

Introduction to Galaxy

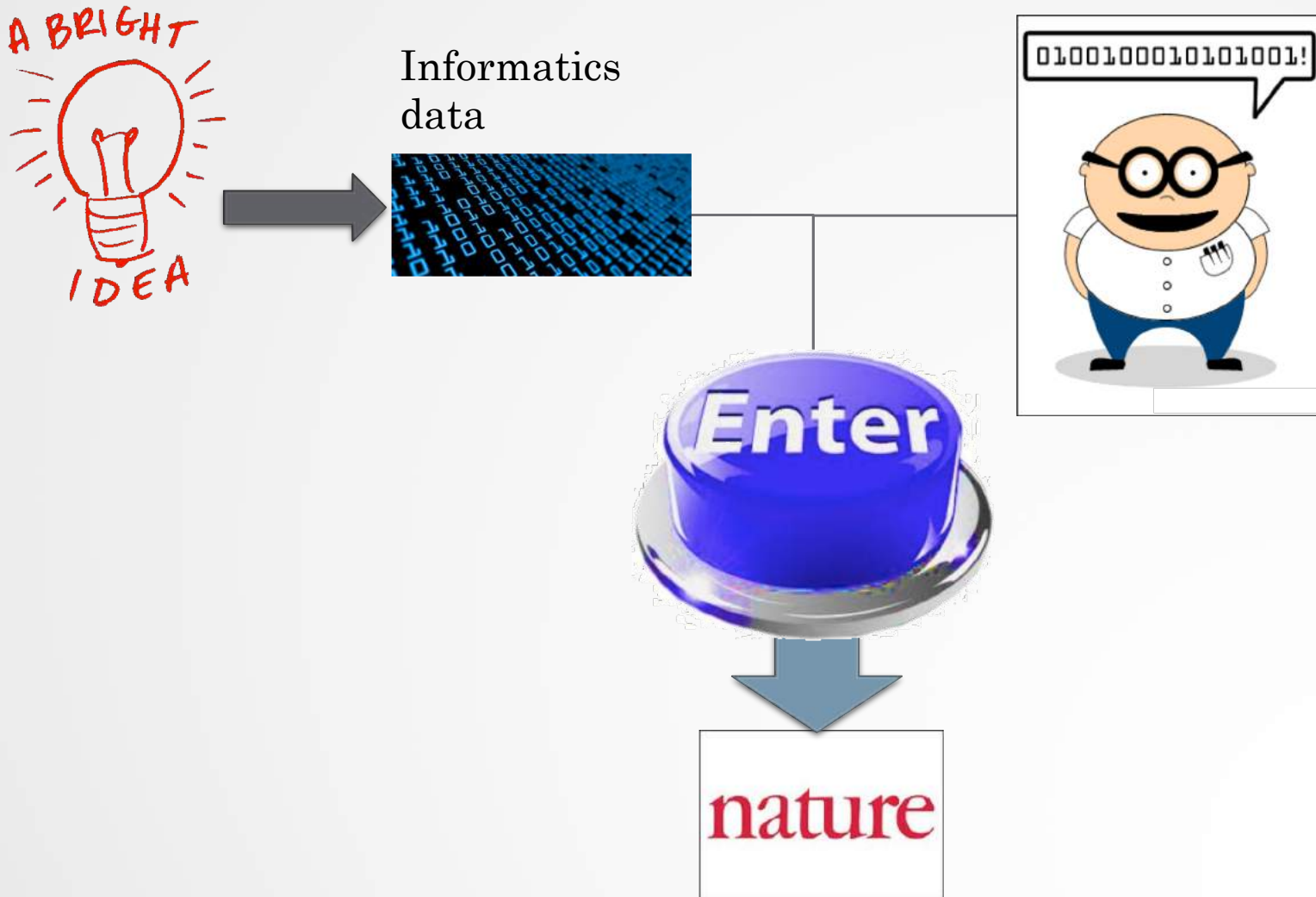
Stéphanie Le Gras
(slegras@igbmc.fr)

Guidelines

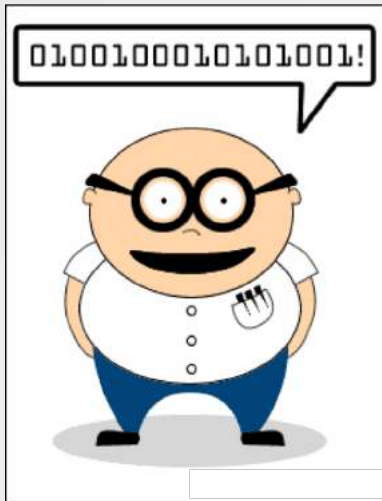
- Analyzing biological data with informatics tools
- Presentation of the Galaxy project
- Description of the main features of the Galaxy platform

Analyzing biological data with informatics tools

Bioinformatics analyses



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 );

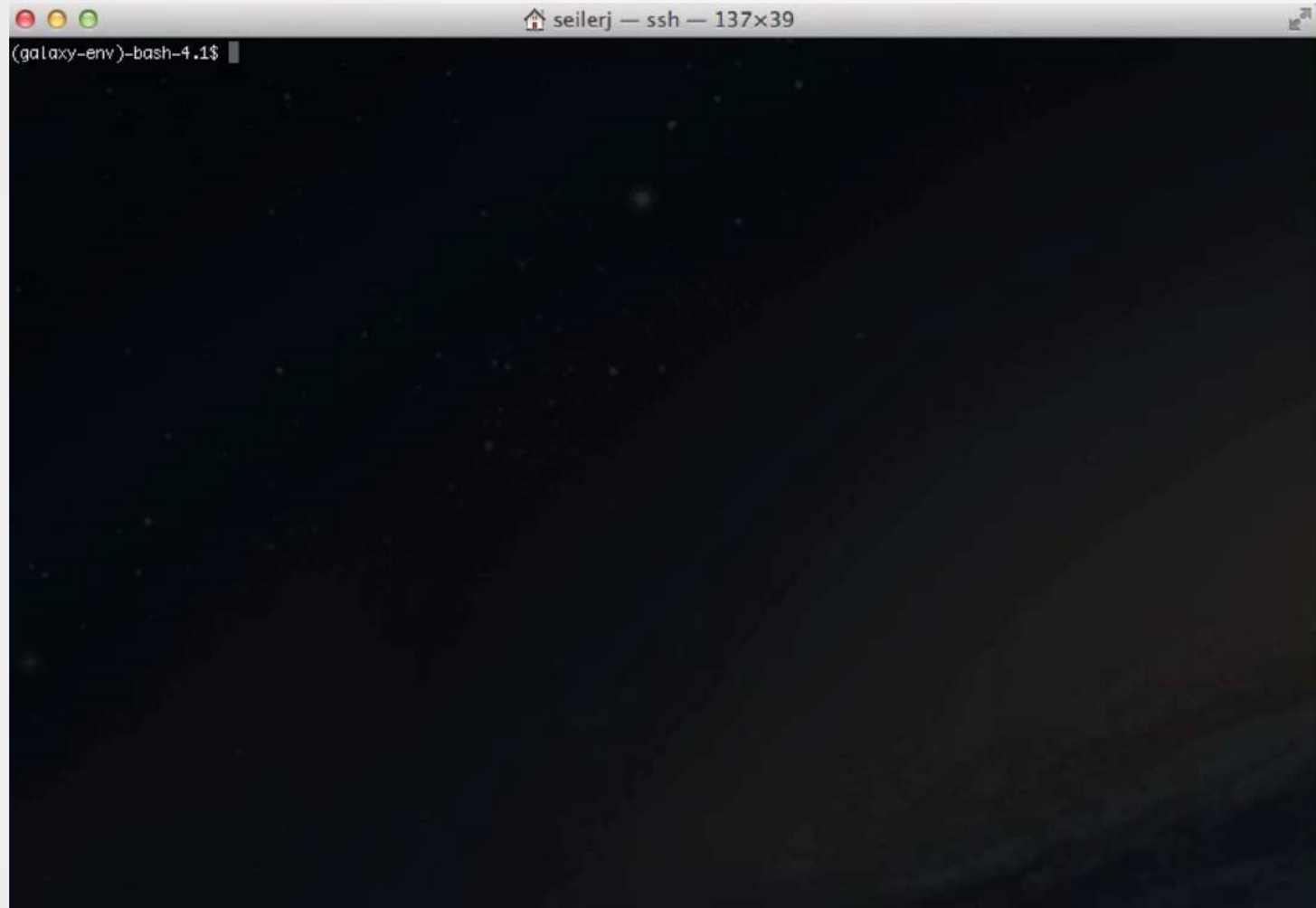
my %ids;
$out = ( defined $out ) ? $out : "results.txt";

## 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.
```

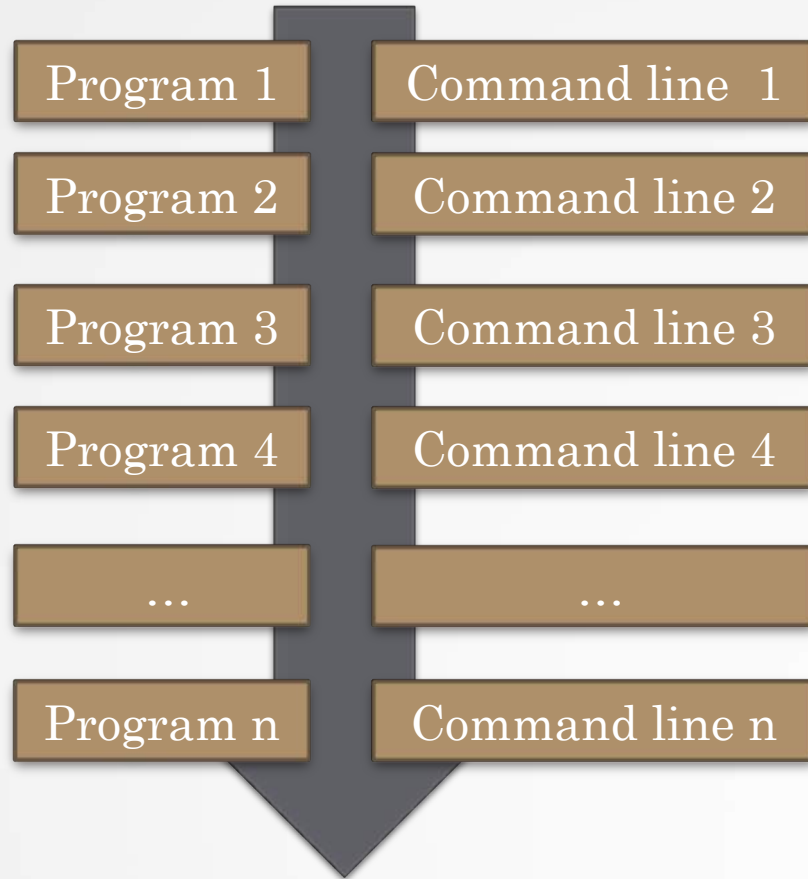
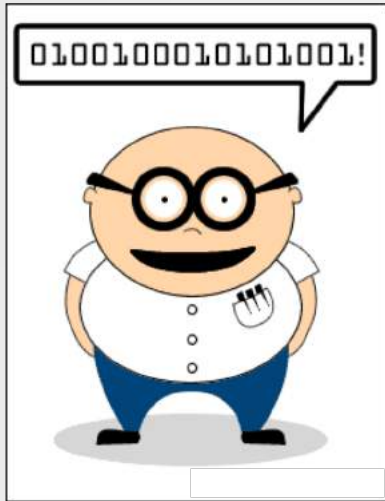
Command line

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

Bioinformatics analyses



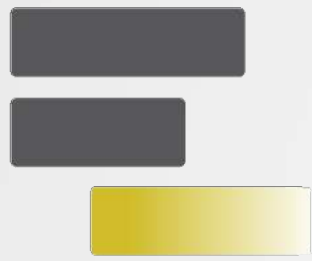
Bioinformatics analyses



**PIPELINE/
WORKFLOW**

Galaxy ?





Galaxy

PROJECT

Galaxy project

What is Galaxy ?

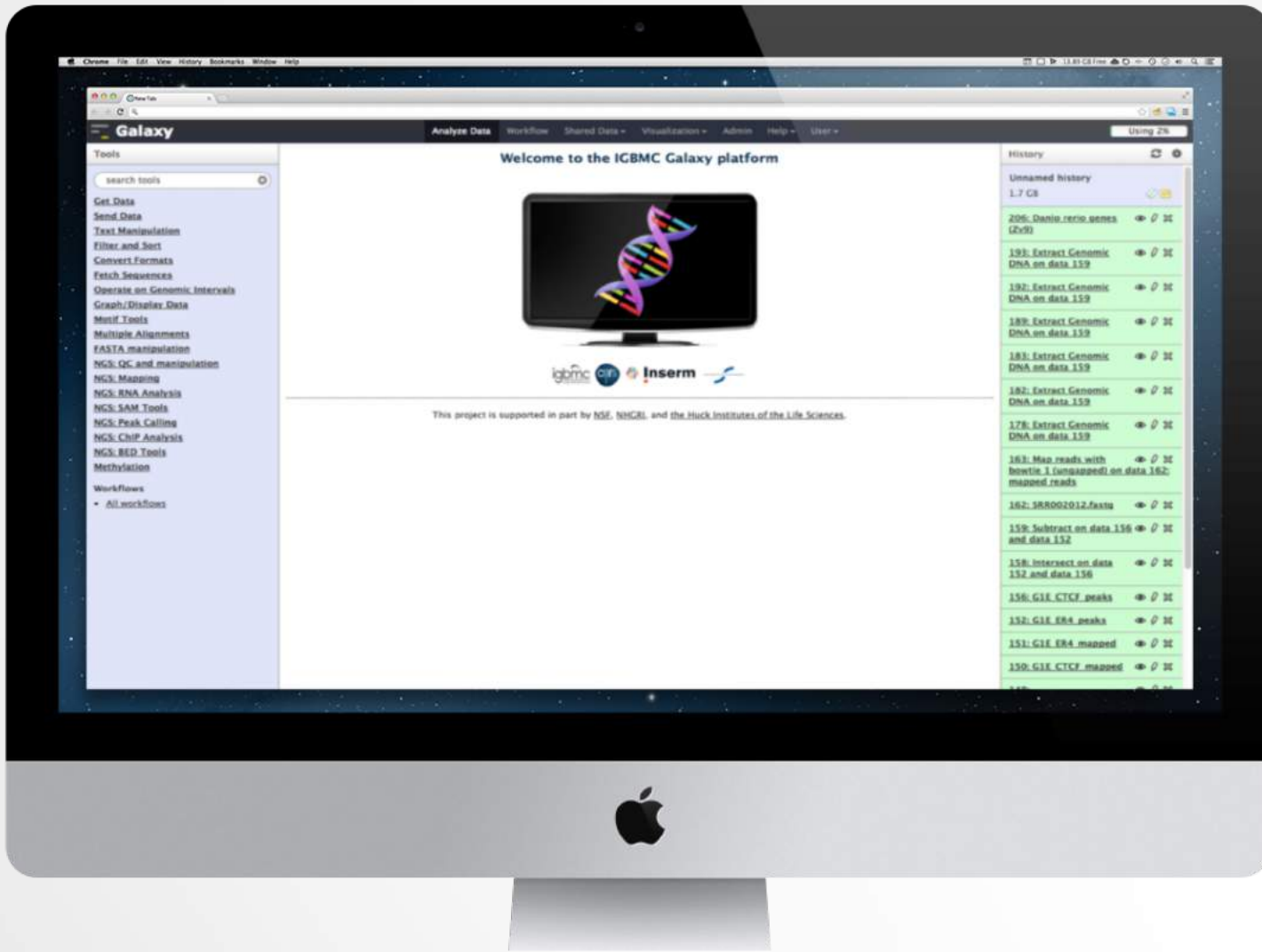
Galaxy is a **computing platform** that enables people to **run complex bioinformatics tools** on a **compute cluster** through a **simple web interface**.



EMORY
UNIVERSITY



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 (116)
 - Genomics Servers
 - Domain Servers
 - Tool Publishing Servers

Public Galaxy Servers list :
<https://galaxyproject.org/use/>
Last Update on: 2019, March 7th

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/>)

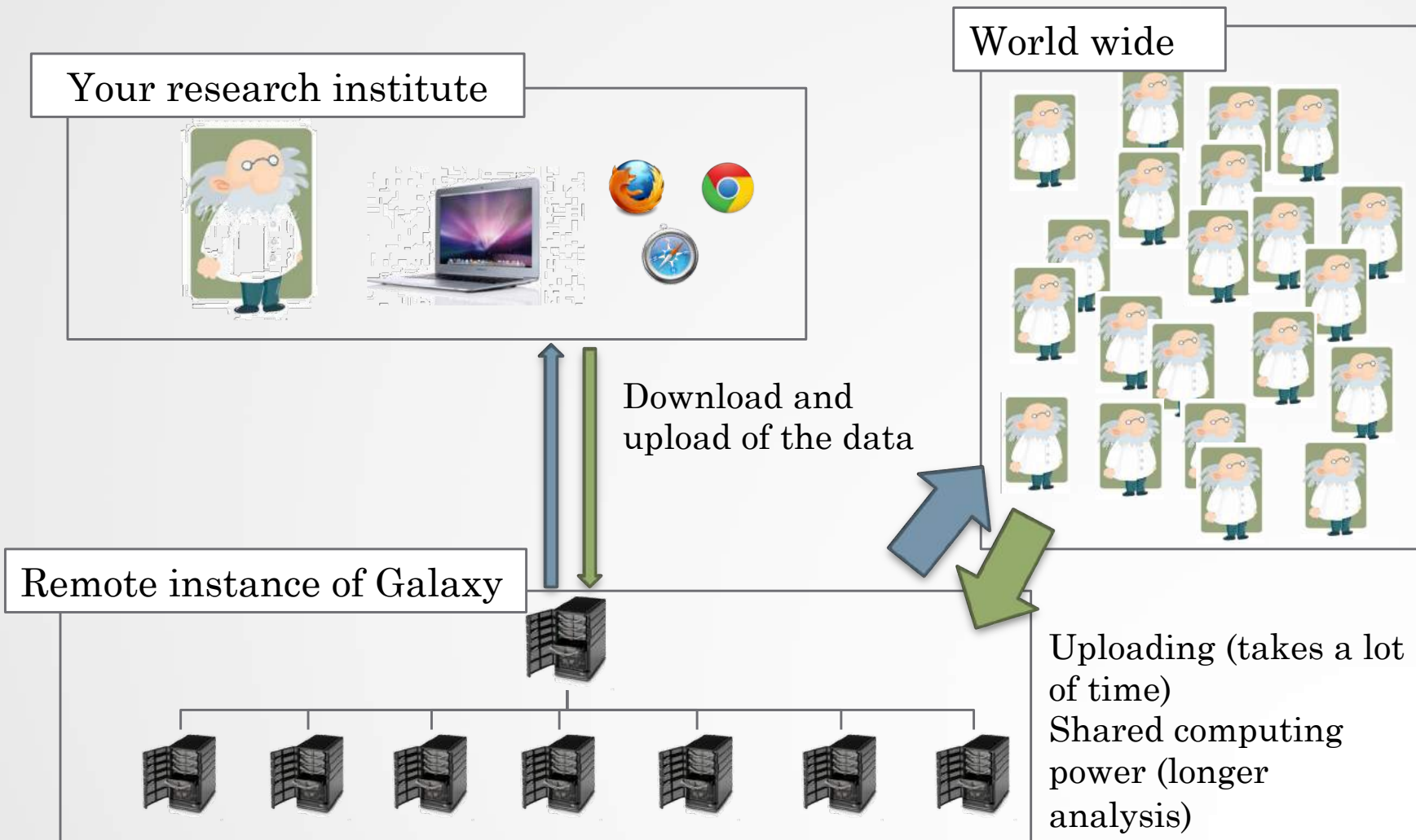
Download and upload of the data

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 top navigation bar containing the 'Galaxy' logo and menu items: 'Analyze Data', 'Workflow', 'Visualize', 'Shared Data', 'Help', 'Login or Register', and a grid icon. A 'Using 0%' indicator is on the right. The interface is divided into three main vertical sections:

- Left Panel (Tool panel):** A sidebar with a search bar and a list of tool categories including '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'.
- Center Panel (Data display and tools dialog window):** Contains a main text area with an introduction to Galaxy, a 'Galaxy Help' dialog box with the text 'Got Questions? Get Answers.' and the URL 'help.galaxyproject.org', and a 'Tweets' section featuring a tweet from '@galaxyproject' about group tags and a GTN tutorial.
- Right Panel (History panel):** Features a search bar for datasets, a title 'Unnamed history', and a message: 'This history is empty. You can load your own data or get data from an external source'.

Three callout boxes with arrows point to these sections: 'Tool panel' on the left, 'Data display and tools dialog window' in the center, and 'History panel' on the right.

Top menu

Run workflows

Visualize your data

Get Help

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** (with a dropdown arrow), **Shared Data** (with a dropdown arrow), **Help** (with a dropdown arrow), **Login or Register** (with a dropdown arrow), and a grid icon. Below the navigation bar, the main content area features a large banner with the text "Try Galaxy on the Cloud" and "Now you can have a personal Galaxy within the infinite Universe". To the left, there is a sidebar with a search bar and a list of tool categories: **Tools**, **Get Data**, **Collection Operations**, **GENERAL TEXT TOOLS** (including Text Manipulation, Filter and Sort, Join, Subtract and Group, and Datamash), **GENOMIC FILE MANIPULATION** (including FASTA/FASTQ, FASTQ Quality Control, SAM/BAM, BED, and VCF/BCF), and **Histor** (with a search bar and "Unnan (empty)").

Callouts with arrows point to the following items in the interface:

- Run workflows**: Points to the **Workflow** menu item.
- Visualize your data**: Points to the **Visualize** menu item.
- Get Help**: Points to the **Help** menu item.
- Run analyses**: Points to the **Tools** sidebar.
- Access public data**: Points to the **Shared Data** menu item.
- 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 is highlighted with a green border and a callout box labeled "Tool panel" with a downward arrow. The panel lists various tool categories such as "Get Data", "Collection Operations", "GENERAL TEXT TOOLS", "GENOMIC FILE MANIPULATION", and "COMMON GENOMICS TOOLS". The main content area features a header with navigation links like "Analyze Data", "Workflow", "Visualize", "Shared Data", "Help", "Login or Register", and a "Using 0%" indicator. Below the header, there is a introductory text about Galaxy, a "Galaxy Help" banner with the text "Got Questions? Get Answers." and the URL "help.galaxyproject.org", and a "Tweets" section by @galaxyproject. On the right side, there is a "History" panel with 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

Upload data

Search a tool

Tool category

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

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

Tool

Tools dialog window

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

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](#)

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 you have a dataset that is not tabular, you can use the [Convert](#) tool to convert it to a tabular dataset.

What it does

This tool computes an expression on every row of a tabular dataset. The expression is defined in the **Add expression** field. The result is displayed as a new column in the dataset.

- Column name: c3-c2
- Expression: c3-c2

Example

If this is your input dataset:

```
chr1 1510 1510 1510 1510
```

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 'Galaxy', 'Analyze Data', 'Workflow', 'Visualize', 'Shared Data', 'Help', and 'Login or Register'. The main content area is divided into a left sidebar, a central tool configuration panel, and a right history panel.

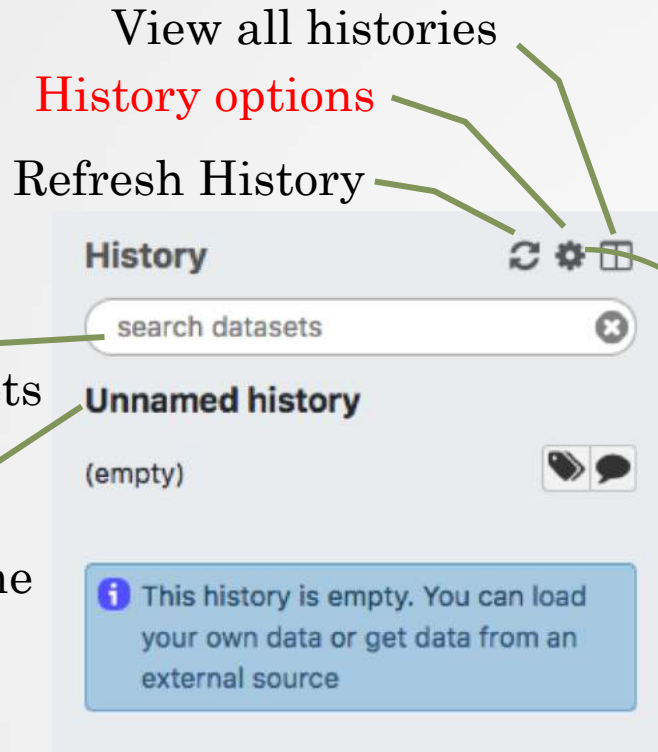
Left Sidebar: Contains 'Tools' with a search bar, 'Get Data', 'Collection Operations', and 'GENERAL TEXT TOOLS'. Under 'Text Manipulation', several tools are listed, including 'Compute an expression on every row'.

Central Tool Configuration Panel: Titled 'Compute an expression on every row (Galaxy Version 1.2.0)'. It features a 'Versions' button and an 'Options' dropdown. The 'Add expression' field contains 'c3-c2'. Below it, the 'as a new column to' dropdown is set to 'No tabular dataset available.' with a note 'Dataset missing? See TIP below'. The 'Round result?' dropdown is set to 'NO', and the 'Skip a header line' dropdown is set to 'no'. A blue 'Execute' button is visible. A tip states: 'TIP: If your data is not TAB delimited, use Text Manipulation->Convert'. 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 input data: 'chr1 151077881 151077918 2 200 -'.

Right History Panel: Titled 'History' with a 'Using 0%' indicator. It has a search bar for datasets. The 'Unnamed history' section is currently empty. A blue information box states: '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



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 menus 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', and 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 in the panels are:

- Panel 2: 6: L1spa_ORF1.1.fastq, 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_ran_dom80_adjacentSeq.fasta
- Panel 3: 5: Correspondance_JASPAR_CORE.txt, 4: fimo.txt
- Panel 4: 2: Brn2_Day2_rtt2_rep2.sort.bed

Exercise 2 : History

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

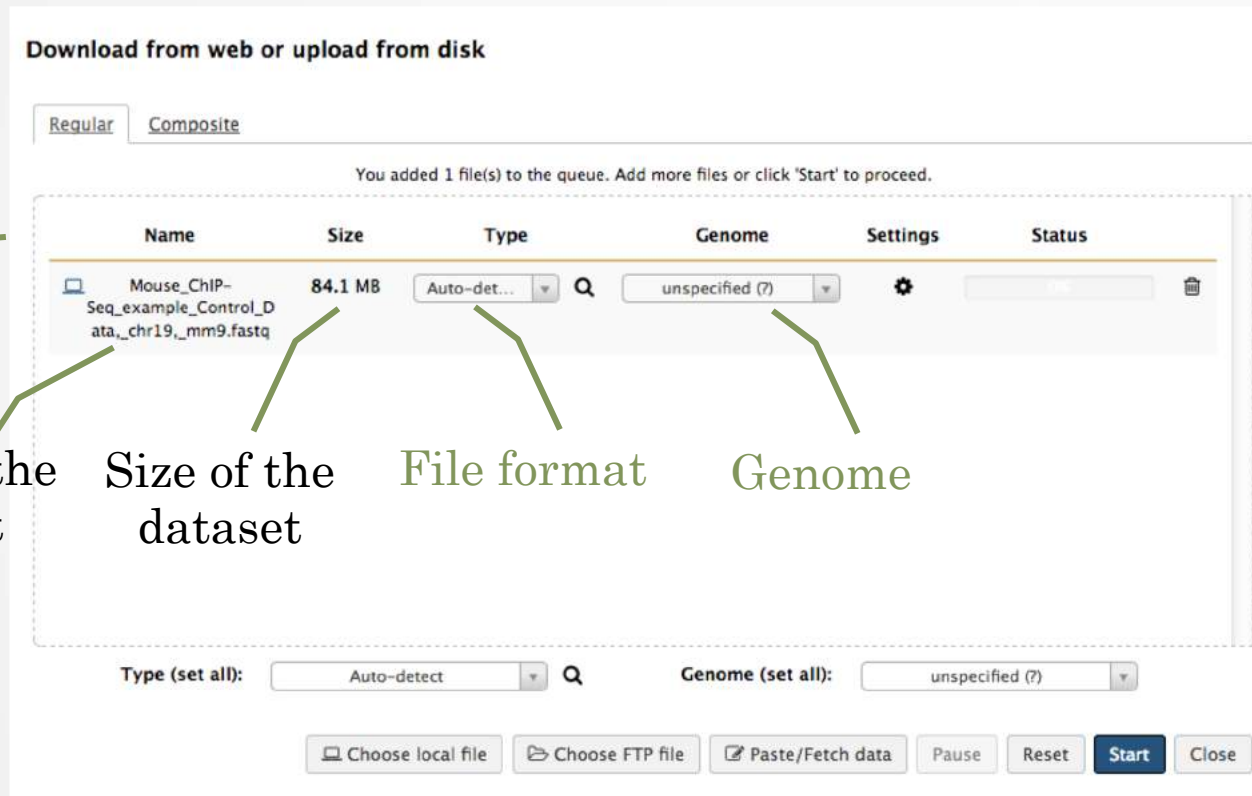
Import data into Galaxy

- 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



Name of the dataset

Size of the dataset

File format

Genome

Import shared data (data libraries)

Shared Data ▾

Visualization

(Top menu)

Data Libraries

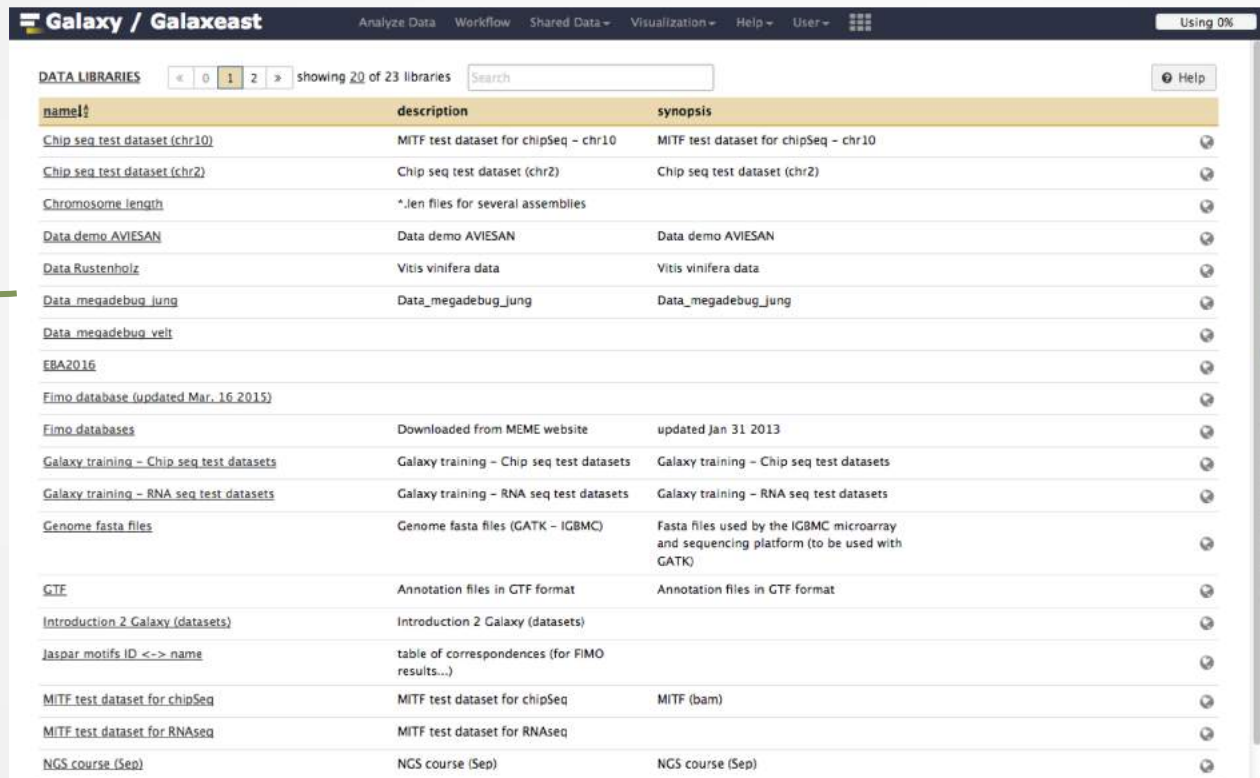
Histories

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_jung	Data_megadebug_jung	Data_megadebug_jung
Data_megadebug_velt		
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 to History Download Delete Details Help

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"/> dm3.len		len	227 bytes	2015-01-08 01:25	
<input type="checkbox"/> hg19.len		len	376 bytes	2015-01-08 01:25	
<input type="checkbox"/> mm10.len		len	1.4 KB	2015-01-08 01:25	
<input type="checkbox"/> mm9.len		len	330 bytes	2015-01-08 01:25	
<input type="checkbox"/> tair10.len		len	75 bytes	2015-01-08 01:25	

< 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 add custom tracks track hubs

table: knownGene describe table schema

region: genome position chr1:121427557-121432936 lookup define regions

identifiers (names/accessions): paste list upload list

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 is : bed

Datasets/Jobs in the History

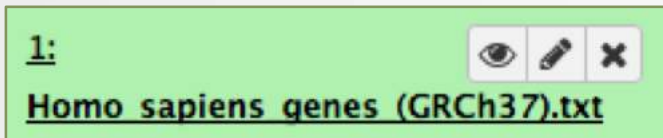
Grey: the job is waiting to run



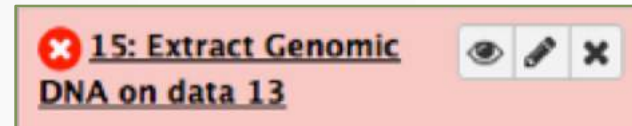
Yellow: the job is running



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

The screenshot shows a web interface for a dataset named '2: sample.bed'. It includes a title bar with an eye icon, an edit icon, and a close icon. Below the title, it displays '32,561 regions' and 'format: bed, database: mm9'. A text input field contains 'uploaded bed file'. Below this are icons for download, information, chart, help, and a folder icon. The interface also provides links to 'display in IGB View', 'display at Ensembl Current', 'display with IGV local Mouse mm9', and 'display at UCSC main test'. At the bottom, a table shows the first five lines of the dataset.

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

The screenshot shows a dataset panel for '2: sample.bed'. At the top right, there are three icons: an eye (view), a pencil (edit), and an 'X' (delete). Below these are the dataset details: '32,561 regions', 'format: bed, database: mm9', and 'uploaded bed file'. A row of icons includes a download icon, an information icon, a bar chart icon, a question mark icon, a magnifying glass icon, and a speech bubble icon. Below the icons are links for 'display in IGB View', 'display at Ensembl Current', 'display with IGV local Mouse mm9', and 'display at UCSC main test'. At the bottom is a table with 4 columns: '1. Chrom', '2. Start', '3. End', and '4. Name'. The table contains 5 rows of genomic data.

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

The screenshot displays 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 grey bar labeled "History" with refresh, settings, and list icons. A search bar with the text "search datasets" is positioned below. The main area shows a list of items under the heading "RNA-seq data analysis" with "2 shown". The first item is "7.23 GB" with a checkmark, folder, and chat icon. The second item is "2: sample.bed" with view, edit, and delete icons. The third item is "1: siLuc3_S12040.fastq" with view, edit, and delete icons. Green lines with labels point to the "Using 20%" indicator (labeled "Quota") and the "7.23 GB" text (labeled "Size of history").

Using 20%

History

search datasets

RNA-seq data analysis
2 shown

7.23 GB

2: sample.bed

1: siLuc3_S12040.fastq

Size of history

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 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar lists various tools and categories, with 'Workflows' and 'All workflows' highlighted at the bottom. The main content area features a banner for 'Public Galaxy Servers and still counting' with a '080+' logo. Below the banner is a tweet from the Galaxy Project. The right sidebar shows a 'History' panel with an 'Unnamed history' section and a message indicating it is empty.

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](#).

080+
Public Galaxy Servers
and still counting

Tweets by @galaxyproject

Galaxy Project @galaxyproject
Did we mention: Galaxy Admin Training early registration ENDS IN 12 HOURS. bit.ly/gat2016

History
search datasets
Unnamed history
0 b
This history is empty. You can [load your own data](#) or [get data from an external source](#)

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

Tools

Workflow Canvas | Chip-seq

search tools

Get Data

Send Data

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

Fetch Sequences

Operate on Genomic Intervals

Graph/Display Data

Motif Tools

Multiple Alignments

FASTA manipulation

NGS: QC and manipulation

NGS: Mapping

NGS: RNA Analysis

NGS: SAM Tools

NGS: GATK Tools (beta)

NGS: Peak Calling

NGS: Miscellaneous

NGS: ChIP Analysis

NGS: BED Tools

NGS: Picard

NGS: Variant Annotation

NGS: S-mart

NGS: DeepTools

Database query

Workflow control

Inputs

FastQC: Read QC

Short read data from your current history

Contaminant list

html_file (html)

Input dataset

output

Convert from BAM to BED

Convert the following BAM file to BED

output (bed)

Map reads with bowtie 1 (ungapped)

FASTQ file

output (sam)

output_suppressed_reads_l (fastq)

output_suppressed_reads_r (fastq)

output_unmapped_reads_l (fastq)

output_unmapped_reads_r (fastq)

SAM-to-BAM

SAM File to Convert

output1 (bam)

Details

Tool: Map reads with bowtie 1 (ungapped)

Version: 0.12.8

Will you select a reference genome from your history or use a built-in index?

Use a built-in index

Select a reference genome: ▼

To be set at runtime

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 running a workflow. The top navigation bar includes 'Galaxy / Galaxeast', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. The main content area is titled 'Running workflow "chip workflow"' and contains four steps:

- Step 1: Input dataset**: Shows an 'Input Dataset' dropdown menu with the value '4: chr10_ctr2_1.fastq.gz' and a search filter 'type to filter'.
- 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)**: Shows 'Homer peaks OR BED format' with 'Output dataset 'output_bed_file' from step 3'. The 'Genome version' dropdown is set to 'tair10'. There is an 'Extra options' checkbox and an 'Action' section with 'Hide output 'out_log''. A checkbox 'Send results to a new history' is also present.

A 'Run workflow' button is located at the bottom of the workflow configuration. The right sidebar shows a 'History' section with a search bar and a list of datasets, including 'test' (1 shown, 3 deleted, 120.7 MB) and '4: chr10_ctr2_1.fastq' (format: fastqsanger, database: hg19).

Set parameters

Run workflow

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