2 >> showing 6 of 6 items include deleted

[Libraries](#) / Chromosome length

<input type="checkbox"/> name ↑	description	data type	size	time updated (UTC)	
<input type="checkbox"/> ..					
<input type="checkbox"/> ce10.len		len	98 bytes	2015-01-08 01:25	<input type="checkbox"/>
<input type="checkbox"/> dm3.len		len	227 bytes	2015-01-08 01:25	<input type="checkbox"/>
<input type="checkbox"/> hg19.len		len	376 bytes	2015-01-08 01:25	<input type="checkbox"/>
<input type="checkbox"/> mm10.len		len	1.4 KB	2015-01-08 01:25	<input type="checkbox"/>
<input type="checkbox"/> mm9.len		len	330 bytes	2015-01-08 01:25	<input type="checkbox"/>
<input type="checkbox"/> tair10.len		len	75 bytes	2015-01-08 01:25	<input type="checkbox"/>

<< 0 1 2 >> showing 6 of 6 items

1. Select dataset

Import public data

Tools

search tools

Get Data

- Upload File from your computer
- UCSC Main table browser
- UCSC Archaea table browser
- EBI SRA ENA SRA
- BioMart Ensembl server
- GrameneMart Central server
- Flymine server
- modENCODE fly server
- modENCODE modMine server
- MouseMine server
- Ratmine server
- YeastMine server
- modENCODE worm server
- WormBase server
- ZebrafishMine server
- EuPathDB server
- GenomeSpace import from file browser

Browse and import external data from public databases

Genomes Genome Browser Tools Mirrors Downloads My Data Help About Us

Table Browser

Use this program to retrieve the data associated with a track in text format, to calculate intersections between tracks, and to retrieve DNA sequence covered by a track. For help in using this application see [Using the Table Browser](#) for a description of the controls in this form, the [User's Guide](#) for general information and sample queries, and the [OpenHelix Table Browser tutorial](#) for a narrated presentation of the software features and usage. For more complex queries, you may want to use [Galaxy](#) or our [public MySQL server](#). To examine the biological function of your set through annotation enrichments, send the data to [GREAT](#). Send data to [GenomeSpace](#) for use with diverse computational tools. Refer to the [Credits](#) page for the list of contributors and usage restrictions associated with these data. All tables can be downloaded in their entirety from the [Sequence and Annotation Downloads](#) page.

clade: Mammal genome: Mouse assembly: Dec. 2011 (GRCm38/mm10)

group: Genes and Gene Predictions track: UCSC Genes

table: knownGene

region: genome position chr1:121427557-121432936

identifiers (names/accessions):

filter: create

intersection: create

correlation: create

output format: BED - browser extensible data Send output to Galaxy GREAT GenomeSpace

output file: (leave blank to keep output in browser)

file type returned: plain text gzip compressed

get output summary/statistics

To reset all user cart settings (including custom tracks), [click here](#).

Using the Table Browser

This section provides brief line-by-line descriptions of the Table Browser controls. For more information on using this program, see the [Table Browser User's Guide](#).

- clade: Specifies which clade the organism is in.
- genome: Specifies which organism data to use.

Exercise 3 : Import data to Galaxy

- 1. Import to Galaxy the file siLuc3_S12040.fastq from the “Shared data > Data Libraries > NGS data analysis training > RNAseq > rawdata” to the history called “RNA-seq data analysis”
- 2. You should be in the history “RNA-seq data analysis” (Switch to it if needed)
- 3. Import to Galaxy the file sample.bed.gz located in the directory galaxy.
 - The **Genome** is : Mouse (mm9)
 - The format (**Type**) is : bed

Datasets/Jobs in the History

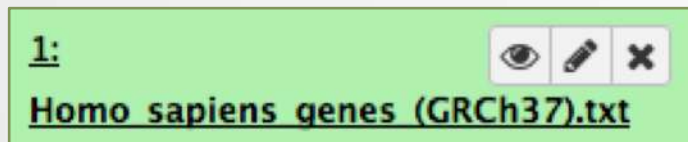
Grey: the job is waiting to run



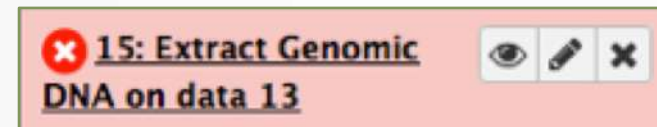
Yellow: the job is running



Green: the job is successfully done



Red: the job encountered a problem

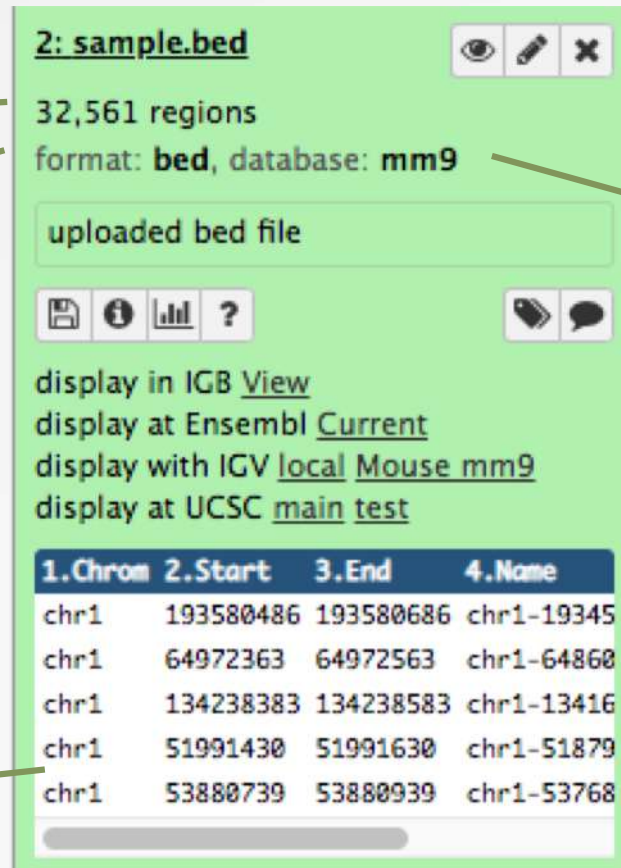


Datasets/Jobs in the History

Number of lines in the file or size of the file

Format

Genome



2: sample.bed

32,561 regions
format: **bed**, database: **mm9**

uploaded bed file

display in IGB [View](#)
display at Ensembl [Current](#)
display with IGV [local](#) [Mouse](#) [mm9](#)
display at UCSC [main](#) [test](#)




1. Chrom	2. Start	3. End	4. Name
chr1	193580486	193580686	chr1-19345
chr1	64972363	64972563	chr1-64860
chr1	134238383	134238583	chr1-13416
chr1	51991430	51991630	chr1-51879
chr1	53880739	53880939	chr1-53768

If the dataset is a text file, the first lines of the file are displayed

Datasets/Jobs in the History







View dataset (if possible) in the middle panel of Galaxy

Download dataset

2: sample.bed   

32,561 regions
format: **bed**, database: **mm9**

uploaded bed file

display in IGB [View](#)
display at Ensembl [Current](#)
display with IGV [local](#) [Mouse](#) [mm9](#)
display at UCSC [main](#) [test](#)

1.Chrom	2.Start	3.End	4.Name
chr1	193580486	193580686	chr1-19345
chr1	64972363	64972563	chr1-64860
chr1	134238383	134238583	chr1-13416
chr1	51991430	51991630	chr1-51879
chr1	53880739	53880939	chr1-53768

Delete dataset

Edit attributes of the dataset (change name, format, genome, permission)


Size of histories and quota

Size of history

The screenshot shows a software interface with a dark blue header bar at the top containing a green progress indicator and the text "Using 20%". Below this is a light blue "History" panel with a search bar labeled "search datasets" and a refresh icon. The panel lists two datasets under the heading "RNA-seq data analysis" with "2 shown". The first dataset is "7.23 GB" with a checkmark, a folder icon, and a speech bubble icon. The second dataset is "2: sample.bed" with an eye icon, a pencil icon, and an 'x' icon. The third dataset is "1: siLuc3 S12040.fastq" with an eye icon, a pencil icon, and an 'x' icon.

Quota

Exercise 4 : remove dataset

- 1. Remove the dataset sample.bed from your history by clicking on the button 
- 2.
 - A. Click on “deleted” in the top of the history panel (below the history name). Remove definitely the file from the disk by clicking on ” Permanently remove it from disk”.
 - B. Click on “hide deleted”

Exercise 5 : Running a tool

- 1. Use the tool “FastQC Read Quality reports” to compute quality analysis on the dataset “siLuc3_S12040.fastq”
 - Use default parameters.

Workflows

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The 'Workflow' tab is selected. On the left, a 'Tools' sidebar lists various categories like NGS: SAMtools, NGS: BamTools, NGS: Picard, etc., with a 'Workflows' section at the bottom containing 'All workflows'. The main content area displays a welcome message, a 'Public Galaxy Servers and still counting' graphic, and a tweet from the Galaxy Project. On the right, a 'History' panel shows an empty history with a message: 'This history is empty. You can load your own data or get data from an external source'. A green arrow points from the 'Workflow' tab to the text 'Create, run, edit (...) workflows'. Another green arrow points from the 'All workflows' link in the sidebar to the text 'Run workflows'.

Create, run,
edit (...) workflows

Run workflows

Workflows

Create new workflow

Your workflows

Name	# of Steps
imported: CloudMap Variant Discovery Mapping (and Variant Calling) workflow _2-7-2014	33
imported: CloudMap Variant Discovery Mapping (and Variant Calling) workflow _2-7-2014	33
imported: metagenomic analysis	16
imported: imported: MACS (mm8)	10
imported: metagenomic analysis	16
imported: CHIP-Seq analysis on BAM files	11
imported: Sort BAM for Peak Calling MACS tool	5
Bergen workshop	0

Workflows shared with you by others
No workflows have been shared with you.

Other options
Configure your workflow menu

Edit, run, share
(...) workflows

Workflows

Save, run (...) workflows

The screenshot displays a workflow editor interface with a central canvas and a left-hand tool palette. The workflow consists of the following steps:

- Input dataset**: Provides an `output` output.
- FastQC: Read QC**: Takes `Short read data from your current history` and a `Contaminant list` as input, producing an `html_file (html)` output.
- Convert from BAM to BED**: Takes a `BAM file` as input and produces an `output (bed)` output.
- Map reads with bowtie 1 (ungapped)**: Takes a `FASTQ file` as input and produces multiple outputs: `output (sam)`, `output_suppressed_reads_l (fastq)`, `output_suppressed_reads_r (fastq)`, `output_unmapped_reads_l (fastq)`, and `output_unmapped_reads_r (fastq)`.
- SAM-to-BAM**: Takes a `SAM File to Convert` as input and produces an `output1 (bam)` output.

The **Details** panel for the 'Map reads with bowtie 1 (ungapped)' tool includes the following configuration options:

- Tool:** Map reads with bowtie 1 (ungapped)
- Version:** 0.12.8
- Reference genome:** Will you select a reference genome from your history or use a built-in index? (Use a built-in index)
- Library type:** Is this library mate-paired? (Single-end)
- FASTQ file:** Data input 'sInput1' (fastqsanger or fastqillumina or fastqsolexa)
- Bowtie settings to use:** Full parameter list
- Skip the first n reads (-s):** 0
- Only align the first n reads (-u):** -1
- Trim n bases from high-quality (left) end of each read before alignment (-5):** 0
- Trim n bases from low-quality (right) end of each read before alignment (-3):** 0

Workflows

Set input file(s)

The screenshot displays the Galaxy web interface for configuring a workflow named "chip workflow". The interface is divided into several sections:

- Tools Panel (Left):** A sidebar containing various tool categories such as "Get Data", "Text Manipulation", "Statistics", and "NGS TOOLBOX BETA".
- Workflow Configuration (Center):** A main area titled "Running workflow 'chip workflow'" with "Expand All" and "Collapse" buttons. It shows four steps:
 - Step 1: Input dataset:** Includes an "Input Dataset" dropdown menu with the selected value "4: chr10_ctr2_1.fastq.gz" and a "type to filter" input field.
 - Step 2: Map with Bowtie for Illumina (version 1.1.3)**
 - Step 3: MACS (version 1.4.2)**
 - Step 4: homer_annotatePeaks (version 0.0.5):** Includes a "Homer peaks OR BED format" section with the text "Output dataset 'output_bed_file' from step 3", a "Genome version" dropdown set to "tair10", and an "Extra options" field with a checkmark icon.
- Action:** A section with the text "Hide output 'out_log'".
- Buttons:** A "Run workflow" button is located at the bottom of the configuration area.
- History Panel (Right):** A sidebar titled "History" with a search bar and a list of datasets. The current dataset is "4: chr10_ctr2_1.fastq" in format "fastqsanger" from database "hg19".

Green arrows point from the text labels to specific elements in the interface:

- "Set input file(s)" points to the "Input Dataset" dropdown menu.
- "Set parameters" points to the "Genome version" dropdown menu.
- "Run workflow" points to the "Run workflow" button.

Privacy

- By default datasets, workflows, histories are private to the user that generated/uploaded them.
- They can be shared across Galaxy users (of the same Galaxy instance) or via links