

Introduction to Galaxy

Stéphanie Le Gras
(slegras@igbmc.fr)

Training material

- Wiki:

<http://genomeast.igbmc.fr/wiki/>

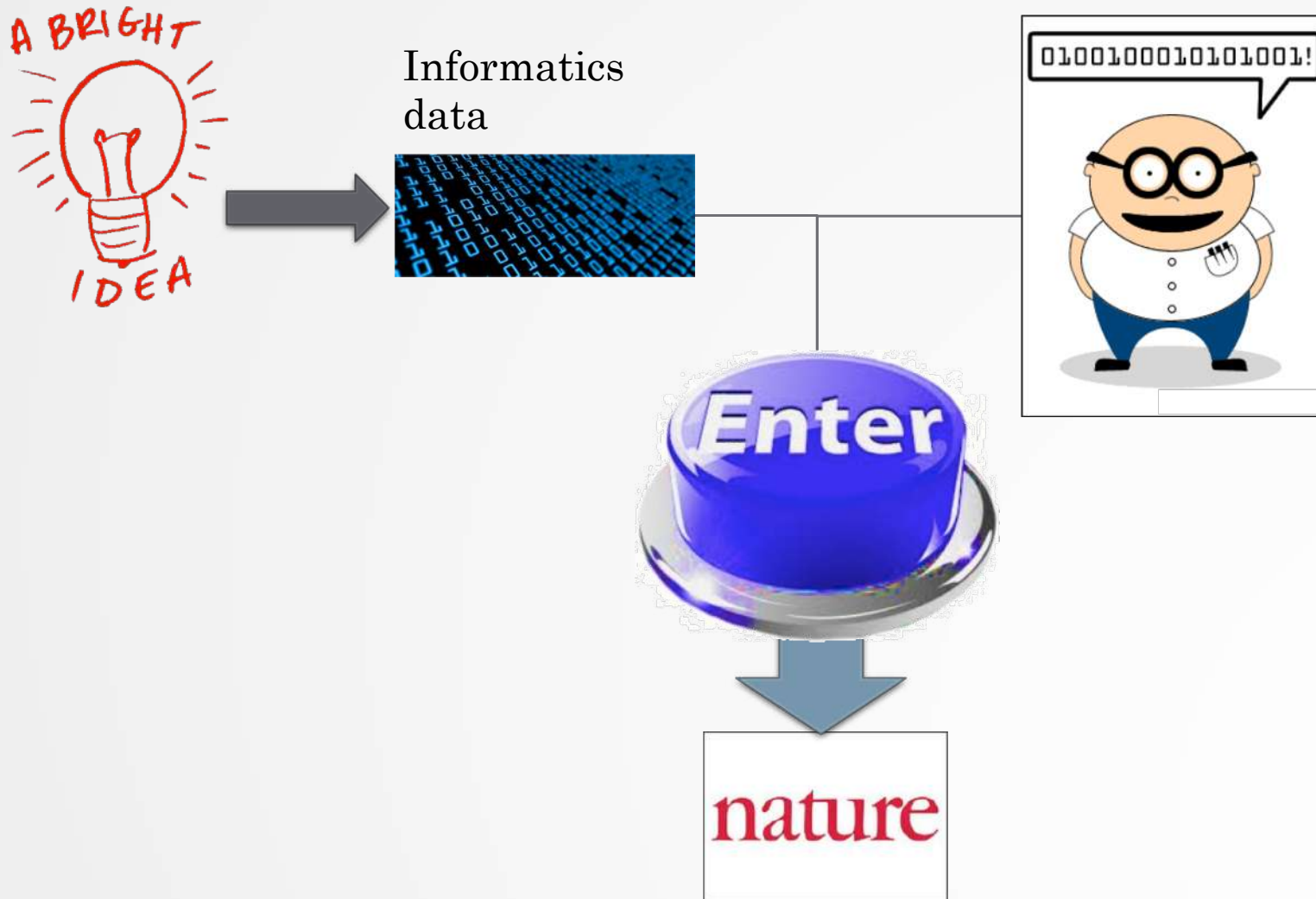
- Go to: Training > DU Dijon > Introduction to Galaxy > Hands-on
- <http://use.galaxeast.fr>
- Login: user[1..22]
- Password: training

Guidelines

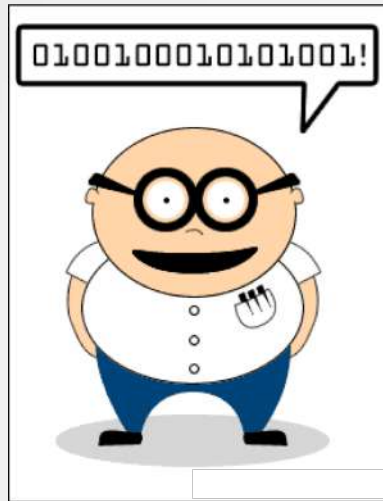
- Analyzing biological data with informatics tools
- Presentation of the Galaxy project
- Description of the main features of the Galaxy platform
- Workflow

Analyzing biological data with informatics tools

Bioinformatics analyses



Bioinformatics analyses



Scripts, software

```
#!/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 );

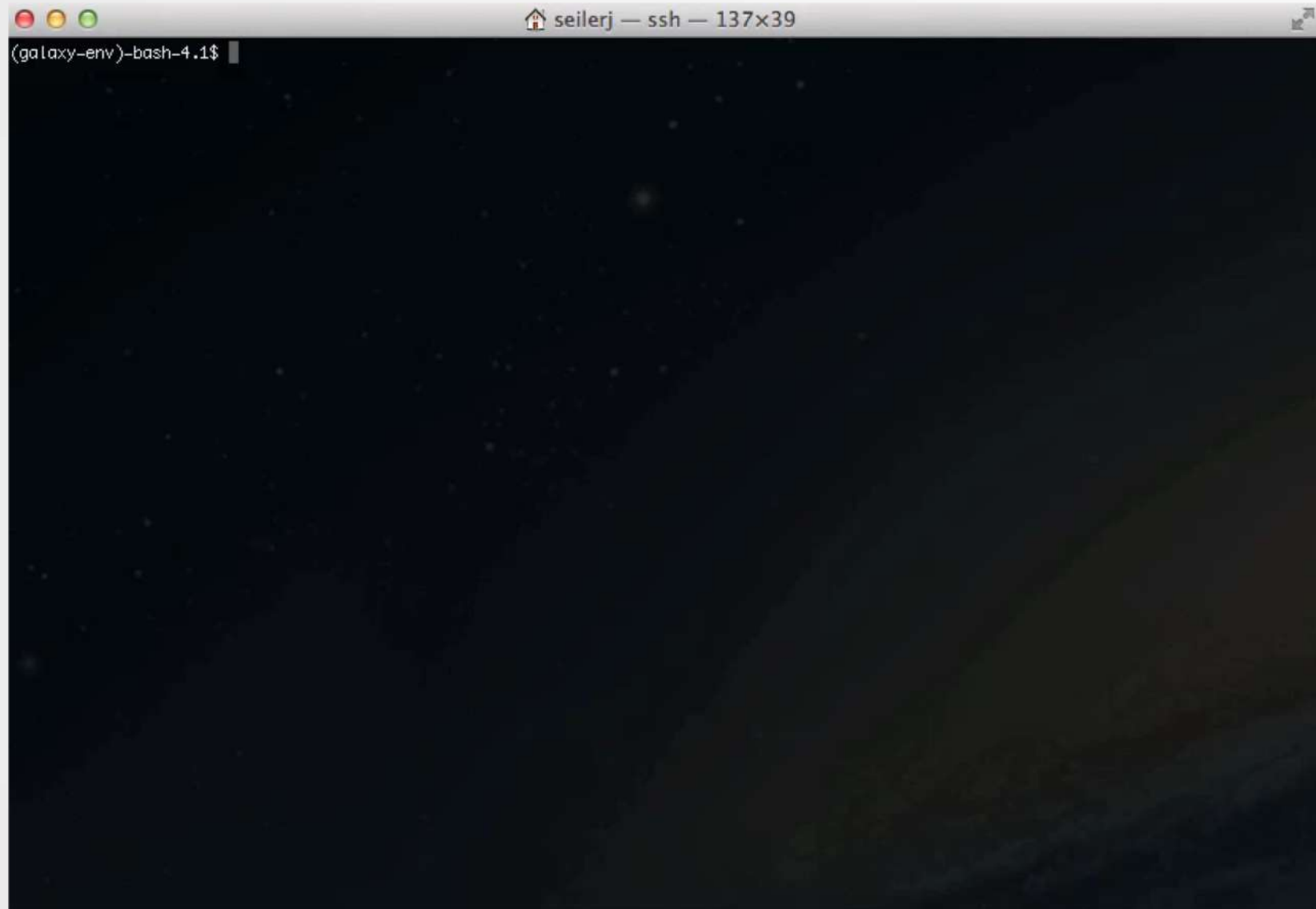
my %ids;
$out = ( defined $out ) ? $out : "results.txt";

## 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.
```

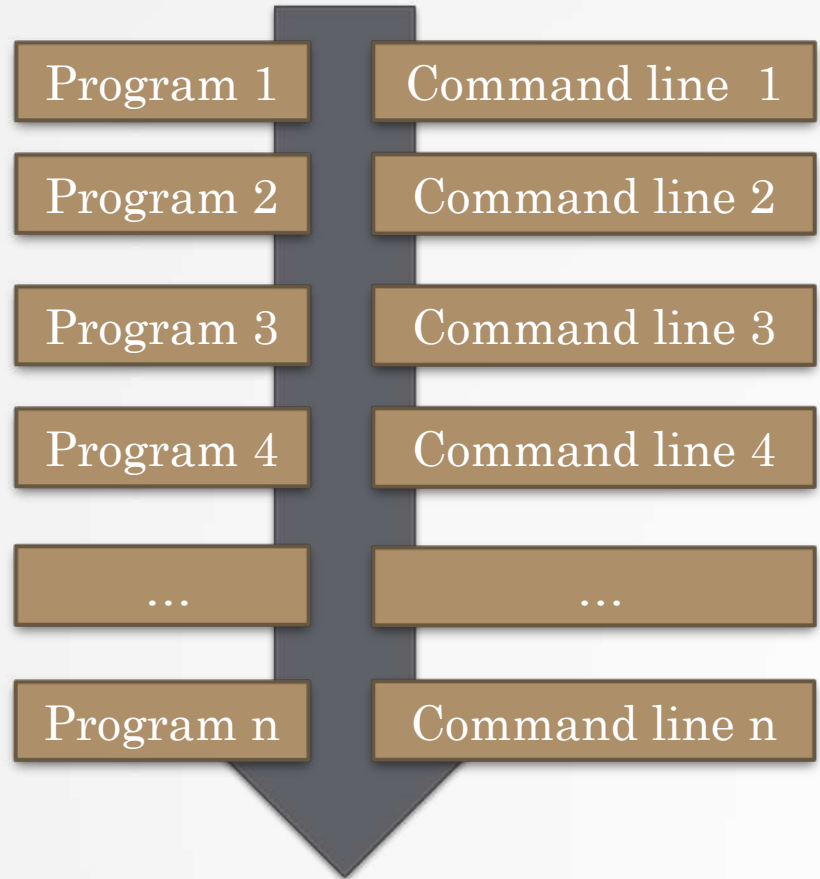
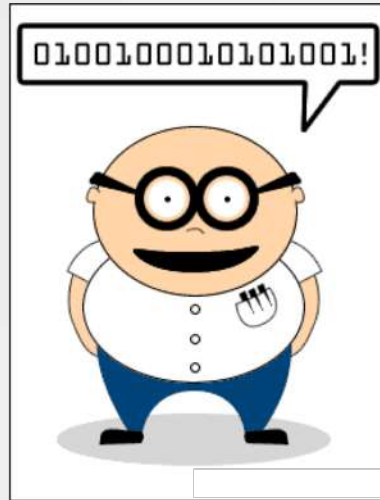
Command line

```
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**.

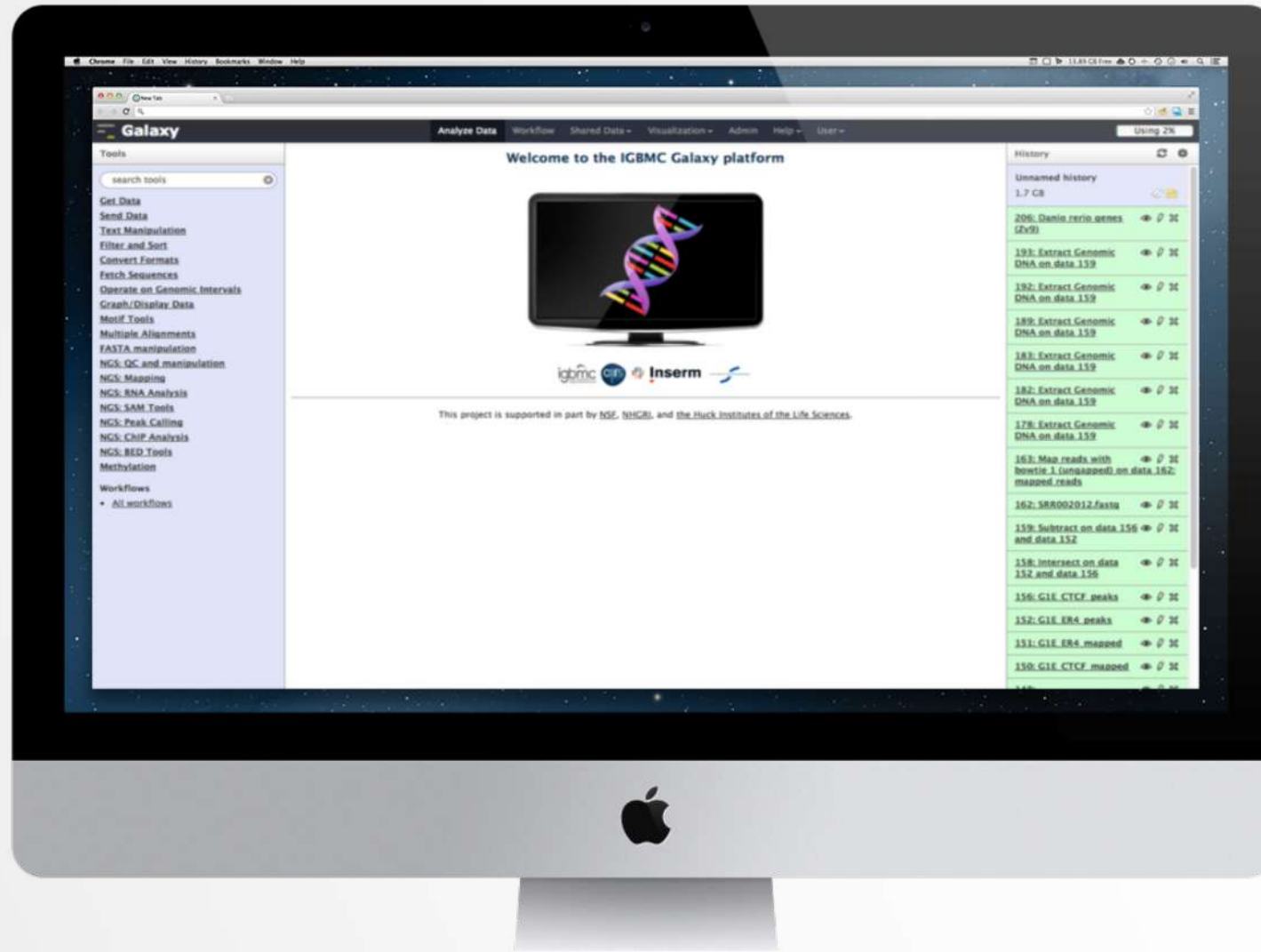


EMORY
UNIVERSITY

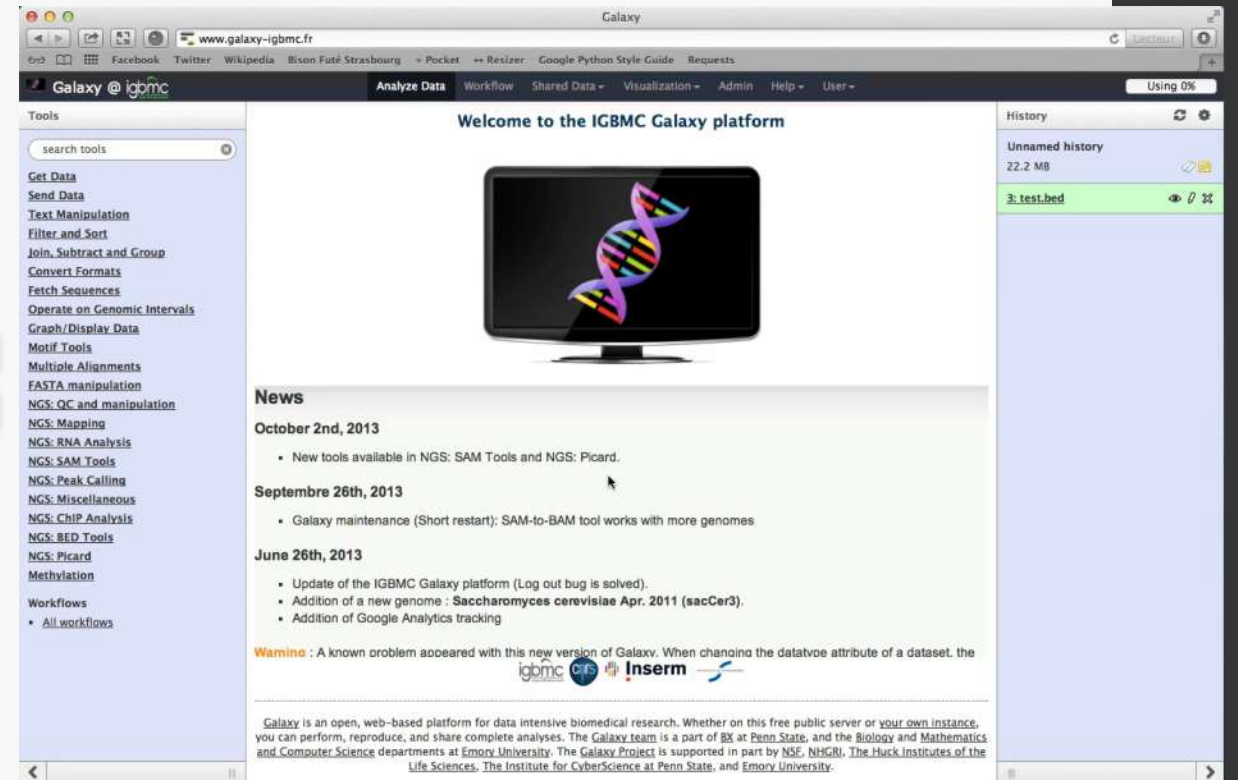
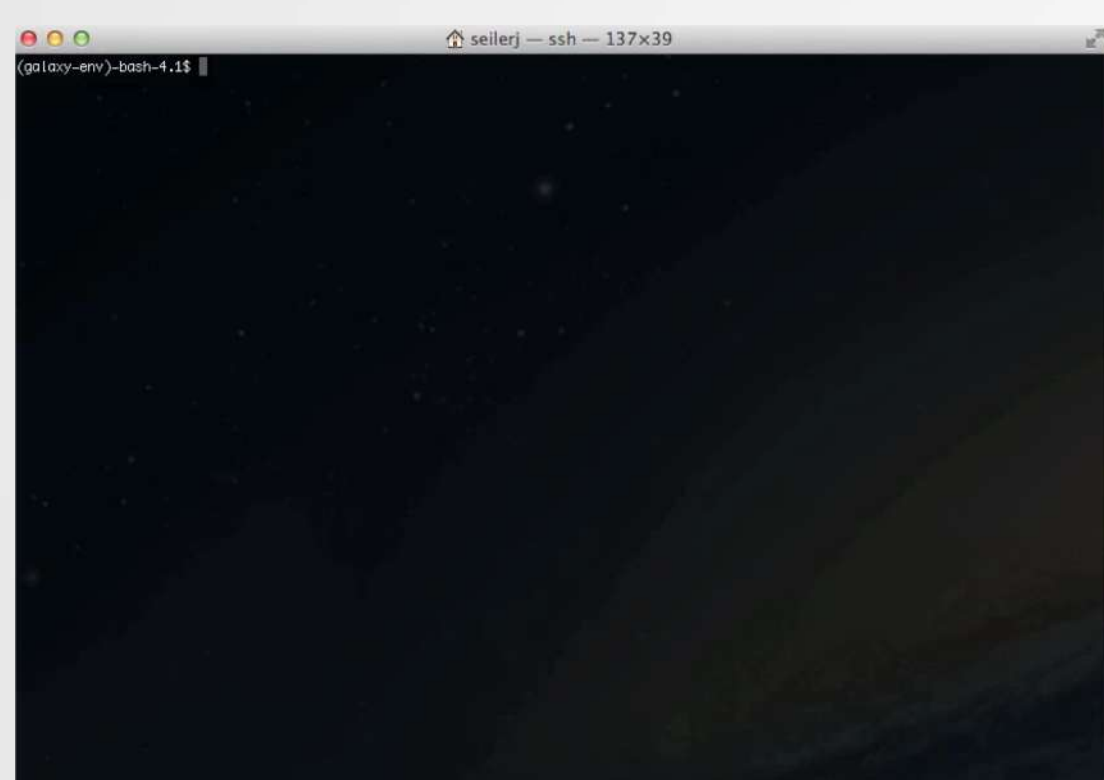


National
Human Genome
Research Institute

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

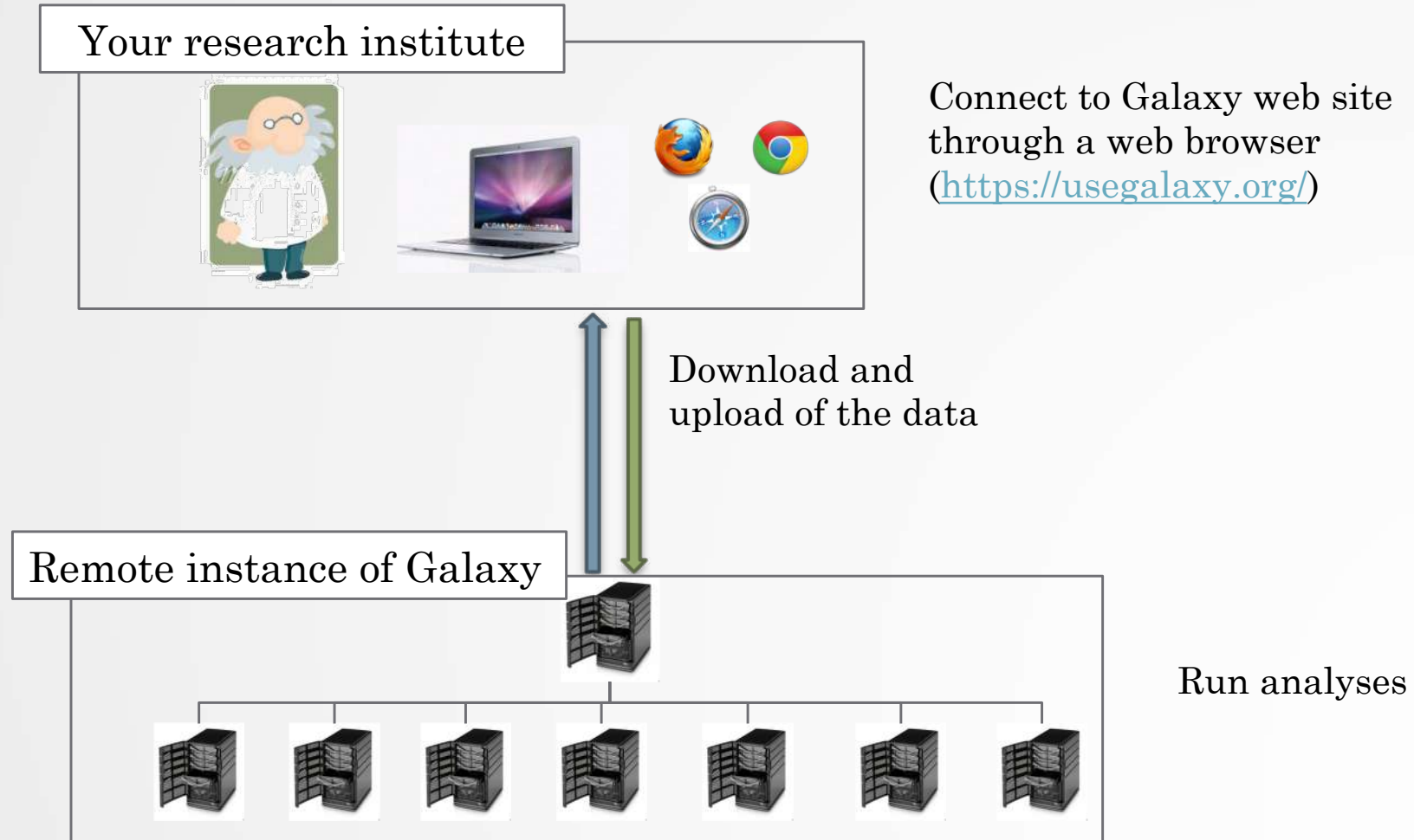
Public Galaxy Servers list :
<https://galaxyproject.org/use/>
Last Update on: 2019, November 14th

Galaxy public servers

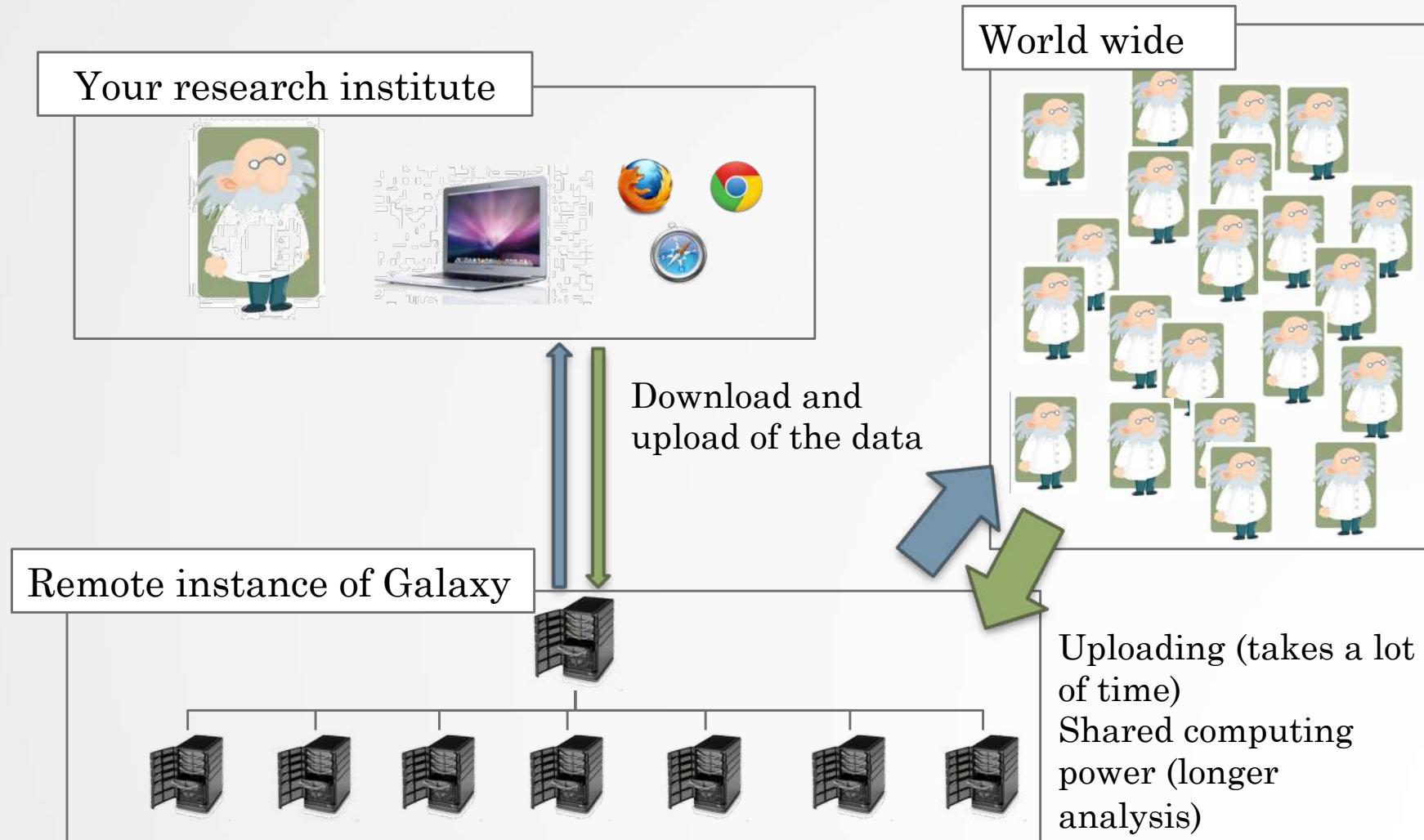
- All analyses are run on remote computing infrastructures
- No need to have a Supercomputer to use Galaxy
- Web browser



Galaxy public servers



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
 - Develop Galaxy tools
 - Develop Galaxy itself



Description of the main features of Galaxy

Galaxy web interface

The image shows a screenshot of the Galaxy web interface. At the top, a dark blue navigation bar contains the 'Galaxy' logo and several menu items: 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. On the far right of this bar, it says 'Using 0 bytes'. A callout box labeled 'Top menu' points to this navigation bar.

On the left side, there is a 'Tools' panel with a search bar and a list of tool categories such as 'Get Data', 'Text Manipulation', and 'NGS: QC and manipulation'. A callout box labeled 'Tool panel' points to this area.

The main central area displays a welcome message for Galaxy, followed by a large graphic showing the number '080+' in a digital font, with the text 'Public Galaxy Servers and still counting' below it. Below this, there is a 'Tweets' section with a tweet from '@galaxyproject'. A callout box labeled 'Data display and tools dialog window' points to this central content area.

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'. A callout box labeled 'History panel' points to this panel.

Top menu

The image shows a screenshot of the Galaxy web interface. The top navigation bar includes the following items: **Analyze Data**, **Workflow**, **Shared Data**, **Visualization**, **Help**, and **User**. The **Help** menu is highlighted with a callout box labeled "Get Help".

Below the navigation bar, the main content area features a central banner with the text: "Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to the platform, you can learn more about Galaxy and its features." This banner is annotated with three callout boxes: "Run workflows" (pointing to the "Analyze Data" menu), "Visualize your data" (pointing to the "Visualization" menu), and "Log in/out, manage your account" (pointing to the "User" menu).

On the left side, there is a "Tools" panel with a search bar and a list of tool categories: **Get Data**, **Lift-Over**, **Text Manipulation**, **Datamash**, **Convert Formats**, **Filter and Sort**, **Join, Subtract and Group**, **Fetch Alignments/Sequences**, **NGS: QC and manipulation**, **NGS: DeepTools**, **NGS: Mapping**, **NGS: RNA Analysis**, **NGS: SAMtools**, and **NGS: BamTools**. A callout box labeled "Run analyses" points to the "Tools" panel.

On the right side, there is a "History" panel with a search bar and a message: "This history is empty. You can load your own data or get data from an external source".

At the bottom center, there is a logo for "Public Galaxy Servers and still counting" featuring the number "080+" in a stylized font.

Hands On

Exercise 1

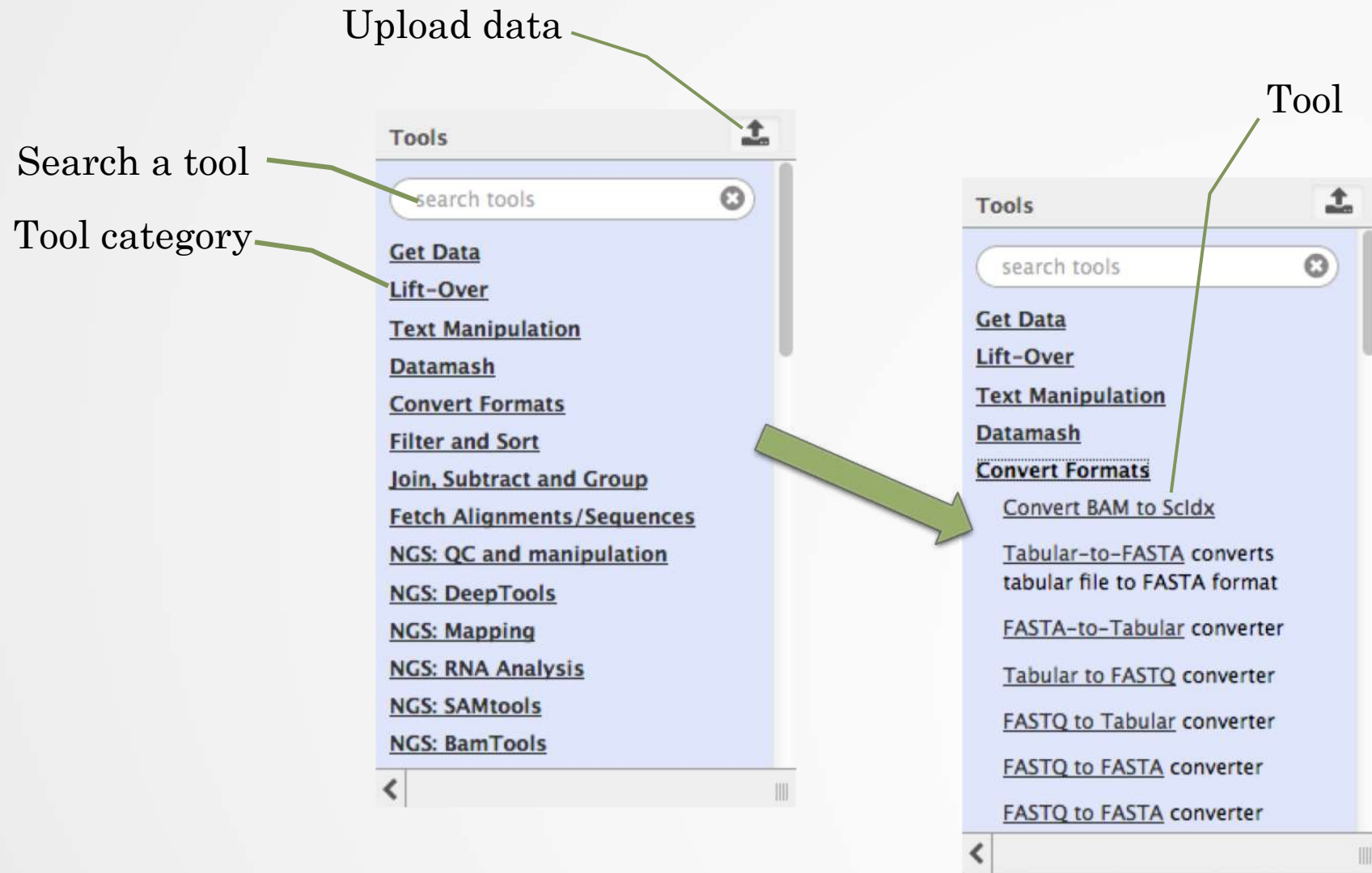
Tool Panel / Run analyses

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". The panel contains a search bar and a list of tool categories: Get Data, Send Data, Lift-Over, Text Manipulation, Datamash, Convert Formats, Filter and Sort, Join, Subtract and Group, Fetch Alignments/Sequences, NGS: QC and manipulation, NGS: DeepTools, NGS: Mapping, NGS: RNA Analysis, NGS: SAMtools, NGS: BamTools, NGS: Picard, NGS: VCF Manipulation, NGS: Peak Calling, NGS: Variant Analysis, NGS: RNA Structure, NGS: Du Novo, NGS: Gemini, Operate on Genomic Intervals, and a partially visible "St...".

The main content area features a header with the Galaxy logo and navigation tabs: Analyze Data, Workflow, Shared Data, Visualization, Help, and User. Below the header is a introductory text: "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](#)." This is followed by a large graphic with the text "080+" and "Public Galaxy Servers and still counting", and a row of seven colored dots. Below this is a "Tweets" section by @galaxyproject, featuring a tweet from Galaxy Project: "Did we mention: Galaxy Admin Training early registration ENDS IN 12 HOURS. bit.ly/gat2016".

On the right is the **History** panel, which includes a search bar for datasets and a section for "Unnamed history" showing "0 b". A blue information box states: "This history is empty. You can [load your own data](#) or [get data from an external source](#)".

Tool Panel / Run analyses



Tools dialog window

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

Tools

search tools

Get Data
Send Data
Lift-Over
Text Manipulation
Datamash
Convert Formats

Convert BAM to Scidx

Tabular-to-FASTA converts tabular file to FASTA format

FASTA-to-Tabular converter

Tabular to FASTQ converter

FASTQ to Tabular converter

FASTQ to FASTA converter

FASTQ to FASTA converter

BED-to-GFF converter

GFF-to-BED converter

MAF to BED Converts a MAF formatted file to the BED format

MAF to Interval Converts a MAF formatted file to the Interval format

MAF to FASTA Converts a MAF formatted file to FASTA format

SFF converter

Wig/BedGraph-to-bigWig converter

BED-to-bigBed converter

Filter and Sort

Citations

Convert BAM to Scidx (Galaxy Version 1.0.0) Options

BAM file
No bam dataset available.

Require proper mate-pairing?
Yes
Required if filtering by insert size for single-end Reqd 1 (below).

Read to output
Read1

Minimum insert size to output
Will not filter out single-end Read 1 unless proper mate-pairing is required (above).

Maximum insert size to output
Will not filter out single-end Read 1 unless proper mate-pairing is required (above).

Execute

What it does
Converts BAM data to Scidx, the Strand-specific coordinate count format, which is used by tools within the Chip-exo Galaxy tool suite. When the strand is not specified, the position of the strand and the strand will be used to create the Scidx file.

Options

- Require proper mate-pairing parameter
- Minimum insert size for single-end reads
- Maximum insert size for single-end reads

History

search datasets

Unnamed history

0 b

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. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar lists various tool categories such as 'Get Data', 'Send Data', 'Text Manipulation', and 'NGS: QC and manipulation'. The main content area features a header for 'Public Galaxy Servers and still counting' with a '080+' logo and a tweet from the Galaxy Project. On the right, the 'History' panel is highlighted with a green border. It contains a search bar for datasets, a section for 'Unnamed history' showing '0 b', and a message: '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 LISTS
 - Saved Histories
 - Histories Shared with Me
- HISTORY ACTIONS
 - Create New
 - Copy History
 - Share or Publish
 - Show Structure
 - Extract Workflow
 - Delete
 - Delete Permanently
- 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 four history panels. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', 'User', and 'Using 0%'. Below the navigation bar are search bars for 'search histories' and 'search all datasets', along with a 'Create new' button. The main area is divided into four panels, each with a 'Switch to' dropdown menu.

- Panel 1 (Left):** 'Current History' panel, 'Unnamed history', 0 b. A message states 'This history is empty'.
- Panel 2:** 'Unnamed history', 127.21 KB, 6 shown. Contains a list of 6 datasets:
 - 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](#)
- Panel 3:** 'Unnamed history', 97.7 KB, 2 shown, 3 deleted. Contains a list of 2 datasets:
 - 5: [Correspondance_JASPAR_CORE.txt](#)
 - 4: [fimo.txt](#)
- Panel 4 (Right):** 'Unnamed history', 1.09 GB, 1 shown, 1 deleted. Contains a list of 1 dataset:
 - 2: [Brn2_Day2_rtta_rep2.sort.bed](#)

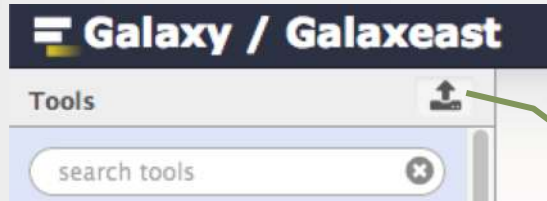
Hands On

Exercise 2

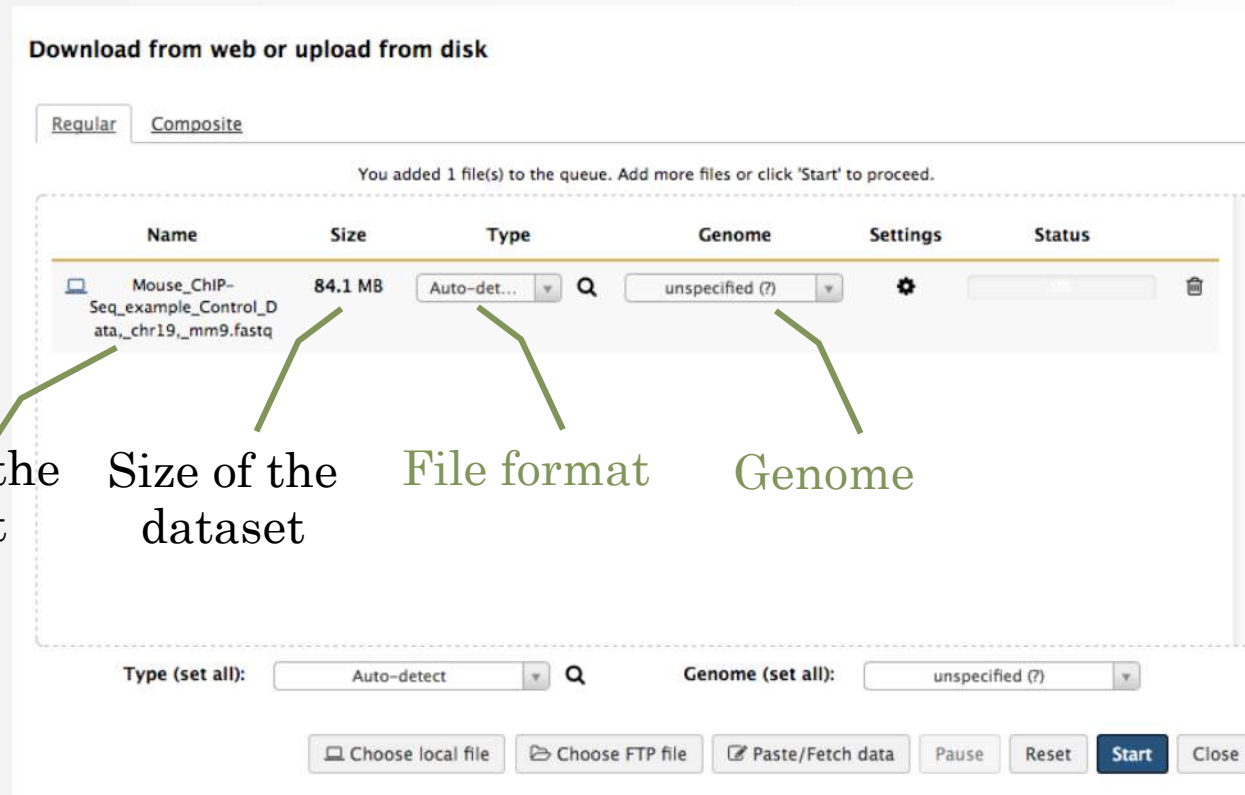
How to import data into Galaxy

1. Your own data (from your computer)
2. Shared data
3. Data from external sources

1. 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

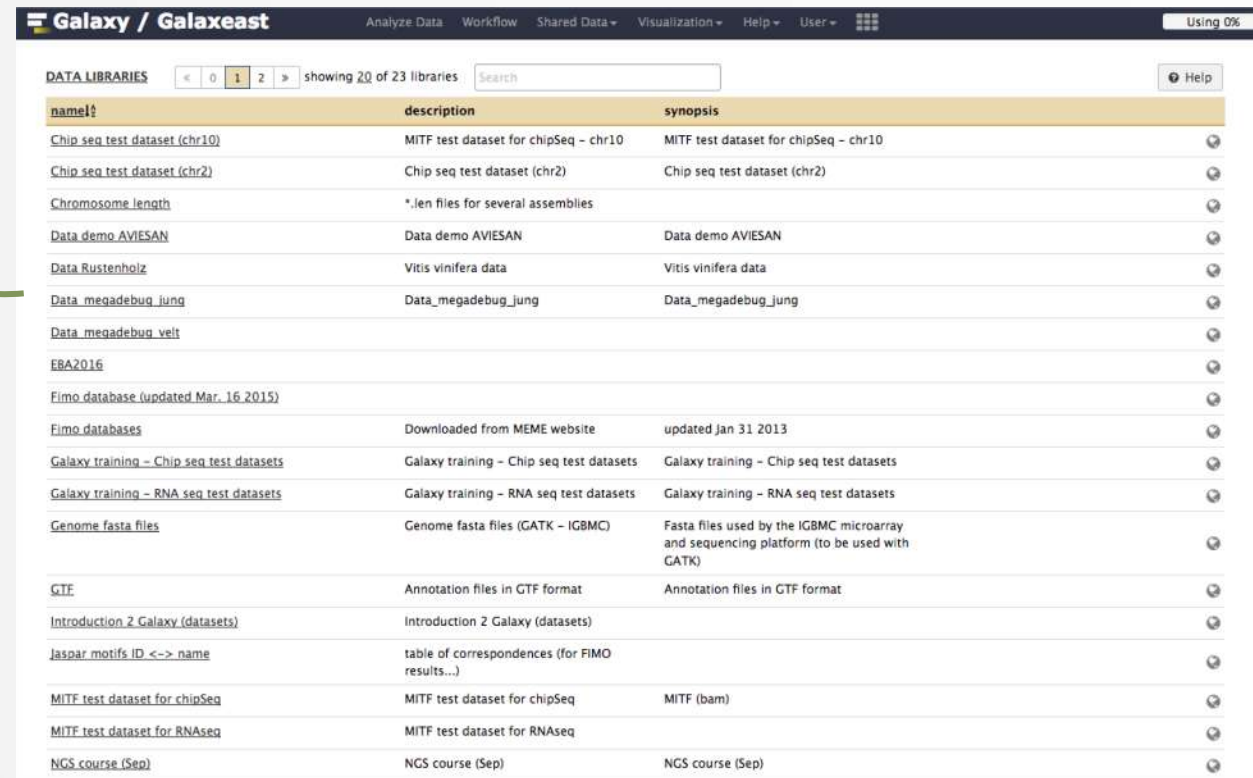
Genome

2. 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_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)

2. 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

3. 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 (GRCh38/mm10)

group: Genes and Gene Predictions track: UCSC Genes add custom tracks track hubs

table: knownGene describe table schema

region: genome position chr1:121427657-121432938 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.

Hands On

Exercise 3.1

Datasets/Jobs in the History

Grey: the job is waiting to run



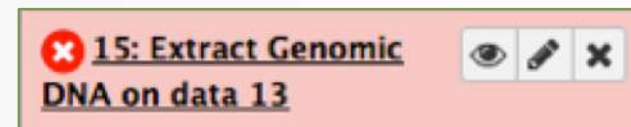
Yellow: the job is running



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

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

2: sample.bed [eye] [edit] [close]

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

uploaded bed file

[save] [info] [chart] [share] [comment]

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	chr1-193
chr1	64972363	64972563	chr1-648
chr1	134238383	134238583	chr1-134
chr1	51991430	51991630	chr1-518
chr1	53880739	53880939	chr1-537

Datasets/Jobs in the History

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

Download dataset

2: sample.bed [eye] [pencil] [x]

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

uploaded bed file

[download] [info] [chart] [share] [comment]

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	chr1-193
chr1	64972363	64972563	chr1-648
chr1	134238383	134238583	chr1-134
chr1	51991430	51991630	chr1-518
chr1	53880739	53880939	chr1-537

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 11%". Below this is a "History" panel with a search bar labeled "search datasets". The panel lists one item: "DNA-seq data analysis", which is "1 shown" and has a size of "1.42 MB". To the right of the size are three icons: a checkmark, a trash can, and a speech bubble. Below the list is a green bar representing the selected item, labeled "1: sample.bed", with three icons: an eye, a pencil, and a close button (X). Two green lines with arrows point from external text labels to the interface: "Size of history" points to the "1.42 MB" text, and "Quota" points to the "Using 11%" header bar.

Using 11%

History

search datasets

DNA-seq data analysis
1 shown

1.42 MB

1: sample.bed

Size of history

Quota

Hands On

Exercise 3.2

Hands On

Exercise 4

Hands On

Exercise 5

Hands On

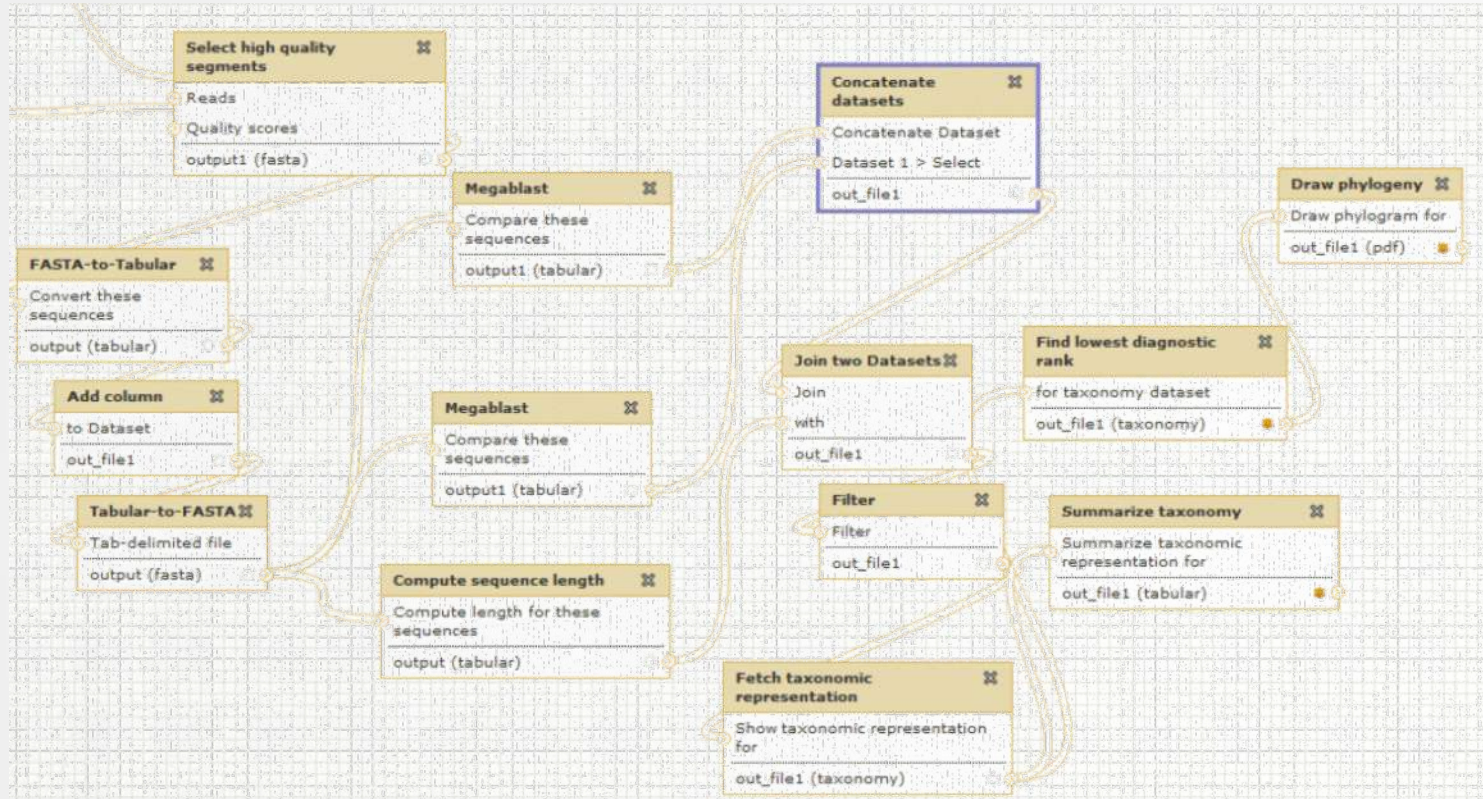
Exercise 6

Workflow

What if we'd mix all together



Galaxy workflow



Galaxy workflows

- Workflow:
 - Analysis protocol with several steps (tools)
 - The output of a step is used as the input of the next next so file formats between two steps should be compatible!
- Workflows are often made general so that they can be run on various datasets
- Some of the parameters are pre-defined while others are set at runtime

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 tool categories, with 'Workflows' at the bottom containing a sub-item 'All workflows'. A green arrow points from the 'Workflow' menu item to the text 'Create, run, edit (...) workflows'. Another green arrow points from the 'All workflows' sub-item to the text 'Run workflows'. The main content area features a header about Galaxy being an open-source platform, a 'Public Galaxy Servers and still counting' graphic with a '080+' counter, and a tweet from the Galaxy Project about an admin training registration deadline.

Create, run,
edit (...) workflows

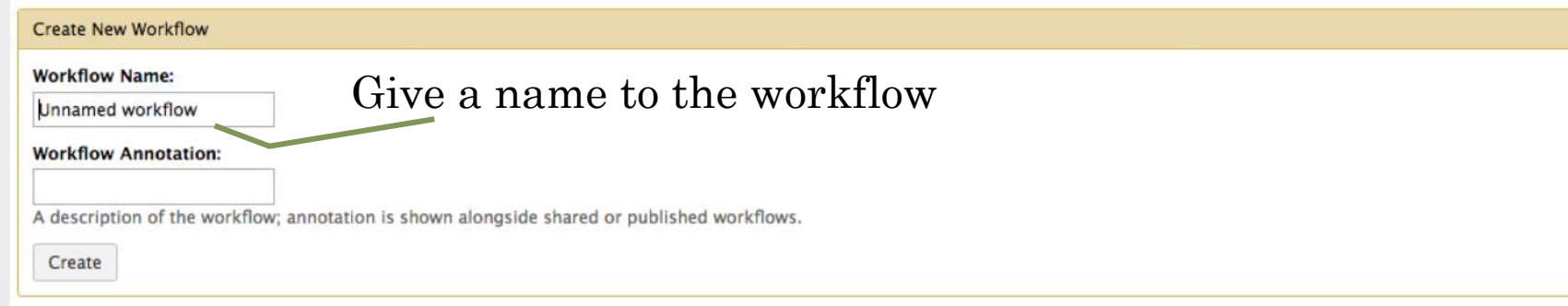
Run workflows

Workflows



The screenshot shows the top navigation bar of the Galaxy / Galaxeast interface. The main menu includes 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right. Below the navigation bar, the 'Your workflows' section indicates that the user has no workflows. The 'Workflows shared with you by others' section also shows no shared workflows. Under 'Other options', there is a button labeled 'Configure your workflow menu'. In the top right corner of the main content area, there are two buttons: 'Create new workflow' (with a plus icon) and 'Upload or import workflow' (with an upload icon). A green arrow points from the text 'Create workflows' to the 'Create new workflow' button.

Create workflows



The 'Create New Workflow' form is displayed with a yellow header. It contains two input fields: 'Workflow Name:' with the text 'Unnamed workflow' and 'Workflow Annotation:'. A green arrow points from the text 'Give a name to the workflow' to the 'Workflow Name' input field. Below the annotation field, there is a descriptive note: 'A description of the workflow; annotation is shown alongside shared or published workflows.' At the bottom left of the form is a 'Create' button.

Give a name to the workflow

Workflow creation

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

Tools

search tools

Inputs

- Get Data
- Send Data
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- Extract Features
- Fetch Sequences
- Statistics
- Graph/Display Data

NGS TOOLBOX BETA

- NGS: QC and manipulation
- NGS: SAM Tools
- Operate on genomic intervals
- Motif tools
- FASTA manipulation
- NGS: GATK Tools (beta)
- NGS: Peak Calling
- NGS: Homer
- NGS: BEDtools
- NGS: Picard
- NGS: Variant Annotation
- NGS: Miscellaneous
- NGS: RNA Analysis
- NGS: Mapping
- NGS: DeepTools
- NGS: RSeQC
- Multiple alignments

Workflow Canvas | Test

Details

Edit Workflow Attributes

Name:
Test

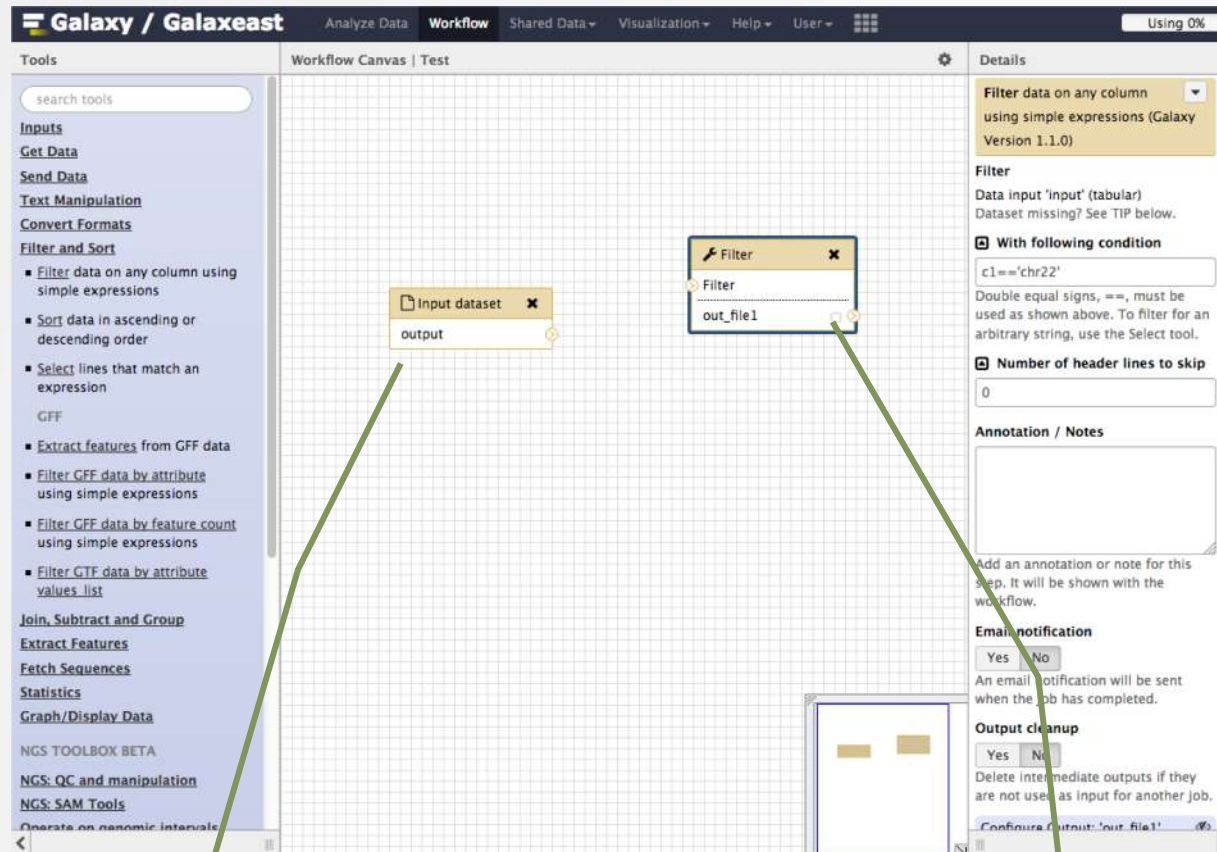
Tags:
[+]

Apply tags to make it easy to search for and find items with the same tag.

Annotation / Notes:
test
Add an annotation or notes to a workflow; annotations are available when a workflow is viewed.

Add tools or input datasets to the workflow

Workflow creation



Input dataset.

Most of the time, a workflow starts with an input dataset to which analyses are applied. In Galaxy, the file format of the input dataset will be limited to the input file format of the subsequent step

Tool to be run

Workflow creation

The screenshot displays the Galaxy / Galaxeast workflow creation interface. The main canvas shows two steps: 'Input dataset' and 'Filter'. A green line connects the 'output' of the 'Input dataset' step to the 'Filter' step, indicating a valid link. The 'Filter' step is configured with the condition 'c1==chr22' and 'Number of header lines to skip' set to 0. The 'Details' panel on the right shows the configuration for the 'Filter' tool, including the condition 'c1==chr22', the number of header lines to skip (0), and options for email notification and output cleanup.

If two steps can be linked together, the link between the two boxes is green

Workflow creation

The screenshot displays the Galaxy workflow editor interface. The top navigation bar includes 'Galaxy / Galaxeast', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', 'User', and 'Using 0%'. The main area is divided into three sections: 'Tools', 'Workflow Canvas | Test', and 'Details'.

Tools: A sidebar on the left contains a search bar and a list of tool categories: 'Inputs', 'Get Data', 'Send Data', 'Text Manipulation', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Extract Features', 'Fetch Sequences', 'Statistics', 'Graph/Display Data', 'NGS TOOLBOX BETA', 'NGS: QC and manipulation', and 'NGS: SAM Tools'. The 'Filter and Sort' category is expanded, showing several options like 'Filter data on any column using simple expressions'.

Workflow Canvas | Test: A central grid shows a workflow with two tools. The first tool is 'Input dataset' with an output named 'output'. A line connects this output to the input of the second tool, 'Filter', which has an output named 'out_file1'.

Details: The right-hand panel shows the configuration for the 'Filter' tool. It includes a dropdown menu for 'Filter data on any column using simple expressions (Galaxy Version 1.1.0)', a 'Filter' section with a text input field containing 'c1==chr22', a 'Number of header lines to skip' input field set to '0', an 'Annotation / Notes' text area, an 'Email notification' section with 'Yes' and 'No' buttons, and an 'Output cleanup' section with 'Yes' and 'No' buttons. A green arrow points from the 'Output cleanup' section to the 'Configure Output: out_file1' link at the bottom of the details panel.

Pre-configure tool parameters and configure parameters to be set at run time

Workflow creation

Click on star to select which datasets will be displayed in the history generated when running of the workflow

Click to get the parameter to be set at runtime

Workflow creation

Save, run workflows

The screenshot displays the Galaxy / Galaxeast workflow editor. The interface includes a top navigation bar with tabs for 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right. On the left, a 'Tools' sidebar lists various categories such as 'Inputs', 'Get Data', 'Send Data', 'Text Manipulation', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Extract Features', 'Fetch Sequences', 'Statistics', and 'Graph/Display Data'. The central 'Workflow Canvas | Test' area shows a workflow with three steps: an initial 'set' tool, a 'Filter' tool with the output 'out_file1', and a 'Sort Dataset' tool also with the output 'out_file1'. A context menu is open over the 'Filter' tool, offering options: 'Save', 'Run', 'Edit Attributes', 'Auto Re-layout', and 'Close'. On the right, a configuration panel for the selected tool is visible, featuring sections for 'With following condition' (with a text input field containing 'c1=='chr22''), 'Number of header lines to skip' (with a numeric input field set to '0'), 'Annotation / Notes' (with a text area), 'Email notification' (with 'Yes' and 'No' radio buttons), and 'Output cleanup' (with 'Yes' and 'No' radio buttons). A 'Configure Output: 'out_file1'' option is at the bottom of the panel.

Run workflows

Set input file(s)

The screenshot displays the Galaxy web interface for running a workflow. The main panel is titled "Running workflow 'chip workflow'" and contains four steps:

- Step 1: Input dataset**: Shows an "Input Dataset" dropdown menu with the selected file "4: chr10_ctr2_1.fastq.gz".
- 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" and "Output dataset 'output_bed_file' from step 3". The "Genome version" dropdown is set to "tair10".

Below the steps, there is an "Action" section with the text "Hide output 'out_log'." and a checkbox for "Send results to a new history". A "Run workflow" button is located at the bottom of the main panel.

The right-hand "History" panel shows a search bar and a list of datasets. The selected dataset is "4: chr10_ctr2_1.fastq" with a format of "fastqsanger" and a database of "hg19".

Set parameters

Run workflow

Hands On

Exercise 7

Hands On

Exercise 8

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