



AB⁴BIMS

14/12/2020



Initiation

Lorraine Guéguen

Erwan Corre

Credits to Gildas Le Corguillé, Galaxy Training Network
v2.6



Learning objectives

- Learning objectives:
 - Familiarize yourself with the basics of Galaxy
 - Learn how to import data
 - Learn how to run tools
 - Learn how histories work
 - Learn how to create a workflow
 - Learn how to share your work
 - Understand and master dataset collections

- Introduction
- Data import
- Tools
- Dataset
- History
- Workflow
- Share

- Introduction
- Data import
- Tools
- Dataset
- History
- Workflow
- Share



Sequence files manipulation



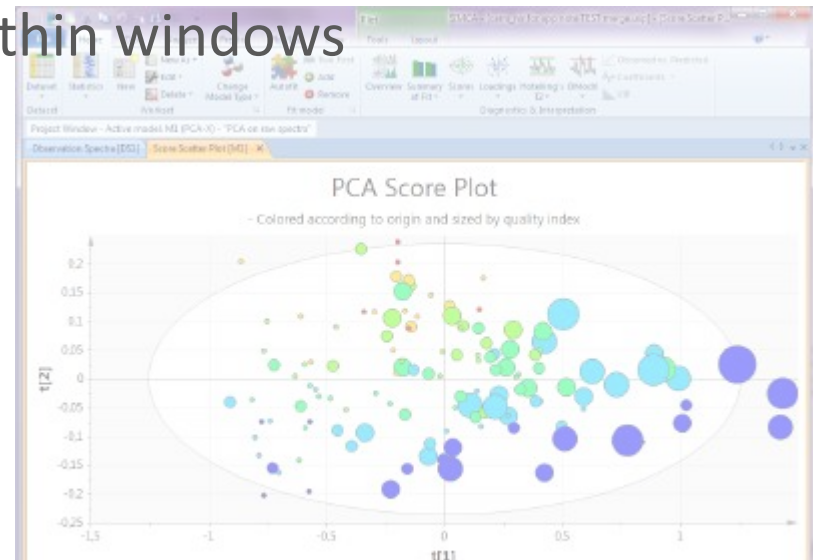
INTRODUCTION / PROBLEMATIC

- In biomedical research, high-throughput technologies produce **large datasets**.
- How to perform analyses of these data **without bioinformatics skills** ?
 - Assemble transcript sequences de novo
 - Determine the gene expression
 - Build a phylogenetic tree
 - Predict subcellular targeting for proteins
 - Identify and quantify metabolites detected by LC-MS
 - ...

Introduction / problematic

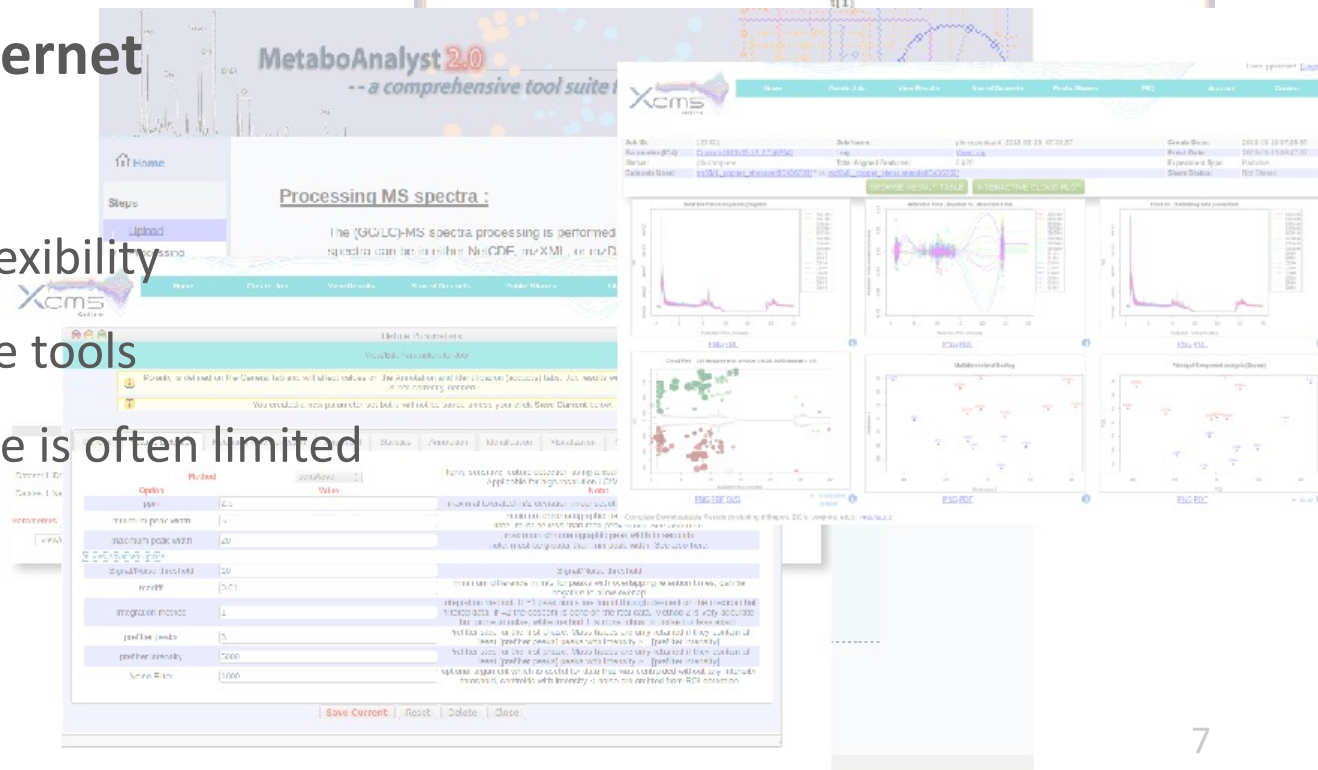
- **Graphical interface click-button tools within windows**

- + very ergonomic
- too ergonomic → lack of flexibility
- few
- paying for it!



- **Tools available on the internet**

- + very ergonomic
- too ergonomic → lack of flexibility
- A small part of the available tools
- the submission size /storage is often limited
- must not be paranoid



Introduction / problematic

```
library(xcms)
loaddata()
polar = "Polar"
```

● Command line tools

```
noise=250000
xset <- xcmsSet(cdfFiles, ppm=ppm, mzdiff=mzwid, peakwidth=peakwidth, noise=noise, snthresh=snth, method="centWave", fitgauss=TRUE, nSlaves=8)
xset2<-retco
+ represent almost the majority of scientific tools
dev.copy2pdf(device = 2, file = paste(pathResult, "/Ret_Cor-Graph", expe, "_", polar, ".pdf", sep=""), paper="a4", height=9, width=14)
xset3<-group(xset2, minfrac = 0.2, bw = bw, minsamp = 1, mzwid = mzwid, max = 50, sleep = 0)
xset5<-fillPeaks(xset3)
+ good parameters completeness
# rapport final avec statistiques de différences entre les deux classes
reporttab <- diffreport(xset5, filebase =paste(pathResult,"/Rapport_",expe,"_",polar, sep=""), mzdec=4, eicmax=5000, metlin = metlin, classeic=levels(xset5@phenoDa
+ can be executed on high performance computers
#écriture du fichier Excel
dir.create(paste(pathResult,"/Rapport_",expe,"_",polar,"_diffreport/", sep=""), showWarnings = FALSE)
write.table(reporttab,paste(pathResult,"/Rapport_",expe,"_",polar,"_diffreport/resultat_",expe,"_",polar,".xls", sep=""),sep="\t")
+ automatable, workflowsable, ...
library(CAMERA)
#annotation version rapide?
an<-annotate(xsg, nSlaves=8, calcIso=TRUE, calcCaS=FALSE, maxcharge=3, maxiso=4, minfrac=0.5,
ppm=15, mzabs=0.015, quick=FALSE, psg_list=NULL, rules=NULL,
polarity=polarity)
diffreport1<-getPeaklist(an)
- minimum linux knowledge is required
- cruel lack of ergonomics
#diffreport <- annotateDiffreport(xsg,pval_th=0.05,fc=0.1, nSlaves=8, calcIso=TRUE, calcCaS=FALSE, maxcharge=3, maxiso=4, minfrac=0.5,
# ppm=15, mzabs=0.015, quick=FALSE, psg_list=NULL, rules=NULL,
# polarity=polarity, sortpval=FALSE)
diffreport<-cbind(reporttab,diffreport1[,c("isotopes", "adduct", "pcgroup")])
write.table(diffreport, file=paste(pathResult,"/result_",expe,"_",polar,"_CAMERA_diffreport-fast.xls",sep=""), row.names=FALSE, sep="\t")

library(FactoMineR)
pca3<-PCA(t(matacp), axes=c(1,2))
pca3<-PCA(t(matacp), axes=c(1,3))
pca3<-PCA(t(matacp), axes=c(2,3))
pca4<-PCA(t(matacplog2))

# -- output png --
# Percentage of variance
png("percentage_of_variance.png", width =800, height = 400);
barplot(resPCA$eig$per,xlab="Components",ylab="percentage of variance");
dev.off()

png("eigenvalue.png", width =800, height = 400);
barplot(resPCA$eig$eig,xlab="Components",ylab="eigenvalue");
dev.off()

library(ctc)
# -- Normalization: logratio --
if (normalization) {
  data=t(scale(t(data)))
```

```
login@sbr4-1042:~$ ssh -Y login@bioinfo.sb-roscoff.fr
```

```
[...]
```

```
[login@n0 ~]$ cdprojet
```

```
[login@n0 login]$ cd 13-07-29-panda/tmp/mapping
```

```
[login@n0 mapping]$ cat tophat.qsub
```

```
#!/bin/bash
```

```
#$ -S /bin/bash
```

```
#$ -M login@sb-roscoff.fr
```

```
#$ -m bea
```

```
#$ -V
```

```
#$ -cwd
```

```
#$ -o qsub.out
```

```
#$ -e qsub.err
```

```
tophat2 panda_v121029 ../input/I11R1-1.fq ../input/I11R1-2.fq  
-GTF ../input/panda_v121029.gtf --b2-sensitive -r 100  
-num-threads 8
```

```
[login@n0 mapping]$ qsub -q long.q -pe thread 8 tophat.qsub
```

```
Your job 5338969 ("tophat.qsub") has been submitted
```

```
[login@n0 mapping]$ ls
```

```
accepted_hits.bam    junctions.bed        qsub.err    unmapped.bam
```

```
deletions.bed        logs                 qsub.out
```

```
insertions.bed       prep_reads.info     tmp
```

```
[login@n0 mapping]$ cd
```

```
login@sbr4-1042:~$ ssh -Y login@bioinfo.sb-roscoff.fr
```

```
[...]
```

```
[login@n0 ~]$ cdprojet
```

```
[login@n0 login]:
```

```
[login@n0 mapping]
```

```
#!/bin/bash
```

```
#$ -S /bin/bash
```

```
#$ -M login@sb-ro
```

```
#$ -m bea
```

```
#$ -V
```

```
#$ -cwd
```

```
#$ -o qsub.out
```

```
#$ -e qsub.err
```

```
tophat2 panda_v12
```

```
-GTF ../input/pa
```

```
-num-threads 8
```

```
[login@n0 mapping]
```

```
Your job 5338969
```

```
[login@n0 mapping]
```

```
accepted_hits.bam
```

```
deletions.bed
```

```
insertions.bed
```

```
[login@n0 mapping]$ cd
```

NOOOOOOOO!

junctions.bed

[logs](#)

prep_reads.info

qsub.err

qsub.out

[tmp](#)

unmapped.bam

.R1-2.fq

:.qsub







Select your level:

Level 1



« I want to know the gene expression »

Level 2



« I want to map my reads on a reference genome and count them »

Level 3



« I want to launch the tools tophat2 and cufflinks.
I have fastq files and my genome in fasta and gtf. »

Level 4



« I want 1TB for my project. I will launch tophat2 through SSH on the cluster in multi-thread mode.

Next I want to submit the bam file to my genome with cufflinks.

Except that, I will manage :P”

Level 5



« I have a bunch of cool tools!
But I'm the only one who can launch them.

Comments? »



- Web-based platform for computational biomedical research (analysis and data integration)
 - Developed at Penn State, Johns Hopkins, OHSU and Cleveland Clinic with substantial outside contributions
 - Open source under Academic Free License
- More than 6,500 citations
- More than 125 [public Galaxy resources](#)
 - 100+ public servers, many more non-public
 - Both general-purpose and domain-specific



INTRODUCTION / GALAXY

Why Galaxy ?

- Accessibility
- Reproductibility
- Transparency

- Galaxy it's ...
 - A web-based interface
 - No need to execute a command line through a terminal
 - Programming or scripting skills are not required
 - Submission of jobs is transparent through a high performance computer cluster
 - Secure histories and data manager
 - A sharing system for data and protocols
 - Tool-boxes of several bioinformatics fields
 - NGS
 - Metabolomics
 - Statistics
 - Chemistry
 - Image analysis
 - Etc ...



Introduction / Galaxy

MR. GEEK



```

[login@n0 ~]$ cdprojet
[login@n0 login]$ cd 13-07-29-panda/tmp/mapping
[login@n0 mapping]$ cat tophat.qsub
#!/bin/bash
#$ -S /bin/bash
#$ -M login@sb-roscoff.fr
#$ -m bea
#$ -V
#$ -cwd
#$ -o qsub.out
#$ -e qsub.err

tophat2 panda_v121029 ../input/IllR1-1.fq ../input/IllR1-2.fq
-GTF ../input/panda_v121029.gtf --b2-sensitive -r 100
-num-threads 8

[login@n0 mapping]$ qsub -q long.q -pe thread 8 tophat.qsub
Your job 5338969 ("tophat.qsub") has been submitted
    
```



Introduction / Galaxy

```

[login@n0 ~]$ cdprojet
[login@n0 login]$ cd 13-07-29-panda/tmp/mapping
[login@n0 mapping]$ cat tophat.qsub
#!/bin/bash
#$ -S /bin/bash
#$ -M login@sb-roscoff.fr
#$ -m bea
#$ -V
#$ -cwd
#$ -o qsub.out
#$ -e qsub.err

tophat2_panda_v121029 ../input/IllR1-1.fq ../input/IllR1-2.fq
-GTF ../input/panda_v121029.gtf --b2-sensitive -r 100
-num-threads 8

[login@n0 mapping]$ qsub -q long.q -pe thread 8 tophat.qsub
Your job 5338969 ("tophat.qsub") has been submitted
    
```



MR. HAPPY

by Roger Hargreaves



Galaxy
PROJECT

Introduction / Galaxy

```
[lecorguille@n0 ~]$ e-PCR --help
e-PCR: invalid option -- -
usage: [-hV] [posix-options] stsfile [fasta ...]
[compat-options]
where posix-options are:
    -m ##      Margin (default 50)
    -w ##      Wordsize (default 7)
    -n ##      Max mismatches allowed (default 0)
    -g ##      Max indels allowed (default 0)
    -f ##      Use ## discontinuous words, slow if
                ##>1
    -o ##      Set output file
    -t ##      Set output format:
                1 - classic, range (pos1..pos2)
                2 - classic, midpoint
                3 - tabular
                4 - tabular with alignment in comments
                    (slow)
    -d##-##    Set default size range
                (default 100-350)
    -p +-      Turn hits postprocess on/off
    -v ##      Verbosity flags
    -a a|f     Use presize alignmens (only if
                gaps>0), slow
                a - Always or f - as Fallback
    -x +-      Use 5'-end lowercase masking of
                primers (default -)
    -u +-      Uppercase all primers (default -)

[...]
```

Galaxy / ABiMS

e-PCR (version 1.0.0)

STS file:

100: (as tabular) Trinity on data 9..Transcripts

format : tabular

Fasta file:

100: Trinity on data 9..Transcripts

format : fasta

Wordsize (W):

7

Set word size for primers hash (nucleotide positions). Longer word size decreases hash collision rate, but increases memory usage. Also no mismatches are allowed within word size near 'inner' boundary of primers unless one uses discontinuous words, and no gaps are ever allowed in that region.

Use ## discontinuous words (F):

1

Set discontinuous word count for primers hash (1 means 'use contiguous words'). Discontinuous words increase number of hash tables and decrease 'effective' word size (thus increasing hash collision rate), so make search significantly slower, but increase sensitivity by allowing mismatches within word size. Reasonable values are 1 (contiguous words) and 3.

Margin (M):

50

Set maximal allowed deviation of hit product size from expected STS size.

Set default sts lower size (D):

100

Set ddefault STS size range - values used for STSs that have no size associated in file.

Set default sts higher size (D):

400

Set ddefault STS size range - values used for STSs that have no size associated in file.

Max mismatches allowed (N):

0

Set maximal number of mismatches allowed in primer-to-sequence alignment (per primer!).

Max indels allowed (G):

0

Set maximal number of gaps allowed in primer-to-sequence alignment (per primer!).

Set output format (T):

tabular

Output formats

Execute



Analyze Data
Workflow
Visualize
Shared Data
Admin
Help
User

Using 34%

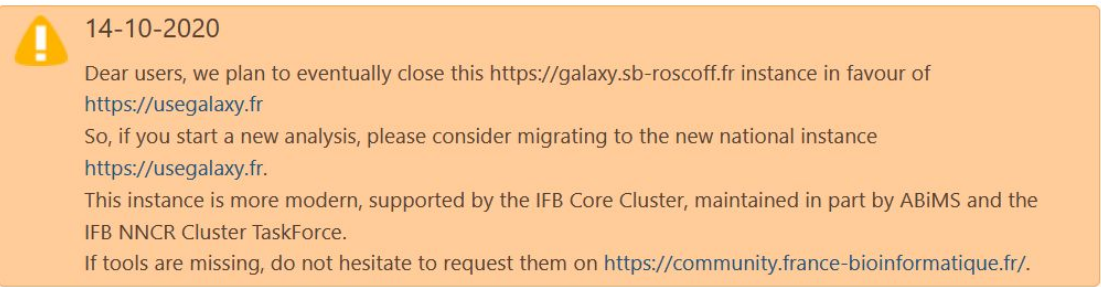
Tools

search tools

Get Data
Send Data
Collection Operations
COMMON TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Statistics
Graph/Display Data
Fasta Fastq Manipulation
Operate on Genomic Intervals
Visualization
NCBI SRA Tools
NGS: deepTools
COMMON NGS TOOLS
NGS: Samtools

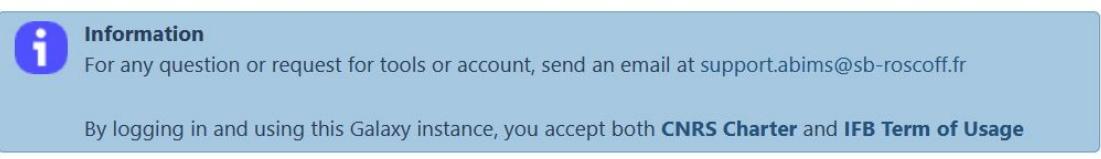


Welcome to galaxy.sb-roscoff.fr



14-10-2020

Dear users, we plan to eventually close this <https://galaxy.sb-roscoff.fr> instance in favour of <https://usegalaxy.fr>. So, if you start a new analysis, please consider migrating to the new national instance <https://usegalaxy.fr>. This instance is more modern, supported by the IFB Core Cluster, maintained in part by ABiMS and the IFB NNCR Cluster TaskForce. If tools are missing, do not hesitate to request them on <https://community.france-bioinformatique.fr/>.



Information

For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

By logging in and using this Galaxy instance, you accept both **CNRS Charter** and **IFB Term of Usage**



Analyses and Bioinformatics for Marine Science

Account

Changelog

- 28-04-20: UPDATE - Trinity suite (2.9.1)
- 26-09-19: NEW - Genome assembler tools: SPAdes (3.12.0) and ABySS (2.2.1)

History

search datasets

Public - Galaxy initiation training 2020 - blast hands-on
13 shown, 12 hidden
154.24 MB

18: NCBI BLAST+ blastn across collection 17
a list with 5 items

17: queries
a list with 5 items

11: blastx query4.fa vs protein BLAST database from data 2

10: megablast query5.fa vs nucleotide BLAST database from data 1

9: protein BLAST database from data 2

8: nucleotide BLAST database from data 1

7: query5.fa

25

Galaxy ABiMS --> Galaxy France

<https://galaxy.sb-roscoff.fr/>



Galaxy / ABiMS Analyze Data Workflow Visualize Shared Data Admin Help User Using 34%

Tools ☆ ↑
 search tools ×

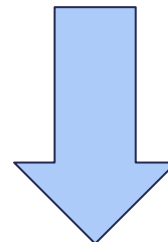
Get Data
 Send Data
 Collection Operations
COMMON TOOLS

Welcome to galaxy.sb-roscoff.fr

14-10-2020
 Dear users, we plan to eventually close this <https://galaxy.sb-roscoff.fr> instance in favour of <https://usegalaxy.fr>. So, if you start a new analysis, please consider migrating to the new national instance <https://usegalaxy.fr>. This instance is more modern, supported by the IFB Core Cluster, maintained in part by ABiMS and the IFB NNCR Cluster TaskForce. If tools are missing, do not hesitate to request them on <https://community.france-bioinformatique.fr/>.

History ↺ + □ ⚙
 search datasets ×

Galaxy initiation training 2019 - tools hands-on
 11 shown, 4 deleted, 7 hidden
 11.73 MB ☑ 🗑 💬



<https://usegalaxy.fr/>



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

Tools ↑
 search tools ×

Get Data
 Send Data
 Collection Operations
GENERAL TEXT TOOLS
 Text Manipulation
 Filter and Sort
 Join, Subtract and Group

Welcome to usegalaxy.fr

By using this Galaxy instance, we assume that you have read and accept the [Term Of Use](#)
 For any questions or support: community.cluster.france-bioinformatique.fr/c/galaxy

History ↺ ⚙
 search datasets ? ×

Unnamed history
 (empty)

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

Galaxy interface

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

Upload File from your computer

Export Data

LC-MS

Format Conversion

Preprocessing

Normalisation

Batch_correction Corrects intensities for signal drift and batch-effects

Determine_batch_correction to choose between linear, lowess and loess methods

Transformation Transforms the dataMatrix intensity values

Quality Control

Statistical Analysis

Annotation

GC-MS

Preprocessing

Normalisation

Quality Control

Statistical Analysis

Annotation

NMR

Preprocessing

Normalisation

Batch_correction (version 2.0.0)

Data Matrix file : 17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv

Sample metadata file : 3: sampleMetadata.tsv

must contain at least the three following columns: 'batch' + 'injectionOrder' + 'sampleType'

Variable metadata file : 16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv

Type of regression model : linear

To select between linear or non-linear (lowess or loess) methods to be used in Van der Kloet algorithm ; when using loess, you can choose to use pools or samples to model batch effect.

Factor of interest : batch

column name of factor of interest (often a biological factor); if none, leave 'batch'

Level of details for plots : basic

Amount of plots in the pdf file output. See Help section for more details.

Execute

Authors

Jean-Francois Martin - PF MetaToul-AXIOM ; INRA ; MetaboHUB (for original version of this tool and overall development of the R script)

Contributors

Melanie Petera - PFEM ; INRA ; MetaboHUB (for R wrapper and R script improvement)

Etienne Thevenot - LIST/LADIS ; CEA ; MetaboHUB (for R script and wrapper concerning "all loess pool" and "all loess sample" methods)

History

search datasets

Sacuri Zip

19 shown

289.7 MB

19: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv (Xdiffreport)

18: xset.group.retcor.group.fillPeaks.annotate.negative.Rdata

17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv

16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv

15: xset.group.retcor.group.fillPeaks.RData

14: xset.group.retcor.group.Rplots.pdf

13: xset.group.retcor.group.RData

12: xset.group.retcor.BPCs_corrected.pdf

11:

27

Galaxy interface

Top menu

The screenshot displays the Galaxy 4 / Metabolomics interface. The top navigation bar includes the Galaxy logo, the text "Galaxy / 4 / Metabolomics", and a series of tabs: "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". The right side of the top bar shows a grid icon and the text "Using -993344424 b".

The left sidebar contains a "Tools" section with a search bar and a list of tool categories: "LC-MS", "Format Conversion", "Preprocessing", "Normalisation", "Quality Control", "Statistical Analysis", "Annotation", "GC-MS", "Preprocessing", "Normalisation", "Quality Control", "Statistical Analysis", "Annotation", "NMR", "Preprocessing", and "Normalisation".

The main content area displays the "Batch_correction (version 2.0.0)" tool configuration. The tool has several input fields and a description:

- Data Matrix file :** 17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv
- Sample metadata file :** 3: sampleMetadata.tsv
- Variable metadata file :** 16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv
- Type of regression model :** linear
- Factor of interest :** batch
- Level of details for plots :** basic

The description of the tool states: "must contain at least the three following columns: 'batch' + 'injectionOrder' + 'sampleType'" and "To select between linear or non-linear (lowess or loess) methods to be used in Van der Kloet algorithm ; when using loess, you can choose to use pools or samples to model batch effect."

The "Execute" button is located at the bottom of the configuration area.

The right sidebar shows the "History" panel, which lists the datasets generated by the tool. The datasets are numbered 11 through 19, and each entry includes a link to the dataset and a small icon indicating its type (e.g., PDF, RData, TSV).

Galaxy interface

Tool list

The screenshot displays the Galaxy 4 Metabolomics web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User. The main content area is titled "Batch_correction (version 2.0.0)".

Tool Configuration:

- Data Matrix file :** 17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv
- Sample metadata file :** 3: sampleMetadata.tsv
must contain at least the three following columns: 'batch' + 'injectionOrder' + 'sampleType'
- Variable metadata file :** 16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv
- Type of regression model :** linear
To select between linear or non-linear (lowess or loess) methods to be used in Van der Kloet algorithm ; when using loess, you can choose to use pools or samples to model batch effect.
- Factor of interest :** batch
column name of factor of interest (often a biological factor); if none, leave 'batch'
- Level of details for plots :** basic
Amount of plots in the pdf file output. See Help section for more details.

Execute

Authors
Jean-Francois Martin - PF MetaToul-AXIOM ; INRA ; MetaboHUB (for original version of this tool and overall development of the R script)

Contributors
Melanie Petera - PFEM ; INRA ; MetaboHUB (for R wrapper and R script improvement)
Etienne Thevenot - LIST/LADIS ; CEA ; MetaboHUB (for R script and wrapper concerning "all loess pool" and "all loess sample" methods)

History Panel:

- search datasets
- Sacuri Zip
19 shown
289.7 MB
- 19: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv (Xdiffreport)
- 18: xset.group.retcor.group.fillPeaks.annotate.negative.Rdata
- 17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv
- 16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv
- 15: xset.group.retcor.group.fillPeaks.RData
- 14: xset.group.retcor.group.Rplots.pdf
- 13: xset.group.retcor.group.RData
- 12: xset.group.retcor.BPCs_corrected.pdf
- 11:

Galaxy interface

Web forms / dataset visualization / miscellaneous information

The screenshot displays the Galaxy web interface for the 'Batch_correction' tool (version 2.0.0). The interface is divided into several sections:

- Header:** 'Galaxy / 4 / Metabolomics' with navigation links for 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. The user is logged in as 'Using -993344424 b'.
- Left Panel (Tools):** A sidebar with a search bar and a list of tool categories: 'LC-MS', 'Format Conversion', 'Preprocessing', 'Normalisation', 'Quality Control', 'Statistical Analysis', 'Annotation', 'GC-MS', 'Preprocessing', 'Normalisation', 'Quality Control', 'Statistical Analysis', 'Annotation', 'NMR', 'Preprocessing', and 'Normalisation'.
- Main Panel (Batch_correction tool):**
 - Data Matrix file:** A dropdown menu showing '17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv'.
 - Sample metadata file:** A dropdown menu showing '3: sampleMetadata.tsv'. Below it, a note states: 'must contain at least the three following columns: 'batch' + 'injectionOrder' + 'sampleType'.
 - Variable metadata file:** A dropdown menu showing '16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv'.
 - Type of regression model:** A dropdown menu showing 'linear'. Below it, a note states: 'To select between linear or non-linear (lowess or loess) methods to be used in Van der Kloet algorithm ; when using loess, you can choose to use pools or samples to model batch effect.'
 - Factor of interest:** A text input field containing 'batch'. Below it, a note states: 'column name of factor of interest (often a biological factor); if none, leave 'batch'.
 - Level of details for plots:** A dropdown menu showing 'basic'. Below it, a note states: 'Amount of plots in the pdf file output. See Help section for more details.'
 - Execute button:** A blue button labeled 'Execute'.
 - Authors:** A section with an information icon and text: 'Jean-Francois Martin - PF MetaToul-AXIOM ; INRA ; MetaboHUB (for original version of this tool and overall development of the R script)'. Below it, a section for 'Contributors' lists: 'Melanie Petera - PFEM ; INRA ; MetaboHUB (for R wrapper and R script improvement)', 'Etienne Thevenot - LIST/LADIS ; CEA ; MetaboHUB (for R script and wrapper concerning "all loess pool" and "all loess sample" methods)'. Below it, a section for 'Contributors' lists: 'Melanie Petera - PFEM ; INRA ; MetaboHUB (for R wrapper and R script improvement)', 'Etienne Thevenot - LIST/LADIS ; CEA ; MetaboHUB (for R script and wrapper concerning "all loess pool" and "all loess sample" methods)'.
- Right Panel (History):** A sidebar with a search bar and a list of datasets. The datasets are listed in a table with columns for dataset name, size, and actions (view, edit, delete). The datasets are: '19: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv (Xdiffreport)', '18: xset.group.retcor.group.fillPeaks.annotate.negative.Rdata', '17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv', '16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv', '15: xset.group.retcor.group.fillPeaks.RData', '14: xset.group.retcor.group.Rplots.pdf', '13: xset.group.retcor.group.RData', '12: xset.group.retcor.BPCs_corrected.pdf', and '11: xset.group.retcor.BPCs_corrected.pdf'.

Galaxy interface

History

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Using -993344424 b

Tools

search tools

Upload File from your computer

Export Data

LC-MS

Format Conversion

Preprocessing

Normalisation

Batch_correction Corrects intensities for signal drift and batch-effects

Determine_batch_correction to choose between linear, lowess and loess methods

Transformation Transforms the dataMatrix intensity values

Quality Control

Statistical Analysis

Annotation

GC-MS

Preprocessing

Normalisation

Quality Control

Statistical Analysis

Annotation

NMR

Preprocessing

Normalisation

Batch_correction (version 2.0.0)

Data Matrix file : 17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv

Sample metadata file : 3: sampleMetadata.tsv

must contain at least the three following columns: 'batch' + 'injectionOrder' + 'sampleType'

Variable metadata file : 16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv

Type of regression model : linear

To select between linear or non-linear (lowess or loess) methods to be used in Van der Kloet algorithm ; when using loess, you can choose to use pools or samples to model batch effect.

Factor of interest : batch

column name of factor of interest (often a biological factor); if none, leave 'batch'

Level of details for plots : basic

Amount of plots in the pdf file output. See Help section for more details.

Execute

Authors

Jean-Francois Martin - PF MetaToul-AXIOM ; INRA ; MetaboHUB (for original version of this tool and overall development of the R script)

Contributors

Melanie Petera - PFEM ; INRA ; MetaboHUB (for R wrapper and R script improvement)

Etienne Thevenot - LIST/LADIS ; CEA ; MetaboHUB (for R script and wrapper concerning "all loess pool" and "all loess sample" methods)

History

search datasets

Sacuri Zip

19 shown

289.7 MB

19:

xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv (Xdiffreport)

18:

xset.group.retcor.group.fillPeaks.annotate.negative.Rdata

17:

xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv

16:

xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv

15:

xset.group.retcor.group.fillPeaks.RData

14:

xset.group.retcor.group.Rplots.pdf

13:

xset.group.retcor.group.RData

12:

xset.group.retcor.BPCs_corrected.pdf

11:

GET HELP


Galaxy / ABiMS

[Analyze Data](#)
[Workflow](#)
[Shared Data](#)
[Visualization](#)
[Admin](#)
[Help](#)
[User](#)

Using 0%

Tools

search tools

[Get Data](#)
[Send Data](#)
[Collection Operations](#)

COMMON TOOLS
[Text Manipulation](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[Convert Formats](#)
[Extract Features](#)
[Fetch Sequences](#)
[Statistics](#)
[Graph/Display Data](#)
[Fasta Fastq Manipulation](#)

COMMON NGS TOOLS
[NGS:Samtools](#)
[NGS:Mapping](#)
[NGS:Bedtools](#)
[NGS:Picard Tools](#)

SEARCHING TOOLS
[Diamond](#)


Welcome to galaxy3.sb-roscoff.fr


Information
 For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr


Analyses and Bioinformatics for Marine Science

Changelog

Tutorials

[Galaxy](#) is an open, web-based platform for data intensive biomedical research. The [Galaxy team](#) is a part of [BX](#) at [Penn State](#), and the [Biology](#) and [Mathematics and Computer Science](#) departments at [Emory University](#). The [Galaxy Project](#) is supported in part by [NHGRI](#), [NSF](#), [The Huck Institutes of the Life Sciences](#), [The Institute for CyberScience](#) at [Penn State](#), and [Emory University](#).

History

search datasets

eba 2016 sartools
 42 shown
 1.59 MB

62: SARTools DESeq2 R objects (.RData)

61: SARTools DESeq2 R log

60: SARTools DESeq2 figures

59: SARTools DESeq2 tables

58: SARTools DESeq2 report

57: SARTools edgeR R objects (.RData)

56: SARTools edgeR R log

55: SARTools edgeR figures

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

Tools

[Get Data](#)
[Send Data](#)
[Collection Operations](#)

COMMON TOOLS
[Text Manipulation](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[Convert Formats](#)
[Extract Features](#)
[Fetch Sequences](#)
[Statistics](#)
[Graph/Display Data](#)
[Fasta Fastq Manipulation](#)

COMMON NGS TOOLS
[NGS:Samtools](#)
[NGS:Mapping](#)
[NGS:Bedtools](#)
[NGS:Picard Tools](#)

SEARCHING TOOLS
[Diamond](#)

Welcome to galaxy3.sb-roscoff.fr

Information
 For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

Analyses and Bioinformatics for Marine Science

[Changelog](#)
[Tutorials](#)

Galaxy is an open, web-based platform for data intensive biomedical research. The [Galaxy team](#) is a part of [BX](#) at [Penn State](#), and the [Biology](#) and [Mathematics and Computer Science](#) departments at [Emory University](#). The [Galaxy Project](#) is supported in part by [NHGRI](#), [NSF](#), [The Huck Institutes of the Life Sciences](#), [The Institute for CyberScience](#) at [Penn State](#), and [Emory University](#).

History

eba 2016 sartools
 42 shown
 1.59 MB

☒
☐
☐

62: SARTools DESeq2 R objects (.RData)

☐
☐
☐

61: SARTools DESeq2 R log

☐
☐
☐

60: SARTools DESeq2 figures

☐
☐
☐

59: SARTools DESeq2 tables

☐
☐
☐

58: SARTools DESeq2 report

☐
☐
☐

57: SARTools edgeR R objects (.RData)

☐
☐
☐

56: SARTools edgeR R log

☐
☐
☐

55: SARTools edgeR figures

☐
☐
☐

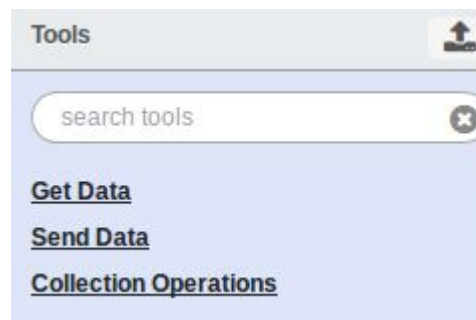
DATA IMPORT

Data import

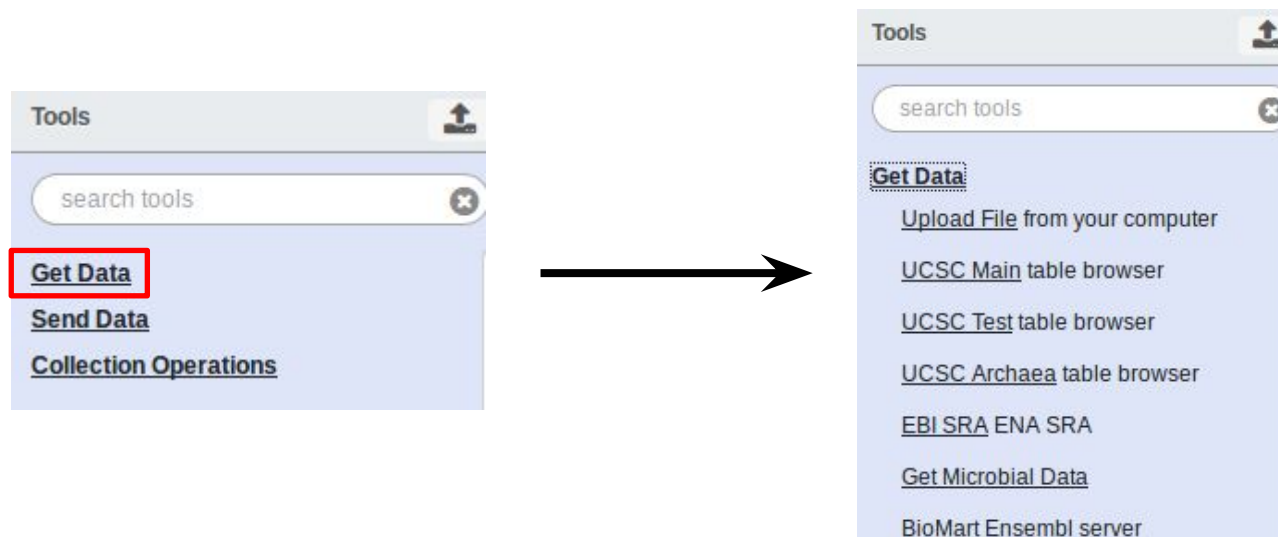
- Copy/paste from a file
- Upload data from your local computer
- Upload data from internet using URL
- Upload data from FTP
- Upload data from online databases: EBI ENA, UCSC, BioMart, etc.
- Import from Shared Data (libraries, histories, workflows)

Data import

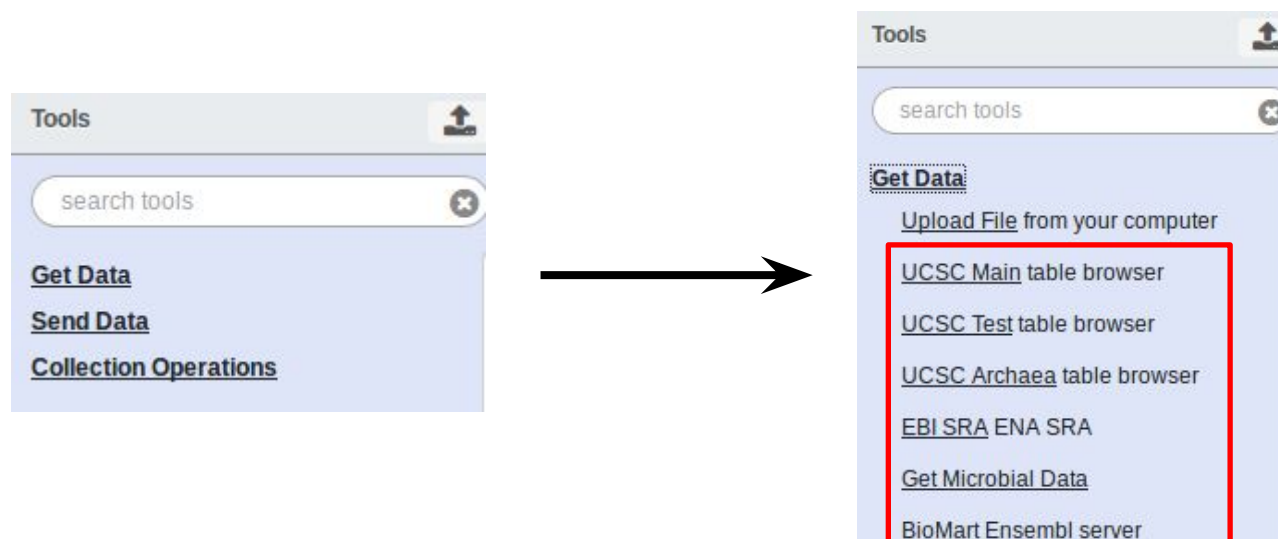
- Copy/paste from a file
- Upload data from your local computer
- Upload data from internet using URL
- Upload data from FTP
- Upload data from online databases: EBI ENA, UCSC, BioMart, etc.
- Import from Shared Data (libraries, histories, workflows)



- Copy/paste from a file
- Upload data from your local computer
- Upload data from internet using URL
- Upload data from FTP
- Upload data from online databases: EBI ENA, UCSC, BioMart, etc.
- Import from Shared Data (libraries, histories, workflows)

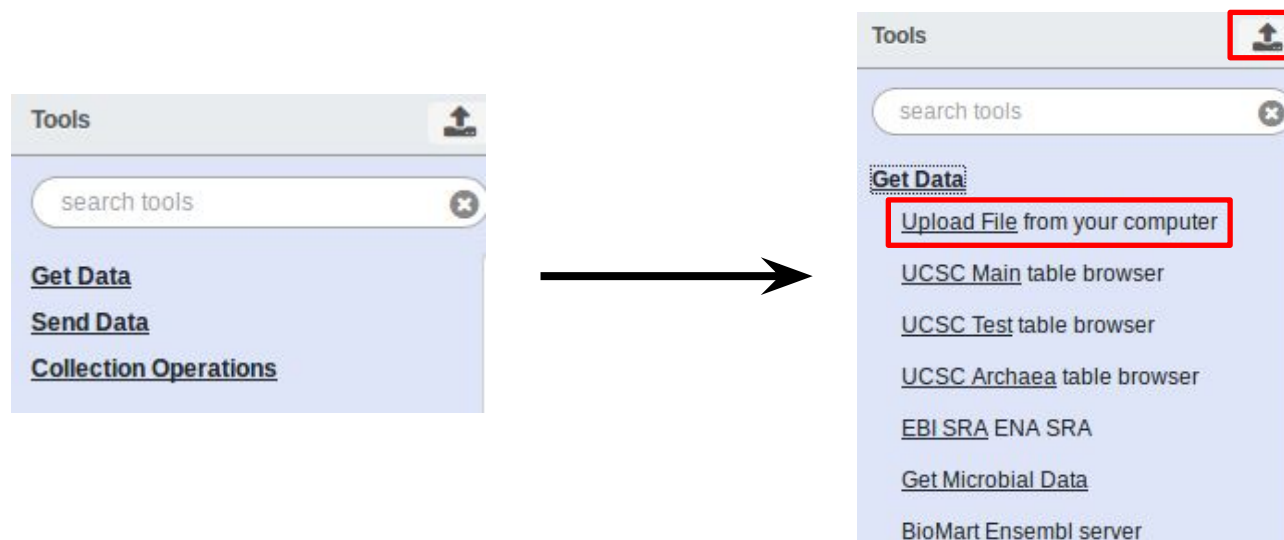


- Copy/paste from a file
- Upload data from your local computer
- Upload data from internet using URL
- Upload data from FTP
- Upload data from online databases: EBI ENA, UCSC, BioMart, etc.
- Import from Shared Data (libraries, histories, workflows)



Data import

- Copy/paste from a file
- Upload data from your local computer
- Upload data from internet using URL
- Upload data from FTP
- Upload data from online databases: EBI ENA, UCSC, BioMart, etc.
- Import from Shared Data (libraries, histories, workflows)



Data import

The screenshot displays the Galaxy 4 / Metabolomics web interface. The browser address bar shows galaxy.workflow4metabolomics.org. The main navigation bar includes links for **Analyze Data**, **Workflow**, **Shared Data**, **Visualization**, **Admin**, **Help**, and **User**. The page features a large green banner with the text "Welcome to workflow4metabolomics.org v2.0" and a publication citation: "Publication: Franck Giacomoni, Gildas Le Corguillé, Mishari Monsoor, Marion Landi, Pierre Pericard, Mélanie Pétéra, Christophe Duperier, Marie Tremblay-Franco, Jean-François Martin, Daniel Jacob, Sophie Goulitquer, Etienne A. Thévenot and Christophe Caron (2014). Workflow4Metabolomics: A collaborative research infrastructure for computational metabolomics. Bioinformatics doi:10.1093/bioinformatics/btu813". Below this, a "Latest news" section lists updates from 2015, including the W4M 2.0 release and a French publication. The left sidebar contains a "Tools" section with a search bar and a list of tool categories: LC-MS, NMR, and COMMON TOOLS. The right sidebar shows a "History" section with a search bar and a message indicating that the history is empty. The bottom of the page features a navigation bar with icons for LC/MS, MS, and Common, and a footer with the text "Active 4" and a timestamp of 10:23 AM 6/5/2015.

Galaxy / 4 / Metabolomics

galaxy.workflow4metabolomics.org

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 2.5 GB

Tools

search tools

Download from URL or upload files from disk

Welcome to workflow4metabolomics.org v2.0

Publication: Franck Giacomoni, Gildas Le Corguillé, Mishari Monsoor, Marion Landi, Pierre Pericard, Mélanie Pétéra, Christophe Duperier, Marie Tremblay-Franco, Jean-François Martin, Daniel Jacob, Sophie Goulitquer, Etienne A. Thévenot and Christophe Caron (2014). **Workflow4Metabolomics: A collaborative research infrastructure for computational metabolomics**. Bioinformatics doi:10.1093/bioinformatics/btu813

Help and support: support@workflow4metabolomics.org

Latest news

01/06/2015 - **Workflow4Metabolomics v2.0 starts today** - Check the changelog section below

01/06/2015 - The W4M 2.0 release is presented in the June 2015 MetaboNews Spotlight [\[link\]](#)

21/09/2015 - [Ecole-chercheurs](#) : Traitement des données métabolomiques sur l'infrastructure online Workflow4Metabolomics (21-25 Sept. 2015) [in French] / Roscoff, France

19/12/2014 - W4M publication in Bioinformatics is now [available](#) - **Workflow4Metabolomics: A collaborative research infrastructure for computational metabolomics**

Changelog

Tutorials

Past events

LC/MS

MS

Common

History

search datasets

Unnamed history

0 bytes

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

javascript:void(0)

Active 4

10:23 AM 6/5/2015

Data import

The screenshot shows a web browser window with the URL `galaxy.workflow4metabolomics.org`. The page title is "Galaxy / 4 / Metabolomics". The navigation bar includes links for "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". The left sidebar lists various tools and categories: "Tools", "Upload File from your computer", "Export Data", "LC-MS", "Preprocessing", "Normalisation", "Statistical Analysis", "Annotation", "GC-MS", "Preprocessing", "Normalisation", "Quality Control", "Statistical Analysis", "Annotation", "NMR", "Preprocessing", "Normalisation", "Quality Control", "Statistical Analysis", "COMMON TOOLS", "Text Manipulation", and "Filter and Sort". A modal dialog box is open in the center, titled "Download data directly from web or upload files from your disk". It features a large dashed rectangular box for file upload, with the text "You can Drag & Drop files into this box." below it. At the bottom of the dialog, there are buttons for "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close". The background shows a workflow diagram with various processing steps.

Galaxy / 4 / Metabolomics

galaxy.workflow4metabolomics.org

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

Upload File from your computer

Export Data

LC-MS

Preprocessing

Normalisation

Statistical Analysis

Annotation

GC-MS

Preprocessing

Normalisation

Quality Control

Statistical Analysis

Annotation

NMR

Preprocessing

Normalisation

Quality Control

Statistical Analysis

COMMON TOOLS

Text Manipulation

Filter and Sort

Download data directly from web or upload files from your disk

You can Drag & Drop files into this box.

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

history

search datasets

unnamed history

bytes

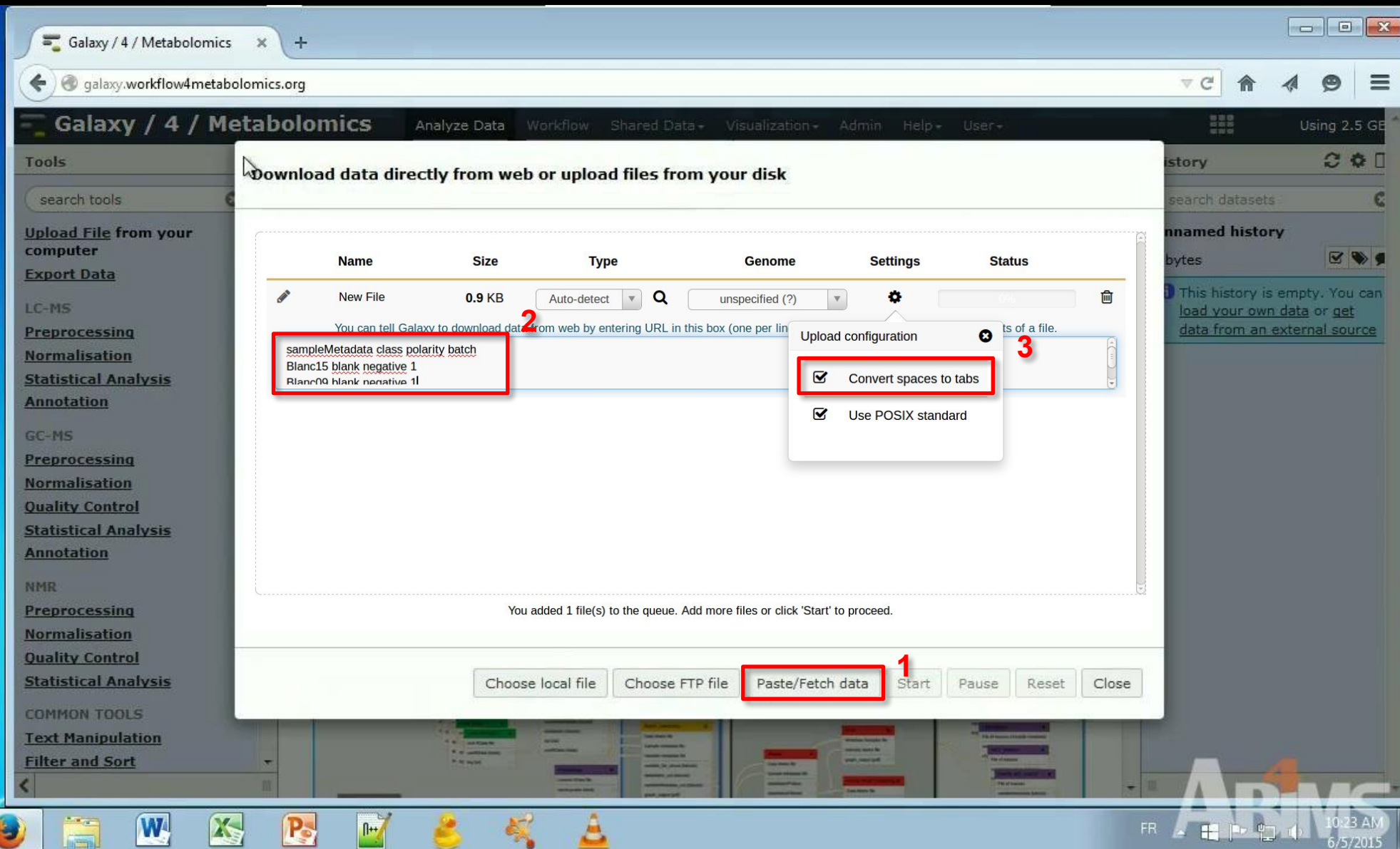
This history is empty. You can [load your own data](#) or [get data from an external source](#)

ABMS

FR 10:23 AM 6/5/2015

Data import

Copy / Paste data



The screenshot displays the Galaxy 4 Metabolomics web interface. The main window shows the 'Download data directly from web or upload files from your disk' dialog. The 'Name' column lists 'New File' with a size of '0.9 KB'. The 'Type' is 'Auto-detect'. The 'Genome' is 'unspecified (?)'. The 'Settings' column shows 'Upload configuration' with 'Convert spaces to tabs' and 'Use POSIX standard' checked. The 'Status' column shows '0.0 KB'.

The data input text area contains the following text:

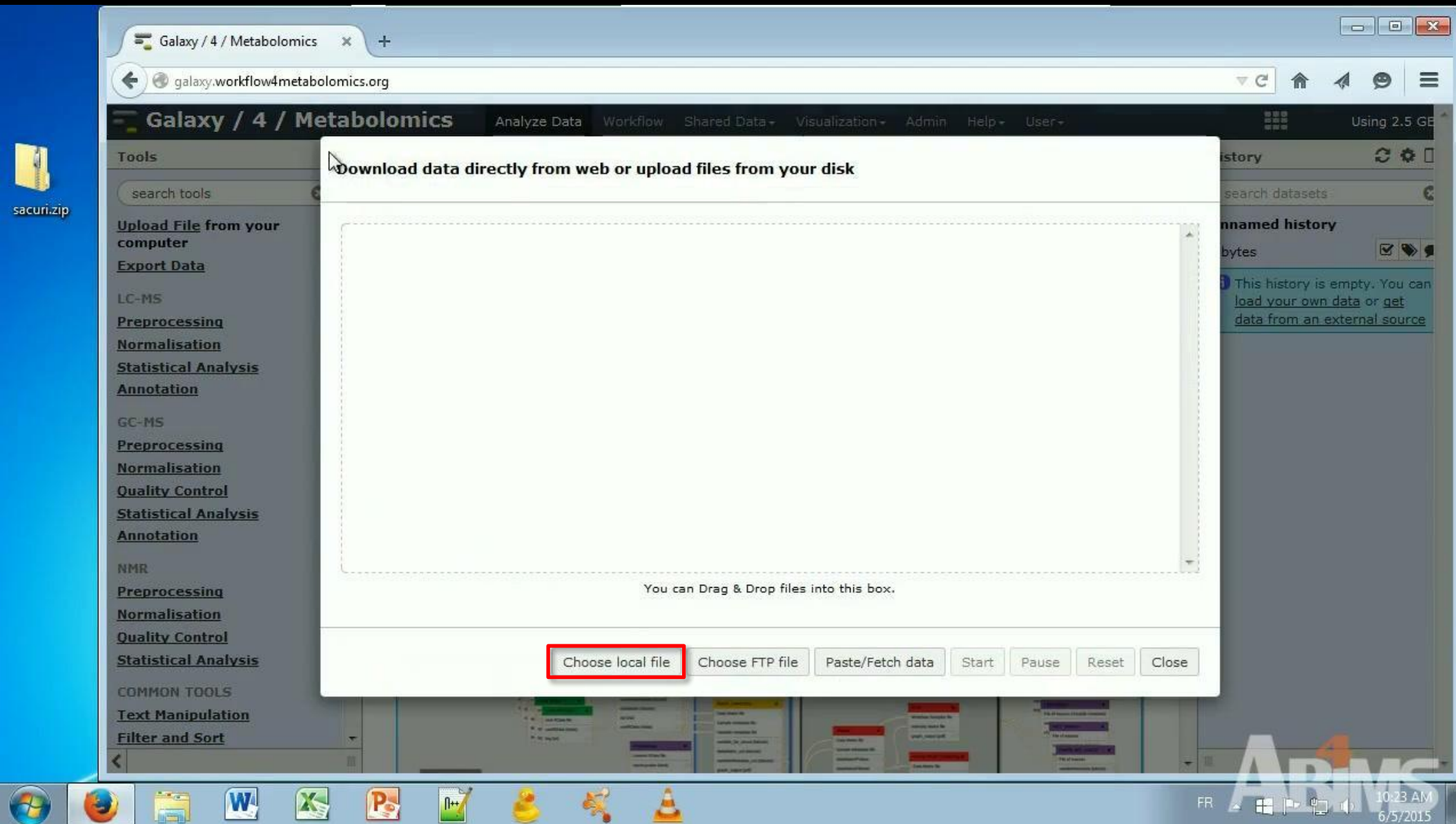
```
sampleMetadata class polarity batch  
Blanc15 blank negative 1  
Blanc09 blank negative 1l
```

The 'Paste/Fetch data' button is highlighted with a red box and the number 1. The data input text area is highlighted with a red box and the number 2. The 'Convert spaces to tabs' checkbox is highlighted with a red box and the number 3.

The bottom taskbar shows the Windows taskbar with various application icons, including Internet Explorer, Word, Excel, PowerPoint, and a yellow duck icon. The system clock shows 10:23 AM on 6/5/2015.

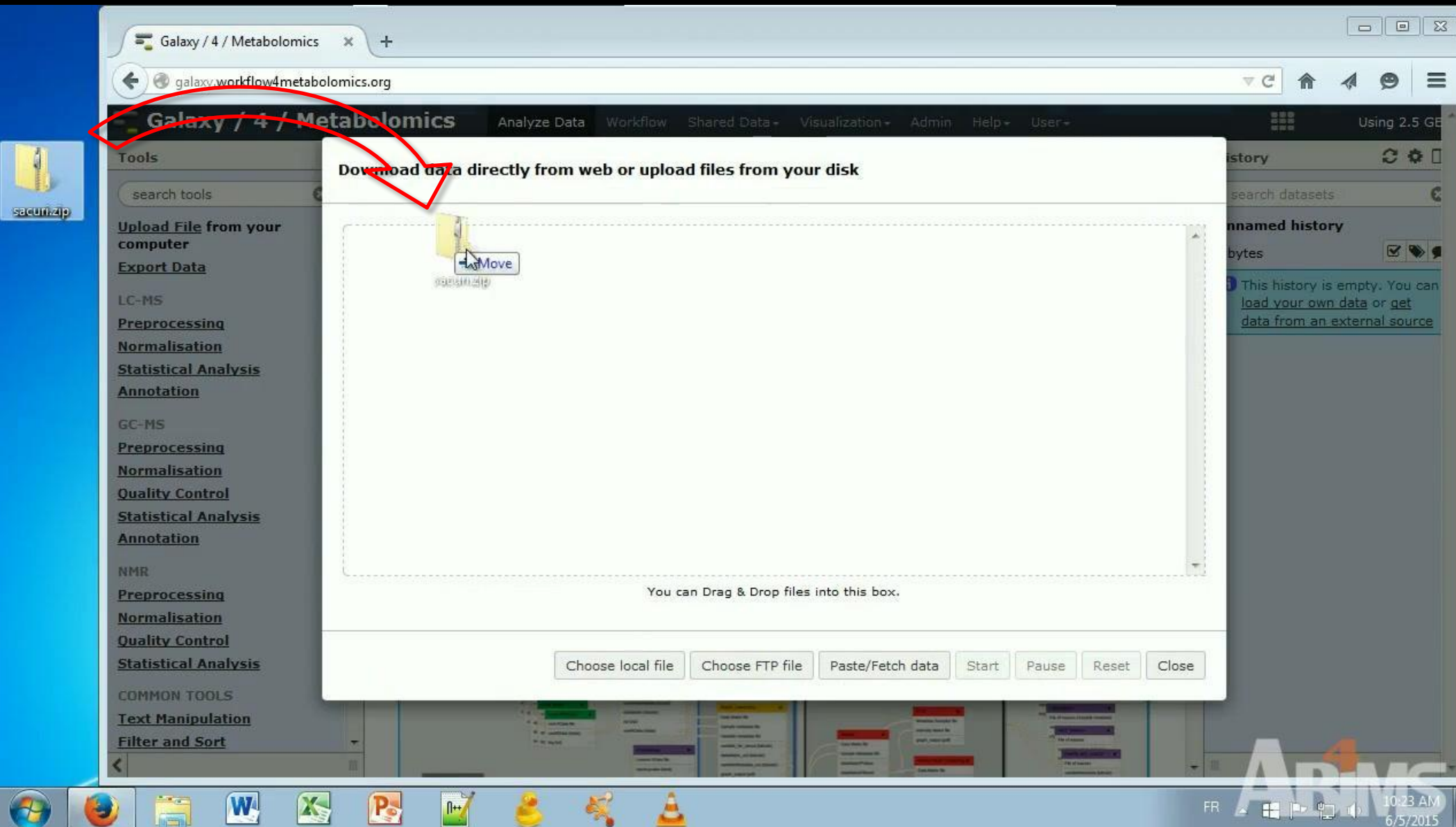
Data import

From local files



Data import

From local files



Data import

From local files

The screenshot shows the Galaxy 4 Metabolomics web interface. A modal dialog titled "Download data directly from web or upload files from your disk" is open. The dialog contains a table with the following columns: Name, Size, Type, Genome, Settings, and Status. A single file, "sacuri.zip", is listed with a size of "0.2 GB" and a type of "Auto-det...". Below the table, a message states: "You added 1 file(s) to the queue. Add more files or click 'Start' to proceed." At the bottom of the dialog, there are buttons for "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close". The "Start" button is highlighted with a mouse cursor. The background shows the Galaxy 4 Metabolomics interface with a sidebar containing links to various tools and a main area displaying a workflow.

Galaxy / 4 / Metabolomics

galaxy.workflow4metabolomics.org

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

Upload File from your computer

Export Data

LC-MS

Preprocessing

Normalisation

Statistical Analysis

Annotation

GC-MS

Preprocessing

Normalisation

Quality Control

Statistical Analysis

Annotation

NMR

Preprocessing

Normalisation

Quality Control

Statistical Analysis

COMMON TOOLS

Text Manipulation

Filter and Sort

Download data directly from web or upload files from your disk

Name	Size	Type	Genome	Settings	Status
sacuri.zip	0.2 GB	Auto-det...	unspecified (?)		

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

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

ABMS

FR 10:23 AM 6/5/2015

Data import

From local files

The screenshot shows the Galaxy 4 Metabolomics web interface. A modal dialog titled "Download data directly from web or upload files from your disk" is open. The dialog contains a table with the following columns: Name, Size, Type, Genome, Settings, and Status. A single file, "sacuri.zip", is listed with a size of "0.2 GB". The "Type" column for this file has a dropdown menu set to "Auto-det..." which is highlighted with a red rectangle. Below the table, a message states: "You added 1 file(s) to the queue. Add more files or click 'Start' to proceed." At the bottom of the dialog, there are several buttons: "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close". The "Start" button is being clicked by a mouse cursor. The background of the interface shows the Galaxy 4 Metabolomics navigation menu on the left and a workflow editor on the right. The browser's address bar shows "galaxy.workflow4metabolomics.org". The Windows taskbar at the bottom displays various application icons and the system clock showing "10:23 AM 6/5/2015".

Galaxy / 4 / Metabolomics

galaxy.workflow4metabolomics.org

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

Upload File from your computer

Export Data

LC-MS

Preprocessing

Normalisation

Statistical Analysis

Annotation

GC-MS

Preprocessing

Normalisation

Quality Control

Statistical Analysis

Annotation

NMR

Preprocessing

Normalisation

Quality Control

Statistical Analysis

COMMON TOOLS

Text Manipulation

Filter and Sort

Download data directly from web or upload files from your disk

Name	Size	Type	Genome	Settings	Status
sacuri.zip	0.2 GB	Auto-det...	unspecified (?)		

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

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

10:23 AM 6/5/2015

Data import

From local files

The screenshot shows the Galaxy 4 Metabolomics web interface. A modal dialog titled "Download data directly from web or upload files from your disk" is open. The dialog contains a table with the following columns: Name, Size, Type, Genome, Settings, and Status. A single file, "sacuri.zip", is listed with a size of 0.2 GB and a status of 50%. Below the table, a progress bar indicates "Please wait...1 out of 1 remaining." At the bottom of the dialog are buttons for "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close". The "Start" button is highlighted with a mouse cursor. The background shows the Galaxy 4 Metabolomics interface with a sidebar containing links to various tools and a main area displaying a workflow.

Galaxy / 4 / Metabolomics

galaxy.workflow4metabolomics.org

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

Upload File from your computer

Export Data

LC-MS

Preprocessing

Normalisation

Statistical Analysis

Annotation

GC-MS

Preprocessing

Normalisation

Quality Control

Statistical Analysis

Annotation

NMR

Preprocessing

Normalisation

Quality Control

Statistical Analysis

COMMON TOOLS

Text Manipulation

Filter and Sort

Download data directly from web or upload files from your disk

Name	Size	Type	Genome	Settings	Status
sacuri.zip	0.2 GB	Auto-det...	unspecified (?)		50

Please wait...1 out of 1 remaining.

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

history

search datasets

unnamed history

bytes

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

ABMS

FR 10:23 AM 6/5/2015

Data import

From local files

The screenshot shows the Galaxy 4 Metabolomics web interface. A modal dialog titled "Download data directly from web or upload files from your disk" is open. The dialog contains a table with the following data:

Name	Size	Type	Genome	Settings	Status
sacuri.zip	0.2 GB	Auto-det...	unspecified (?)		100% ✓

Below the table, it says "You can Drag & Drop files into this box." At the bottom of the dialog are buttons: "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close". The background shows the Galaxy 4 Metabolomics interface with a sidebar containing links like "Upload File from your computer", "Export Data", "LC-MS", "Preprocessing", "Normalisation", "Statistical Analysis", "Annotation", "GC-MS", "NMR", "COMMON TOOLS", "Text Manipulation", and "Filter and Sort". The desktop background is blue with a "sacuri.zip" icon. The taskbar at the bottom shows various application icons and the system clock indicating 10:23 AM on 6/5/2015.

Data import

From local files

The screenshot displays the Galaxy 4 / Metabolomics web interface. The browser address bar shows `galaxy.workflow4metabolomics.org`. The main navigation bar includes links for **Analyze Data**, **Workflow**, **Shared Data**, **Visualization**, **Admin**, **Help**, and **User**. The left sidebar contains a **Tools** section with a search bar and categories for **LC-MS**, **GC-MS**, **NMR**, and **COMMON TOOLS**. The **LC-MS** category is expanded, showing sub-options like **Preprocessing**, **Normalisation**, **Statistical Analysis**, and **Annotation**. The **COMMON TOOLS** category also shows **Text Manipulation** and **Filter and Sort**. The main content area features a green banner for **Welcome to workflow4metabolomics.org v2.0**, a blue box for **Latest news** with dates and links, and a section for **Changelog**, **Tutorials**, and **Past events**. At the bottom, there are three panels for **LC/MS**, **MS**, and **Common** workflows. The right sidebar shows the **History** panel with a search bar and a list of datasets. The first dataset is **1: sacuri.zip**, which is highlighted by a mouse cursor. The status bar at the bottom right indicates **Using 2.5 GB** and the system clock shows **10:23 AM 6/5/2015**.

Data import

From local files

The screenshot displays the Galaxy 4 / Metabolomics web interface in a browser window. The address bar shows the URL galaxy.workflow4metabolomics.org. The main navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User. The interface is divided into three main sections:

- Left Panel (Tools):** Contains a search bar and a list of tool categories: LC-MS, NMR, and COMMON TOOLS. Under LC-MS, there are links for Preprocessing, Normalisation, Statistical Analysis, and Annotation. Under NMR, there are links for Preprocessing, Normalisation, Quality Control, Statistical Analysis, and Annotation. Under COMMON TOOLS, there are links for Text Manipulation and Filter and Sort.
- Center Panel:** Features a large green banner with a checkmark icon and the text "Welcome to workflow4metabolomics.org v2.0". Below this is a publication notice: "Publication: Franck Giacomoni, Gildas Le Corguillé, Mishari Monsoor, Marion Landi, Pierre Pericard, Mélanie Pétéra, Christophe Duperier, Marie Tremblay-Franco, Jean-François Martin, Daniel Jacob, Sophie Goulitquer, Etienne A. Thévenot and Christophe Caron (2014). Workflow4Metabolomics: A collaborative research infrastructure for computational metabolomics. Bioinformatics doi:10.1093/bioinformatics/btu813". Below the banner is a "Latest news" section with three items: "01/06/2015 - Workflow4Metabolomics v2.0 starts today", "21/09/2015 - The W4M 2.0 release is presented in the June 2015 MetaboNews Spotlight", and "19/12/2014 - W4M publication in Bioinformatics is now available". At the bottom of the center panel are links for Changelog, Tutorials, and Past events.
- Right Panel (History):** Shows a "History" section with a search bar and a list of datasets. The first entry is "1: sacuri.zip" with a status of "This job is currently running".

The bottom of the image shows a Windows taskbar with various application icons, including Internet Explorer, Word, Excel, PowerPoint, and a file explorer. The system clock in the bottom right corner indicates the date and time: 10:23 AM, 6/5/2015.

Step 1: Choose a FTP Client

DATA IMPORT USING FTP

STEP 1: CHOOSE A FTP CLIENT



STEP 1: CHOOSE A FTP CLIENT



Avoid:
Malwares inside



Cyberduck



WinSCP

STEP 1: CHOOSE A FTP CLIENT



Step 2: Easy!

DATA IMPORT USING FTP

Data import using FTP

The screenshot displays the Galaxy / ABiMS web interface in a browser window. The address bar shows the URL `galaxy.sb-roscoff.fr`. The page features a dark navigation bar with tabs for **Analyze Data**, **Workflow**, **Shared Data**, **Visualization**, **Admin**, **Help**, and **User**. A status indicator on the right shows **Using 46%**.

On the left, a sidebar lists various tool categories under the heading **Tools**:

- Get Data**
- COMMON TOOLS**
- Convert Formats**
- FASTA manipulation**
- Filter and Sort**
- Join, Subtract and Group**
- Text Manipulation**
- Graphics**
- Statistics**
- EMBOSS 5 Suite**
- SEARCHING TOOLS**
- NCBI BLAST+**
- Diamond**
- Primer/Microsatellite**
- NGS TOOLS**
- NGS: BedTools**
- NGS: Mapping**
- NGS: Picard**
- NGS: QC and manipulation**

A tooltip is visible over the upload icon, stating: **Download from URL or upload files from disk**.

The main content area has a green header: **Welcome to galaxy.sb-roscoff.fr**. Below this is an **Information** box with a blue 'i' icon, containing the text: **Information** For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr.

The center of the page features the **ABiMS** logo (with a large orange '4' above the 'i') and the text **Analyses and Bioinformatics for Marine Science**. To the right of the logo is the logo for **CNRS UPRC Station Biologique Roscoff**.

Below the logo, there are two expandable sections: **Changelog** and **Tutorials**.

A paragraph of text describes the platform: **Galaxy** is an open, web-based platform for data intensive biomedical research. The **Galaxy team** is a part of **BX** at Penn State, and the **Biology** and **Mathematics and Computer Science** departments at Emory University. The **Galaxy Project** is supported in part by **NHGRI**, **NSF**, **The Huck Institutes of the Life Sciences**, **The Institute for CyberScience** at Penn State, and **Emory University**.

On the right side, there is a **History** panel with a search bar for **search datasets**. It shows **Unnamed history** with **0 bytes**. A message box states: **This history is empty. You can load your own data or get data from an external source**.

The Windows taskbar at the bottom shows various application icons, including Internet Explorer, Google Chrome, and several office applications. The system clock in the bottom right corner displays **11:19 AM 7/31/2015**.

Data import using FTP

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The navigation bar includes links for "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". A modal dialog box is open in the center, titled "Download data directly from web or upload files from your disk". Inside the dialog, there is a large dashed box for file upload with the text "You can Drag & Drop files into this box." Below the box, there are two dropdown menus: "Type (set all):" with "Auto-detect" selected, and "Genome (set all):" with "unspecified (?)" selected. At the bottom of the dialog, there are buttons for "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close". The background of the web page shows a sidebar with "Tools" and "Get Data" sections, and a main content area with a message about an empty history.

Galaxy / ABiMS

galaxy.sb-roscoff.fr

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

Get Data

COMMON TOOLS

[Convert Formats](#)

[FASTA manipulation](#)

[Filter and Sort](#)

[Join, Subtract and C](#)

[Text Manipulation](#)

[Graphics](#)

[Statistics](#)

[EMBOSS 5 Suite](#)

SEARCHING TOOLS

[NCBI BLAST+](#)

[Diamond](#)

[Primer/Microsatelli](#)

NGS TOOLS

[NGS: BedTools](#)

[NGS: Mapping](#)

[NGS: Picard](#)

[NGS: QC and manip](#)

Download data directly from web or upload files from your disk

You can Drag & Drop files into this box.

Type (set all): Auto-detect

Genome (set all): unspecified (?)

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

history

history is empty. You can
own data or get
an external source

FR 11:19 AM 7/31/2015

Data import using FTP

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The main navigation bar includes links for "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". A sidebar on the left lists various tools under categories like "Tools", "Get Data", "COMMON TOOLS", "SEARCHING TOOLS", and "NGS TOOLS".

A modal dialog box is open in the center of the screen with the title "Download data directly from web or upload files from your disk". Inside the dialog, there is a large dashed box for file upload with the text "You can Drag & Drop files into this box." Below this box, there are two dropdown menus: "Type (set all):" with "Auto-detect" selected, and "Genome (set all):" with "unspecified (?)" selected. At the bottom of the dialog, there are several buttons: "Choose local file", "Choose FTP file" (which is being clicked by a mouse cursor), "Paste/Fetch data", "Start", "Pause", "Reset", and "Close".

The desktop background is blue with icons for "lecorguille" and "left_kept_r...". The taskbar at the bottom shows various application icons including Internet Explorer, Google Chrome, Word, Excel, PowerPoint, and a clock showing 11:19 AM on 7/31/2015.

Data import using FTP

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The main content area is titled "Download data directly from web or upload files from your disk". Below this title, there is a dashed box indicating where files can be dragged and dropped. A modal dialog box titled "FTP files" is open, displaying the following text: "This Galaxy server allows you to upload files via FTP. To upload some files, log in to the FTP server at **galaxy.sb-roscoff.fr** using your Galaxy credentials (email address and password)." Below the text, there is an orange warning box that says "Your FTP directory does not contain any files." At the bottom of the modal, there are several buttons: "Choose local file", "Choose FTP file" (which is highlighted by a mouse cursor), "Paste/Fetch data", "Start", "Pause", "Reset", and "Close". The background of the web interface shows a sidebar with various tools categorized under "Tools", "COMMON TOOLS", "SEARCHING TOOLS", and "NGS TOOLS". The desktop background is blue with icons for "lecorguille" and "left_kept_r...". The taskbar at the bottom shows various application icons and the system clock indicating 11:19 AM on 7/31/2015.

Galaxy / ABiMS

galaxy.sb-roscoff.fr

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

Get Data

COMMON TOOLS

Convert Formats

FASTA manipulation

Filter and Sort

Join, Subtract and

Text Manipulation

Graphics

Statistics

EMBOSS 5 Suite

SEARCHING TOOLS

NCBI BLAST+

Diamond

Primer/Microsatelli

NGS TOOLS

NGS: BedTools

NGS: Mapping

NGS: Picard

NGS: QC and manip

Download data directly from web or upload files from your disk

You can Drag & Drop files into this box.

FTP files

This Galaxy server allows you to upload files via FTP. To upload some files, log in to the FTP server at **galaxy.sb-roscoff.fr** using your Galaxy credentials (email address and password).

Your FTP directory does not contain any files.

Type (set a

specified (?)

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

11:19 AM 7/31/2015

Data import using FTP

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The main navigation bar includes links for "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". A sidebar on the left lists various tools under categories like "Tools", "Get Data", "COMMON TOOLS", "SEARCHING TOOLS", "NGS TOOLS", and "NGS: QC and manipulation".

A large modal window is open in the center with the title "Download data directly from web or upload files from your disk". It contains a dashed box for file upload and a smaller "FTP files" dialog box. The "FTP files" dialog box has the following text:

FTP files

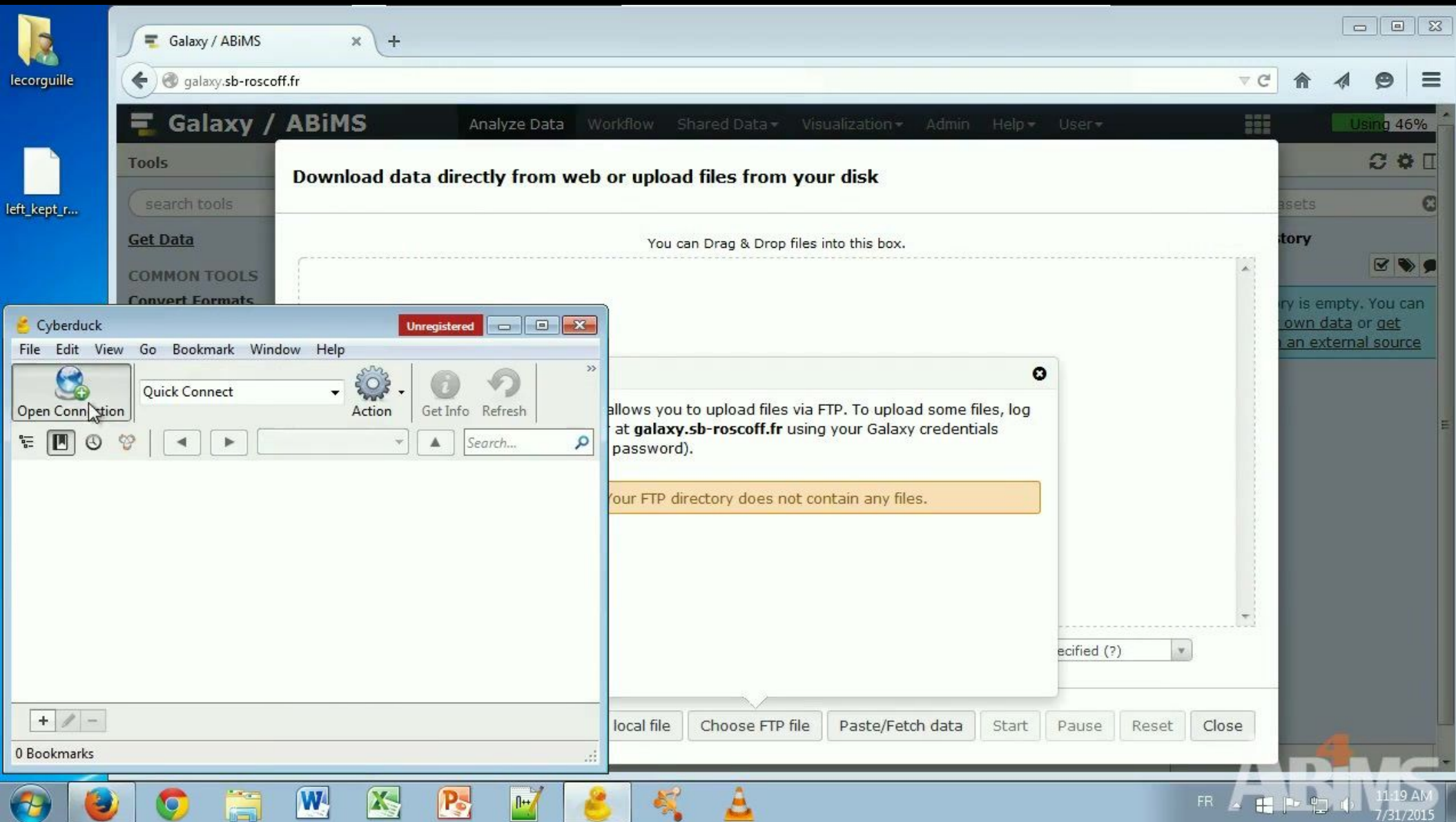
This Galaxy server allows you to upload files via FTP. To upload some files, log in to the FTP server at **galaxy.sb-roscoff.fr** using your Galaxy credentials (email address and password).

Your FTP directory does not contain any files.

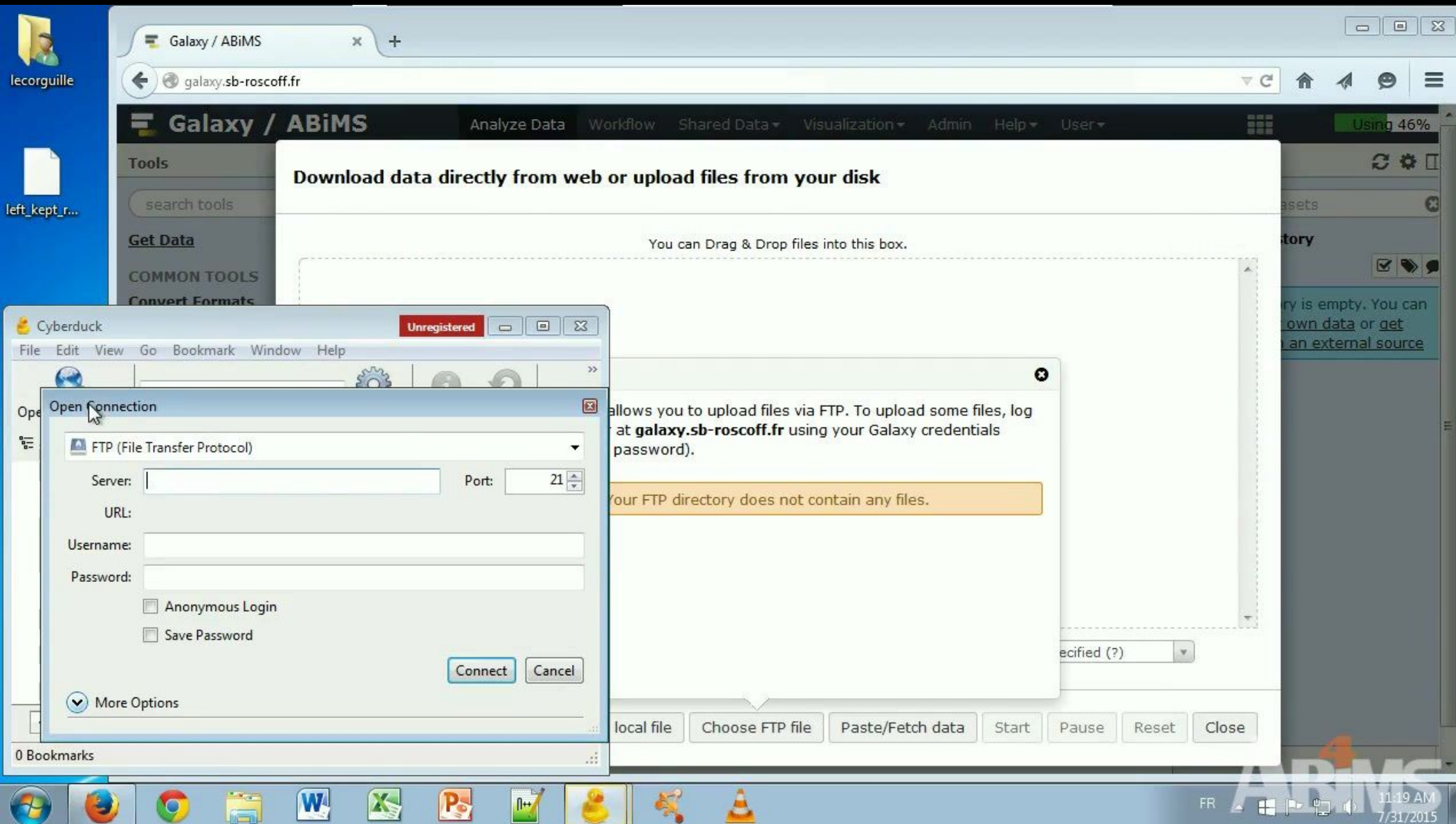
At the bottom of the "FTP files" dialog box, there are buttons: "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close".

The desktop background is blue with icons for "lecorguille", "left_kept_r...", and a taskbar at the bottom with various application icons. The system clock in the bottom right corner shows "11:19 AM 7/31/2015".

Data import using FTP



Data import using FTP



Data import using FTP

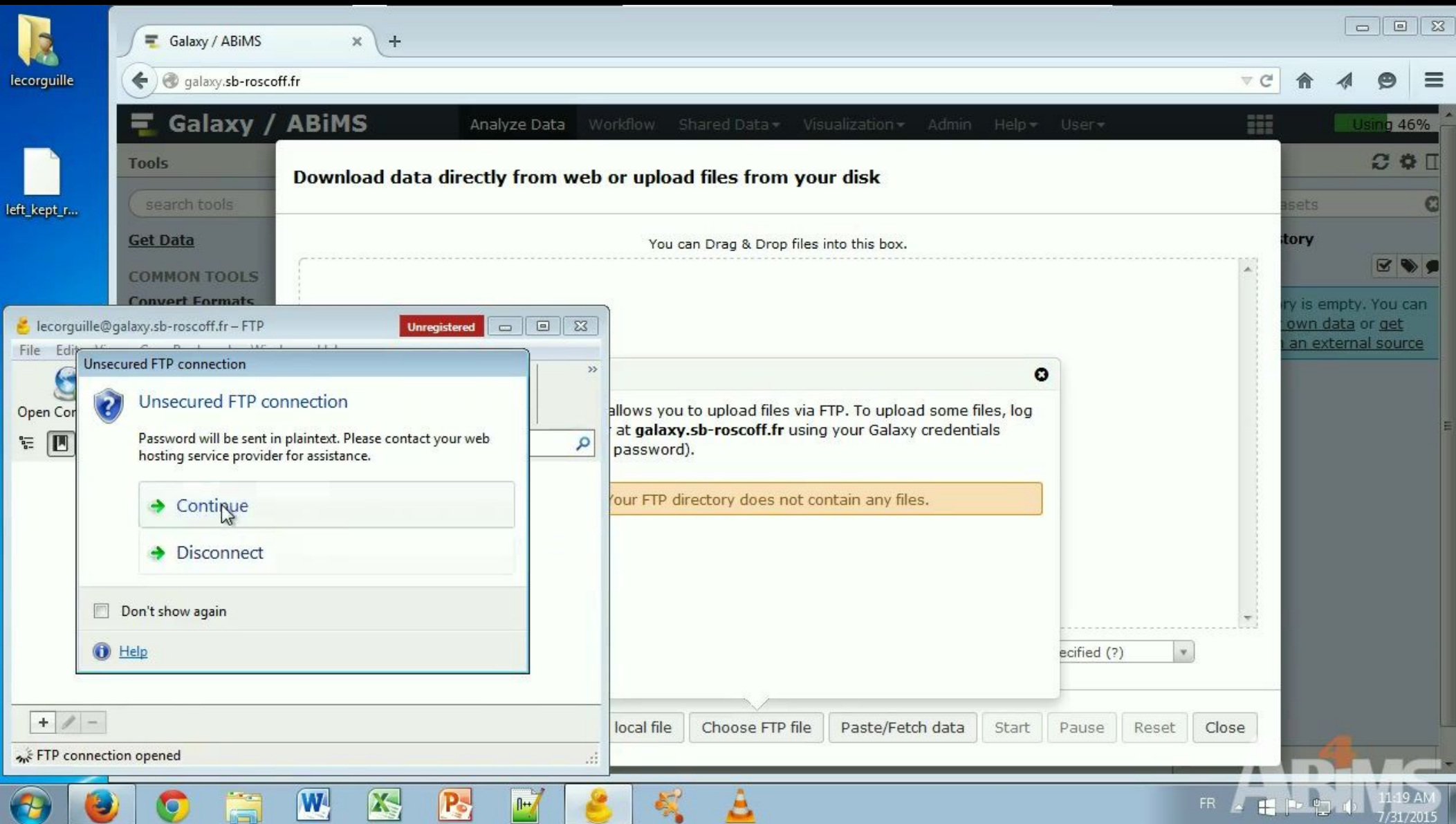
The screenshot shows a desktop environment with a Windows taskbar at the bottom. The taskbar includes icons for Internet Explorer, Google Chrome, File Explorer, Microsoft Word, Microsoft Excel, Microsoft PowerPoint, a calculator, a rubber duck, a traffic cone, and system icons for network, volume, and date/time (11:19 AM 7/31/2015).

The main application window is a web browser displaying the Galaxy / ABiMS interface. The address bar shows `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The navigation menu includes "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". The main content area has a header "Download data directly from web or upload files from your disk" and a sub-header "You can Drag & Drop files into this box." Below this is a large dashed box for file upload. A sidebar on the left contains "Tools", "Get Data", and "COMMON TOOLS". A status bar at the bottom right shows "Using 46%".

Overlaid on the browser is the Cyberduck application window. The title bar says "Cyberduck" and "Unregistered". The menu bar includes "File", "Edit", "View", "Go", "Bookmark", "Window", and "Help". The main area shows an "Open Connection" dialog box. The dialog has a dropdown menu set to "FTP (File Transfer Protocol)". The "Server" field contains `galaxy.sb-roscoff.fr` and the "Port" field contains `21`. The "URL" field contains `ftp://lecorguille@galaxy.sb-roscoff.fr:21/`. The "Username" field contains `lecorguille` and the "Password" field contains a series of dots. There are checkboxes for "Anonymous Login" and "Save Password", both of which are unchecked. At the bottom of the dialog are "Connect" and "Cancel" buttons. A "More Options" link is also visible.

Below the "Open Connection" dialog, a smaller dialog box is visible. It contains the text: "allows you to upload files via FTP. To upload some files, log at **galaxy.sb-roscoff.fr** using your Galaxy credentials (password)." Below this text is a yellow box with the message: "your FTP directory does not contain any files." At the bottom of this dialog are buttons for "local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close".

Data import using FTP



Data import using FTP

The screenshot shows a desktop environment with a Windows taskbar at the bottom. The taskbar includes icons for Internet Explorer, Google Chrome, File Explorer, Microsoft Word, Excel, PowerPoint, a calculator, a rubber duck, a traffic cone, and system tray icons for network, volume, and date/time (11:19 AM 7/31/2015).

The main application window is a web browser displaying the Galaxy / ABiMS interface at galaxy.sb-roscoff.fr. The browser's address bar shows the URL. The page has a dark header with navigation links: Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User. A sidebar on the left lists 'Tools' and 'COMMON TOOLS' with a search bar. The main content area features a large box with the text 'Download data directly from web or upload files from your disk' and 'You can Drag & Drop files into this box.' Below this is a message: 'allows you to upload files via FTP. To upload some files, log at galaxy.sb-roscoff.fr using your Galaxy credentials (username and password)'. An orange banner states 'your FTP directory does not contain any files.' At the bottom of the main area are buttons: 'local file', 'Choose FTP file', 'Paste/Fetch data', 'Start', 'Pause', 'Reset', and 'Cancel'.

Overlaid on the bottom left is an FTP client window titled 'lecorguille@galaxy.sb-roscoff.fr - FTP'. It has a menu bar (File, Edit, View, Go, Bookmark, Window, Help) and a toolbar with 'Open Connection', 'Quick Connect', 'Action', 'Get Info', and 'Refresh'. Below the toolbar is a search bar and a table with columns 'Filename', 'Size', and 'Modified'. The table is currently empty, showing '0 Files'. A mouse cursor is hovering over a 'Copy to /' button. The window also displays 'Unregistered' in the top right corner.

In the bottom right corner, a system notification bubble says 'Connection opened galaxy.sb-roscoff.fr'.

Data import using FTP

The screenshot displays a web browser window with the Galaxy / ABiMS interface at galaxy.sb-roscoff.fr. A 'Transfers' window is open, showing the upload of a file named 'left_kept_reads.bam'. The progress bar indicates that 50.6 MiB (53,018,624 bytes) of 91.6 MiB has been uploaded (55% complete) at a rate of 70.1 MB/sec, with 1 second remaining. The local file path is C:\Users\lecorguille\Desktop\left_kept_reads.bam, and the URL is ftp://galaxy.sb-roscoff.fr/left_kept_reads.bam. A notification bubble in the bottom right corner states 'Connection opened galaxy.sb-roscoff.fr'. In the background, a 'from your disk' panel is visible, along with a message about logging in via FTP.

Galaxy / ABiMS

galaxy.sb-roscoff.fr

Transfers

Resume Reload Stop Remove Open Show

left_kept_reads.bam

50.6 MiB (53,018,624 bytes) of 91.6 MiB (55%, 70.1 MB/sec, 1 seconds remaining)
Uploading left_kept_reads.bam

URL: ftp://galaxy.sb-roscoff.fr/left_kept_reads.bam
Local File: C:\Users\lecorguille\Desktop\left_kept_reads.bam

Connection opened
galaxy.sb-roscoff.fr

from your disk

Drop files into this box.

Files via FTP. To upload some files, log in via FTP using your Galaxy credentials

does not contain any files.

FTP file Paste/Fetch data Start Pause Reset

Data import using FTP

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The main content area displays a message: "Download data directly from web or upload files from your disk". Below this, there is a large dashed box with the text "You can Drag & Drop files into this box." and a smaller text "allows you to upload files via FTP. To upload some files, log at **galaxy.sb-roscoff.fr** using your Galaxy credentials (username and password)." A yellow banner below the dashed box states "Your FTP directory does not contain any files." At the bottom of the page, there are buttons: "local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Cancel".

Overlaid on the browser window is a file manager window titled "lecorguille@galaxy.sb-roscoff.fr - FTP". The window shows a list of files with columns "Filename", "Size", and "Modified". The file "left_kept_reads.bam" is listed with a size of "91.6 MiB" and a modification date of "7/31/2015 9:19:00 AM". The status bar at the bottom of the file manager indicates "1 Files".

In the bottom right corner, a notification bubble says "Upload complete" and "left_kept_reads.bam".

The desktop background is blue with icons for "lecorguille" and "left_kept_r...". The taskbar at the bottom shows various application icons including Internet Explorer, Google Chrome, File Explorer, Word, Excel, PowerPoint, and a yellow duck icon.

Data import using FTP

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The main navigation bar includes links for "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". A sidebar on the left lists various tools under categories like "Tools", "Get Data", "COMMON TOOLS", "SEARCHING TOOLS", and "NGS TOOLS".

A modal dialog titled "Download data directly from web or upload files from your disk" is open. It contains a large dashed box for file upload. Below this box, there is a section for "FTP files" with the following text:

This Galaxy server allows you to upload files via FTP. To upload some files, log in to the FTP server at **galaxy.sb-roscoff.fr** using your Galaxy credentials (email address and password).

An orange warning box below the text states: "Your FTP directory does not contain any files."

At the bottom of the modal, there are several buttons: "Choose local file", "Choose FTP file" (which is highlighted by a mouse cursor), "Paste/Fetch data", "Start", "Pause", "Reset", and "Close".

In the bottom right corner of the browser window, there is a small notification that says "Upload complete" with a close button.

Data import using FTP

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The main navigation bar includes links for "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". A sidebar on the left lists various tools under categories like "Get Data", "COMMON TOOLS", "SEARCHING TOOLS", and "NGS TOOLS". A modal window is open in the center with the title "Download data directly from web or upload files from your disk". Inside the modal, there is a large dashed box for drag-and-drop uploads, with the text "You can Drag & Drop files into this box." above it. Below the box, there are two dropdown menus: "Type (set all):" with "Auto-detect" selected, and "Genome (set all):" with "unspecified (?)" selected. At the bottom of the modal, there are buttons for "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close". The Windows taskbar at the bottom shows several application icons, including Internet Explorer, Google Chrome, and various office applications. The system clock in the bottom right corner indicates the date and time as 11:19 AM on 7/31/2015.

Galaxy / ABiMS

galaxy.sb-roscoff.fr

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

Get Data

COMMON TOOLS

Convert Formats

FASTA manipulation

Filter and Sort

Join, Subtract and C

Text Manipulation

Graphics

Statistics

EMBOSS 5 Suite

SEARCHING TOOLS

NCBI BLAST+

Diamond

Primer/Microsatelli

NGS TOOLS

NGS: BedTools

NGS: Mapping

NGS: Picard

NGS: QC and manip

Download data directly from web or upload files from your disk

You can Drag & Drop files into this box.

Type (set all): Auto-detect

Genome (set all): unspecified (?)

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

Upload complete

11:19 AM 7/31/2015

Data import using FTP

The screenshot shows the Galaxy / ABiMS web interface in a browser window. The address bar displays `galaxy.sb-roscoff.fr`. The main navigation bar includes links for **Analyze Data**, **Workflow**, **Shared Data**, **Visualization**, **Admin**, **Help**, and **User**. A sidebar on the left lists various tools under categories like **Get Data**, **COMMON TOOLS**, **SEARCHING TOOLS**, and **NGS TOOLS**.

A central dialog box titled "Download data directly from web or upload files from your disk" is active. It contains a dashed box for file uploads and a smaller "FTP files" sub-dialog. The "FTP files" dialog explains that the server allows file uploads via FTP and provides a table of available files.

FTP files

This Galaxy server allows you to upload files via FTP. To upload some files, log in to the FTP server at **galaxy.sb-roscoff.fr** using your Galaxy credentials (email address and password).

Available files: 1 files 96 MB

☐	Name	Size	Created
☒	left_kept_reads.bam	96 MB	07/31/2015 11:19:45 AM

Buttons at the bottom of the dialog include: **Choose local file**, **Choose FTP file** (highlighted), **Paste/Fetch data**, **Start**, **Pause**, **Reset**, and **Close**.

Data import using FTP

The screenshot shows the Galaxy / ABiMS web interface in a browser window. The address bar shows `galaxy.sb-roscoff.fr`. The main navigation bar includes links for **Analyze Data**, **Workflow**, **Shared Data**, **Visualization**, **Admin**, **Help**, and **User**. A sidebar on the left lists various tools under categories like **Get Data**, **COMMON TOOLS**, **SEARCHING TOOLS**, and **NGS TOOLS**.

A modal window titled "Download data directly from web or upload files from your disk" is open. It contains a table with the following data:

Name	Size	Type	Genome	Settings	Status
left_kept_reads.bam	96 MB	Auto-det...	unspecified (?)		

Below the table, a "Type (set a)" dropdown is visible. At the bottom of the modal are buttons: **Choose local file**, **Choose FTP file**, **Paste/Fetch data**, **Start**, **Pause**, **Reset**, and **Close**.

An "FTP files" dialog box is also open, providing instructions on how to upload files via FTP. It lists the available files:

Name	Size	Created
left_kept_reads.bam	96 MB	07/31/2015 11:19:45 AM

The dialog box also includes a checkbox for selecting files and a close button.

Data import using FTP

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The navigation bar includes links for "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". A sidebar on the left lists various tools under categories like "Get Data", "COMMON TOOLS", "SEARCHING TOOLS", and "NGS TOOLS".

A modal dialog box is open in the center, titled "Download data directly from web or upload files from your disk". It contains the text: "You added 1 file(s) to the queue. Add more files or click 'Start' to proceed." Below this is a table with the following data:

Name	Size	Type	Genome	Settings	Status
left_kept_reads.bam	96 MB	Auto-det...	unspecified (?)	⚙️	🗑️

Below the table, there are two sections for "Type (set all):" and "Genome (set all):", each with a dropdown menu and a search icon. At the bottom of the dialog, there are buttons for "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close". A mouse cursor is hovering over the "Start" button.

The Windows taskbar at the bottom shows the time as 11:20 AM on 7/31/2015. The system tray includes icons for network, volume, and power.

Data import using FTP

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The main navigation bar includes links for "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". A sidebar on the left lists various tools under categories like "Tools", "Get Data", "COMMON TOOLS", "SEARCHING TOOLS", and "NGS TOOLS".

A modal dialog box is open in the center, titled "Download data directly from web or upload files from your disk". It contains a table with the following data:

Name	Size	Type	Genome	Settings	Status
left_kept_reads.bam	96 MB	Auto-det...	unspecified (?)		100%

Below the table, there are two dropdown menus: "Type (set all):" with "Auto-detect" selected, and "Genome (set all):" with "unspecified (?)" selected. At the bottom of the dialog, there are buttons for "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close".

Data import using FTP

Galaxy / ABiMS

galaxy.sb-roscoff.fr

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 46%

Tools

search tools

Get Data

COMMON TOOLS

Convert Formats

FASTA manipulation

Filter and Sort

Join, Subtract and Group

Text Manipulation

Graphics

Statistics

EMBOSS 5 Suite

SEARCHING TOOLS

NCBI BLAST+

Diamond

Primer/Microsatellite

NGS TOOLS

NGS: BedTools

NGS: Mapping

NGS: Picard

NGS: QC and manipulation

✓ Welcome to galaxy.sb-roscoff.fr

Information

For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS 4

Analyses and Bioinformatics for Marine Science

Changelog

Tutorials

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History

search datasets

Unnamed history

1 shown

0 bytes

1: left_kept_reads.bam

FR 11:20 AM 7/31/2015

Data import using FTP

Galaxy / ABiMS

galaxy.sb-roscoff.fr

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 46%

Tools

search tools

Get Data

COMMON TOOLS

Convert Formats

FASTA manipulation

Filter and Sort

Join, Subtract and Group

Text Manipulation

Graphics

Statistics

EMBOSS 5 Suite

SEARCHING TOOLS

NCBI BLAST+

Diamond

Primer/Microsatellite

NGS TOOLS

NGS: BedTools

NGS: Mapping

NGS: Picard

NGS: QC and manipulation

✓ Welcome to galaxy.sb-roscoff.fr

Information

For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS 4

Analyses and Bioinformatics for Marine Science

Station Biologique Roscoff

Changelog

Tutorials

Galaxy is an open, web-based platform for data intensive biomedical research. The [Galaxy team](#) is a part of [BX](#) at Penn State, and the [Biology](#) and [Mathematics and Computer Science](#) departments at [Emory University](#). The [Galaxy Project](#) is supported in part by [NHGRI](#), [NSF](#), [The Huck Institutes of the Life Sciences](#), [The Institute for CyberScience at Penn State](#), and [Emory University](#).

History

search datasets

Unnamed history

1 shown

0 bytes

1: left_kept_reads.bam

FR 11:20 AM 7/31/2015

Data import using FTP

The screenshot displays the Galaxy / ABiMS web interface in a browser window. The address bar shows the URL `galaxy.sb-roscoff.fr`. The interface includes a top navigation bar with links for **Analyze Data**, **Workflow**, **Shared Data**, **Visualization**, **Admin**, **Help**, and **User**. A status bar on the right indicates **Using 46%**.

Tools Panel (Left): Contains a search bar and a list of tool categories: **Get Data**, **COMMON TOOLS**, **Convert Formats**, **FASTA manipulation**, **Filter and Sort**, **Join, Subtract and Group**, **Text Manipulation**, **Graphics**, **Statistics**, **EMBOSS 5 Suite**, **SEARCHING TOOLS**, **NCBI BLAST+**, **Diamond**, **Primer/Microsatellite**, **NGS TOOLS**, **NGS: BedTools**, **NGS: Mapping**, **NGS: Picard**, and **NGS: QC and manipulation**.

Main Content Area: Features a green welcome banner for `galaxy.sb-roscoff.fr`. Below it is an information box with contact details for `support.abims@sb-roscoff.fr`. The center displays the **ABiMS** logo and the text **Analyses and Bioinformatics for Marine Science**. A sidebar on the left lists **Changelog** and **Tutorials**. A paragraph at the bottom describes the Galaxy platform and its affiliations.

History Panel (Right): Titled **History**, it shows a search bar and a list of datasets. The first entry is **1: left_kept_reads.bam**, which is 91.6 MB, in **bam** format, and is associated with the database `2`. It includes options to **display in IGB** and **View**, and a description as a **Binary bam alignments file**.

The Windows taskbar at the bottom shows various application icons, including Internet Explorer, Google Chrome, and several office and utility programs. The system clock in the bottom right corner shows the date **7/31/2015** and time **11:20 AM**.

~~DATA IMPORT~~

Data import

For HUGE public resources: genome, databank ...

--> Make a request to the support team

Galaxy / ABiMS Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

[Get Data](#)
[Send Data](#)
[Collection Operations](#)

COMMON TOOLS
[Text Manipulation](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[Convert Formats](#)
[Extract Features](#)
[Fetch Sequences](#)
[Statistics](#)
[Graph/Display Data](#)
[Fasta Fastq Manipulation](#)

COMMON NGS TOOLS
[NGS:Samtools](#)
[NGS:Mapping](#)
[NGS:Bedtools](#)
[NGS:Picard Tools](#)

SEARCHING TOOLS
[Diamond](#)

NCBI BLAST+ blastn Search nucleotide database with nucleotide query sequence(s) (Galaxy Version 0.1.08) Options

Nucleotide query sequence(s)

No fasta dataset available.

Subject database/sequences

Locally installed BLAST database

Nucleotide BLAST database

☐ Select/Unselect all

- nt
- genbank
- genbank Bacterial
- genbank Environmental sampling
- genbank EST (expressed sequence tag)
- genbank GSS (genome survey sequence)
- genbank HTC (high throughput cDNA sequencing)
- genbank HTGS (high throughput genomic sequencing)

Set expectation value cutoff

0.001

Output format

History

search datasets

eba 2016 sartools
42 shown
1.59 MB

- 62: SARTools DESeq2 R objects (.RData)**
- 61: SARTools DESeq2 R log**
- 60: SARTools DESeq2 figures**
- 59: SARTools DESeq2 tables**
- 58: SARTools DESeq2 report**
- 57: SARTools edgeR R objects (.RData)**
- 56: SARTools edgeR R log**
- 55: SARTools edgeR figures**

Hands-on
DATA IMPORT





1. Fetch the files with your internet browser:
<https://tinyurl.com/galaxy-initiation-datasets>
2. Upload files into Galaxy (exons.bed, snps.bed, repeats.bed)
 - a. From disk
 - b. Using FTP

TOOLS

Tools - panel

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 0%

Tools

search tools

Get Data

Send Data

Collection Operations

COMMON TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

Extract Features

Fetch Sequences

Statistics

Graph/Display Data

Fasta Fastq Manipulation

Filter sequences by ID from a tabular file

FastQC Read Quality reports

FASTQ Groomer convert between various FASTQ quality formats

COMMON NGS TOOLS

NGS-Samtools

✓ Welcome to galaxy3.sb-roscoff.fr

i

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr



Analyses and Bioinformatics for Marine Science

Changelog

Tutorials

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History

search datasets

Trinity example

2 shown, 3 deleted

40.02 KB

4: reads.left.fq

3: reads.right.fq

83

Tools - panel

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User, along with a 'Using 0%' status indicator. The left sidebar, titled 'Tools', contains a search bar and a list of tool categories: Get Data, Send Data, Collection Operations, COMMON TOOLS, Text Manipulation, Filter and Sort, Join, Subtract and Group, Convert Formats, Extract Features, Fetch Sequences, Statistics, Graph/Display Data, and Fasta Fastq Manipulation. The main content area features a green 'Welcome to galaxy3.sb-roscoff.fr' message, an information box with contact details, and the ABiMS logo with the tagline 'Analyses and Bioinformatics for Marine Science'. The right sidebar, titled 'History', shows a search bar and a list of datasets, including 'Trinity example' and '4: reads.left.fq'. A large white box with a black border is overlaid on the center of the page, containing the text 'What tools are available?'.

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 0%

Tools

search tools

Get Data

Send Data

Collection Operations

COMMON TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

Extract Features

Fetch Sequences

Statistics

Graph/Display Data

Fasta Fastq Manipulation

Filter sequences by ID from a tabular file

FastQC Read Quality reports

FASTQ Groomer convert between various FASTQ quality formats

COMMON NGS TOOLS

NGS-Samtools

Welcome to galaxy3.sb-roscoff.fr

Information

For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS

Analyses and Bioinformatics for Marine Science

Station Biologique Roscoff

History

search datasets

Trinity example

2 shown, 3 deleted

40.02 KB

4: reads.left.fq

3: reads.right.fq

What tools are available?

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

Tools - panel

125+ platforms for using Galaxy: <https://galaxyproject.org/use>
[Servers, clouds, deployable resources]

125+ platforms for using Galaxy: <https://galaxyproject.org/use>
[Servers, clouds, deployable resources]

GENERAL PURPOSE:

usegalaxy.org, usegalaxy.eu, usegalaxy.fr, usegalaxy.org.au

125+ platforms for using Galaxy: <https://galaxyproject.org/use>
 [Servers, clouds, deployable resources]

GENERAL PURPOSE:

usegalaxy.org, usegalaxy.eu, usegalaxy.fr, usegalaxy.org.au

DOMAIN SPECIFIC:



RNAseq: <http://galaxy.sb-roscoff.fr>

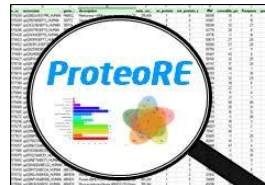
SBR tools: <http://webtools.sb-roscoff.fr>

Metagenomics: <http://galaxy4metab.sb-roscoff.fr>

Metabolomics:



Proteomics:



ChIP-seq:



Tools - panel

125+ platforms for using Galaxy: <https://galaxyproject.org/use>
 [Servers, clouds, deployable resources]

GENERAL PURPOSE:

usegalaxy.org, usegalaxy.eu, usegalaxy.fr, usegalaxy.org.au

DOMAIN SPECIFIC:



RNAseq: <http://galaxy.sb-roscoff.fr>

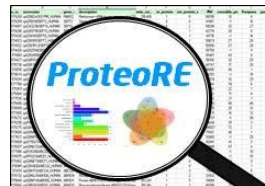
SBR tools: <http://webtools.sb-roscoff.fr>

Metagenomics: <http://galaxy4metab.sb-roscoff.fr>

Metabolomics:



Proteomics:



ChIP-seq:



Catalog of French Galaxy tools:

<http://galaxycat.france-bioinformatique.fr>



Tools - panel

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

Tools
trinity|
Trinity suite
1- ASSEMBLY:
[Trinity](#) de novo assembly of RNA-Seq data
[Trinity Statistics](#) Obtain basic stats for the number of genes and isoforms and contiguity of the assembly
[Generate gene to transcript map](#) for Trinity assembly
2- COUNTING:
[Align reads and estimate abundance](#) on a de novo assembly of RNA-Seq data
[Build expression matrix](#) for a de novo assembly of RNA-Seq data by Trinity
3- DIFFERENTIAL EXPRESSION:
[RNASeq samples quality check](#) for transcript quantification

✓ **Welcome to galaxy3.sb-roscoff.fr**

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr


Analyses and Bioinformatics for Marine Science

[Changelog](#)
[Tutorials](#)

[Galaxy](#) is an open, web-based platform for data intensive biomedical research. The [Galaxy team](#) is a part of [BX](#) at [Penn State](#), and the [Biology](#) and [Mathematics and Computer Science](#) departments at [Emory University](#). The [Galaxy Project](#) is supported in part by [NHGRI](#), [NSF](#), [The Huck Institutes of the Life Sciences](#), [The Institute for CyberScience](#) at [Penn State](#), and [Emory University](#).

History
search datasets
Trinity example
2 shown, 3 deleted
40.02 KB
4: reads.left.fq
3: reads.right.fq

Tools - form

Galaxy / ABiMS

Analyze DataWorkflowShared DataVisualizationAdminHelpUser

Using 0%

Tools

trinity

Trinity suite

1- ASSEMBLY:

[Trinity](#) de novo assembly of RNA-Seq data

[Trinity Statistics](#) Obtain basic stats for the number of genes and isoforms and contiguity of the assembly

[Generate gene to transcript map](#) for Trinity assembly

2- COUNTING:

[Align reads and estimate abundance](#) on a de novo assembly of RNA-Seq data

[Build expression matrix](#) for a de novo assembly of RNA-Seq data by Trinity

3- DIFFERENTIAL EXPRESSION:

[RNASeq samples quality check](#) for transcript quantification

Trinity de novo assembly of RNA-Seq data (Galaxy Version 2.2.0.0) Options

Paired or Single-end data?

Single

Single-end reads



4: reads.left.fq
3: reads.right.fq

(--single)

Strand specific data

YesNo

Run in silico normalization of reads

YesNo

Defaults to max. read coverage of 50. (--normalize_reads)

[Additional Options](#)

Execute

[Trinity](#) assembles transcript sequences from Illumina RNA-Seq data.

Citations [Show BibTeX](#)

Grabherr, Manfred G and Haas, Brian J and Yassour, Moran and Levin, Joshua Z and Thompson, Dawn A and Amit, Ido and Adiconis, Xian and Fan, Lin and Raychowdhury, Raktim and Zeng, Qiandong and et al. (2011). Full length

History

search datasets

Trinity example

2 shown, 2 deleted

37.53 KB

4: reads.left.fq

3: reads.right.fq

90

Tools - form

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 0%

Tools

trinity

Trinity suite

1- ASSEMBLY:

Trinity de novo assembly of RNA-Seq data

Trinity Statistics Obtain basic stats for the number of genes and isoforms and contiguity of the assembly

Generate gene to transcript map for Trinity assembly

2- COUNTING:

Align reads and estimate abundance on a de novo assembly of RNA-Seq data

Build expression matrix for a de novo assembly of RNA-Seq data by Trinity

3- DIFFERENTIAL EXPRESSION:

RNASeq samples quality check for transcript quantification

Trinity de novo assembly of RNA-Seq data (Galaxy Version 2.2.0.0)

Options

Paired or Single-end data?

Single

Single-end reads

4: reads.left.fq

3: reads.right.fq

(--single)

Strand specific data

Yes No

Run in silico normalization of reads

Yes No

Defaults to max. read coverage of 50. (--normalize_reads)

Additional Options

Execute

Trinity assembles transcript sequences from Illumina RNA-Seq data.

Citations Show BibTeX

Grabherr, Manfred G and Haas, Brian J and Yassour, Moran and Levin, Joshua Z and Thompson, Dawn A and Amit, Ido and Adiconis, Xian and Fan, Lin and Raychowdhury, Raktima and Zeng, Qiandong and et al. (2011). Full length

History

search datasets

Trinity example

2 shown, 2 deleted

37.53 KB

4: reads.left.fq

3: reads.right.fq

Tools - form

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 0%

Tools

trinity

Trinity suite

1- ASSEMBLY:

Trinity de novo assembly of RNA-Seq data

Trinity Statistics Obtain basic stats for the number of genes and isoforms and contiguity of the assembly

Generate gene to transcript map for Trinity assembly

2- COUNTING:

Align reads and estimate abundance on a de novo assembly of RNA-Seq data

Build expression matrix for a de novo assembly of RNA-Seq data by Trinity

3- DIFFERENTIAL EXPRESSION:

RNASeq samples quality check for transcript quantification

Trinity de novo assembly of RNA-Seq data (Galaxy Version 2.2.0.0)

Options

Paired or Single-end data?

Single

Single-end reads

4: reads.left.fq

3: reads.right.fq

(--single)

Strand specific data

Yes No

Run in silico normalization of reads

Yes No

Defaults to max. read coverage of 50. (--normalize_reads)

Additional Options

Execute

Trinity assembles transcript sequences from Illumina RNA-Seq data.

Citations Show BibTeX

Grabherr, Manfred G and Haas, Brian J and Yassour, Moran and Levin, Joshua Z and Thompson, Dawn A and Amit, Ido and Adiconis, Xian and Fan, Lin and Raychowdhury, Raktima and Zeng, Qiandong and et al. (2011). Full length

History

search datasets

Trinity example

2 shown, 2 deleted

37.53 KB

4: reads.left.fq

3: reads.right.fq

92

Tools - form

Tools can have some advanced options

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

Tools trinity

Trinity suite

1- ASSEMBLY:

[Trinity](#) de novo assembly of RNA-Seq data

[Trinity Statistics](#) Obtain basic stats for the number of genes and isoforms and contiguity of the assembly

[Generate gene to transcript map](#) for Trinity assembly

2- COUNTING:

[Align reads and estimate abundance](#) on a de novo assembly of RNA-Seq data

[Build expression matrix](#) for a de novo assembly of RNA-Seq data by Trinity

3- DIFFERENTIAL EXPRESSION:

[RNASeq samples quality check](#) for transcript quantification

Run in silico normalization of reads

Yes No

Defaults to max. read coverage of 50. (--normalize_reads)

Additional Options

Minimum Contig Length

200

All contigs shorter than this will be discarded (--min_contig_length)

Use the genome guided mode?

No

If you already mapped the reads to the genome, Trinity can use this information

Error-corrected or circular consensus (CCS) pac bio reads

Nothing selected

Experimental feature! Long reads must be in the same orientation as short reads if they are strand specific (--long_reads)

Minimum count for K-mers to be assembled

1

(--min_kmer_cov)

Execute

[Trinity](#) assembles transcript sequences from Illumina RNA-Seq data.

Citations [Show BibTeX](#)

History search datasets

Trinity example
2 shown, 2 deleted

37.53 KB

4: reads.left.fq

3: reads.right.fq

Tools - form

Galaxy / ABiMS

Analyze DataWorkflowShared DataVisualizationAdminHelpUser

Using 0%

Tools

trinity

Trinity suite

1- ASSEMBLY:

Trinity de novo assembly of RNA-Seq data

Trinity Statistics Obtain basic stats for the number of genes and isoforms and contiguity of the assembly

Generate gene to transcript map for Trinity assembly

2- COUNTING:

Align reads and estimate abundance on a de novo assembly of RNA-Seq data

Build expression matrix for a de novo assembly of RNA-Seq data by Trinity

3- DIFFERENTIAL EXPRESSION:

RNASeq samples quality check for transcript quantification

Trinity de novo assembly of RNA-Seq data (Galaxy Version 2.2.0.0)

Options

Paired or Single-end data?

Paired

Left/Forward strand reads

4: reads.left.fq
3: reads.right.fq

(--left)

Right/Reverse strand reads

4: reads.left.fq
3: reads.right.fq

(--right)

Strand specific data

Yes No

Run in silico normalization of reads

Yes No

Defaults to max. read coverage of 50. (--normalize_reads)

Additional Options

Execute

History

search datasets

Trinity example

2 shown, 3 deleted

40.02 KB

4: reads.left.fq

3: reads.right.fq

Tools - form

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 0%

Tools

trinity

Trinity suite

1- ASSEMBLY:

[Trinity](#) de novo assembly of RNA-Seq data

[Trinity Statistics](#) Obtain basic stats for the number of genes and isoforms and contiguity of the assembly

[Generate gene to transcript map](#) for Trinity assembly

2- COUNTING:

[Align reads and estimate abundance](#) on a de novo assembly of RNA-Seq data

[Build expression matrix](#) for a de novo assembly of RNA-Seq data by Trinity

3- DIFFERENTIAL EXPRESSION:

[RNASeq samples quality check](#) for transcript quantification

✓

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

5: Trinity on data 3 and data 4: Assembled Transcripts

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

Trinity example

3 shown, 2 deleted

37.53 KB

🕒 5: Trinity on data 3 and data 4: Assembled Transcripts

4: reads.left.fq

3: reads.right.fq



Job is waiting to run

= the job is in the scheduler « queue »

Duration time of this status depends on the amount of actual queued jobs and on the requested number of processors



Job is currently running

= the job is being executed on the computing cluster

Duration time of this status depends on the job's attributes and the computing resources allocated.

Some programs are executed with several processors (using 4, 8 or 16 Gb of RAM).

And others are mono-threaded 😞

Tools - Job status

2: xset.RData



Job is finished and status is OK

But warnings or errors can be hidden behind!

Tools - Job status



Job is finished but with an error status

= the program sends an error

The error is often explained by the program but sometimes ...
not.



Job is finished but with an error status

= the program sends an error

Possible causes of error :

— ~~The user~~ :P

- Bad usage : input file, format or option
- Bad integration of the program into Galaxy ... sorry :/
- Non anticipated crash of the program

Tools - Handle errors

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

Tools
search tools
[Get Data](#)
[Send Data](#)
[Collection Operations](#)

COMMON TOOLS
[Text Manipulation](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[Convert Formats](#)
[Extract Features](#)
[Fetch Sequences](#)
[Statistics](#)
[Graph/Display Data](#)
[Fasta Fastq Manipulation](#)

COMMON NGS TOOLS
[NGS:Samtools](#)
[NGS:Mapping](#)
[NGS:Bedtools](#)
[NGS:Picard Tools](#)

SEARCHING TOOLS
[Diamond](#)

 **Welcome to galaxy3.sb-roscoff.fr**

 **Information**
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr


Analyses and Bioinformatics for Marine Science
[Changelog](#)
[Tutorials](#)

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History
search datasets
eba 2016 sartools
28 shown, 14 [deleted](#)
1.59 MB
48: [group2_count2.txt](#)
47: [group2_count1.txt](#)
46: [group1_count2.txt](#)
45: [group1_count1.txt](#)
44: [SARTools edgeR R objects \(.RData\)](#)
43: [SARTools edgeR R log](#)
42: [SARTools edgeR figures](#)
41: [SARTools edgeR tables](#)
40: [SARTools edgeR](#)

Tools - Handle errors

The screenshot displays the Galaxy/ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User, along with a 'Using 0%' status indicator. The left sidebar contains a 'Tools' section with a search bar and various tool categories: Get Data, Send Data, Collection Operations, COMMON TOOLS, Text Manipulation, Filter and Sort, Join, Subtract and Group, Convert Formats, Extract Features, Fetch Sequences, Statistics, Graph/Display Data, Fasta Fastq Manipulation, COMMON NGS TOOLS, NGS:Samtools, NGS:Mapping, NGS:Bedtools, NGS:Picard Tools, SEARCHING TOOLS, and Diamond.

The main content area is titled 'Dataset generation errors'. It shows details for 'Dataset 48: group2_count2.txt', stating that the Galaxy framework encountered an error while attempting to run the tool. A traceback is displayed, showing the error occurred in 'local.py' and 'tempfile.py', resulting in an 'OSError: [Errno 2] No such file or directory: '/w/gala:'. Below the traceback, a message box indicates 'failure running job'. A red arrow points from this message to a 'View or report this error' button in the History panel.

The History panel on the right shows a list of datasets. The top entry is 'eba 2016 sartools' (1.59 MB). Below it is '48: group2_count2.txt', which is marked as a 'tool error' and 'failure running job'. This entry is highlighted in red. Below it are '47: group2_count1.txt', '46: group1_count2.txt', and '45: group1_count1.txt', all marked as 'group1_count1.txt' and highlighted in green. Each entry has icons for viewing, editing, and deleting.

At the bottom of the main content area, there is a section titled 'Report this error to the local Galaxy administrators'. It explains that local Galaxy administrators regularly review errors and provides a form to report the error. The form includes a field for 'Your email' (e.g., e.gueguen@sb-roscoff.fr) and a field for the error report (e.g., galaxy3.sb-roscoff.fr/dataset/errors?id=53eb0ef5c8056a28).

Tools - Handle errors

Sent to the support team

The screenshot shows the Galaxy / ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User. A status bar on the right indicates 'Using 0%'. The left sidebar contains a 'Tools' section with a search bar and a list of tool categories: Get Data, Send Data, Collection Operations, COMMON TOOLS, Text Manipulation, Filter and Sort, Join, Subtract and Group, Convert Formats, Extract Features, Fetch Sequences, Statistics, Graph/Display Data, Fasta Fastq Manipulation, COMMON NGS TOOLS, NGS:Samtools, NGS:Mapping, NGS:Bedtools, NGS:Picard Tools, SEARCHING TOOLS, and Diamond.

The main content area displays an error message: 'Tool execution generated the following error message: failure running job'. Below this, a section titled 'Report this error to the local Galaxy administrators' provides instructions: 'Usually the local Galaxy administrators regularly review errors that occur on the server. However, if you would like to provide additional information (such as what you were trying to do when the error occurred) and a contact e-mail address, we will be better able to investigate your problem and get back to you.' This section contains an 'Error Report' form with fields for 'Your email' (loraine.gueguen@sb-roscoff.fr) and 'Message' (a large text area). A 'Report' button is at the bottom of the form.

The right sidebar shows a 'History' panel with a search bar and a list of datasets. The first dataset is 'eba 2016 sartools' (28 shown, 14 deleted, 1.59 MB). Below it, a red box highlights an error for dataset '48: group2_count2.txt', stating 'tool error' and 'An error occurred with this dataset: failure running job'. Below this, a table shows gene counts for '1.gene0' (2.1813) and 'gene1' (1353), 'gene10' (72), 'gene100' (496), and 'gene1000' (50). Other datasets listed include '47: group2_count1.txt', '46: group1_count2.txt', and '45: group1_count1.txt'.

DATASET

Dataset

Both inputs and outputs

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User, along with a 'Using 0%' status indicator. The left sidebar contains a 'Tools' section with a search bar and a list of categories: Get Data, Send Data, Collection Operations, COMMON TOOLS, and COMMON NGS TOOLS. The main content area features a green welcome banner for 'galaxy3.sb-roscoff.fr', an information box with contact details, and the ABiMS logo with the tagline 'Analyses and Bioinformatics for Marine Science'. The right sidebar shows a 'History' section with a search bar and a list of datasets. The dataset '5: Trinity on data 3 and data 4: Assembled Transcripts' is highlighted with a red box. Below it are '4: reads.left.fq' and '3: reads.right.fq'. The word 'Dataset' is written in red text over the main content area.

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

Tools search tools

Get Data
Send Data
Collection Operations

COMMON TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Statistics
Graph/Display Data
Fasta Fastq Manipulation

COMMON NGS TOOLS
NGS:Samtools
NGS:Mapping
NGS:Bedtools
NGS:Picard Tools

SEARCHING TOOLS
Diamond

Welcome to galaxy3.sb-roscoff.fr

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS
Analyses and Bioinformatics for Marine Science

Station Biologique Roscoff

History search datasets

Trinity example
3 shown, 3 deleted
40.3 KB

5: Trinity on data 3 and data 4: Assembled Transcripts

4: reads.left.fq

3: reads.right.fq

Dataset

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

Dataset

Informations

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

Tool: Trinity

Number:	5
Name:	Trinity on data 3 and data 4: Assembled Transcripts
Created:	Wed 01 Mar 2017 03:52:36 PM (UTC)
Filesize:	2.5 KB
Dbkey:	?
Format:	fasta
Galaxy Tool ID:	toolshed.g2.bx.psu.edu/repos/iuc/trinity/trinity/2.2.0.0
Galaxy Tool Version:	2.2.0.0
Tool Version:	
Tool Standard Output:	stdout
Tool Standard Error:	stderr
Tool Exit Code:	0

Input Parameter	Value	Note for rerun
Paired or Single-end data?	paired	
Left/Forward strand reads	4: reads.left.fq	
Right/Reverse strand reads	3: reads.right.fq	
Strand specific data	true	
Strand-specific Library Type	Reverse-Forward	
Jaccard Clip options	Not available.	
Run in silico normalization of reads	True	
additional_params		
Minimum Contig Length	200	
Use the genome guided mode?	no	

[galaxy3.sb-roscoff.fr/datasets/c10ec933dc50450a/show_params](#)

History

search datasets

Trinity example
3 shown, 3 deleted
40.3 KB

5: Trinity on data 3 and data 4: Assembled Transcripts
7 sequences
format: **fasta**, database: ?
View details 0_g1_i1 len=541 path=[519:0-546
GTCTGAATTCGCATGTAATGCAGCTTTCCAGACACAAGTATGG
TCGCCATTGTGCAAAATATGTGTCTGATAGACCCGCGCTTTCA
TGACATGAGCGTGGCACCCTGAAGACAGGGTGTGGGTGAGAGGCT
TGAGTTGTCTTGTATCATCAATAGATGCAAAATTAGATGTAAGAAC

4: reads.left.fq

3: reads.right.fq

Dataset

Informations

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

Tool: Trinity

Number:	5
Name:	Trinity on data 3 and data 4: Assembled Transcripts
Created:	Wed 01 Mar 2017 03:52:36 PM (UTC)
Filesize:	2.5 KB
Dbkey:	?
Format:	fasta
Galaxy Tool ID:	toolshed.g2.bx.psu.edu/repos/iuc/trinity/trinity/2.2.0.0
Galaxy Tool Version:	2.2.0.0
Tool Version:	
Tool Standard Output:	stdout
Tool Standard Error:	stderr
Tool Exit Code:	0

Input Parameter	Value	Note for rerun
Paired or Single-end data?	paired	
Left/Forward strand reads	4: reads.left.fq	
Right/Reverse strand reads	3: reads.right.fq	
Strand specific data	true	
Strand-specific Library Type	Reverse-Forward	
Jaccard Clip options	Not available.	
Run in silico normalization of reads	True	
additional_params		
Minimum Contig Length	200	
Use the genome guided mode?	no	

[galaxy3.sb-roscoff.fr/datasets/c10ec933dc50450a/show_params](#)

History

search datasets

Trinity example
3 shown, 3 deleted
40.3 KB

5: Trinity on data 3 and data 4: Assembled Transcripts
7 sequences
format: **fasta**, database: ?
[View details](#) 0_g1_i1 len=541 path=[519:0-546
GTCTGAATTCGCATGTAATGCAGCTTTCCAGACACAAGTATGG
TCGCCATTGTGCAAAATATGTGTCTGATAGACCCGCGCTTTCA
TGACATGAGCGTGGCACCCTGAAGACAGGGTGTGGGTGAGAGGCT
TGAGTTGTCTTGTATCATCAATAGATGCAAAATTAGATGTAAGAAC

4: reads.left.fq

3: reads.right.fq

Dataset

Download

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

Tool: Trinity

Number:	5
Name:	Trinity on data 3 and data 4: Assembled Transcripts
Created:	Wed 01 Mar 2017 03:52:36 PM (UTC)
Filesize:	2.5 KB
Dbkey:	?
Format:	fasta
Galaxy Tool ID:	toolshed.g2.bx.psu.edu/repos/iuc/trinity/trinity/2.2.0.0
Galaxy Tool Version:	2.2.0.0
Tool Version:	
Tool Standard Output:	stdout
Tool Standard Error:	stderr
Tool Exit Code:	0

Input Parameter	Value	Note for rerun
Paired or Single-end data?	paired	
Left/Forward strand reads	4: reads.left.fq	
Right/Reverse strand reads	3: reads.right.fq	
Strand specific data	true	
Strand-specific Library Type	Reverse-Forward	
Jaccard Clip options	Not available.	
Run in silico normalization of reads	True	
additional_params		
Minimum Contig Length	200	
Use the genome guided mode?	no	

[galaxy3.sb-roscoff.fr/datasets/c10ec933dc50450a/display?to_ext=fasta](#)

History

search datasets

Trinity example
3 shown, 3 deleted
40.3 KB

5: Trinity on data 3 and data 4: Assembled Transcripts
7 sequences
format: **fasta**, database: ?
[Download](#) [Info](#) [Refresh](#)
DN0_c0_g1_i1 len=541 path=[519:0-546
GTCTGAATTCGCATGTAATGCAGCTTTCCAGACACAAGTATGG
TCGCCATTGTGCAAAATATGTGTCTGATAGACCCGCGCTTTCA
TGACATGAGCGTGGCACCTGAAGACAGGCTGTGGGTGAGAGGCT
TGAGTTGTCTTGTATCATCAATAGATGCAAAATTAGATGTAAGAA

4: reads.left.fq

3: reads.right.fq

Dataset

Re-run a job

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

Tools search tools

Get Data
Send Data
Collection Operations
COMMON TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Statistics
Graph/Display Data
Fasta Fastq Manipulation
COMMON NGS TOOLS
NGS:Samtools
NGS:Mapping
NGS:Bedtools
NGS:Picard Tools
SEARCHING TOOLS
Diamond

Trinity de novo assembly of RNA-Seq data (Galaxy Version 2.2.0.0) Options

Paired or Single-end data?
Paired

Left/Forward strand reads
5: Trinity on data 3 and data 4: Assembled Transcripts
4: reads.left.fq
3: reads.right.fq
(--left)

Right/Reverse strand reads
5: Trinity on data 3 and data 4: Assembled Transcripts
4: reads.left.fq
3: reads.right.fq
(--right)

Strand specific data
Yes No

Strand-specific Library Type
Reverse-Forward
(--SS_lib_type)

Jaccard Clip options
Yes No

History search datasets

Trinity example
3 shown, 3 deleted
40.3 KB

5: Trinity on data 3 and data 4: Assembled Transcripts
7 sequences
format: fasta, database: ?
Run this job again len=541 path=[519:0-541]
GTCTGAATTCGCATGTAATGCAGCTTTCCGAGACACAAGTATGGA
TCGCCATTGTGCAAAATATGTGTCTGATAGACCGCAGGCTTTCA
TGACATGAGCGTGGCACCTGAAGACAGGCTGTGGGTGAGAGGCT
TGAGTTGTCTTGTATCATCAATAGATGCAAAATTAGATGTAAGAAC

4: reads.left.fq
3: reads.right.fq

galaxy3.sb-roscoff.fr/tool_runner/rerun?id=c10ec933dc50450a

Dataset

Dataset display : text, tabular, pdf, picture, html ...

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

Tools

search tools

[Get Data](#)
[Send Data](#)
[Collection Operations](#)

COMMON TOOLS
[Text Manipulation](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[Convert Formats](#)
[Extract Features](#)
[Fetch Sequences](#)
[Statistics](#)
[Graph/Display Data](#)
[Fasta Fastq Manipulation](#)

COMMON NGS TOOLS
[NGS:Samtools](#)
[NGS:Mapping](#)
[NGS:Bedtools](#)
[NGS:Picard Tools](#)

SEARCHING TOOLS
[Diamond](#)

1	2	3	4	5	6
	wt_37_2	wt_37_3	wt_37_1	wt_GSNO_3	wt_GSNO_1
TR24 c0_g1_i1	90.00	67.00	85.00	36.00	35.00
TR2779 c0_g1_i1	186.00	137.00	217.00	147.00	186.00
TR127 c1_g1_i1	9.00	23.00	16.00	2.00	0.00
TR2107 c1_g1_i1	59.00	65.00	47.00	6.00	6.00
TR2011 c5_g1_i1	11.00	4.00	4.00	8.00	5.00
TR4163 c0_g1_i1	368.00	422.00	425.00	172.00	216.00
TR5055 c0_g2_i1	36.00	17.00	27.00	4.00	7.00
TR1449 c0_g1_i1	196.00	230.00	207.00	66.00	113.00
TR1982 c2_g1_i1	7.00	7.00	6.00	4.00	3.00
TR1859 c3_g1_i1	0.00	0.00	1.00	0.00	0.00
TR1492 c0_g1_i2	1895.00	1906.00	1921.00	1104.00	1263.00
TR1122 c0_g1_i1	2.00	3.00	0.00	3.00	0.00
TR2278 c0_g1_i1	497.00	610.00	598.00	333.00	406.00
TR4084 c0_g1_i1	95.00	148.00	86.00	77.00	111.00
TR4761 c0_g1_i1	2089.00	1746.00	1875.00	155.00	174.00
TR3638 c0_g1_i1	647.00	676.00	712.00	117.00	184.00
TR2090 c0_g1_i1	0.00	0.00	0.00	22.00	0.00
TR3854 c0_g1_i1	1878.00	1734.00	1864.00	1775.00	2173.00
TR131 c0_g1_i1	32.00	28.00	31.00	1001.00	1233.00
TR5075 c0_g1_i1	13.00	22.00	21.00	6.00	8.00
TR2182 c3_g2_i6	1.44	2.70	3.84	3.35	0.00
TR3788 c0_g1_i1	17.00	30.00	22.00	91.00	132.00
TR4859 c0_g1_i1	6.00	12.00	8.00	4.00	1.00
TR2487 c0_g1_i1	386.00	383.00	424.00	689.00	866.00

History

1: samples.txt

2:
input.matrix.wt 37 vs wt GS
NO.DESeq2.DE results

3:
input.matrix.wt 37 vs wt ph8
.DESeq2.DE results

4:
input.matrix.wt GSNO vs wt
ph8.DESeq2.DE results

5:
matrix.counts.matrix

6: de results
a list of 3 datasets

7: Extract and cluster differentially expressed transcripts on data 2, data 3, and others: extracted differentially expressed genes
a list of datasets

8: Extract and cluster differentially expressed transcripts on data 2, data 3, and others

galaxy3.sb-roscoff.fr/datasets/4437d546e349a03a/display/?preview=True

Dataset

Renaming and annotation

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes links for 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User', along with a 'Using 0%' indicator. The left sidebar contains a 'Tools' section with a search bar and a list of tool categories: 'Get Data', 'Send Data', 'Collection Operations', 'COMMON TOOLS', 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Convert Formats', 'Extract Features', 'Fetch Sequences', 'Statistics', 'Graph/Display Data', 'Fasta Fastq Manipulation', 'COMMON NGS TOOLS', 'NGS:Samtools', 'NGS:Mapping', 'NGS:Bedtools', 'NGS:Picard Tools', 'SEARCHING TOOLS', and 'Diamond'. The main content area is titled 'Edit Attributes' and contains the following fields:

- Name:** A text input field containing 'matrix.counts.matrix'.
- Info:** A text input field containing 'uploaded tabular file'.
- Annotation / Notes:** A text input field containing 'This is my expression matrix.' Below this field is a note: 'Add an annotation or notes to a dataset; annotations are available when a history is viewed.'
- Database/Build:** A dropdown menu currently set to 'unspecified (?)'. Below it are 'Save' and 'Auto-detect' buttons. A note below the buttons states: 'This will inspect the dataset and attempt to correct the above column values if they are not accurate.'

The right sidebar shows the 'History' panel with a list of dataset entries. The top entry is '6: Extract and cluster differentially expressed transcripts on data 2, data 3, and others'. Below it is '7: Extract and cluster differentially expressed transcripts on data 2, data 3, and others: extracted differentially expressed genes'. The third entry is '6: de results'. The fourth entry is '5: matrix.counts.matrix', which is highlighted and has an 'Edit attributes' tooltip. Below this entry, it shows '41 lines', 'format: txt, database: ?', 'uploaded tabular file', and 'Tags: trinity'. The 'Annotation' field for this entry contains 'This is my expression matrix.' and a list of column names: 'wt_37_2 wt_37_3 wt_37_1 wt_GSMO_3'.

galaxy3.sb-roscoff.fr/datasets/4437d546e349a03a/edit

Dataset

Renaming and annotation

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes links for 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User', along with a 'Using 0%' indicator. The left sidebar contains a 'Tools' section with a search bar and a list of tool categories: 'Get Data', 'Send Data', 'Collection Operations', 'COMMON TOOLS', 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Convert Formats', 'Extract Features', 'Fetch Sequences', 'Statistics', 'Graph/Display Data', 'Fasta Fastq Manipulation', 'COMMON NGS TOOLS', 'NGS:Samtools', 'NGS:Mapping', 'NGS:Bedtools', 'NGS:Picard Tools', 'SEARCHING TOOLS', and 'Diamond'. The main content area is titled 'Edit Attributes' and contains the following fields:

- Name:** A text input field containing 'matrix.counts.matrix'.
- Info:** A text input field containing 'uploaded tabular file'.
- Annotation / Notes:** A text input field containing 'This is my expression matrix.' Below this field is a note: 'Add an annotation or notes to a dataset; annotations are available when a history is viewed.'
- Database/Build:** A dropdown menu currently set to 'unspecified (?)'. Below it are 'Save' and 'Auto-detect' buttons. A note below the buttons states: 'This will inspect the dataset and attempt to correct the above column values if they are not accurate.'

The right sidebar shows the 'History' panel with a list of dataset entries. The entry '5: matrix.counts.matrix' is highlighted, and a tooltip 'Edit attributes' is visible. Below the history panel, a red box highlights the 'Tags' section, which contains a tag 'x trinity'. The 'Annotation' section below the tags also contains the text 'This is my expression matrix.' The bottom of the interface shows the URL 'galaxy3.sb-roscoff.fr/datasets/4437d546e349a03a/edit'.

Dataset

Change the Datatype of the Dataset

The screenshot shows the Galaxy / ABiMS interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User. The 'Tools' panel on the left lists various categories of tools, including 'COMMON TOOLS' and 'COMMON NGS TOOLS'. The main panel displays the 'Datatype' tab for a dataset, with a 'Change data type' section. A dropdown menu for 'New Type' is open, showing a list of data types: txt, supermatcher, svg, swiss, syco, tabix, table, tabular (highlighted), tagseq, and tandem. The 'History' panel on the right shows a list of datasets, including '1: samples.txt', '2: input.matrix.wt 37 vs wt GS NO.DESeq2.DE results', '3: input.matrix.wt 37 vs wt ph8 .DESeq2.DE results', '4: input.matrix.wt GSNO vs wt ph8.DESeq2.DE results', '5: matrix.counts.matrix', '6: de results', and '7: Extract and cluster differentially expressed transcripts on data 2, data 3, and others'.

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 0%

Tools

search tools

Get Data

Send Data

Collection Operations

COMMON TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

Extract Features

Fetch Sequences

Statistics

Graph/Display Data

Fasta Fastq Manipulation

COMMON NGS TOOLS

NGS:Samtools

NGS:Mapping

NGS:Bedtools

NGS:Picard Tools

SEARCHING TOOLS

Diamond

Attributes Convert Format **Datatype** Permissions

Change data type

New Type:

txt

supermatcher

svg

swiss

syco

tabix

table

tabular

tagseq

tandem

ing dataset but *not* modify its contents.
d the type of your dataset.

History

7: Extract and cluster differentially expressed transcripts on data 2, data 3, and others

7: Extract and cluster differentially expressed transcripts on data 2, data 3, and others: extracted differentially expressed genes

a list of datasets

6: de results

a list of 3 datasets

5: matrix.counts.matrix

4: input.matrix.wt GSNO vs wt ph8.DESeq2.DE results

3: input.matrix.wt 37 vs wt ph8 .DESeq2.DE results

2: input.matrix.wt 37 vs wt GS NO.DESeq2.DE results

1: samples.txt







1 Visualize this data

uc002zsw.2_cds_0_0_chr22_2
uc003bhh.3 cds 0 0 chr22 4



Trackster
Fast, interactive visualization for large, NGS/HTS datasets using only a web browser.



Bar Horizontal (NVD3)
Renders a horizontal bar diagram using NVD3 hosted at <http://www.nvd3.org>.



Bar Horizontal Stacked (NVD3)
Renders a stacked horizontal bar diagram using NVD3 hosted at <http://www.nvd3.org>.



Bar Stacked (NVD3)
Renders a scatter plot using NVD3 hosted at <http://www.nvd3.org>.



Bar diagram (NVD3)
Renders a regular bar diagram using NVD3 hosted at <http://www.nvd3.org>.



Bar diagram (jqPlot)
Renders a bar diagram using jqPlot hosted at <http://www.jqplot.com>.



Box plot (jqPlot)
Processes tabular data using R and renders a box plot using jqPlot hosted at

7: Select first on data 6




5 lines

format: **tabular**, database: **hg38**









1	Tool Help	2
uc010gqp.3_cds_0_0_chr22_15690078_f	63	
uc011agd.3_cds_0_0_chr22_15528159_f	48	
uc062bek.1_cds_0_0_chr22_15690246_f	46	
uc003bhh.4_cds_0_0_chr22_46256561_r	30	
uc062bej.1_cds_1_0_chr22_15690426_f	26	

Tool help for Select first

What it does

This tool outputs specified number of lines from the **beginning** of a dataset

Example

Selecting 2 lines from this:

```
chr7 56632 56652 D17003_CTCF_R6
chr7 56736 56756 D17003_CTCF_R7
chr7 56761 56781 D17003_CTCF_R4
chr7 56772 56792 D17003_CTCF_R7
chr7 56775 56795 D17003_CTCF_R4
```

will produce:

```
chr7 56632 56652 D17003_CTCF_R6
chr7 56736 56756 D17003_CTCF_R7
```


HISTORY

History panel

Both inputs and outputs

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User, along with a 'Using 0%' status indicator. The left sidebar contains a 'Tools' section with a search bar and various tool categories like 'Get Data', 'Send Data', 'Collection Operations', 'COMMON TOOLS', 'COMMON NGS TOOLS', and 'SEARCHING TOOLS'. The main content area features a 'Welcome to galaxy3.sb-roscoff.fr' message, an 'Information' box with contact details, the ABiMS logo, and links to Changelog and Tutorials. A large red 'History' text is overlaid on the main content area. On the right, the 'History' panel is visible, showing a list of datasets. The dataset '48: group2_count2.txt' is highlighted with a red border, indicating it is the current selection. Other datasets listed include '47: group2_count1.txt', '46: group1_count2.txt', '45: group1_count1.txt', '44: SARTools edgeR R objects (.RData)', '43: SARTools edgeR R log', '42: SARTools edgeR figures', '41: SARTools edgeR tables', and '40: SARTools edgeR'.

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

Tools

search tools

Get Data
Send Data
Collection Operations

COMMON TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Statistics
Graph/Display Data
Fasta Fastq Manipulation

COMMON NGS TOOLS
NGS:Samtools
NGS:Mapping
NGS:Bedtools
NGS:Picard Tools

SEARCHING TOOLS
Diamond

Welcome to galaxy3.sb-roscoff.fr

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS
Analyses and Bioinformatics for Marine Science

Station Biologique Roscoff

Changelog
Tutorials

History

search datasets

eba 2016 sartools
28 shown, 14 [deleted](#)
1.59 MB

48:
group2_count2.txt

47:
group2_count1.txt

46:
group1_count2.txt

45:
group1_count1.txt

44: SARTools edgeR R objects (.RData)

43: SARTools edgeR R log

42: SARTools edgeR figures

41: SARTools edgeR tables

40: SARTools edgeR

History

search datasets

Unnamed History
28 shown, 14 deleted
1.59 MB

Click to rename history

48: group2_count2.txt

47: group2_count1.txt

46: group1_count2.txt

45: group1_count1.txt

44: SARTools edgeR R objects (.RData)

43: SARTools edgeR R log

42: SARTools edgeR figures

41: SARTools edgeR tables

40: SARTools edgeR

History

search datasets

eba 2016 sartools
28 shown, 14 deleted
1.59 MB

48: group2_count2.txt

47: group2_count1.txt

46: group1_count2.txt

45: group1_count1.txt

44: SARTools edgeR R objects (.RData)

43: SARTools edgeR R log

42: SARTools edgeR figures

41: SARTools edgeR tables

40: SARTools edgeR

History

search datasets

eba 2016 sartools
28 shown, 14 deleted
1.59 MB

Tags:
 x RNAseq
 x differentialanalysis

Annotation:
 bla bla bla

48: group2_count2.txt

47: group2_count1.txt

46: group1_count2.txt

45: group1_count1.txt

44: SARTools edgeR R objects (.RData)

43: SARTools edgeR R log

History panel

History menu: Create new, Rename, Delete, **Delete Permanently**

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User, along with a 'Using 0%' status indicator. The left sidebar contains a 'Tools' section with a search bar and various tool categories like 'Get Data', 'Send Data', 'Collection Operations', 'COMMON TOOLS', 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Convert Formats', 'Extract Features', 'Fetch Sequences', 'Statistics', 'Graph/Display Data', 'Fasta Fastq Manipulation', 'COMMON NGS TOOLS', 'NGS:Samtools', 'NGS:Mapping', 'NGS:Bedtools', 'NGS:Picard Tools', 'SEARCHING TOOLS', and 'Diamond'. The main content area features a green welcome banner for 'galaxy3.sb-roscoff.fr', an information box with contact details, the ABiMS logo (Analyses and Bioinformatics for Marine Science), and links to Changelog and Tutorials. A paragraph describes Galaxy as an open, web-based platform for data intensive biomedical research, supported by various institutions. On the right, the 'History' panel is open, showing a list of actions: 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), and OTHER ACTIONS (Import from File). The 'Delete Permanently' option is highlighted with a red box.

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 0%

Tools

search tools

Get Data

Send Data

Collection Operations

COMMON TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

Extract Features

Fetch Sequences

Statistics

Graph/Display Data

Fasta Fastq Manipulation

COMMON NGS TOOLS

NGS:Samtools

NGS:Mapping

NGS:Bedtools

NGS:Picard Tools

SEARCHING TOOLS

Diamond

Welcome to galaxy3.sb-roscoff.fr

Information

For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS

Analyses and Bioinformatics for Marine Science

Changelog

Tutorials

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History

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

History panel

Saved histories

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

Tools search tools

Get Data
Send Data
Collection Operations
COMMON TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Statistics
Graph/Display Data
Fasta Fastq Manipulation
COMMON NGS TOOLS
NGS:Samtools
NGS:Mapping
NGS:Bedtools
NGS:Picard Tools
SEARCHING TOOLS
Diamond

search history names and tags

Advanced Search

Name	Datasets	Tags	Sharing	Size on Disk	Created	Last
eba2016 deseq2	8	0 Tags		829.2 KB	Nov 14, 2016	Nov 14, 2016
Copy of TP_ITMO2016 shared by xi.liu@sb- roscoff.fr (active items only)	32 2	0 Tags		28.5 MB	Nov 10, 2016	Nov 10, 2016
eba 2016 macs2	3 2	0 Tags		602.8 KB	Nov 09, 2016	Nov 09, 2016
eba 2016 sickle	10	0 Tags		4.2 MB	Oct 04, 2016	Nov 04, 2016

Page: 1 2 | Show All

For 0 selected histories: Rename Delete Delete Permanently Undelete

Histories that have been deleted for more than a time period specified by the Galaxy administrator(s) may be permanently deleted.

History

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

galaxy3.sb-roscoff.fr/history/list

History panel

Saved histories: Switch histories

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

Tools search tools

Get Data
Send Data
Collection Operations
COMMON TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Statistics
Graph/Display Data
Fastq Manipulation
COMMON NGS TOOLS
NGS:Samtools
NGS:Mapping
NGS:Bedtools
NGS:Picard Tools
SEARCHING TOOLS
Diamond

search history names and tags

Advanced Search

☐	Name	Datasets	Tags	Sharing	Size on Disk	Created	Last
☐	eba2016 deseq2	8	0 Tags		829.2 KB	Nov 14, 2016	Nov 1
☐	Copy of TP_ITMO2016 shared by xi.liu@sb- roscoff.fr (active items only)	32 2	0 Tags		28.5 MB	Nov 10, 2016	Nov 1
☐	eba 2016 macs2	3 2	0 Tags		602.8 KB	Nov 09, 2016	Nov C
☐	eba 2016 sickle	10	0 Tags		4.2 MB	Oct 04, 2016	Nov C

Page: 1 2 | Show All

For 0 selected histories: Rename Delete Delete Permanently Undelete

Histories that have been deleted for more than a time period specified by the Galaxy administrator(s) may be permanently deleted.

History

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

History panel

History menu: Create new, Rename, Delete, Delete Permanently

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User, along with a 'Using 0%' status indicator. The left sidebar contains a 'Tools' section with a search bar and various tool categories like 'Get Data', 'Send Data', 'Collection Operations', 'COMMON TOOLS', 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Convert Formats', 'Extract Features', 'Fetch Sequences', 'Statistics', 'Graph/Display Data', 'Fasta Fastq Manipulation', 'COMMON NGS TOOLS', 'NGS:Samtools', 'NGS:Mapping', 'NGS:Bedtools', 'NGS:Picard Tools', 'SEARCHING TOOLS', and 'Diamond'. The main content area features a welcome message for 'galaxy3.sb-roscoff.fr', an information box with contact details, the ABiMS logo (Analyses and Bioinformatics for Marine Science), and links to Changelog and Tutorials. A paragraph describes Galaxy as an open, web-based platform for data intensive biomedical research, mentioning its affiliation with Penn State and Emory University. On the right, the 'History' panel is open, showing a list of actions. The 'HISTORY ACTIONS' section is highlighted with a red box, and the 'Create New' option is selected.

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

Tools

search tools

Get Data
Send Data
Collection Operations

COMMON TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Statistics
Graph/Display Data
Fasta Fastq Manipulation

COMMON NGS TOOLS
NGS:Samtools
NGS:Mapping
NGS:Bedtools
NGS:Picard Tools

SEARCHING TOOLS
Diamond

Welcome to galaxy3.sb-roscoff.fr

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS
Analyses and Bioinformatics for Marine Science

Changelog
Tutorials

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History

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

History panel

[illegible]

History panel

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

Done search histories search all datasets Create new

Current History

Switch to

eba 2016 sartools
28 shown, 14 [deleted](#)
1.59 MB

search datasets

Drag datasets here to copy them to the current history

- ✖ 48: [group2_count2.txt](#)
- 47: [group2_count1.txt](#)
- 46: [group1_count2.txt](#)
- 45: [group1_count1.txt](#)
- 44: [SARTools edgeR R objects \(.RData\)](#)
- 43: [SARTools edgeR R log](#)
- 42: [SARTools edgeR figures](#)
- 41: [SARTools edgeR tables](#)
- 40: [SARTools edgeR report](#)

Switch to

Trinity example
3 shown, 3 [deleted](#)
40.3 KB

search datasets

- 5: [Trinity on data 3 and data 4: Assembled Transcripts](#)
- 4: [reads.left.fq](#)
- 3: [reads.right.fq](#)

Switch to

trinity_contig_exn50_statistic
12 shown, 15 [deleted](#)
47.01 KB

search datasets

- 14: [Build expression matrix on data 7 and data 6: matrix of UpperQuartile-normalized expression values](#)
- 13: [Build expression matrix on data 7 and data 6: matrix of TPM expression values \(not cross-sample normalized\)](#)
- 12: [Build expression matrix on data 7 and data 6: estimated RNA-Seq fragment counts \(raw counts\)](#)
- 9: [Build expression matrix on data 7 and data 6: matrix of TPM expression values \(not cross-sample normalized\)](#)
- 8: [Build expression matrix on data 7 and data 6: estimated](#)

Switch to

eba 2016 tr
16 shown
21.92 KB

search datasets

- ✖ 16: [Extra](#)
- 15: [Extract](#)
- 14: [Extract](#)
- ✖ 13: [Extra](#)
- ✖ 12: [Extra](#)

History panel

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

Done search histories search all datasets Create new

Current History Switch to

eba 2016 sartools
28 shown, 14 [deleted](#)
1.59 MB

search datasets

Drag datasets here to copy them to the current history

- ✖ 48: [group2_count2.txt](#)
- 47: [group2_count1.txt](#)
- 46: [group1_count2.txt](#)
- 45: [group1_count1.txt](#)
- 44: [SARTools edgeR R objects \(.RData\)](#)
- 43: [SARTools edgeR R log](#)
- 42: [SARTools edgeR figures](#)
- 41: [SARTools edgeR tables](#)
- 40: [SARTools edgeR report](#)

Trinity example
3 shown, 3 [deleted](#)
40.3 KB

search datasets

- 5: [Trinity on data 3 and data 4: Assembled Transcripts](#)
- 4: [reads.left.fq](#)
- 3: [reads.right.fq](#)

trinity_contig_exn50_statistic
12 shown, 15 [deleted](#)
47.01 KB

search datasets

- 14: [Build expression matrix on data 7 and data 6: matrix of UpperQuartile-normalized expression values](#)
- 13: [Build expression matrix on data 7 and data 6: matrix of TPM expression values \(not cross-sample normalized\)](#)
- 12: [Build expression matrix on data 7 and data 6: estimated RNA-Seq fragment counts \(raw counts\)](#)
- 9: [Build expression matrix on data 7 and data 6: matrix of TPM expression values \(not cross-sample normalized\)](#)
- 8: [Build expression matrix on data 7 and data 6: estimated](#)

eba 2016 tr
16 shown
21.92 KB

search datasets

- ✖ 16: [Extra](#)
- 15: [Extract](#)
- 14: [Extract](#)
- ✖ 13: [Extra](#)
- ✖ 12: [Extra](#)

History panel

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

Done search histories search all datasets Create new

Current History

Switch to

eba 2016 sartools
28 shown, 14 [deleted](#)
1.59 MB

search datasets

Drag datasets here to copy them to the current history

- 48: group2_count2.txt
- 47: group2_count1.txt
- 46: group1_count2.txt
- 45: group1_count1.txt
- 44: SARTools edgeR R objects (.RData)
- 43: SARTools edgeR R log
- 42: SARTools edgeR figures
- 41: SARTools edgeR tables
- 40: SARTools edgeR report

Switch to

Trinity example
3 shown, 3 [deleted](#)
40.3 KB

search datasets

- 5: Trinity on data 3 and data 4: Assembled Transcripts
- 4: reads.left.fq
- 3: reads.right.fq

Switch to

trinity_contig_exn50_statistic
12 shown, 15 [deleted](#)
47.01 KB

search datasets

- 14: Build expression matrix on data 7 and data 6: matrix of UpperQuartile-normalized expression values
- 13: Build expression matrix on data 7 and data 6: matrix of TPM expression values (not cross-sample normalized)
- 12: Build expression matrix on data 7 and data 6: estimated RNA-Seq fragment counts (raw counts)
- 9: Build expression matrix on data 7 and data 6: matrix of TPM expression values (not cross-sample normalized)
- 8: Build expression matrix on data 7 and data 6: estimated

Switch to

eba 2016 tr
16 shown
21.92 KB

search datasets

- 16: Extra differentially transcripts of RData file
- 15: Extract expressed to data 3, and depleted cat a list of dataset
- 14: Extract expressed to data 3, and a list of dataset
- 13: Extra differentially transcripts of
- 12: Extra differentially

History panel

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

Done search histories search all datasets Create new

Current History Switch to

eba 2016 sartools
28 shown, 14 [deleted](#)
1.59 MB

search datasets

Drag datasets here to copy them to the current history

- ✖ 48: [group2_count2.txt](#)
- 47: [group2_count1.txt](#)
- 46: [group1_count2.txt](#)
- 45: [group1_count1.txt](#)
- 44: [SARTools edgeR R objects \(.RData\)](#)
- 43: [SARTools edgeR R log](#)
- 42: [SARTools edgeR figures](#)
- 41: [SARTools edgeR tables](#)
- 40: [SARTools edgeR report](#)

Trinity example
3 shown, 3 [deleted](#)
40.3 KB

search datasets

- 5: [Trinity on data 3 and data 4: Assembled Transcripts](#)
- 4: [reads.left.fq](#)
- 3: [reads.right.fq](#)

trinity_contig_exn50_statistic
12 shown, 15 [deleted](#)
47.01 KB

search datasets

- 14: [Build expression matrix on data 7 and data 6: matrix of UpperQuartile-normalized expression values](#)
- 13: [Build expression matrix on data 7 and data 6: matrix of TPM expression values \(not cross-sample normalized\)](#)
- 12: [Build expression matrix on data 7 and data 6: estimated RNA-Seq fragment counts \(raw counts\)](#)
- 9: [Build expression matrix on data 7 and data 6: matrix of TPM expression values \(not cross-sample normalized\)](#)
- 8: [Build expression matrix on data 7 and data 6: estimated](#)

eba 2016 tr
16 shown
21.92 KB

search datasets

- ✖ 16: [Extra](#)
- 15: [Extract](#)
- 14: [Extract](#)
- ✖ 13: [Extra](#)
- ✖ 12: [Extra](#)

History panel

Galaxy / ABiMS Analyze Data Workflow Shared Data Visualization Admin Help User

Using 0%

Done search histories search all datasets

Create new

Current History

Switch to

eba 2016 sartools
28 shown, 14 [deleted](#)
1.59 MB

search datasets

Drag datasets here to copy them to the current history

- ✖ 48: [group2_count2.txt](#)
- 47: [group2_count1.txt](#)
- 46: [group1_count2.txt](#)
- 45: [group1_count1.txt](#)
- 44: [SARTools edgeR R objects \(.RData\)](#)
- 43: [SARTools edgeR R log](#)
- 42: [SARTools edgeR figures](#)
- 41: [SARTools edgeR tables](#)
- 40: [SARTools edgeR report](#)

Trinity example
3 shown, 3 [deleted](#)
40.3 KB

search datasets

- 5: [Trinity on data 3 and data 4: Assembled Transcripts](#)
- 4: [reads.left.fq](#)
- 3: [reads.right.fq](#)

trinity_contig_exn50_statistic
12 shown, 15 [deleted](#)
47.01 KB

search datasets

- 14: [Build expression matrix on data 7 and data 6: matrix of UpperQuartile-normalized expression values](#)
- 13: [Build expression matrix on data 7 and data 6: matrix of TPM expression values \(not cross-sample normalized\)](#)
- 12: [Build expression matrix on data 7 and data 6: estimated RNA-Seq fragment counts \(raw counts\)](#)
- 9: [Build expression matrix on data 7 and data 6: matrix of TPM expression values \(not cross-sample normalized\)](#)
- 8: [Build expression matrix on data 7 and data 6: estimated](#)

eba 2016 tr
16 shown
21.92 KB

search datasets

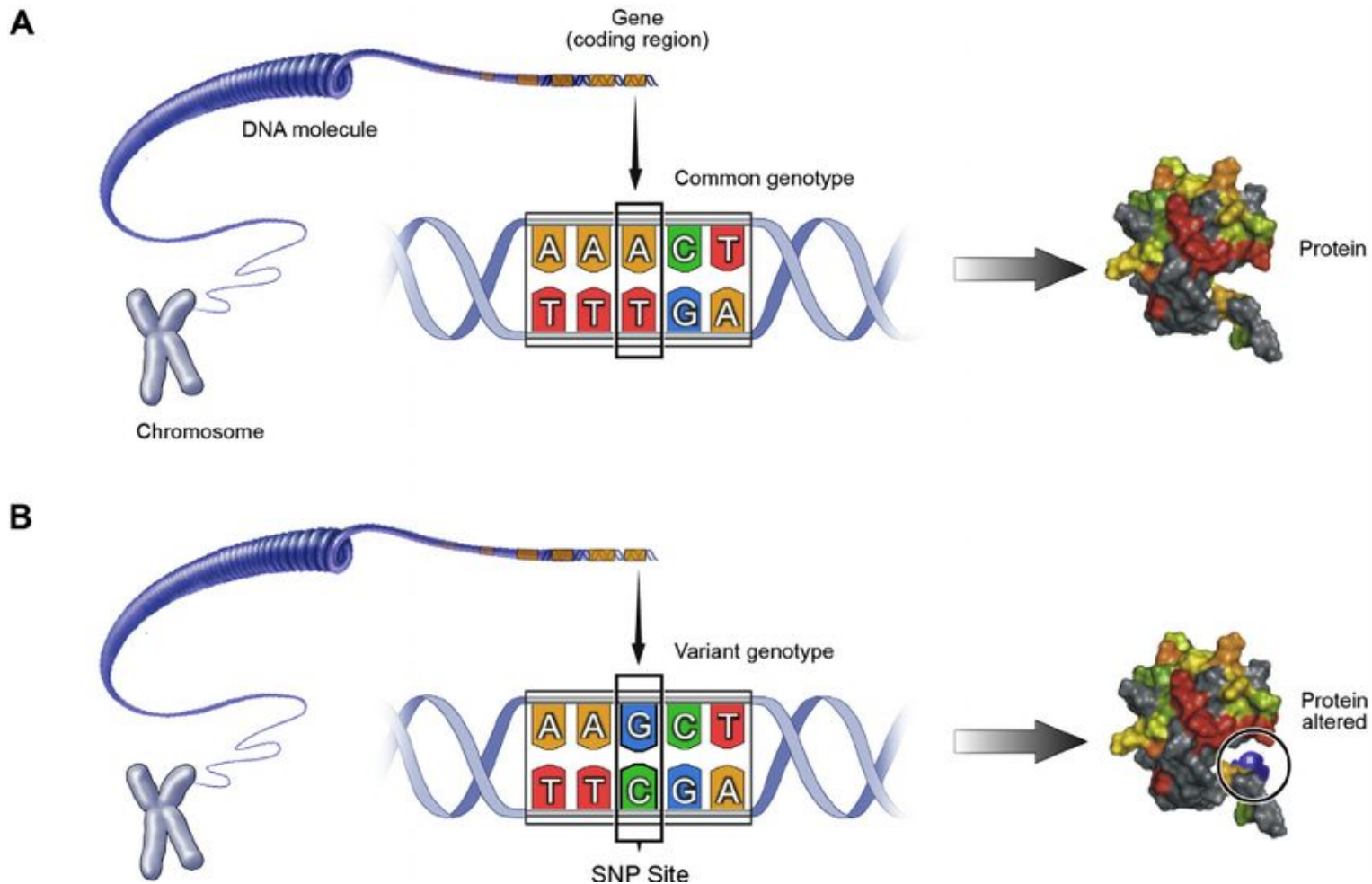
- ✖ 16: [Extract differentially expressed transcripts from RData file](#)
- 15: [Extract expressed transcripts from data 3, and depleted cat](#)
- 14: [Extract expressed transcripts from data 3, and](#)
- ✖ 13: [Extract differentially expressed transcripts](#)
- ✖ 12: [Extract differentially](#)

Hands-on **TOOLS**





Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?





Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

exons.bed

Chrom	Start	End	Name	Score	Strand
chr22	16258185	16258303	uc002zlh.1_cds_1_0_chr22_16258186_r	0	-
chr22	16266928	16267095	uc002zlh.1_cds_2_0_chr22_16266929_r	0	-
chr22	16268136	16268181	uc002zlh.1_cds_3_0_chr22_16268137_r	0	-

snps.bed

Chrom	Start	End	Name	Score	Strand
chr22	16266919	16266920	rs757764551	0	+
chr22	16266920	16266920	rs375488594	0	+
chr22	16267014	16267015	rs544633418	0	+
chr22	16267029	16267030	rs563306354	0	+
chr22	16267081	16267082	rs568292779	0	+



Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

exons.bed

Chrom	Start	End	Name	Score	Strand
chr22	16258185	16258303	uc002zlh.1_cds_1_0_chr22_16258186_r	0	-
chr22	16266928	16267095	uc002zlh.1_cds_2_0_chr22_16266929_r	0	-
chr22	16268136	16268181	uc002zlh.1_cds_3_0_chr22_16268137_r	0	-

snps.bed

Chrom	Start	End	Name	Score	Strand
chr22	16266919	16266920	rs757764551	0	+
chr22	16266920	16266920	rs375488594	0	+
chr22	16267014	16267015	rs544633418	0	+
chr22	16267029	16267030	rs563306354	0	+
chr22	16267081	16267082	rs568292779	0	+

exons



snps





Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

exons.bed

Chrom	Start	End	Name	Score	Strand
chr22	16258185	16258303	uc002zlh.1_cds_1_0_chr22_16258186_r	0	-
chr22	16266928	16267095	uc002zlh.1_cds_2_0_chr22_16266929_r	0	-
chr22	16268136	16268181	uc002zlh.1_cds_3_0_chr22_16268137_r	0	-

snps.bed

Chrom	Start	End	Name	Score	Strand
chr22	16266919	16266920	rs757764551	0	+
chr22	16266920	16266920	rs375488594	0	+
chr22	16267014	16267015	rs544633418	0	+
chr22	16267029	16267030	rs563306354	0	+
chr22	16267081	16267082	rs568292779	0	+

exons



snps



intersect





Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

exons.bed

Chrom	Start	End	Name	Score	Strand
chr22	16258185	16258303	uc002zlh.1_cds_1_0_chr22_16258186_r	0	-
chr22	16266928	16267095	uc002zlh.1_cds_2_0_chr22_16266929_r	0	-
chr22	16268136	16268181	uc002zlh.1_cds_3_0_chr22_16268137_r	0	-

snps.bed

Chrom	Start	End	Name	Score	Strand
chr22	16266919	16266920	rs757764551	0	+
chr22	16266920	16266920	rs375488594	0	+
chr22	16267014	16267015	rs544633418	0	+
chr22	16267029	16267030	rs563306354	0	+
chr22	16267081	16267082	rs568292779	0	+

exons



snps



intersect



exon	count
x	3
z	1
...	...



Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

exons.bed

Chrom	Start	End	Name	Score	Strand
chr22	16258185	16258303	uc002zlh.1_cds_1_0_chr22_16258186_r	0	-
chr22	16266928	16267095	uc002zlh.1_cds_2_0_chr22_16266929_r	0	-
chr22	16268136	16268181	uc002zlh.1_cds_3_0_chr22_16268137_r	0	-

snps.bed

Chrom	Start	End	Name	Score	Strand
chr22	16266919	16266920	rs757764551	0	+
chr22	16266920	16266920	rs375488594	0	+
chr22	16267014	16267015	rs544633418	0	+
chr22	16267029	16267030	rs563306354	0	+
chr22	16267081	16267082	rs568292779	0	+

exons



snps



intersect



exon	count
x	3
z	1
...	...

sort



Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

exons.bed

Chrom	Start	End	Name	Score	Strand
chr22	16258185	16258303	uc002zlh.1_cds_1_0_chr22_16258186_r	0	-
chr22	16266928	16267095	uc002zlh.1_cds_2_0_chr22_16266929_r	0	-
chr22	16268136	16268181	uc002zlh.1_cds_3_0_chr22_16268137_r	0	-

snps.bed

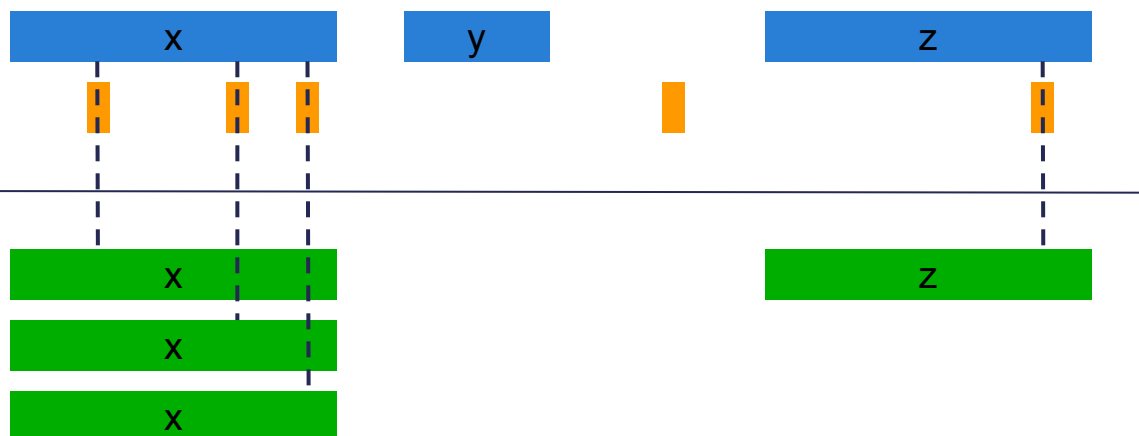
Chrom	Start	End	Name	Score	Strand
chr22	16266919	16266920	rs757764551	0	+
chr22	16266920	16266920	rs375488594	0	+
chr22	16267014	16267015	rs544633418	0	+
chr22	16267029	16267030	rs563306354	0	+
chr22	16267081	16267082	rs568292779	0	+

exons

snps

intersect

1



exon	count
x	3
z	1
...	...

sort

3



Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

1. Get the data in a new history
2. Intersect exons with SNPs
3. Count the number of SNPs per exon
4. Sort exons by SNP count
5. Select top five
6. Recover exon info
7. Display data in UCSC genome browser



Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

1. Get the data in a new history
2. Intersect exons with SNPs *Bedtools -> Intersect intervals*
3. Count the number of SNPs per exon
4. Sort exons by SNP count
5. Select top five
6. Recover exon info
7. Display data in UCSC genome browser

Q : For the first 3 exons in your file, what is the number of SNPs that fall into that exon?



Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

1. Get the data in a new history
2. Intersect exons with SNPs *Bedtools -> Intersect intervals*
3. Count the number of SNPs per exon *Join, Subtract, and Group -> Group*
4. Sort exons by SNP count
5. Select top five
6. Recover exon info
7. Display data in UCSC genome browser

Q : How many exons are there in total in your file?



Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

1. Get the data in a new history
2. Intersect exons with SNPs *Bedtools -> Intersect intervals*
3. Count the number of SNPs per exon *Join, Subtract, and Group -> Group*
4. Sort exons by SNP count *Filter and Sort -> Sort*
5. Select top five
6. Recover exon info
7. Display data in UCSC genome browser

Q : Which exon has the highest number of SNPs in your file?



Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

1. Get the data in a new history
2. Intersect exons with SNPs *Bedtools -> Intersect intervals*
3. Count the number of SNPs per exon *Join, Subtract, and Group -> Group*
4. Sort exons by SNP count *Filter and Sort -> Sort*
5. Select top five *Text Manipulation -> Select First*
6. Recover exon info
7. Display data in UCSC genome browser



Which coding exon has the highest number of single nucleotide polymorphisms on chromosome 22?

1. Get the data in a new history
2. Intersect exons with SNPs *Bedtools -> Intersect intervals*
3. Count the number of SNPs per exon *Join, Subtract, and Group -> Group*
4. Sort exons by SNP count *Filter and Sort -> Sort*
5. Select top five *Text Manipulation -> Select First*
6. Recover exon info *Join, Subtract and Group -> Compare two Datasets*
7. Display data in UCSC genome browser

BED file

Chrom	Start	End	Name	Score	Strand
chr22	16258185	16258303	uc002zlh.1_cds_1_0_chr22_16258186_r	0	-
chr22	16266928	16267095	uc002zlh.1_cds_2_0_chr22_16266929_r	0	-

Cleanup **DATASET**



Dataset

Delete a dataset

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

Tools
search tools
[Get Data](#)
[Send Data](#)
[Collection Operations](#)

COMMON TOOLS
[Text Manipulation](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[Convert Formats](#)
[Extract Features](#)
[Fetch Sequences](#)
[Statistics](#)
[Graph/Display Data](#)
[Fasta Fastq Manipulation](#)

COMMON NGS TOOLS
[NGS:Samtools](#)
[NGS:Mapping](#)
[NGS:Bedtools](#)
[NGS:Picard Tools](#)

SEARCHING TOOLS
[Diamond](#)

✓ **Welcome to galaxy3.sb-roscoff.fr**

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr


Analyses and Bioinformatics for Marine Science

[Changelog](#)
[Tutorials](#)

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History
search datasets
eba 2016 sartools
41 shown, 1 [deleted](#)
1.59 MB
62: SARTools DESeq2 R objects (.RData)
61: SARTools DESeq2 R log
60: SARTools DESeq2 figures
59: SARTools DESeq2 tables
58: SARTools DESeq2 report
57: SARTools edgeR R objects (.RData)
56: SARTools edgeR R log
55: SARTools edgeR figures

Dataset



The dataset isn't really deleted. It's in the Trash

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

Tools
search tools
[Get Data](#)
[Send Data](#)
[Collection Operations](#)

COMMON TOOLS
[Text Manipulation](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[Convert Formats](#)
[Extract Features](#)
[Fetch Sequences](#)
[Statistics](#)
[Graph/Display Data](#)
[Fasta Fastq Manipulation](#)

COMMON NGS TOOLS
[NGS:Samtools](#)
[NGS:Mapping](#)
[NGS:Bedtools](#)
[NGS:Picard Tools](#)

SEARCHING TOOLS
[Diamond](#)

✓ **Welcome to galaxy3.sb-roscoff.fr**

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr


Analyses and Bioinformatics for Marine Science

[Changelog](#)
[Tutorials](#)

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History
search datasets
eba 2016 sartools
42 shown, [hide deleted](#)
1.59 MB

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

62: SARTools DESeq2 R objects (.RData)			
61: SARTools DESeq2 R log			
60: SARTools DESeq2 figures			
59: SARTools DESeq2 tables			
58: SARTools DESeq2 report			
57: SARTools edgeR R objects (.RData)			
56: SARTools edgeR			

javascript:void(0);

Dataset



“Empty Trash” : to free up disk space

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

Tools search tools

Get Data
Send Data
Collection Operations

COMMON TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Statistics
Graph/Display Data
Fasta Fastq Manipulation

COMMON NGS TOOLS
NGS:Samtools
NGS:Mapping
NGS:Bedtools
NGS:Picard Tools

SEARCHING TOOLS
Diamond

Welcome to galaxy3.sb-roscoff.fr

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS 4
Analyses and Bioinformatics for Marine Science

[Changelog](#)
[Tutorials](#)

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History

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

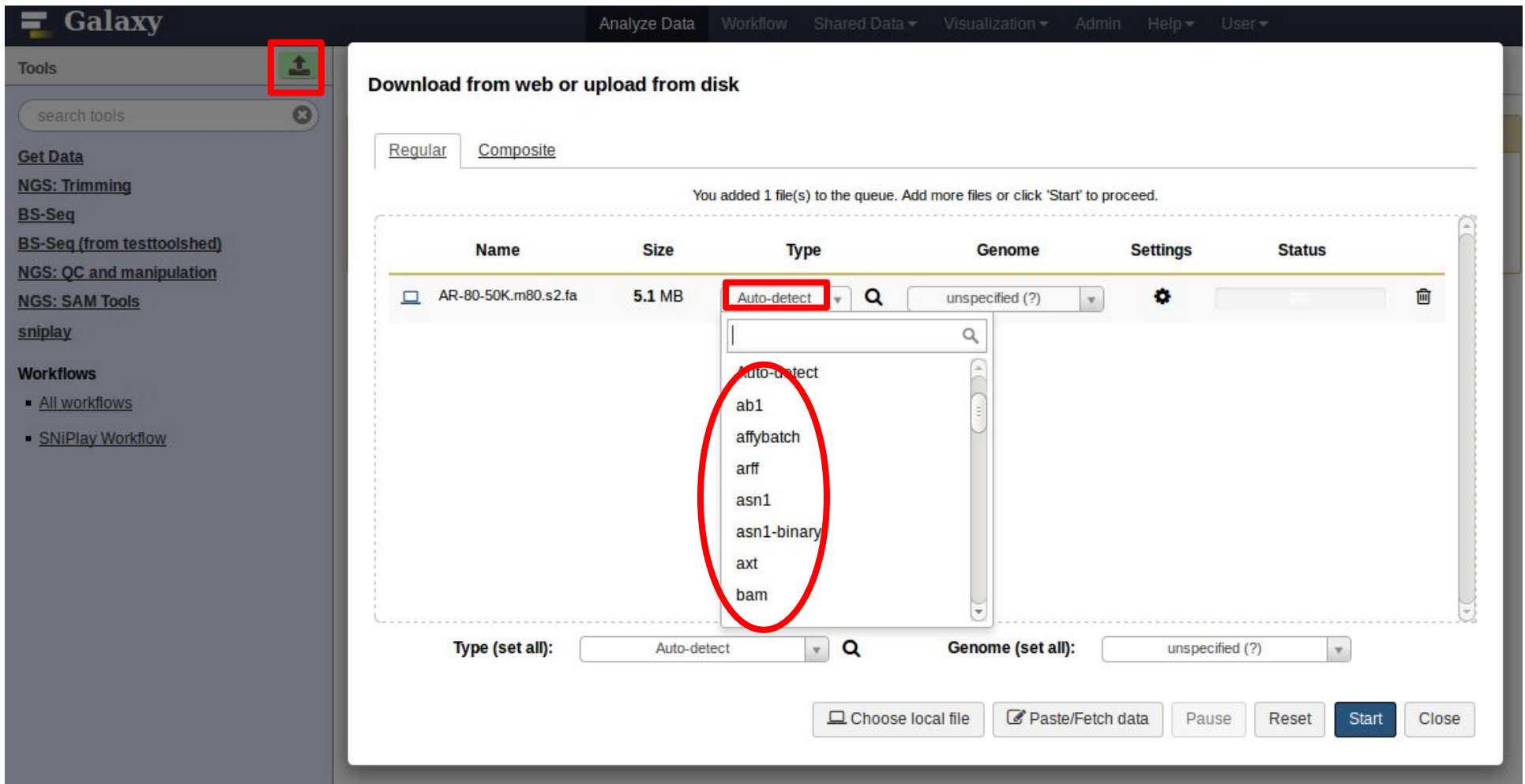
galaxy3.sb-roscoff.fr/history/purge_deleted_datasets

Datatypes

DATASET

Dataset - Datatypes

- Every Galaxy dataset is associated with a datatype.
- Datatype can be detected or user specified.



Galaxy

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

Get Data

NGS: Trimming

BS-Seq

BS-Seq (from testtoolshed)

NGS: QC and manipulation

NGS: SAM Tools

sniplay

Workflows

- All workflows
- SNIPlay Workflow

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
AR-80-50K.m80.s2.fa	5.1 MB	Auto-detect	unspecified (?)		

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

Choose local file Paste/Fetch data Pause Reset Start Close

Common text formats:

- ***txt***: plain text ('.txt')
- ***tabular***: tab delimited ('.tab', '.txt', etc.)

	wt_37_2	wt_37_3	wt_37_1
TR24 c0_g1_i1	90.00	67.00	85.00
TR2779 c0_g1_i1	186.00	137.00	217.00
TR127 c1_g1_i1	9.00	23.00	16.00

- ***csv***: comma-separated values ('.csv')
- ***html***: standard language for web pages

```
Year,Make,Model
1997,Ford,E350
2000,Mercury,Cougar
```

```
<!DOCTYPE html>
<html>
  <head>
    <title>This is a
title</title>
  </head>
  <body>
    <p>Hello world!</p>
  </body>
</html>
```

Dataset - Datatypes

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

Tools

search tools

Get Data
Send Data
Collection Operations

COMMON TOOLS

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

COMMON NGS TOOLS

NGS:Samtools
NGS:Mapping
NGS:Bedtools
NGS:Picard Tools

SEARCHING TOOLS

Diamond

Welcome to galaxy3.sb-roscoff.fr

Information
For any support, send an email at

Text Manipulation

- Add column to an existing dataset
- Concatenate datasets tail-to-head
- Cut columns from a table
- Merge Columns together
- Convert delimiters to TAB
- Create single interval as a new dataset
- Change Case of selected columns
- Paste two files side by side
- Remove beginning of a file
- Select random lines from a file

History

search datasets

eba 2016 sartools
42 shown
1.59 MB

62: SARTools DESeq2 R objects (.RData)

61: SARTools DESeq2 R log

60: SARTools DESeq2 figures

59: SARTools DESeq2 tables

58: SARTools DESeq2 report

57: SARTools edgeR R objects (.RData)

56: SARTools edgeR R log

55: SARTools edgeR figures

151

Dataset - Datatypes


Galaxy / ABiMS

[Analyze Data](#)
[Workflow](#)
[Shared Data](#)
[Visualization](#)
[Admin](#)
[Help](#)
[User](#)

Using 0%

Tools

search tools

[Get Data](#)
[Send Data](#)
[Collection Operations](#)

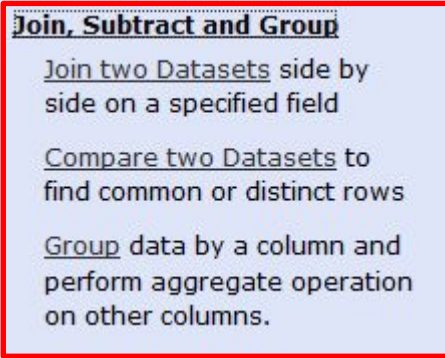
COMMON TOOLS
[Text Manipulation](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[Convert Formats](#)
[Extract Features](#)
[Fetch Sequences](#)
[Statistics](#)
[Graph/Display Data](#)
[Fasta Fastq Manipulation](#)

COMMON NGS TOOLS
[NGS:Samtools](#)
[NGS:Mapping](#)
[NGS:Bedtools](#)
[NGS:Picard Tools](#)

SEARCHING TOOLS
[Diamond](#)




Information
 For any support, send an email at [support@galaxy3.sb-roscoff.fr](#)



Join, Subtract and Group
[Join two Datasets](#) side by side on a specified field
[Compare two Datasets](#) to find common or distinct rows
[Group](#) data by a column and perform aggregate operation on other columns.



Analyses and Bioinformatics for Marine Science

[Changelog](#)
[Tutorials](#)

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History

search datasets

eba 2016 sartools
 42 shown
 1.59 MB

62: SARTools DESeq2 R objects (.RData)

61: SARTools DESeq2 R log

60: SARTools DESeq2 figures

59: SARTools DESeq2 tables

58: SARTools DESeq2 report

57: SARTools edgeR R objects (.RData)

56: SARTools edgeR R log

55: SARTools edgeR figures

Common binary formats:

- ***data***: generic binary format
- ***zip, tar***: archives
- ***pdf, png, jpg, bmp, tiff, gif***: images
- ***rdata***: statistical computing program R
- ***bam, wig, bigwig***: sequence alignment

Sequence file formats:

- ***fasta***: a single-line description with '>', followed by lines of sequence data ('.fasta', '.fas')

```
>sequence1
atgcgtttgcgatgcgcttgcgctgcgcttgcgctgcgcttgcgctg
c
atgcgtttgcgctgc
>sequence2
tttcgtgcgtatagtttcgtgcgtatagtttcgtgcgtatagtttcgtgcgtata
g
tggcgcggt
```

- ***fastq***: sequence + quality score ('.fastq', '.fq')

```
@SEQ_ID
GATTGTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTT
T
+
!'*( ((( (**+) )%%++) (%%%) .1***-+*'') ) **55CCF>>>>>CCCCCCC6
5
@SEQ_ID2
GATTGTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTT
T
+
!'*( ((( (**+) )%%++) (%%%) .1***-+*'') ) **55CCF>>>>>CCCCCCC6
```

Dataset - Datatypes

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

Tools

search tools

Get Data
Send Data
Collection Operations

COMMON TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Statistics
Graph/Display Data
Fasta Fastq Manipulation

COMMON NGS TOOLS
NGS:Samtools
NGS:Mapping
NGS:Bedtools
NGS:Picard Tools

SEARCHING TOOLS
Diamond

Welcome to galaxy3.sb-roscoff.fr

Information
 For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS
 Analyses and Bioinformatics for Marine Science

Fasta Fastq Manipulation
[faSplit](#) Split a FASTA file
[FastQC](#) Read Quality reports
[Filter sequences by length](#)
[FASTQ Groomer](#) convert between various FASTQ quality formats
[Filter sequences by ID](#) from a tabular file

History

search datasets

eba 2016 sartools
 42 shown
 1.59 MB

62: SARTools DESeq2 R objects (.RData)

61: SARTools DESeq2 R log

60: SARTools DESeq2 figures

59: SARTools DESeq2 tables

58: SARTools DESeq2 report

57: SARTools edgeR R objects (.RData)

56: SARTools edgeR R log

55: SARTools edgeR figures

Sequence file formats:

- ***gff3*, *bed*, *genbank***: sequence + annotations

bed

```
track name=pairedReads description="Clone Paired Reads"
useScore=1
chr22 1000 5000 cloneA 960 + 1000 5000 0 2 567,488, 0,3512
chr22 2000 6000 cloneB 900 - 2000 6000 0 2 433,399, 0,3601
```

gff3

```
##gff-version 3
ctg123 . exon 1300 1500 . + .
ID=exon00001
ctg123 . exon 1050 1500 . + .
ID=exon00002
ctg123 . exon 3000 3902 . + .
ID=exon00003
##FASTA
>ctg123
cttctgggcgtacccgattctcggagaacttgccgcaccattccgcctt
g
tgttcattgctgcctgcatgttcattgtctacctcggctacgtgtggct
```

<https://genome.ucsc.edu/FAQ/FAQformat.html#format3>

...

Dataset - Datatypes

- Every Galaxy dataset is associated with a datatype.
- Datatype can be detected or user specified.

The screenshot shows the Galaxy web interface. On the left, the 'Tools' sidebar is visible with a search bar and various tool categories. The main panel displays the 'Download from web or upload from disk' dialog. The 'Regular' tab is selected, and a message indicates that 1 file has been added to the queue. A table lists the file 'AR-80-50K.m80.s2.fa' with a size of 5.1 MB. The 'Type' column for this file shows 'Auto-detect', which is highlighted with a red box. A dropdown menu is open for this entry, showing a list of datatypes: 'Auto-detect', 'ab1', 'affybatch', 'arff', 'asn1', 'asn1-binary', 'axt', and 'bam'. This list is circled in red. At the bottom of the dialog, there are buttons for 'Choose local file', 'Paste/Fetch data', 'Pause', 'Reset', 'Start', and 'Close'.

Dataset - Datatypes

- Every Galaxy dataset is associated with a datatype.
- Datatype can be detected or user specified.

The screenshot shows the Galaxy / ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User. The left sidebar contains a 'Tools' section with a search bar and various tool categories like 'Get Data', 'Send Data', 'Collection Operations', 'COMMON TOOLS', 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Convert Formats', 'Extract Features', 'Fetch Sequences', 'Statistics', 'Graph/Display Data', 'Fasta Fastq Manipulation', 'COMMON NGS TOOLS', 'NGS:Samtools', 'NGS:Mapping', 'NGS:Bedtools', 'NGS:Picard Tools', 'SEARCHING TOOLS', and 'Diamond'.

The main panel displays the 'Change data type' dialog. The 'New Type:' dropdown menu is open, showing a list of datatypes: txt, supermatcher, svg, swiss, syco, tabix, table, **tabular** (highlighted), tagseq, and tandem. A text box next to the dropdown states: 'Changing dataset but *not* modify its contents. and the type of your dataset.'

On the right, the 'History' panel shows a list of datasets. A red arrow points from the 'tabular' datatype in the dialog to the '5: matrix.counts.matrix' dataset in the history. The history list includes:

- 1: samples.txt
- 2: input.matrix.wt 37 vs wt GS NO.DESeq2.DE results
- 3: input.matrix.wt 37 vs wt ph8 .DESeq2.DE results
- 4: input.matrix.wt GSNO vs wt ph8.DESeq2.DE results
- 5: matrix.counts.matrix
- 6: de results
- 7: Extract and cluster differentially expressed transcripts on data 2, data 3, and others

Dataset - Datatypes

- Many tools will only accept input datasets with the appropriate datatype assigned.

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

Tools

trinity

Trinity suite

1- ASSEMBLY:

[Trinity](#) de novo assembly of RNA-Seq data

[Trinity Statistics](#) Obtain basic stats for the number of genes and isoforms and contiguity of the assembly

[Generate gene to transcript map](#) for Trinity assembly

2- COUNTING:

[Align reads and estimate abundance](#) on a de novo assembly of RNA-Seq data

[Build expression matrix](#) for a de novo assembly of RNA-Seq data by Trinity

4- ANNOTATION:

[Filter low expression transcripts](#) from a Trinity assembly

[TransDecoder](#) Find coding

Trinity de novo assembly of RNA-Seq data (Galaxy Version 2.2.0.0) Options

Paired or Single-end data?

Paired

Left/Forward strand reads

No fasta or fastqsanger dataset available.

(--left)

Right/Reverse strand reads

No fasta or fastqsanger dataset available.

(--right)

Strand specific data

Yes No

Jaccard Clip options

Yes No

set if you expect high gene density with UTR overlap (--jaccard_clip)

Run in silico normalization of reads

Yes No

Defaults to max. read coverage of 50. (--normalize_reads)

History

search datasets

eba 2016 sartools
40 shown, 2 deleted
1.59 MB

61: SARTools DESeq2 R log

60: SARTools DESeq2 figures

59: SARTools DESeq2 tables

58: SARTools DESeq2 report

57: SARTools edgeR R objects (.RData)

56: SARTools edgeR R log

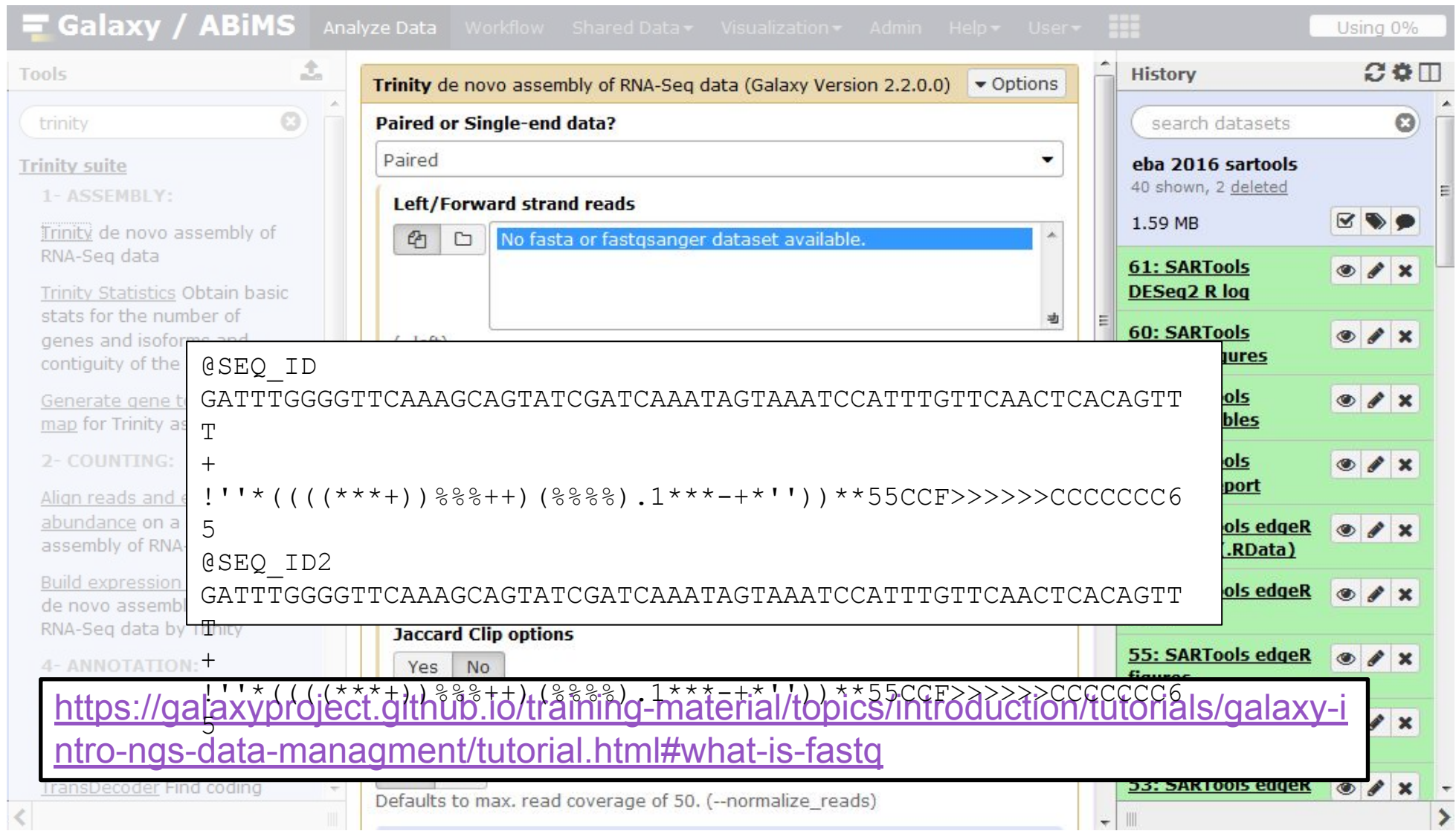
55: SARTools edgeR figures

54: SARTools edgeR tables

53: SARTools edgeR

Dataset - Datatypes

- Many tools will only accept input datasets with the appropriate datatype assigned.



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

Tools

Trinity

Trinity suite

1- ASSEMBLY:

Trinity de novo assembly of RNA-Seq data

Trinity Statistics Obtain basic stats for the number of genes and isoforms and contiguity of the

Generate gene t map for Trinity as

2- COUNTING:

Align reads and abundance on a assembly of RNA

Build expression de novo assembly RNA-Seq data by Trinity

4- ANNOTATION:

TransDecoder Find coding

Trinity de novo assembly of RNA-Seq data (Galaxy Version 2.2.0.0) Options

Paired or Single-end data?

Paired

Left/Forward strand reads

No fasta or fastqsanger dataset available.

Jaccard Clip options

Yes No

Defaults to max. read coverage of 50. (--normalize_reads)

History

search datasets

eba 2016 sartools
40 shown, 2 deleted
1.59 MB

61: SARTools
DESeq2 R log

60: SARTools

55: SARTools edgeR

53: SARTools edgeR

```

@SEQ_ID
GATTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTT
T
+
!''*((( (**+)) %%%++) (%%%) .1***-+*'')) **55CCF>>>>>CCCCCCC6
5
@SEQ_ID2
GATTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTT
T

```

<https://galaxyproject.github.io/training-material/topics/introduction/tutorials/galaxy-intro-ngs-data-managment/tutorial.html#what-is-fastq>

Hands-on

TOOLS (sequence files manipulation)





Part 1:

You have sequencing data from your favorite species. You want to check the quality of your sequences and to map on the reference genome.

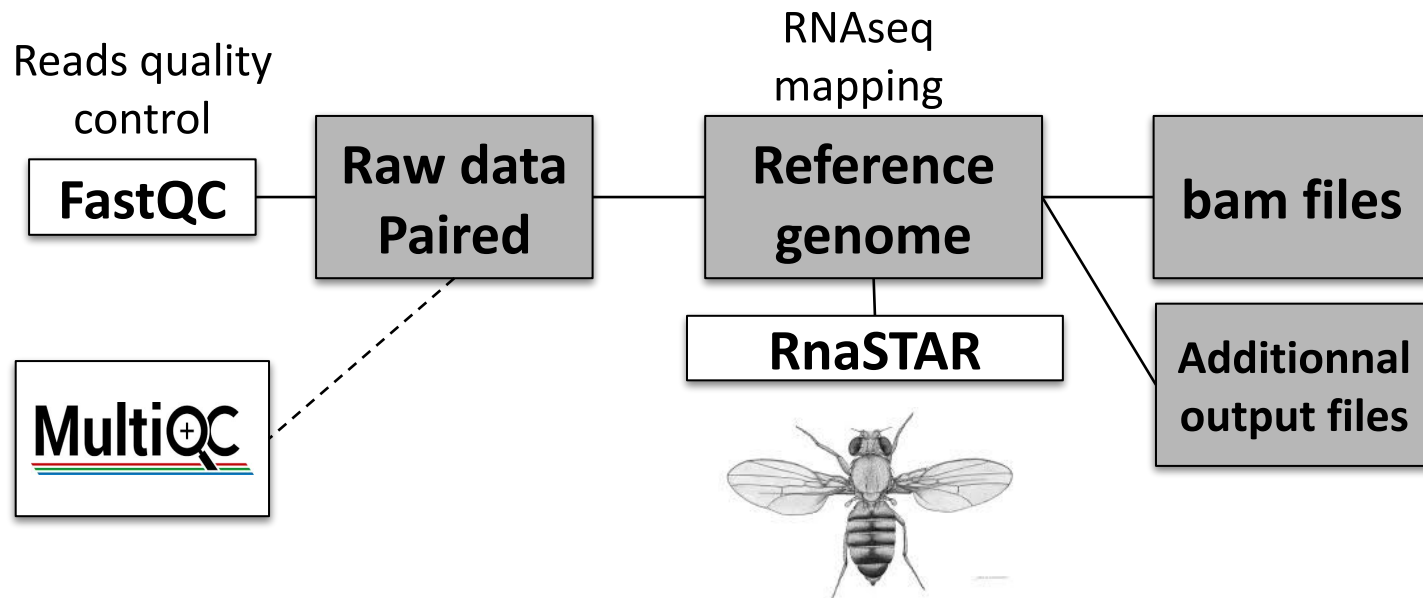
Case of *Drosophila melanogaster* (dm)



Part 1:

You have sequencing data from your favorite species. You want to check the quality of your sequences and to map on the reference genome.

Case of *Drosophila melanogaster* (dm)





Part 1:

You have sequencing data from your favorite species. You want to check the quality of your sequences and to map on the reference genome.

Case of *Drosophila melanogaster* (dm)

1. Look into the description of the tool FastQC, what is its purpose ?
2. Does it takes compressed files ? ".gz"
3. Which encoding of the file (in FastQC results, basic statistics) ? => edit if needed the datatype of your reads files
4. Look at mapping tools, map your reads on genome with RNA STAR.
5. What is the default parameter of the option "Maximum ratio of mismatches to mapped length" in RNA STAR ?
6. How to map on another reference genome or assembly release ?



Part 2:

You have new sequences that you want to compare with the gene and protein databases from your favorite species (BLAST).

Case of *Drosophila melanogaster* (dm)

1. Create new history
2. Import CDS and peptide sequences databases
3. Import query sequences
4. Make BLAST databases
5. Run BLAST against the CDS database
6. Run BLAST against the protein database



Part 2:

You have new sequences that you want to compare with the gene and protein databases from your favorite species (BLAST).

Case of *Drosophila melanogaster* (dm)

1. Create new history
2. Import CDS and peptide sequences databases
3. Import query sequences
4. Make BLAST databases *Blast --> NCBI BLAST+ makeblastdb*
5. Run BLAST against the CDS database
6. Run BLAST against the protein database



Part 2:

You have new sequences that you want to compare with the gene and protein databases from your favorite species (BLAST).

Case of *Drosophila melanogaster* (dm)

1. Create new history
2. Import CDS and peptide sequences databases
3. Import query sequences
4. Make BLAST databases *Blast --> NCBI BLAST+ makeblastdb*
5. Run BLAST against the CDS database *Blast --> NCBI BLAST+ blastn*
6. Run BLAST against the protein database



Part 2:

You have new sequences that you want to compare with the gene and protein databases from your favorite species (BLAST).

Case of *Drosophila melanogaster* (dm)

1. Create new history
2. Import CDS and peptide sequences databases
3. Import query sequences
4. Make BLAST databases *Blast --> NCBI BLAST+ makeblastdb*
5. Run BLAST against the CDS database *Blast --> NCBI BLAST+ blastn*
6. Run BLAST against the protein database *Blast --> NCBI BLAST+ blastx*

DATASET COLLECTION

Dataset collection

- Problematic: you have a large numbers of datasets to send through the same analysis

- Problematic: you have a large numbers of datasets to send through the same analysis
- Solution 1: select multiple datasets as input

Dataset collection

Select multiple datasets as input

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](#)
[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](#)

Workflow: Find exons with highest number of SNPs Run workflow

History Options
Send results to a new history
Yes No

1: Input dataset
2: Exons

2: Input dataset
4: Repeats
3: SNPs
Multiple datasets Exons

This is a batch mode input field. Separate jobs will be triggered for each dataset selection.

Batch options:

3: Join the intervals of two datasets side-by-side (Galaxy Version 1.0.0)

4: Group data by a column and perform aggregate operation on other columns. (Galaxy Version 2.1.1)

History search datasets

Galaxy initiation - multiple datasets
3 shown, 1 deleted
10.41 MB

4: Repeats

3: SNPs

2: Exons

Dataset collection

Select multiple datasets as input

The screenshot displays the Galaxy web interface. At the top, the navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A status indicator on the right shows 'Using 0%'. On the left, a 'Tools' sidebar lists various categories like 'Get Data', 'Send Data', 'Text Manipulation', and 'NGS: DeepTools'. The main workspace features a green notification box with a checkmark icon, stating: 'Successfully invoked workflow **Find exons with highest number of SNPs 2 times.** 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.' To the right, the 'History' pane shows a search bar and a list of datasets. The top entry is 'Galaxy initiation - multiple datasets' (7 shown, 1 deleted, 6 hidden). Below it, a red box highlights four datasets: '14: Top exon genetic location', '13: Top exons', '12: Top exon genetic location', and '10: Top exons'. Further down are '4: Repeats', '3: SNPs', and '2: Exons'. Each dataset entry includes an eye icon, a pencil icon, and a close icon.

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](#)
[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](#)

Successfully invoked workflow **Find exons with highest number of SNPs 2 times.**
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

Galaxy initiation - multiple datasets
7 shown, 1 [deleted](#), 6 [hidden](#)
11.68 MB

14: Top exon genetic location

13: Top exons

12: Top exon genetic location

10: Top exons

4: Repeats

3: SNPs

2: Exons

Dataset collection

- Problematic: you have a large numbers of datasets to send through the same analysis
- Solution 1: select multiple datasets as input
- Solution 2: create a dataset collection

- Problematic: you have a large numbers of datasets to send through the same analysis
- Solution 1: select multiple datasets as input
- Solution 2: create a dataset collection
 - Dataset list: set of files of the same type
 - Dataset pairs: pairs of read files (forward, reverse)
 - List of dataset pairs

- Problematic: you have a large numbers of datasets to send through the same analysis
- Solution 1: select multiple datasets as input
- Solution 2: create a dataset collection
 - Dataset list: set of files of the same type
 - Dataset pairs: pairs of read files (forward, reverse)
 - List of dataset pairs
- Galaxy runs the tool automatically on each dataset in the collection using the same settings

Dataset collection

Create a dataset collection

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User, along with a 'Using 0%' status indicator. The left sidebar contains a 'Tools' section with a search bar and a list of tool categories: Get Data, Send Data, Collection Operations, COMMON TOOLS, Text Manipulation, Filter and Sort, Join, Subtract and Group, Convert Formats, Extract Features, Fetch Sequences, Statistics, Graph/Display Data, Fasta Fastq Manipulation, COMMON NGS TOOLS, NGS:Samtools, NGS:Mapping, NGS:Bedtools, NGS:Picard Tools, SEARCHING TOOLS, and Diamond. The main content area features a green welcome banner for 'galaxy3.sb-roscoff.fr', an information box with contact details, the ABiMS logo (Analyses and Bioinformatics for Marine Science), and links to Changelog and Tutorials. A paragraph describes Galaxy as an open, web-based platform for data intensive biomedical research, supported by various institutions. The right sidebar shows the 'History' section with a search bar and a list of dataset collections: 'Galaxy initiation - collection' (3 shown, 11.32 MB), '3: Repeats', '2: SNPs', and '1: Exons'. A red box highlights the 'Operations on multiple datasets' button next to the '3: Repeats' collection.

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

Tools

search tools

[Get Data](#)
[Send Data](#)
[Collection Operations](#)

COMMON TOOLS

[Text Manipulation](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[Convert Formats](#)
[Extract Features](#)
[Fetch Sequences](#)
[Statistics](#)
[Graph/Display Data](#)
[Fasta Fastq Manipulation](#)

COMMON NGS TOOLS

[NGS:Samtools](#)
[NGS:Mapping](#)
[NGS:Bedtools](#)
[NGS:Picard Tools](#)

SEARCHING TOOLS

[Diamond](#)

Welcome to galaxy3.sb-roscoff.fr

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS
Analyses and Bioinformatics for Marine Science

[Changelog](#)
[Tutorials](#)

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of [BX](#) at [Penn State](#), and the [Biology](#) and [Mathematics and Computer Science](#) departments at [Emory University](#). The Galaxy Project is supported in part by [NHGRI](#), [NSF](#), [The Huck Institutes of the Life Sciences](#), [The Institute for CyberScience](#) at [Penn State](#), and [Emory University](#).

History

search datasets

Galaxy initiation - collection
3 shown
11.32 MB

[All](#) [None](#) [Operations on multiple datasets](#)

☒ **3: Repeats**
☒ **2: SNPs**
☐ **1: Exons**

Dataset collection

Create a dataset collection

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User, along with a 'Using 0%' status indicator. The left sidebar lists various tools and categories, including 'Get Data', 'Send Data', 'Collection Operations', 'COMMON TOOLS', 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Convert Formats', 'Extract Features', 'Fetch Sequences', 'Statistics', 'Graph/Display Data', 'Fasta Fastq Manipulation', 'COMMON NGS TOOLS', 'NGS:Samtools', 'NGS:Mapping', 'NGS:Bedtools', 'NGS:Picard Tools', 'SEARCHING TOOLS', and 'Diamond'. The main content area features a welcome message for 'galaxy3.sb-roscoff.fr', an information box with contact details, the ABiMS logo (Analyses and Bioinformatics for Marine Science), and links to Changelog and Tutorials. The right sidebar shows the 'History' panel with a search bar and a list of datasets. A context menu is open over the 'Galaxy initiation - collection' dataset, highlighting the 'Build Dataset List' option. Other options in the menu include 'Hide datasets', 'Unhide datasets', 'Delete datasets', 'Undelete datasets', 'Permanently delete datasets', 'Build Dataset Pair', and 'Build List of Dataset Pairs'.

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

Tools search tools

Get Data
Send Data
Collection Operations

COMMON TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Statistics
Graph/Display Data
Fasta Fastq Manipulation

COMMON NGS TOOLS
NGS:Samtools
NGS:Mapping
NGS:Bedtools
NGS:Picard Tools

SEARCHING TOOLS
Diamond

Welcome to galaxy3.sb-roscoff.fr

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS
Analyses and Bioinformatics for Marine Science

Changelog
Tutorials

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History search datasets

Galaxy initiation - collection
3 shown
11.32 MB

All None For all selected...

Hide datasets
Unhide datasets
Delete datasets
Undelete datasets
Permanently delete datasets
Build Dataset List
Build Dataset Pair
Build List of Dataset Pairs

Dataset collection

Create a dataset collection

The screenshot shows the Galaxy / ABiMS web interface. A modal dialog box titled "Create a collection from a list of datasets" is open. The dialog contains a text box with the description: "Collections of datasets are permanent, ordered lists of datasets that can be passed to tools and workfl... [More help](#)". Below this, there is a "Start over" link. Two input fields are shown: "Repeats" and "SNPs", each with a "Discard" button to its right. A "Name" label is followed by a text input field containing "Collection of different features". At the bottom left of the dialog is a "Cancel" button, and at the bottom right is a "Create list" button. The background shows the Galaxy navigation menu with categories like "Get Data", "Send Data", "Collection Operations", "COMMON TOOLS", "Text Manipulation", "Filter and Sort", "Join, Subtract and Merge", "Convert Formats", "Extract Features", "Fetch Sequences", "Statistics", "Graph/Display", "Fastq Manipulation", "COMMON NGS TOOLS", "NGS:Samtools", "NGS:Mapping", "NGS:Bedtools", "NGS:Picard Tools", "SEARCHING TOOLS", and "Diamond".

Create a collection from a list of datasets

Collections of datasets are permanent, ordered lists of datasets that can be passed to tools and workfl... [More help](#)

[Start over](#)

Name

Dataset collection

Create a dataset collection

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User, along with a 'Using 0%' status indicator. The left sidebar lists various tool categories: Get Data, Send Data, Collection Operations, COMMON TOOLS (Text Manipulation, Filter and Sort, Join, Subtract and Group, Convert Formats, Extract Features, Fetch Sequences, Statistics, Graph/Display Data, Fasta Fastq Manipulation), COMMON NGS TOOLS (NGS:Samtools, NGS:Mapping, NGS:Bedtools, NGS:Picard Tools), and SEARCHING TOOLS (Diamond). The main content area features a welcome message for galaxy3.sb-roscoff.fr, an information box with contact details, the ABiMS logo (Analyses and Bioinformatics for Marine Science), and links to Changelog and Tutorials. A paragraph describes Galaxy as an open, web-based platform for data intensive biomedical research, supported by various institutions. The right sidebar shows the History panel with a search bar and a list of datasets. The first dataset, '4: Collection of different features', is highlighted with a red box. It shows a size of 11.32 MB and a description 'a list of datasets'. Other datasets listed are '3: Repeats', '2: SNPs', and '1: Exons'.

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

Tools

search tools

Get Data
Send Data
Collection Operations

COMMON TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Statistics
Graph/Display Data
Fasta Fastq Manipulation

COMMON NGS TOOLS
NGS:Samtools
NGS:Mapping
NGS:Bedtools
NGS:Picard Tools

SEARCHING TOOLS
Diamond

Welcome to galaxy3.sb-roscoff.fr

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS
Analyses and Bioinformatics for Marine Science

Changelog
Tutorials

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History

search datasets

Galaxy initiation - collection
4 shown

11.32 MB

All None For all selected...

☐ **4: Collection of different features**
a list of datasets

☐ **3: Repeats**

☐ **2: SNPs**

☐ **1: Exons**

Dataset collection

Create a dataset collection

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User, along with a 'Using 0%' status indicator. The left sidebar lists various tool categories: Get Data, Send Data, Collection Operations, COMMON TOOLS (Text Manipulation, Filter and Sort, Join, Subtract and Group, Convert Formats, Extract Features, Fetch Sequences, Statistics, Graph/Display Data, Fasta Fastq Manipulation), COMMON NGS TOOLS (NGS:Samtools, NGS:Mapping, NGS:Bedtools, NGS:Picard Tools), and SEARCHING TOOLS (Diamond). The main content area features a green welcome banner for 'galaxy3.sb-roscoff.fr', an information box with contact details, and the ABiMS logo with the tagline 'Analyses and Bioinformatics for Marine Science'. Below the logo are links for Changelog and Tutorials. A paragraph describes Galaxy as an open, web-based platform for data intensive biomedical research, mentioning its affiliation with Penn State and Emory University. The right sidebar shows a History section with a 'Back to Galaxy initiation - collection' link and a red-bordered box containing 'Collection of different features' (a list of datasets), 'Repeats', and 'SNPs', each with view and edit icons.

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

Tools

search tools

[Get Data](#)
[Send Data](#)
[Collection Operations](#)

COMMON TOOLS
[Text Manipulation](#)
[Filter and Sort](#)
[Join, Subtract and Group](#)
[Convert Formats](#)
[Extract Features](#)
[Fetch Sequences](#)
[Statistics](#)
[Graph/Display Data](#)
[Fasta Fastq Manipulation](#)

COMMON NGS TOOLS
[NGS:Samtools](#)
[NGS:Mapping](#)
[NGS:Bedtools](#)
[NGS:Picard Tools](#)

SEARCHING TOOLS
[Diamond](#)

Welcome to galaxy3.sb-roscoff.fr

Information
For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr

ABiMS
Analyses and Bioinformatics for Marine Science

[Changelog](#)
[Tutorials](#)

Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University.

History

[Back to Galaxy initiation - collection](#)

Collection of different features
a list of datasets

Repeats

SNPs

Dataset collection

Use a collection as input

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 under categories like 'Get Data', 'Send Data', 'Text Manipulation', and 'Convert Formats'. The main panel displays the 'BED-to-GFF converter (Galaxy Version 2.0.0)' tool. In the 'Convert this dataset' section, the 'Dataset collection' icon (a folder) is highlighted with a red box, and a tooltip 'Dataset collection' is shown. The dropdown menu shows '5: Collection of different features'. Below this is an 'Execute' button. The 'What it does' section explains that the tool converts BED format data to GFF format. The 'Example' section shows a sample BED format table and its corresponding GFF output. The right sidebar shows the 'History' panel with a search bar and a list of datasets, including 'Galaxy initiation - collection' and '5: Collection of different features'.

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

BED-to-GFF converter (Galaxy Version 2.0.0) Options

Convert this dataset

5: Collection of different features

Dataset collection

Execute

What it does

This tool converts data from BED format to GFF format (scroll down for format description).

Example

The following data in BED format:

chr28	346187	388197	BC114771	0	+	346187	388197	0
chr28	346187	388197	BC114771	0	+	346187	388197	0

Will be converted to GFF (note that the start coordinate is incremented by 1):

chr28	bed2gff	mRNA	346188	388197	0	+	.	mRNA	BC1147
chr28	bed2gff	exon	346188	346331	0	+	.	exon	BC1147
chr28	bed2gff	exon	370283	370363	0	+	.	exon	BC1147
chr28	bed2gff	exon	372378	372492	0	+	.	exon	BC1147
chr28	bed2gff	exon	377194	377256	0	+	.	exon	BC1147
chr28	bed2gff	exon	378319	378473	0	+	.	exon	BC1147
chr28	bed2gff	exon	379722	379817	0	+	.	exon	BC1147
chr28	bed2gff	exon	383182	383315	0	+	.	exon	BC1147
chr28	bed2gff	exon	387981	388085	0	+	.	exon	BC1147

History

search datasets

Galaxy initiation - collection
4 shown, 5 deleted, 6 hidden

11.68 MB

5: Collection of different features
a list of 2 datasets

4: Repeats

2: SNPs

1: Exons

Dataset collection

Use a collection as input

The screenshot displays the Galaxy web interface. At the top, the navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A 'Using 0%' indicator is on the right. The left sidebar contains a 'Tools' section with a search bar and a list of tool categories: 'Get Data', 'Send Data', 'Lift-Over', 'Text Manipulation', 'Datamash', and 'Convert Formats'. Under 'Convert Formats', several tools are listed, including 'Convert BAM to ScIdx', 'Tabular-to-FASTA', 'FASTA-to-Tabular', 'Tabular to FASTQ', 'FASTQ to Tabular', 'FASTQ to FASTA', 'FASTQ to FASTA', 'BED-to-GFF', 'GFF-to-BED', 'MAF to BED', and 'MAF to Interval'.

The main content area features a large green notification box with a checkmark icon. It states: '2 jobs have been successfully added to the queue - resulting in the following datasets:'. Below this, it lists '16: BED-to-GFF on data 4' and '17: BED-to-GFF on data 2'. A paragraph follows: '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.'

The right sidebar shows the 'History' pane. It has a search bar and displays a list of dataset collections. The first collection is 'Galaxy initiation - collection' with 7 shown, 5 deleted, and 6 hidden datasets, totaling 11.68 MB. Below it, three collections are highlighted with a red box: '18: BED-to-GFF on collection 5' (a list of datasets), '17: BED-to-GFF on data 2', and '16: BED-to-GFF on data 4'. Each of these three collections has a green background and a close button. Below the red box, there are three more collections: '5: Collection of different features' (a list of 2 datasets), '4: Repeats', '2: SNPs', and '1: Exons'. Each of these also has a green background and a close button.

Hands-on **COLLECTION**





Part 2:

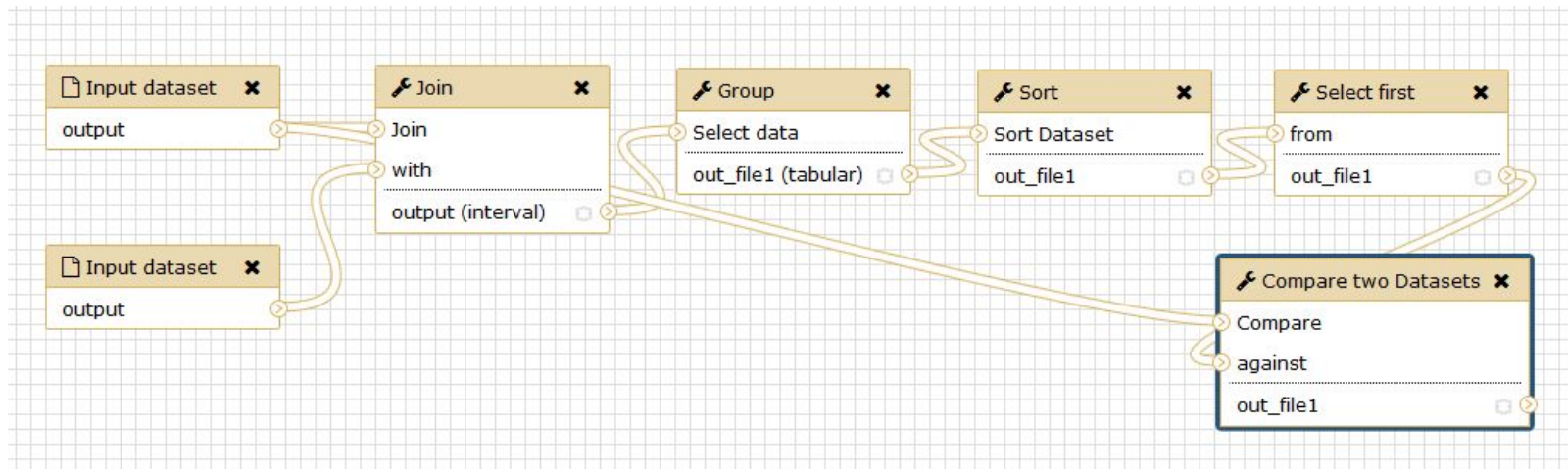
You have new sequences that you want to compare with the gene and protein databases from your favorite species (BLAST).

Case of *Drosophila melanogaster* (dm)

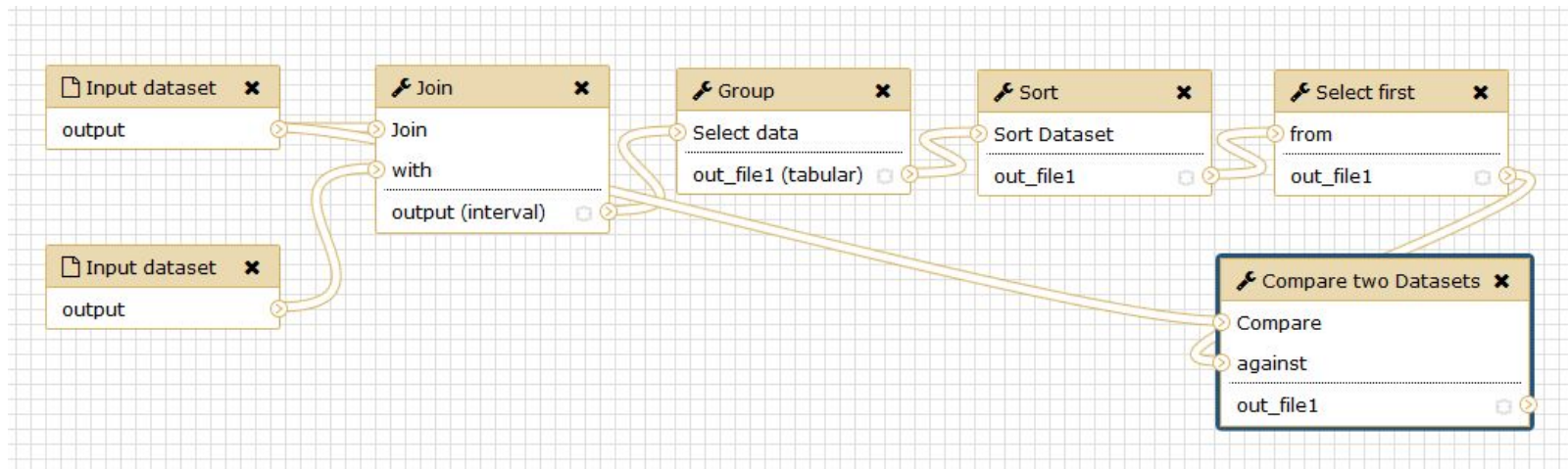
1. Create new history
2. Import CDS and peptide sequences databases
3. Import query sequences
4. Make BLAST databases
5. Run BLAST against the CDS database
6. Run BLAST against the protein database
7. Create a dataset list with all the query sequences
8. Run BLAST against the CDS database on the dataset list

WORKFLOW

Workflow



- A workflow is a sequence of tool operations and parameters
- A workflow is built to be replayed (more or less strict)

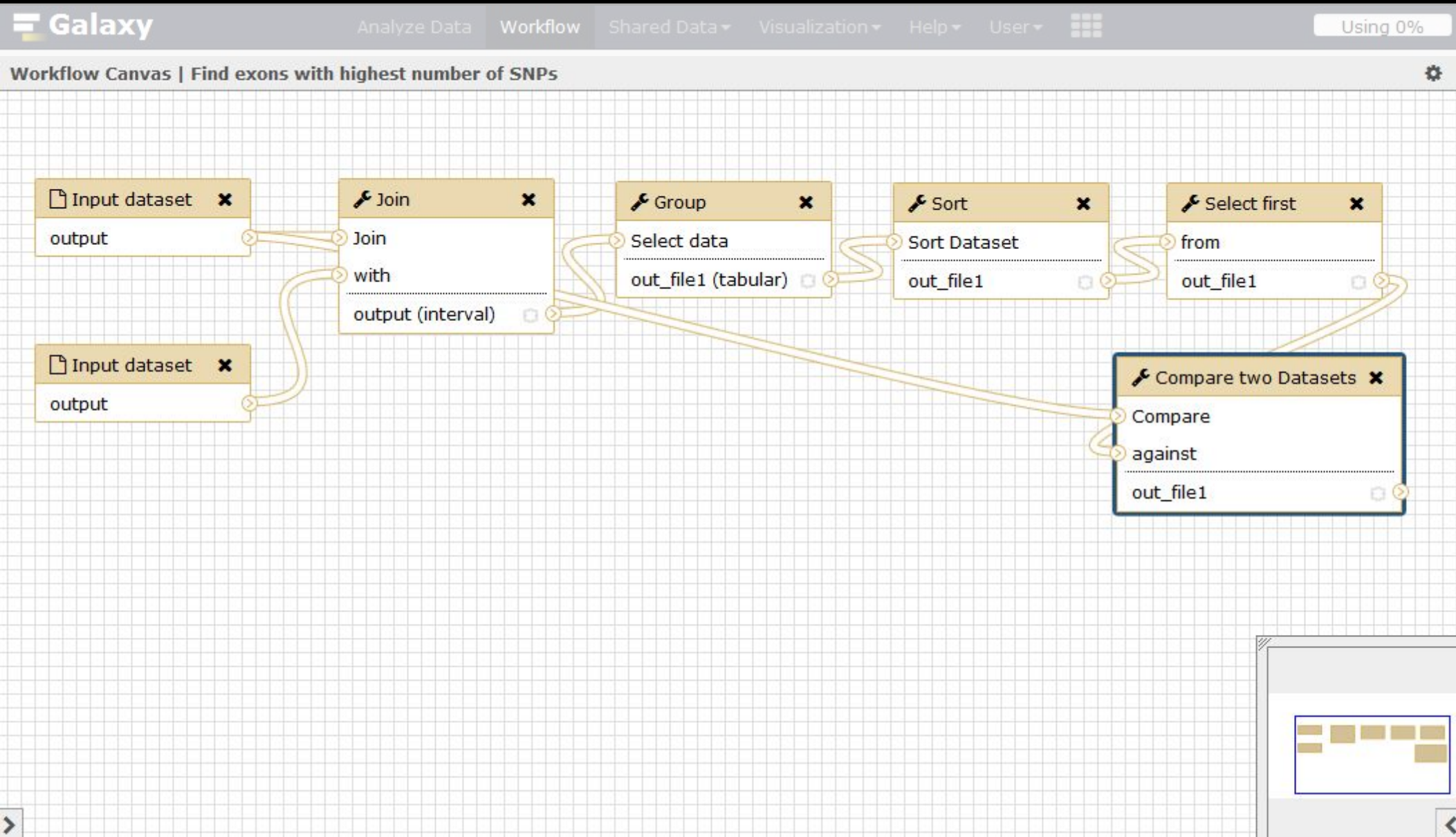


Why would you want to create workflows?

- Re-run the same analysis on different input data sets
- Change parameters before re-running a similar analysis
- Make use of the workflow job scheduling (jobs are submitted as soon as their inputs are ready)
- Share workflows for publication and with the community

Workflow

Our workflow with Galaxy



Workflow

From history

 Galaxy

Analyze Data Workflow Shared Data Visualization Help User

Using 0%

The following list contains each tool that was run to create the datasets in your current history. Please select those that you wish to include in the workflow.

Tools which cannot be run interactively and thus cannot be incorporated into a workflow will be shown in gray.

Workflow name

Workflow constructed from history 'Galaxy initiation'

Create Workflow Check all Uncheck all

Tool	History items created
UCSC Main <i>This tool cannot be used in workflows</i>	1 Exons ☒ Treat as input dataset Exons
UCSC Main <i>This tool cannot be used in workflows</i>	2 SNPs ☒ Treat as input dataset SNPs
Join ☒ Include "Join" in workflow	3 Join on data 2 and data 1
Group ☒ Include "Group" in workflow	5 Group on data 3
Sort ☒ Include "Sort" in workflow	6 Sort on data 5

https://usegalaxy.org/workflow/build_from_current_history

History



HISTORY LISTS

Saved Histories

Histories Shared with Me

CURRENT HISTORY

Create New

Copy History

Share or Publish

Show Structure

7: Extract Workflow

6: Delete

6: Delete Permanently

DATASET ACTIONS

5: Copy Datasets

3: Dataset Security

da: Resume Paused Jobs

2: Collapse Expanded Datasets

1: Unhide Hidden Datasets

Delete Hidden Datasets

Purge Deleted Datasets

DOWNLOADS

Export Tool Citations

Export History to File

OTHER ACTIONS

Import from File

Workflow

From history

Galaxy

Analyze DataWorkflowShared DataVisualizationHelpUser

Using 0%

The following list contains each tool that was run to create the datasets in your current history. Please select those that you wish to include in the workflow.

Tools which cannot be run interactively and thus cannot be incorporated into a workflow will be shown in gray.

Workflow name

Find exons with the highest SNPs

Create WorkflowCheck allUncheck all

Tool

History items created

UCSC Main

This tool cannot be used in workflows

UCSC Main

This tool cannot be used in workflows

Join

☒ Include "Join" in workflow

Group

☒ Include "Group" in workflow

Sort

☒ Include "Sort" in workflow

1 Exons

☒ Treat as input datasetExons

2 SNPs

☒ Treat as input datasetSNPs

3 Join on data 2 and data 1

5 Group on data 3

6 Sort on data 5

History

search datasets

Galaxy initiation
7 shown, 1 deleted, 1 hidden
8.77 MB

8: Compare two Datasets on data 7 and data 1

7: Select first on data 6

6: Sort on data 5

5: Group on data 3

3: Join on data 2 and data 1

2: SNPs

1: Exons

Workflow

From history

Galaxy

Analyze Data Workflow Shared Data Visualization Help User

Using 0%

The following list contains each tool that was run to create the datasets in your current history. Please select those that you wish to include in the workflow.

Tools which cannot be run interactively and thus cannot be incorporated into a workflow will be shown in gray.

Workflow name

Find exons with the highest SNPs

Create Workflow Check all Uncheck all

Tool	History items created
UCSC Main This tool cannot be used in workflows	1 Exons ☒ Treat as input dataset Exons
UCSC Main This tool cannot be used in workflows	2 SNPs ☒ Treat as input dataset SNPs
Join ☒ Include "Join" in workflow	3 Join on data 2 and data 1
Group ☒ Include "Group" in workflow	5 Group on data 3
Sort ☒ Include "Sort" in workflow	6 Sort on data 5

History

search datasets

Galaxy initiation
7 shown, 1 deleted, 1 hidden
8.77 MB

8: Compare two Datasets on data 7 and data 1

7: Select first on data 6

6: Sort on data 5

5: Group on data 3

3: Join on data 2 and data 1

2: SNPs

1: Exons

192

Workflow

Workflow manager

Your workflows

 Create new workflow

 Upload or import workflow

Name	# of Steps
Find exons with highest number of SNPs ▾	7
Convert to tab (imported from API) ▾	2
imported: ChIP-seq workflow ▾	3

Workflows shared with you by others

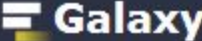
No workflows have been shared with you.

Other options

Configure your workflow menu ▾

Workflow

Workflow manager

 Galaxy

Analyze Data

Workflow

Shared Data ▾

Visualization ▾

Help ▾

User ▾



Using 0%

Your workflows

Create new workflow

Upload or import workflow

Name	# of Steps
Find exons with highest number of SNPs ▾	7
Convert to tab (imported from ...)	2
imported: ChIP-seq workflow	3

Workflows shared by others

No workflows have been shared by others

Other options

Configure your workflow menu

Edit

Run

Share or Download

Copy

Rename

View

Delete

<https://usegalaxy.org/workflow/editor?id=17b7895387cc2214>

Workflow

Edit a workflow: add tags and annotation

Galaxy

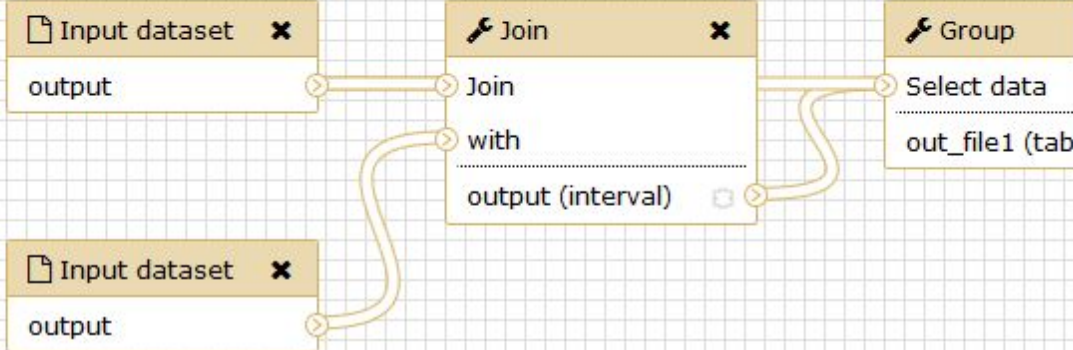
Analyze DataWorkflowShared DataVisualizationHelpUserUsing 0%

Tools

search tools

Inputs
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

Workflow Canvas | Find exons with highest number of SNPs



```
graph LR; A[Input dataset  
output] --> C[Join  
with  
output (interval)]; B[Input dataset  
output] --> C; C --> D[Group  
Select data  
out_file1 (tab)];
```

Details

Edit Workflow Attributes

Name:
Find exons with highest number of SNPs

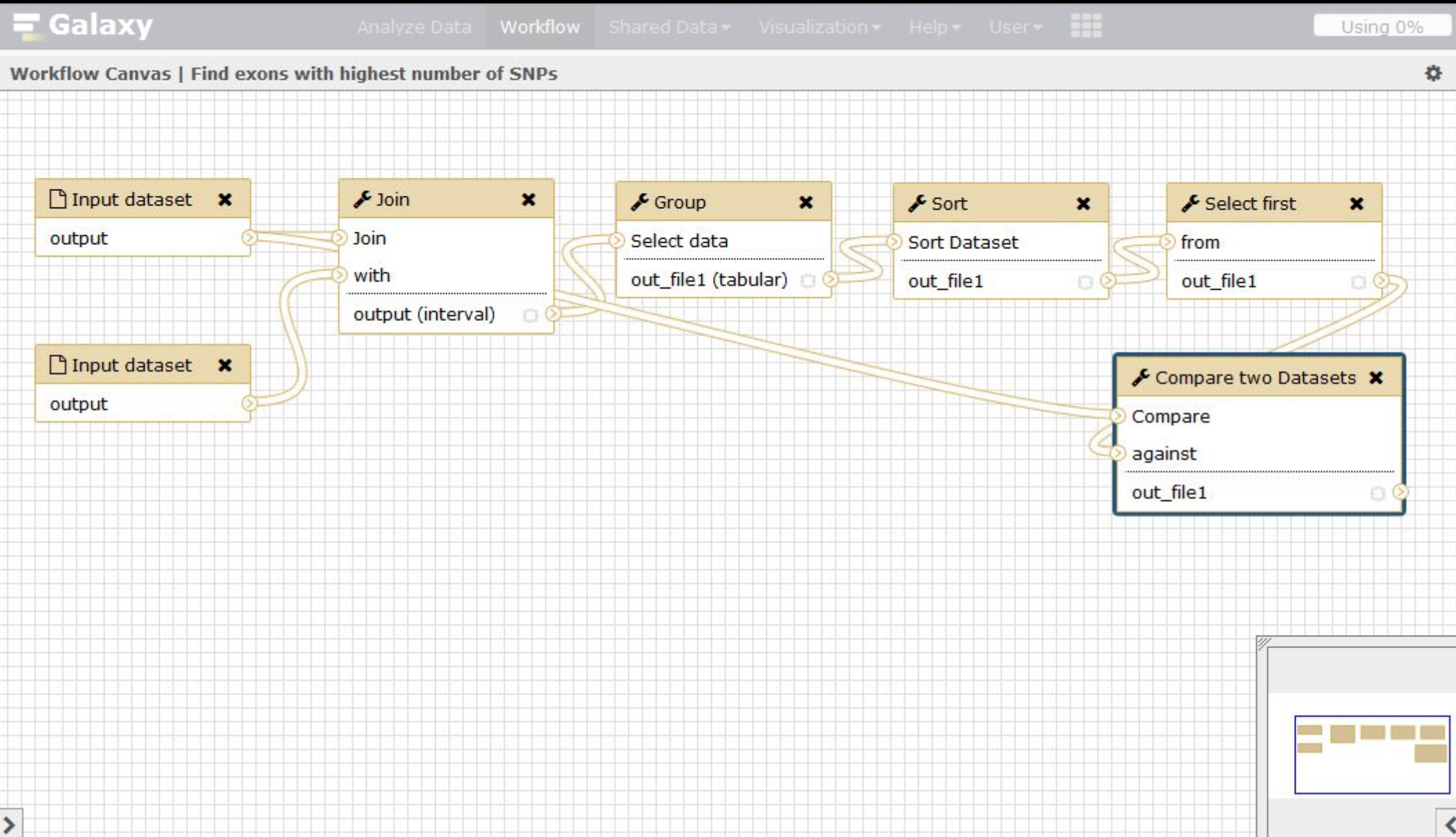
Tags:

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

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

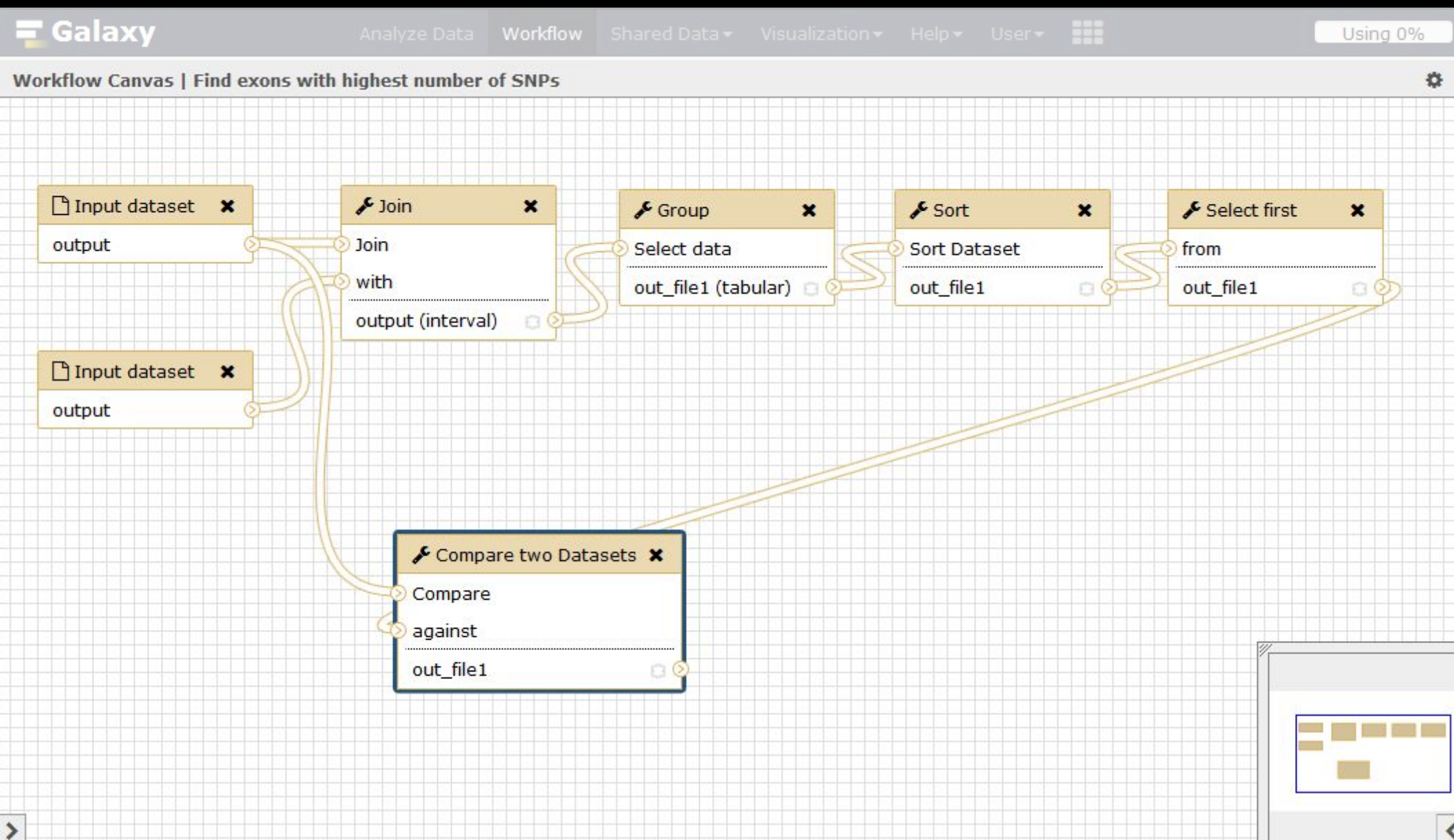
Workflow

Edit a workflow



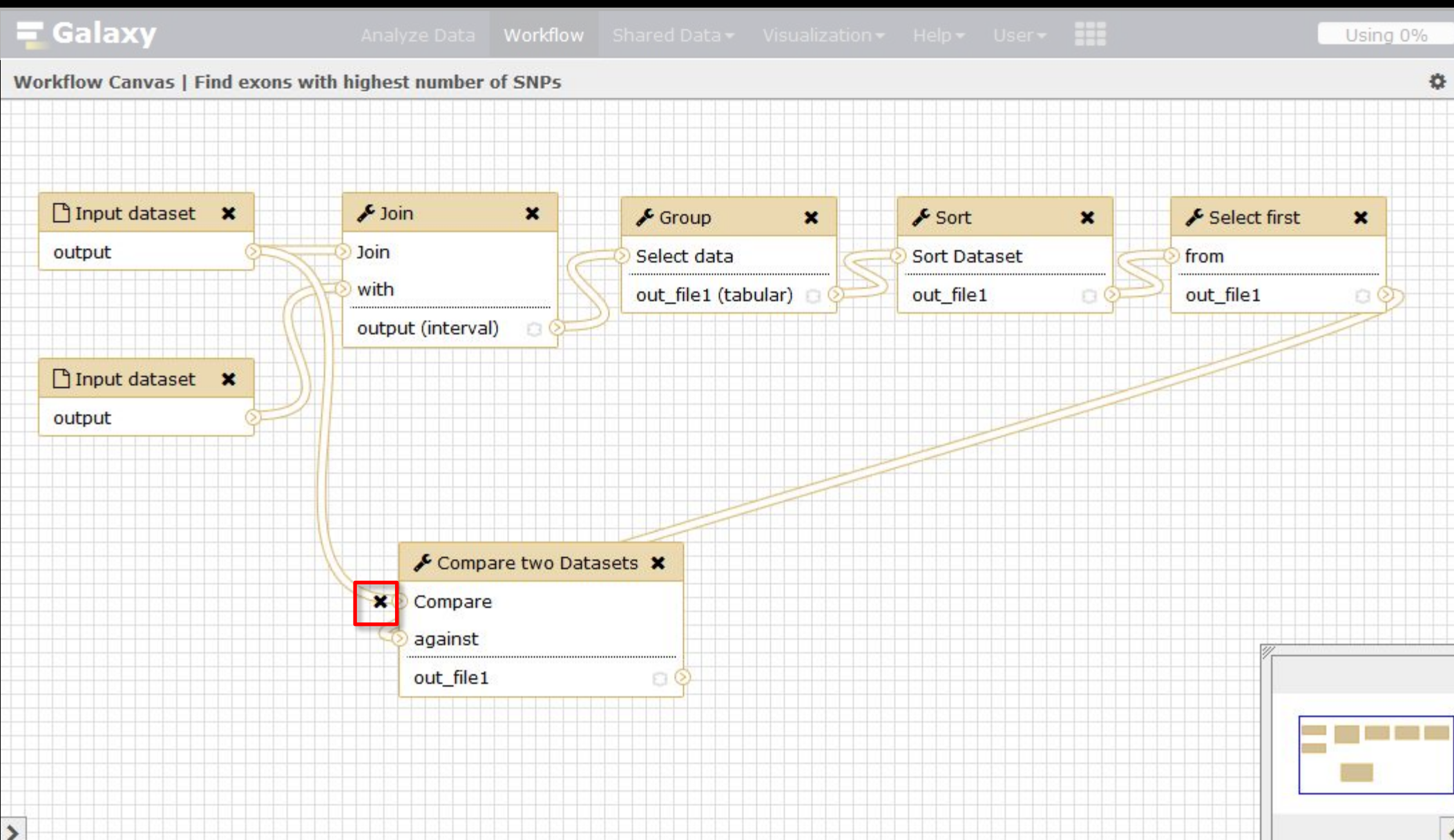
Workflow

Edit a workflow: drag and drop



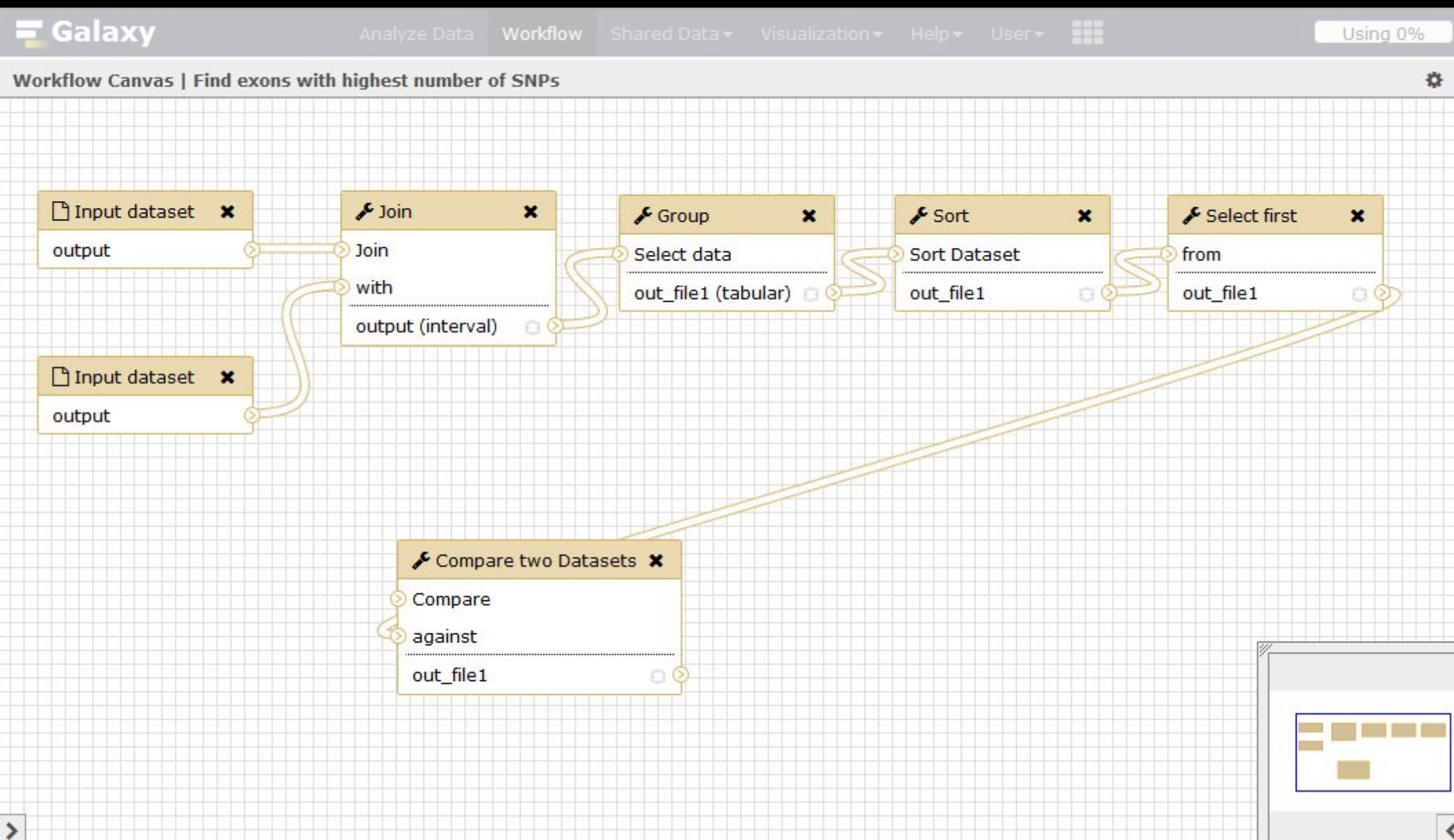
Workflow

Edit a workflow: delete a noodle



Workflow

Edit a workflow: delete a noodle



Workflow

Edit a workflow: add a tool

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

Tools

search tools

Get Data

Send Data

Lift-Over

Text Manipulation

- UniProt ID mapping and retrieval
- Compute an expression on every row
- Concatenate datasets tail-to-head
- Add column to an existing dataset
- Concatenate datasets tail-to-head (cat)
- tac reverse a file (reverse cat)
- Condense consecutive characters
- Cut columns from a table
- Convert delimiters to TAB
- Merge Columns together**
- Remove beginning of a file

Workflow Canvas | Find exons with highest number of SNPs

```
graph LR; I1[Input dataset] --> J[Join]; I2[Input dataset] --> J; J --> G[Group]; G --> S[Sort]; S --> MC[Merge Columns]; MC --> C[Compare two Datasets];
```

Workflow steps:

- Input dataset (output)
- Join (with output (interval))
- Group (Select data out_file1 (tabular))
- Sort (Sort Dataset out_file1)
- Merge Columns** (Select data out_file1 (tabular))
- Compare two Datasets (Compare against out_file1)

Workflow

Edit a workflow: add a noodle

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

Tools

search tools

Get Data

Send Data

Lift-Over

Text Manipulation

- UniProt ID mapping and retrieval
- Compute an expression on every row
- Concatenate datasets tail-to-head
- Add column to an existing dataset
- Concatenate datasets tail-to-head (cat)
- tac reverse a file (reverse cat)
- Condense consecutive characters
- Cut columns from a table
- Convert delimiters to TAB
- Merge Columns together
- Remove beginning of a file

Workflow Canvas | Find exons with highest number of SNPs

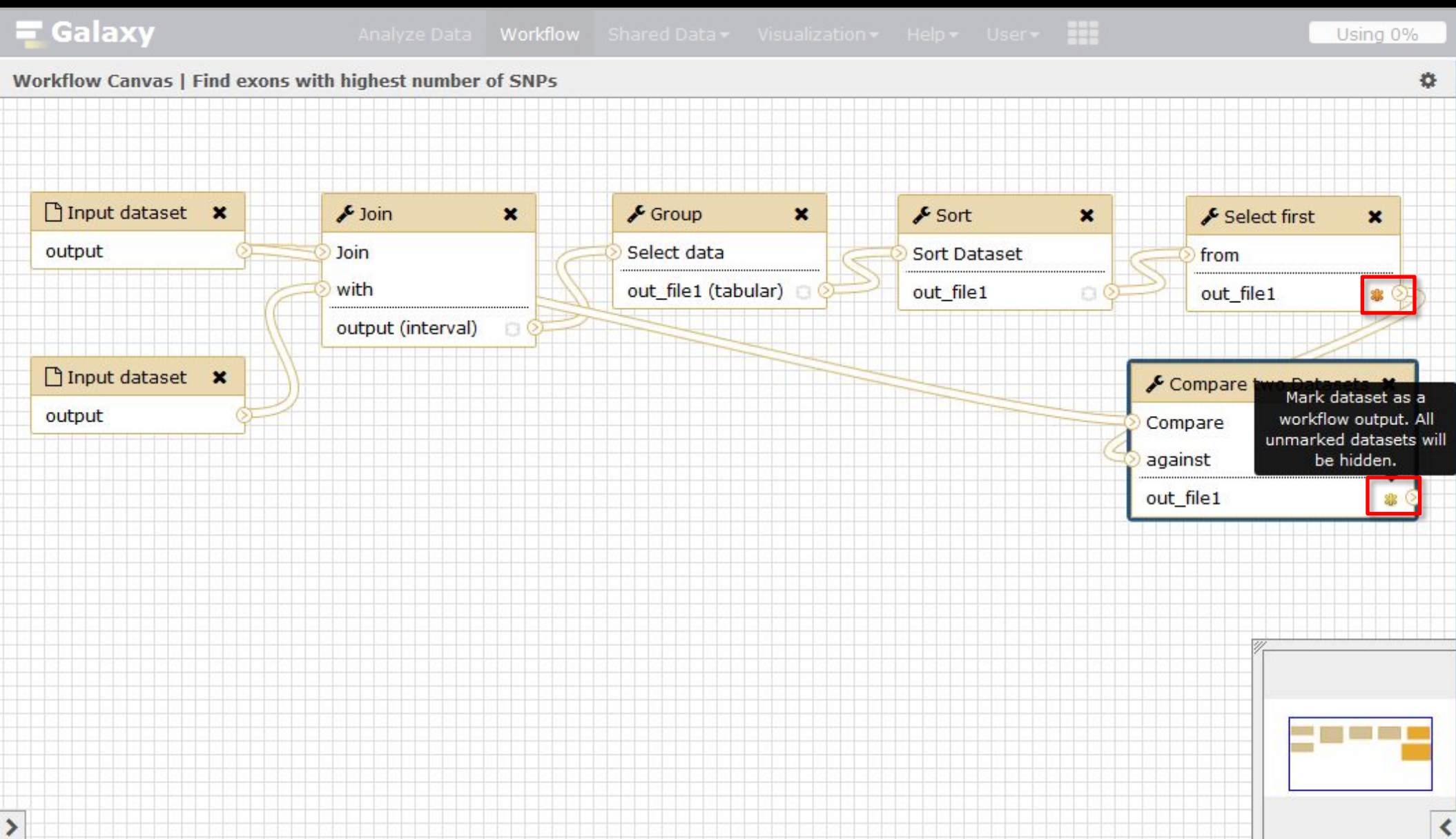
```
graph LR; A[Input dataset] --> B[Join]; C[Input dataset] --> B; B --> D[Merge Columns]; D --> E[Group]; E --> F[Sort]; F --> G[Compare two Datasets];
```

Workflow Steps:

- Input dataset (output)
- Join (with output (interval))
- Merge Columns (Select data: out_file1 (tabular))
- Group (Select data: out_file1 (tabular))
- Sort (Sort Dataset: out_file1)
- Compare two Datasets (Compare: out_file1)

Workflow

Edit a workflow: hide intermediate steps



Workflow

Edit a workflow: set or release a parameter

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

Workflow Canvas | Find exons with highest number of SNPs

```
graph LR; A[Group] --> B[Sort]; B --> C[Select first]; C --> D[Compare two Datasets];
```

The workflow consists of the following steps:

- Group**: Select data from `out_file1 (tabular)`.
- Sort**: Sort Dataset from `out_file1`.
- Select first** (highlighted): Select first lines from a dataset (Galaxy Version 1.0.0). The 'from' parameter is set to `out_file1`.
- Compare two Datasets**: Compare against `out_file1`.

Details

Select first (Galaxy Version 1.0.0)

Select first

5 lines

from

Data input 'input' (txt)

Annotation / Notes

Add an annotation or note for this step. It will be shown with the workflow.

Email notification

Yes No

An email notification will be sent when the job has completed.

Output cleanup

Yes No

Workflow

Edit a workflow: set or release a parameter

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

Workflow Canvas | Find exons with highest number of SNPs

```
graph LR; A[Select data  
out_file1 (tabular)] --> B[Sort Dataset  
out_file1]; B --> C[Select first  
from  
out_file1]; C --> D[Compare two Datasets  
Compare  
against  
out_file1];
```

The workflow canvas displays four steps connected by arrows. The 'Select first' step is highlighted with a blue border. A red box highlights the 'Select first' icon in the details panel. A tooltip 'Set at Runtime' is visible next to the icon.

Details

Select first lines from a dataset (Galaxy Version 1.0.0)

☒ **Select first**

Set at Runtime

lines

from

Data input 'input' (txt)

Annotation / Notes

Add an annotation or note for this step. It will be shown with the workflow.

Email notification

Yes No

An email notification will be sent when the job has completed.

Output cleanup

Yes No

Workflow

Edit a workflow: rename the outputs

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

Workflow Canvas | Find exons with highest number of SNPs

Group Select data
out_file1 (tabular)

Sort Sort Dataset
out_file1

Select first from
out_file1

Compare two Datasets Compare
against
out_file1

Details
step. It will be shown with the workflow.

Email notification
Yes No
An email notification will be sent when the job has completed.

Output cleanup
Yes No
Upon completion of this step, delete non-starred outputs from completed workflow steps if they are no longer required as inputs.

Configure Output: 'out_file1'

Label

This will provide a short name to describe the output - this must be unique across workflows.

Rename dataset

This action will rename the output dataset. Click [here](#) for more information. Valid inputs are: **input**.

Workflow

Save

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

Workflow Canvas | Find exons with highest number of SNPs

```
graph LR; A[Group: Select data  
out_file1 (tabular)] --> B[Sort: Sort Dataset  
out_file1]; B --> C[Select first: from  
out_file1]; C --> D[Compare two Datasets: Compare  
against  
out_file1];
```

Details

Save
Save As
Run
Edit Attributes
Auto Re-layout
Close

from
Data input 'input' (txt)
Annotation / Notes
Add an annotation or note for this step. It will be shown with the workflow.
Email notification
Yes No
An email notification will be sent when the job has completed.
Output cleanup
Yes No

Workflow

Run a workflow

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

Your workflows

Create new workflow Upload or import workflow

Name	# of Steps
Find exons with highest number of SNPs	7
Convert to tab (imported from ...)	2
imported: ChIP-seq workflow	3

Workflows shared by others
No workflows have been shared by others

Other options
Configure your workflow menu

https://usegalaxy.org/root?workflow_id=17b7895387cc2214

Workflow

Run a workflow

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
- 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

Workflow: Find exons with highest number of SNPs Run workflow

History Options

Send results to a new history
Yes No

1: Input dataset
2: Exons

2: Input dataset
1: Repeats

3: Join the intervals of two datasets side-by-side (Galaxy Version 1.0.0)

4: Group data by a column and perform aggregate operation on other columns. (Galaxy Version 2.1.1)

5: Sort data in ascending or descending order (Galaxy Version 1.0.3)

6: Select first lines from a dataset (Galaxy Version 1.0.0)
Select first
20
lines

History search datasets

Galaxy initiation - workflow
2 shown
2.77 MB

- 2: Exons
- 1: Repeats

Workflow

Run a workflow

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](#)
[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](#)

Workflow: Find exons with highest number of SNPs [Run workflow](#)

History Options

Send results to a new history
Yes No

1: Input dataset
2: Exons

2: Input dataset
1: Repeats

3: Join the intervals of two datasets side-by-side (Galaxy Version 1.0.0)

Join
Output dataset 'output' from step 1
with
Output dataset 'output' from step 2
☒ **with min overlap**
[Edit](#)
Return
Only records that are joined (INNER JOIN)

History

search datasets

Galaxy initiation - workflow
2 shown, 2 deleted, 3 hidden
2.92 MB

2: Exons

1: Repeats

Workflow

Run a workflow

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
- 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

Workflow: Find exons with highest number of SNPs **Run workflow**

History Options

Send results to a new history
Yes No

1: Input dataset
2: Exons

2: Input dataset
1: Repeats

3: Join the intervals of two datasets side-by-side (Galaxy Version 1.0.0)

4: Group data by a column and perform aggregate operation on other columns. (Galaxy Version 2.1.1)

5: Sort data in ascending or descending order (Galaxy Version 1.0.3)

6: Select first lines from a dataset (Galaxy Version 1.0.0)

Select first
20
lines

History search datasets

Galaxy initiation - workflow
2 shown
2.77 MB

- 2: Exons**
- 1: Repeats**

Workflow

Run a workflow

tools

[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](#)

Analyze Data

Workflow

Shared Data

Visualization

Help

User

Using 0%

Successfully invoked workflow **Find exons with highest number of SNPs.**
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

Galaxy initiation - workflow
7 shown
2.77 MB

7: Top exon genetic location

6: Top exons

5: Sort on data 4

4: Group on data 3

3: Join on data 1 and data 2

2: Exons

1: Repeats

Workflow

Run a workflow

The screenshot displays the Galaxy web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Help, and User. A 'Using 0%' status indicator is present in the top right. The left sidebar contains a 'Tools' section with a search bar and a list of tool categories such as 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, and NGS: Gemini.

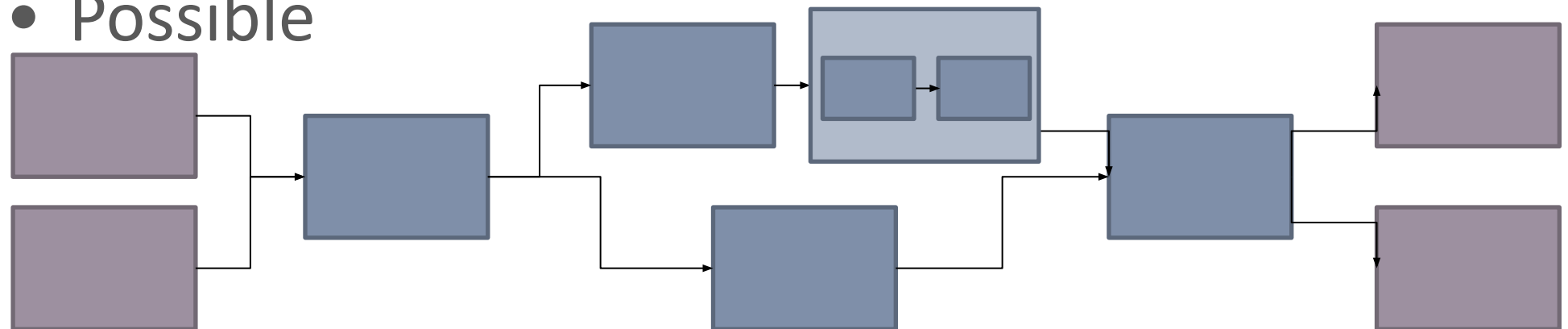
The main content area features a green notification box with a checkmark icon, stating: 'Successfully invoked workflow **Find exons with highest number of SNPs.** You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.'

The right sidebar shows the 'History' pane, which includes a search bar and a list of datasets. The datasets are listed as follows:

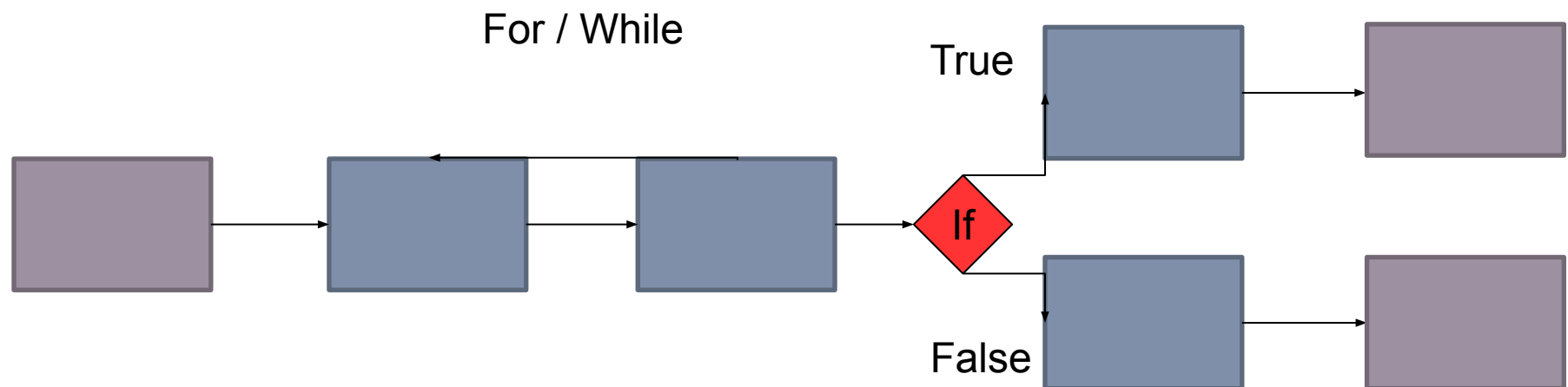
- Galaxy initiation - workflow (4 shown, 3 hidden)
- 2.92 MB
- 7: Top exon genetic location
- 6: Top exons
- 2: Exons
- 1: Repeats

Each dataset entry includes an eye icon, a pencil icon, and a close icon. The '7: Top exon genetic location' and '6: Top exons' entries are highlighted with a red box.

- Possible



- Impossible (until now)



SHARE

biologist $\leftrightarrow$ biologist

- Sharing histories or datasets
 - With or without linked workflow

bioanalyst ↔ biologist

- Sharing workflows
 - Pre-configured parameters
 - With or without release parameters (set at runtime)
- According to the user-end knowledge

bioinformatician ↔ bioinformatician

- Sharing tools ,scripts and wrappers
 - Toolshed



Share

History

The screenshot shows 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 like 'Get Data', 'Send Data', 'Text Manipulation', etc. The main panel displays 'Saved Histories' with a table of saved workflows. The 'History' sidebar on the right shows a list of datasets for the selected workflow.

Saved Histories Table:

Name	Datasets	Tags	Sharing	Size on Disk	Created	Last Upd
Galaxy initiation - workflow	4	0 Tags		2.9 MB	~5 hours ago	~4 hours
		0 Tags		8.8 MB	~10 hours ago	~5 hours
		0 Tags		247.7 MB	~5 hours ago	~5 hours
imported: Galaxy 101 (2015)	7	0 Tags		8.8 MB	~11 hours ago	~11 hour ago
Unnamed history		0 Tags		0 bytes	Jun 27, 2016	Jun 27, 2016

History Sidebar:

- Galaxy initiation - workflow (4 shown, 3 hidden)
- 2.92 MB
- 7: Top exon genetic location
- 6: Top exons
- 2: Exons
- 1: Repeats

Context Menu for 'Galaxy initiation - workflow':

- Switch
- View
- Share or Publish**
- Copy
- Rename
- Delete
- Delete Permanently

URL: https://usegalaxy.org/history/list?f-sharing=All&sort=-update_time&f-name=All&f-tags=All&f-deleted=False&operation=Share+or+Publish&id=99569b6f012ffc3c

Workflow

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

Your workflows

Create new workflow Upload or import workflow

Name	# of Steps
Find exons with highest number of SNPs	7
Convert to tab (imported from ...)	2
imported: ChIP-seq workflow	3

Workflows shared by others

No workflows have been shared by others

Other options

Configure your workflow menu

Context menu for the first workflow:

- Edit
- Run
- Share or Download
- Copy
- Rename
- View
- Delete

Share

Mode

Galaxy

Analyze DataWorkflowShared DataVisualizationHelpUser

Using 0%

[Go back to Workflows List](#)

Workflow ' Find exons with highest number of SNPs '

Share

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

Make Workflow Accessible via Link

→ Restricted community

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

Make Workflow Accessible and Publish

→ All the Galaxy server users

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

You have not shared this workflow with any users yet.

Share with a user

→ Designated community
(login@sb-roscoff.fr)

Export

Download

workflow as a file so that it can be saved or imported into another Galaxy server.

This workflow must be accessible. Please use the option above to "Make Workflow Accessible and Publish" before receiving a URL for importing to another Galaxy.

Create image

of workflow in SVG format

Export to the www.myexperiment.org site.

- Get shared histories

Galaxy / METABO Analyze Data Workflow Shared Data Visualization Help User Using 216.1 MB

Tools search tools

Get Data

WORKFLOW 4 METABO

2-Preprocessing

3-Normalisation

4-Quality Control

5-Statistical Analysis

Histories shared with you by others

Name	Datasets	Created	Last Updated	Shared by
mmonsoor	6	Apr 28, 2014	~2 days ago	mmonsoor@sb-roscoff.fr

For 0 selected histories: Copy Unshare

History

- HISTORY LISTS
- Saved Histories
- Histories Shared with Me
- CURRENT HISTORY
- Create New
- Copy History
- Copy Datasets
- Share or Publish

Individual

Galaxy / METABO Analyze Data Workflow Shared Data Visualization Help User Using 70.9 MB

Published Histories

search name, annotation, owner, and tags

Advanced Search

Name	Annotation	Owner	Community Tags	Last Updated
Preprocessing		mlandi		~14 seconds ago
TP1 xcms sacuri		mmonsoor	★★★★★	~1 day ago
TP1 xcms sacuri		jfmartin	★★★★★	Apr 28, 2014

Shared Data

- Data Libraries
- Data Libraries Beta
- Published Histories
- Published Workflows
- Published Visualizations
- Published Pages

Public

- Get shared workflows

Galaxy / METABO Analyze Data **Workflow** Shared Data Visualization Help User Using 216.1 MB

Your workflows

[Create new workflow](#) [Upload or import workflow](#)

Name	# of Steps
complete_workflow_RFMF	17

Individual

Workflows shared with you by others

Name	Owner	# of Steps
Workflow mmonsoor	mmonsoor@sb-roscoff.fr	7

Galaxy / METABO Analyze Data Workflow **Shared Data** Visualization Help User Using 111.4 MB

Published Workflows

[Advanced Search](#)

- Data Libraries
- Data Libraries Beta
- Published Histories
- Published Workflows
- Published Visualizations
- Published Pages

Name	Annotation	Owner	Rating	Community Tags	Last Updated
complete_workflow_RFMF		mmand			~17 hours ago

Public

- Import shared

Galaxy / METABO Analyze Data Workflow **Shared Data** Visualization Help User Using 216.1 MB

Published Histories | [mmonsoor](#) | TP1 xcms sacuri Import history About this History

TP1 xcms sacuri

65.4 MB

search datasets

Dataset	Annotation
1: xset.RData	
2: sampleMetadata.tsv	
3: xset.TICs_raw.pdf	
4: xset.log.txt	

Author
mmonsoor

Related Histories
[All published histories](#)
[Published histories by mmonsoor](#)

Rating
 Community (0 ratings, 0.0 average)
 Yours

Tags
Community: none

Histories

Galaxy / METABO Analyze Data **Workflow** Shared Data Visualization Help User Using 216.1 MB

Your workflows

[Create new workflow](#) [Upload or import workflow](#)

Name	# of Steps
complete_workflow_RFMF	17

Workflows shared with you by others

Name	Owner	# of Steps
Workflow mmonsoor	mmonsoor@sb-roscoff.fr	7

Other

View

Run

Copy

Remove

Workflows



Level 5

- Share tools and descriptions in the ToolShed

Level 4



- Launch tools autonomously
- Use advanced parameters
- Use the Galaxy API
- Provide workflow for colleagues Level 1-3



Level 3

- Launch tools autonomously
- Use workflow more or less preset



Level 2

- Use preset workflow



Level 1

- Share his data to colleagues Level 2-5

Hands-on **WORKFLOW**





Which coding exon has the highest number of single
~~nucleotide polymorphisms~~ on chromosome 22?
repeats

1. Extract a workflow from your history
2. Edit the workflow (hide intermediate steps, rename inputs/outputs, set parameters at runtime, save)
3. Create a new history with the input data
4. Run the workflow
5. Share your history/workflow with your neighbour

CONCLUSION

Key points on Galaxy

- Easy-to-use graphical user interface for often complex command-line tools
- Keeps a full record of your analysis in a history
- Workflows enable you to repeat your analysis on different data
- Galaxy can connect to external sources for data import and visualization purposes
- Galaxy provides ways to share your results and methods with others

Tools - panel

125+ platforms for using Galaxy: <https://galaxyproject.org/use>
 [Servers, clouds, deployable resources]

GENERAL PURPOSE:

usegalaxy.org, usegalaxy.eu, usegalaxy.fr, usegalaxy.org.au

DOMAIN SPECIFIC:



RNAseq: <http://galaxy.sb-roscoff.fr>

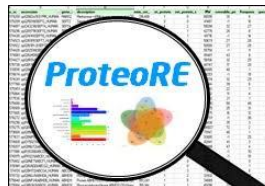
SBR tools: <http://webtools.sb-roscoff.fr>

Metagenomics: <http://galaxy4metab.sb-roscoff.fr>

Metabolomics:



Proteomics:



ChIP-seq:



Catalog of French Galaxy tools:

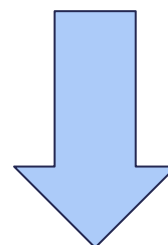
<http://galaxycat.france-bioinformatique.fr>



Galaxy ABiMS --> Galaxy France

<https://galaxy.sb-roscoff.fr/>





<https://usegalaxy.fr/>




To go further

- On your own:

- Training materials:



- <https://galaxyproject.github.io/training-material>

- Interactive tours of Galaxy:

- <http://galaxy.sb-roscoff.fr/tours>

- Training courses:

Training	What ?	Where ?	When ?
RNAseq analysis with Galaxy	RNAseq	Roscoff, France	first semester 2021
Galaxy Community Conference (GCC)	General purpose (data-intensive biology and Galaxy)	Ghent, Belgium	July 2021
Workflow4Experimenters	Metabolomics	Toulouse, France	October 2021

END

Thank you for completing the training evaluation questionnaire:

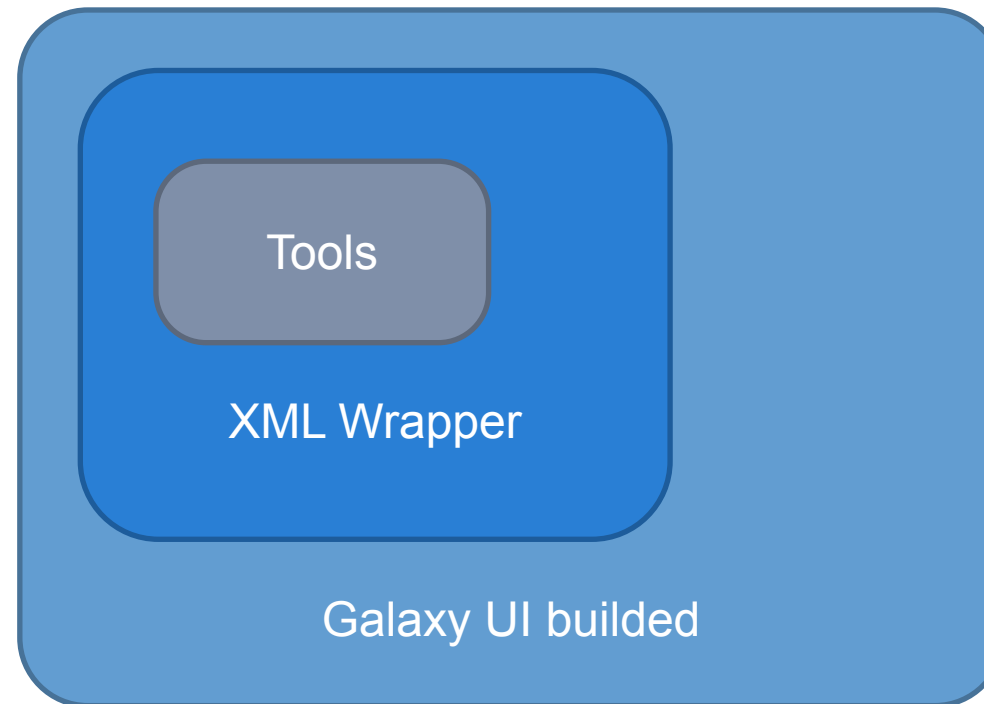
http://abims.sb-roscoff.fr/evaluation_formation

BONUS

How are tools born?

BONUS

- How to import a tool in Galaxy?



• How to import a tool in Galaxy?

```
[lecorguille@n0 ~]$ e-PCR --help
e-PCR: invalid option -- -
usage: [-hV] [posix-options] stsfile [fasta ...]
[compat-options]
where posix-options are:
-m ##      Margin (default 50)
-w ##      Wordsize (default 7)
-n ##      Max mismatches allowed (default 0)
-g ##      Max indels allowed (default 0)
-f ##      Use ## discontiguos words, slow if
           ##>1
-o ##      Set output file
-t ##      Set output format:
           1 - classic, range (pos1..pos2)
           2 - classic, midpoint
           3 - tabular
           4 - tabular with alignment in comments
              (slow)
-d##-##    Set default size range
           (default 100-350)
-p +-      Turn hits postprocess on/off
-v ##      Verbosity flags
-a a|f     Use presize alignmens (only if
           gaps>0), slow
           a - Allways or f - as Fallback
-x +-      Use 5'-end lowercase masking of
           primers (default -)
-u +-      Uppercase all primers (default -)
```

Galaxy / ABiMS

e-PCR (version 1.0.0)

STS file:

 format : tabular

Fasta file:

 format : fasta

Wordsize (W):

 Set word size for primers hash (nucleotide positions). Longer word size decreases hash collision rate, but increases memory usage. Also no mismatches are allowed within word size near 'inner' boundary of primers unless one uses discontinuous words, and no gaps are ever allowed in that region.

Use ## discontinuos words (F):

 Set discontinuous word count for primers hash (1 means 'use contiguous words'). Discontinuous words increase number of hash tables and decrease 'effective' word size (thus increasing hash collision rate), so make search significantly slower, but increase sensitivity by allowing mismatches within word size. Reasonable values are 1 (contiguous words) and 3.

Margin (M):

 Set maximal allowed deviation of hit product size from expected STS size.

Set default sts lower size (D):

 Set ddefault STS size range - values used for STSs that have no size associated in file.

Set default sts higher size (D):

 Set ddefault STS size range - values used for STSs that have no size associated in file.

Max mismatches allowed (N):

 Set maximal number of mismatches allowed in primer-to-sequence alignment (per primer!).

Max indels allowed (G):

 Set maximal number of gaps allowed in primer-to-sequence alignment (per primer!).

Set output format (T):

 Output formats

• How to import a tool in Galaxy?

```

<tool id="abims_epcr" name="e-PCR">
  <!-- author : lecorguille@sb-roscoff.fr -->
  <!-- date : 11-05-12 -->
  <description>e-PCR parses stsfile in unists format, then reads nucleotide sequence data in FASTA format from files listed in commandline if
any, or from stdin otherwise. For input sequences e-PCR finds matches and prints output in one of three formats.</description>
  <command>e-PCR -w $wordsize -f $wordcnt -m $margin -d$sts_size_lo-$sts_size_hi -n $max_mismatch -g $max_gap -t $output_format $infile_stsfile
$infile_fasta > $output</command>
  <inputs>
    <param name="infile_stsfile" type="data" label="STS file" format="tabular" help="format : tabular" />
    <param name="infile_fasta" type="data" label="Fasta file" format="fasta" help="format : fasta" />
    <param name="wordsize" type="integer" label="Wordsize (W)" value="7" help="Set word size for primers hash (nucleotide positions).
Longer word size decreases hash collision rate, but increases memory usage. Also no mismatches are allowed within word size near
'inner' boundary of primers unless one uses discontinuous words, and no gaps are ever allowed in that region." />
    <param name="wordcnt" type="integer" label="Use ## discontinuos words (F)" value="1" help="Set discontinuous word count for primers
hash (1 means 'use contiguous words'). Discontinuous words increase number of hash tables and decrease 'effective' word size (thus
increasing hash collision rate), so make search significantly slower, but increase sencitivity by allowing mismatches within word
size. Reasonable values are 1 (contiguous words) and 3." />
    <param name="margin" type="integer" label="Margin (M)" value="50" help="Set maximal allowed deviation of hit product size from
expected STS size." />
    <param name="sts_size_lo" type="integer" label="Set default sts lower size (D)" value="100" help="Set ddefault STS size range - values
used for STSs that have no size associated in file." />
    <param name="sts_size_hi" type="integer" label="Set default sts higher size (D)" value="400" help="Set ddefault STS size range -
values used for STSs that have no size associated in file." />
    <param name="max_mismatch" type="integer" label="Max mismatches allowed (N)" value="0" help="Set maximal number of mismatches allowed
in primer-to-sequence alignment (per primer!)." />
    <param name="max_gap" type="integer" label="Max indels allowed (G)" value="0" help="Set maximal number of gaps allowed in primer-to-
sequence alignment (per primer!)." />
    <param name="output_format" type="select" help="Output formats">
      <label>Set output format (T)</label>
      <option value="1">classic, range (pos1..pos2)</option>
      <option value="2">classic, midpoint</option>
      <option value="3" selected="true">tabular</option>
      <option value="4">tabular with alignment in comments (slow)</option>
    </param>
  </inputs>
  <outputs>
    <data name="output" format="tabular" />
  </outputs>

```


• How to import a tool in Galaxy?

```
[lecorguille@n0 ~]$ e-PCR --help
e-PCR: invalid option -- -
usage: [-hV] [posix-options] stsfile [fasta ...]
[compat-options]
where posix-options are:
-m ##      Margin (default 50)
-w ##      Wordsize (default 7)
-n ##      Max mismatches allowed (default 0)
-g ##      Max indels allowed (default 0)
-f ##      Use ## discontinuos words, slow if
           ##>1
```

Galaxy / ABIMS

e-PCR (version 1.0.0)

STS file:

100: (as tabular) Trinity on data 9..Transcripts

format : tabular

Fasta file:

100: Trinity on data 9..Transcripts

format : fasta

Wordsize (W):

7

Set word size for primers hash (nucleotide positions). Longer word size decreases hash collision rate, but increases memory usage. Also no mismatches are allowed within word size near 'inner' boundary of primers unless one uses discontinuous words, and no gaps are ever allowed in that region.

Use ## discontinuos words (F):

1

```
<tool id="abims_epcr" name="e-PCR">
  <!-- author : lecorguille@sb-roscoff.fr -->
  <!-- date : 11-05-12 -->
  <description>e-PCR parses stsfile in unists format, then reads nucleotide sequence data in FASTA format from files listed in commandline if
any, or from stdin otherwise. For input sequences e-PCR finds matches and prints output in one of three formats.</description>
  <command>e-PCR -w $wordsize -f $wordcnt -m $margin -d$sts_size_lo-$sts_size_hi -n $max_mismatch -g $max_gap -t $output_format $infile_stsfile
$infile_fasta > $output</command>
  <inputs>
    <param name="infile_stsfile" type="data" label="STS file" format="tabular" help="format : tabular" />
    <param name="infile_fasta" type="data" label="Fasta file" format="fasta" help="format : fasta" />
    <param name="wordsize" type="integer" label="Wordsize (W)" value="7" help="Set word size for primers hash (nucleotide positions).
Longer word size decreases hash collision rate, but increases memory usage. Also no mismatches are allowed within word size near
'inner' boundary of primers unless one uses discontinuous words, and no gaps are ever allowed in that region." />
    <param name="wordcnt" type="integer" label="Use ## discontinuos words (F)" value="1" help="Set discontinuous word count for primers
hash (1 means 'use contiguous words'). Discontinuous words increase number of hash tables and decrease 'effective' word size (thus
increasing hash collision rate), so make search significantly slower, but increase sencitivity by allowing mismatches within word
size. Reasonable values are 1 (contiguous words) and 3." />
    <param name="margin" type="integer" label="Margin (M)" value="50" help="Set maximal allowed deviation of hit product size from
expected STS size." />
    <param name="sts_size_lo" type="integer" label="Set default sts lower size (D)" value="100" help="Set ddefault STS size range - values
used for STSs that have no size associated in file." />
    <param name="sts_size_hi" type="integer" label="Set default sts higher size (D)" value="400" help="Set ddefault STS size range -
values used for STSs that have no size associated in file." />
    <param name="max_mismatch" type="integer" label="Max mismatches allowed (N)" value="0" help="Set maximal number of mismatches allowed
```