



Biogenouest
BIOGÉOINFORMATIQUE



OCEANOMICS



AB⁴BIMS

15/03/2017

The Galaxy logo consists of three horizontal bars (two grey, one yellow) to the left of the word 'Galaxy' in a bold, white, sans-serif font, all on a dark blue background.

Galaxy

Initiation

Lorraine Guéguen



Annie Lebreton

Credit to Gildas Le Corguillé – V2.3



- Slides available
- <http://galaxy3.sb-roscoff.fr>
- Login:
 - login@sb-roscoff.fr
 - *****

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

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



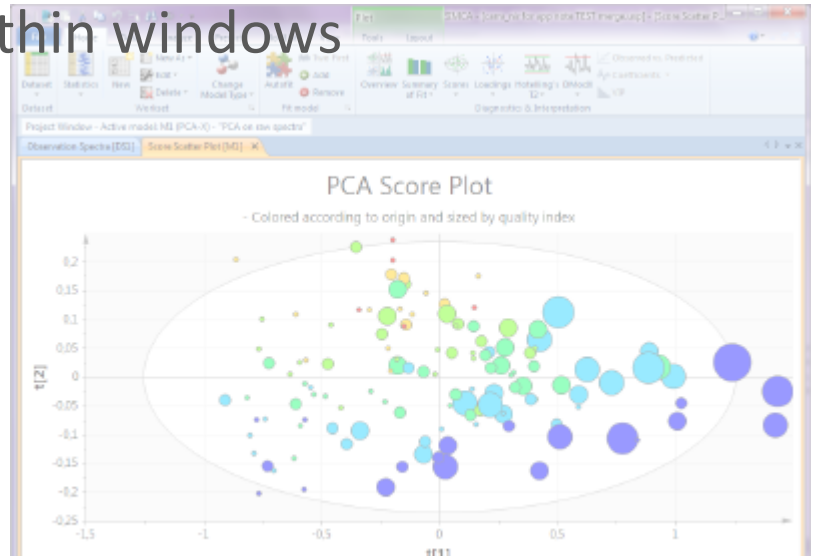
INTRODUCTION / PROBLEMATIC

- 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

Option	Value	Description
minimum peak width	1	Minimum peak width (in number of data points)
maximum peak width	10	Maximum peak width (in number of data points)
Signal-to-noise threshold	3.0	Signal-to-noise threshold
smoothing	500	Smoothing window size
degree of polynomial	1	Degree of polynomial fit
peak list length	5000	Number of peaks to return
peak list format	1000	Number of peaks to return (if 0, all peaks are returned)

Introduction / problematic

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

noise=250000
xset <- xcmsSet(cdfFiles, ppm=ppm, mzdiff=mzwid, peakwidth=peakwidth, noise=noise, snthresh=snth, method="centWave", fitgauss=TRUE, nSlaves=8)
xset2 <- retcor(xset)
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)

# 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@phenoData))

#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")

library(CAMERA)
#annotation version rapide?
an <- annotate(xset, ppm=15, mzabs=0.015, quick=FALSE, psg_list=NULL, rules=NULL, maxcharge=3, maxiso=4, minfrac=0.5,
polarity=polarity)
diffreport1 <- getPeaklist(an)

#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)))

```

● Command line tools

+ represent almost the majority of scientific tools

+ good parameters completeness

+ can be executed on high performance computers

+ automatable, workflowsable, ...

- minimum linux knowledge is required

- crucial lack of ergonomics


```
login@sbr4-1042:~$ ssh -Y login@bioinfo.sb-roscoff.fr
[...]
[login@n0 ~]$ cd projet
[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@n0 mapping]$ mkdir cufflinks
```

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

```
[login@n0 ~]$ cd projet
[login@n0 login]$
[login@n0 mapping]$
#! /bin/bash
#$ -S /bin/bash
#$ -M login@sb-roscoff.fr
#$ -m bea
#$ -V
#$ -cwd
#$ -o qsub.out
#$ -e qsub.err
```



```
tophat2 panda_v12
-GTF ../input/pa
-num-threads 8
```

R1-2.fq

```
[login@n0 mapping]$
Your job 5338969
[login@n0 mapping]$
accepted_hits.bam
deletions.bed
insertions.bed
[login@n0 mapping]$ cd ..
[login@n0 mapping]$ mkdir cufflinks
```

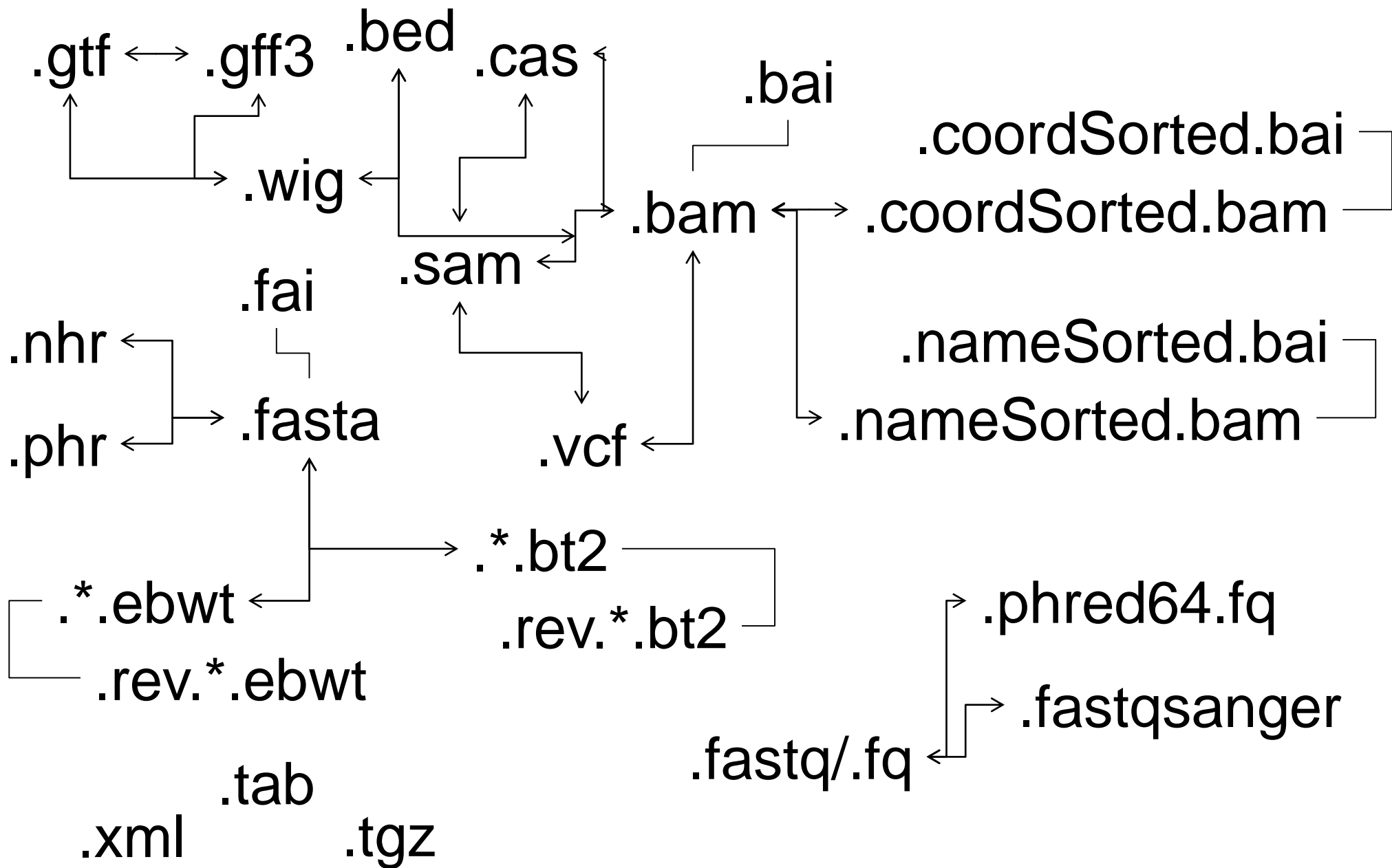
qsub

NOOOOOOOO!

1.bam

prep_reads.info tmp

Introduction / problematic





Select your level:

Level I



« 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? »



Tools



search tools

[Upload File from your computer](#)[Export Data](#)

LC-MS

[Format Conversion](#)[Preprocessing](#)[Normalisation](#)[Quality Control](#)[Statistical Analysis](#)[Annotation](#)

GC-MS

[Preprocessing](#)[Normalisation](#)[Quality Control](#)[Statistical Analysis](#)[Annotation](#)

NMR

[Preprocessing](#)[Normalisation](#)[Quality Control](#)[Statistical Analysis](#)

COMMON TOOLS

[Data Handling](#)[Text Manipulation](#)[Filter and Sort](#)[Join, Subtract and Group](#)

xcms.xcmsSet version 2.0.1

**Choose your inputs method:**

Zip file from your history containing your chromatograms

Zip file:

1: sacuri.zip

Extraction method for peaks detection:

matchedFilter

[method] See the help section below

Step size to use for profile generation:

0.01

[step] The peak detection algorithm creates extracted ion base peak chromatograms (EIBPC) on a fixed step size

Full width at half maximum of matched filtration gaussian model peak:

30

[fwhm] Only used to calculate the actual sigma

Advanced options:

hide

Execute

Authors Colin A. Smith csmith@scripps.edu, Ralf Tautenhahn rtautenh@gmail.com, Steffen Neumann sneumann@ipb-halle.de, Paul Benton hpaul.benton08@imperial.ac.uk and Christopher Conley cjconley@ucdavis.edu

If you use this tool, please cite: Smith, C.A. et al. (2006). XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.*, 78, 779–787.

For details about this tool, please go to <http://www.bioconductor.org/packages/release/bioc/html/xcms.html>

Galaxy integration ABIMS TEAM, Station biologique de Roscoff.

Contact support@workflow4metabolomics.org for any questions or concerns about the Galaxy implementation of this tool.

Galaxy

History

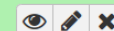
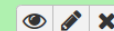
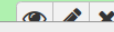


search datasets

**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:**

INTRODUCTION / GALAXY

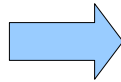
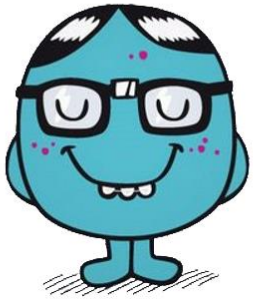
Why Galaxy ?

- 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 data and protocols sharing system
 - Tool-boxes of several bioinformatics fields
 - NGS
 - Metabolomics
 - Statistics
 - Chemistry
 - Image analysis
 - Etc ...



MR. GEEK



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

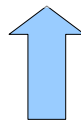
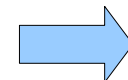


Introduction / Galaxy

```
[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/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
```



MR. HAPPY

by Roger Hargreaves



Galaxy
PROJECT

```
[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
                    (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:

 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 ## discontinuous 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

Galaxy interface

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 :

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

Variable metadata file :

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

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

Execute

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

1 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:

Galaxy interface

Menu

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

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

1 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:

Galaxy interface

Tool list

The screenshot displays the Galaxy web interface. At the top, the navigation bar shows 'Galaxy / 4 / Metabolomics' and various menu items like 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. The main content area is divided into three panels:

- Tools Panel (Left):** Contains a search bar and a list of tool categories: 'Upload File from your computer', 'Export Data', 'LC-MS', 'Format Conversion', 'Preprocessing', 'Normalisation', 'Quality Control', 'Statistical Analysis', 'Annotation', 'GC-MS', 'Preprocessing', 'Normalisation', 'Quality Control', 'Statistical Analysis', 'Annotation', 'NMR', 'Preprocessing', and 'Normalisation'.
- Tool Configuration Panel (Center):** Shows the 'Batch_correction (version 2.0.0)' tool. It includes input fields for:
 - 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 :** basicAn 'Execute' button is located at the bottom of this panel. Below the configuration, there is a section for '1 Authors' (Jean-Francois Martin) and '1 Contributors' (Melanie Petera, Etienne Thevenot).
- History Panel (Right):** Shows a list of datasets with search and view options. The datasets listed include:
 - 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: (partially visible)

Galaxy interface

Web forms / dataset visualization / diverse information

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

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

1 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:

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 :

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

Variable metadata file :

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

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

Execute

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

1 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**

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

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

< 2 GO

Data import < 2 Go

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)

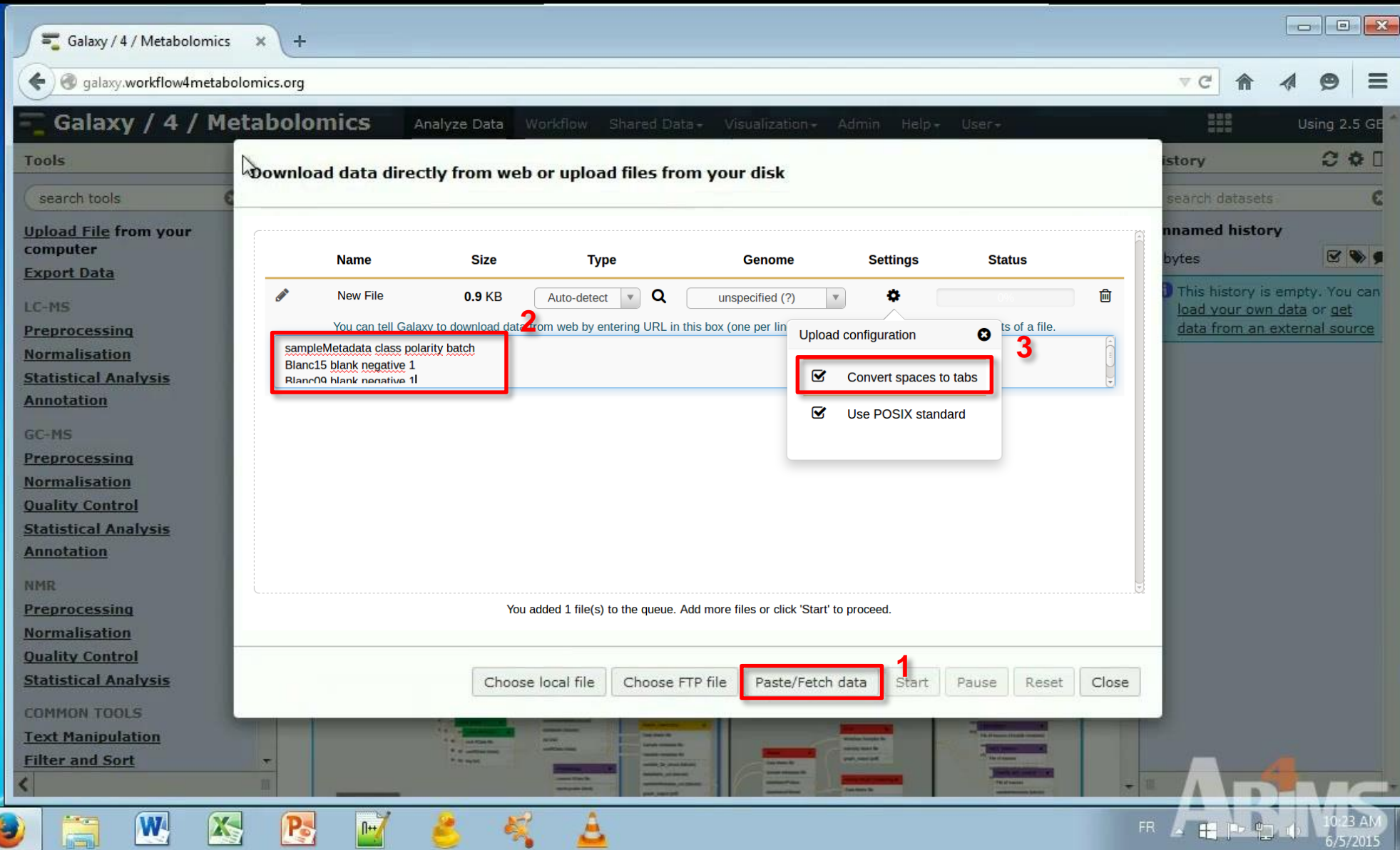
10:23 AM 6/5/2015

Data import < 2 Go

The screenshot shows a web browser window with the URL `galaxy.workflow4metabolomics.org`. The page title is "Galaxy / 4 / Metabolomics". A modal dialog box is open in the center, titled "Download data directly from web or upload files from your disk". The dialog contains a large dashed rectangular area for file upload, with the text "You can Drag & Drop files into this box." below it. 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 background shows the Galaxy interface with a sidebar of tool categories like "Upload File from your computer", "Export Data", "LC-MS", "GC-MS", and "NMR". The Windows taskbar at the bottom shows various application icons and the system clock indicating 10:23 AM on 6/5/2015.

Data import < 2 Go

Copy / Paste data



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
New File	0.9 KB	Auto-detect	unspecified (?)		

You can tell Galaxy to download data from web by entering URL in this box (one per line)

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

Upload configuration

- Convert spaces to tabs
- Use POSIX standard

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

ARMS

FR 10:23 AM 6/5/2015

Data import < 2 Go

From local files

The screenshot shows a web browser window with the URL `galaxy.workflow4metabolomics.org`. The page title is "Galaxy / 4 / Metabolomics". A modal dialog box is open in the center, titled "Download data directly from web or upload files from your disk". The dialog contains a large dashed rectangular area for file upload, with the text "You can Drag & Drop files into this box." below it. At the bottom of the dialog, there are several buttons: "Choose local file" (highlighted with a red box), "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close". The background shows the Galaxy interface with a sidebar of tool categories like "Upload File from your computer", "Export Data", "LC-MS", "GC-MS", and "NMR". The Windows taskbar at the bottom shows various application icons and the system clock indicating 10:23 AM on 6/5/2015.

Data import < 2 Go

From local files

The screenshot shows a web browser window with the URL `galaxy.workflow4metabolomics.org`. The browser's address bar and the page's navigation bar both display `Galaxy / 4 / Metabolomics`, which is circled in red. 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**. The dialog contains a large dashed-line box for file uploads, with a folder icon and a **Move** button. Below the box, it says "You can Drag & Drop files into this box." At the bottom of the dialog are several buttons: **Choose local file**, **Choose FTP file**, **Paste/Fetch data**, **Start**, **Pause**, **Reset**, and **Close**. The background shows the Galaxy interface with a sidebar of tool categories like **Upload File from your computer**, **Export Data**, **LC-MS**, **GC-MS**, and **NMR**. The Windows taskbar at the bottom shows various application icons and the system clock indicating 10:23 AM on 6/5/2015.

Data import < 2 Go

From local files

The screenshot shows the Galaxy 4 Metabolomics web interface. A modal dialog box 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. One 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 interface with a sidebar of tools and a main workspace area.

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

ARMS

Data import < 2 Go

From local files

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

Using 2.5 GB

history

search datasets

unnamed history

bytes

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

ARMS

FR 10:23 AM 6/5/2015

Data import < 2 Go

From local files

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

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

Below the table, it says "Please wait...1 out of 1 remaining." At the bottom of the dialog are buttons: "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start" (highlighted with a mouse cursor), "Pause", "Reset", and "Close".

The background shows the Galaxy interface with a sidebar of tools categorized by LC-MS, NMR, and COMMON TOOLS. The top navigation bar includes "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". The browser address bar shows "galaxy.workflow4metabolomics.org".



Data import < 2 Go

From local files

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 (?)		100% ✓

You can Drag & Drop files into this box.

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

Using 2.5 GB

history

search datasets

unnamed history

bytes

1: sacuri.zip

FR 10:23 AM 6/5/2015

ARMS

Data import < 2 Go

From local files

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

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

✓ 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

i 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

1: sacuri.zip

FR 10:23 AM 6/5/2015

Data import < 2 Go

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 'Galaxy / 4 / Metabolomics', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. The 'Tools' sidebar on the left lists various analysis categories: LC-MS, NMR, and COMMON TOOLS, each with sub-options like 'Preprocessing', 'Normalisation', 'Statistical Analysis', and 'Annotation'. A 'sacuri.zip' file icon is visible on the desktop to the left of the browser window.

The central content area features a green banner with a checkmark icon and the text: **Welcome to workflow4metabolomics.org v2.0**. Below this, a 'Publication' section lists authors and the title: **Workflow4Metabolomics: A collaborative research infrastructure for computational metabolomics**. A 'Latest news' section contains three entries with dates and brief descriptions of updates and releases.

On the right side, the 'History' panel shows a search bar and a list of datasets. The first entry is '1: sacuri.zip', which is highlighted in yellow and has a status of 'This job is currently running'. The system status at the top right indicates 'Using 2.5 GB'.

Step 1: Choose a FTP Client

DATA IMPORT

> 2 GO

STEP 1: CHOOSE A FTP CLIENT



STEP 1: CHOOSE A FTP CLIENT



Avoid:
Malwares inside



Cyberduck



WinSCP

STEP 1: CHOOSE A FTP CLIENT



FileZilla



Cyberduck



WinSCP

Step 2: Easy!

DATA IMPORT

> 2 GO

Data import > 2 Go

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 status bar in the top right corner indicates **Using 46%**.

The **Tools** section on the left contains a search bar and a list of 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**.

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

The main content area features a green banner: **Welcome to galaxy.sb-roscoff.fr**. Below it is an **Information** box with the text: **For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr**. The center of the page displays the **ABiMS** logo (Analyses and Bioinformatics for Marine Science) and the **Station Biologique Roscoff** logo. Navigation links for **Changelog** and **Tutorials** are provided. A paragraph at the bottom describes the Galaxy platform and its affiliations.

The right sidebar shows the **History** section, which is currently empty (0 bytes). A message 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, Chrome, Word, Excel, PowerPoint, and a duck icon. The system tray in the bottom right corner displays the date and time: **FR 11:19 AM 7/31/2015**.

Data import > 2 Go

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The navigation menu includes "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". The dialog contains a large dashed box for file upload with the text "You can Drag & Drop files into this box." Below the box are two dropdown menus: "Type (set all):" with "Auto-detect" selected, and "Genome (set all):" with "unspecified (?)" selected. At the bottom of the dialog are buttons for "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close". The background interface shows a sidebar with "Tools" and "Get Data" sections, and a main content area with a search bar and a "History" section.

Data import > 2 Go

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The navigation menu includes "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". A sidebar on the left lists various tools under categories like "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 can Drag & Drop files into this box." and a large dashed-line box for file upload. Below the box are two dropdown menus: "Type (set all):" with "Auto-detect" selected, and "Genome (set all):" with "unspecified (?)" selected. At the bottom of the dialog are several buttons: "Choose local file", "Choose FTP file" (with a mouse cursor over it), "Paste/Fetch data", "Start", "Pause", "Reset", and "Close".

The Windows taskbar at the bottom shows the system tray with the date "7/31/2015" and time "11:19 AM".

Data import > 2 Go

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

Using 46%

Assets

History

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

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

FR 11:19 AM 7/31/2015

Data import > 2 Go

Galaxy / ABiMS

galaxy.sb-roscoff.fr

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

Get Data

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

Using 46%

Assets

History

History is empty. You can download data or get data from an external source

lecorguille

left_kept_r...

11:19 AM 7/31/2015

ABiMS

Data import > 2 Go

The screenshot shows a Windows desktop environment. In the background, a web browser window displays the Galaxy/ABiMS interface at galaxy.sb-roscoff.fr. The browser's address bar shows the URL. The page content includes a navigation menu with options like 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. A central dialog box titled 'Download data directly from web or upload files from your disk' is active, featuring a dashed border and the text 'You can Drag & Drop files into this box.' Below this, there is a text area with instructions: 'allows you to upload files via FTP. To upload some files, log in at galaxy.sb-roscoff.fr using your Galaxy credentials (username and password). Your FTP directory does not contain any files.' At the bottom of the dialog, there are buttons for 'local file', 'Choose FTP file', 'Paste/Fetch data', 'Start', 'Pause', 'Reset', and 'Close'. In the foreground, a Cyberduck window is open, showing the 'Open Connection' button and a search bar. The desktop taskbar at the bottom contains icons for various applications, including Internet Explorer, Chrome, File Explorer, Word, Excel, PowerPoint, and a duck icon. The system tray in the bottom right corner shows the date and time as 'FR 11:19 AM 7/31/2015'.

Data import > 2 Go

The screenshot shows a web browser window displaying the Galaxy/ABiMS interface. The browser's address bar shows the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The main content area features a large white box with the heading "Download data directly from web or upload files from your disk" and the instruction "You can Drag & Drop files into this box." Below this, there is a text box containing the message: "allows you to upload files via FTP. To upload some files, log in at **galaxy.sb-roscoff.fr** using your Galaxy credentials (username and password). Your FTP directory does not contain any files." At the bottom of this box are buttons for "local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close".

Overlaid on the browser is the Cyberduck application window. The title bar reads "Cyberduck" and "Unregistered". The menu bar includes "File", "Edit", "View", "Go", "Bookmark", "Window", and "Help". The "Open connection" dialog is active, showing "FTP (File Transfer Protocol)" as the selected protocol. The "Server:" field is empty, and the "Port:" is set to "21". There are also fields for "URL:", "Username:", and "Password:". At the bottom of the dialog are "Connect" and "Cancel" buttons, and a "More Options" dropdown menu.

The Windows taskbar at the bottom shows the system tray with the date and time "11:19 AM 7/31/2015" and the language "FR". The taskbar includes icons for Internet Explorer, Google Chrome, File Explorer, Microsoft Word, Microsoft Excel, Microsoft PowerPoint, a yellow duck icon, a traffic cone icon, and the system tray.

Data import > 2 Go

The screenshot shows a web browser window displaying the Galaxy/ABiMS interface. The browser's address bar shows the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The navigation menu includes "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". A central panel titled "Download data directly from web or upload files from your disk" contains a large dashed box with the text "You can Drag & Drop files into this box." Below this, there is a message: "allows you to upload files via FTP. To upload some files, log at **galaxy.sb-roscoff.fr** using your Galaxy credentials (password)." A yellow warning box states "Your FTP directory does not contain any files." At the bottom of the panel are buttons for "local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close".

Overlaid on the browser is a "Cyberduck" window titled "Open Connection". The window is "Unregistered" and has a menu bar with "File", "Edit", "View", "Go", "Bookmark", "Window", and "Help". The "FTP (File Transfer Protocol)" protocol is selected. The "Server" field contains `galaxy.sb-roscoff.fr` and the "Port" is set to `21`. The "URL" field shows `ftp://lecorguille@galaxy.sb-roscoff.fr:21/`. The "Username" is `lecorguille` and the "Password" is masked with dots. There are checkboxes for "Anonymous Login" and "Save Password", both of which are unchecked. "Connect" and "Cancel" buttons are at the bottom right. A "More Options" section is expanded at the bottom left.

The Windows taskbar at the bottom shows various application icons, including Internet Explorer, Chrome, File Explorer, Word, Excel, PowerPoint, and a yellow duck icon. The system tray on the right shows the date and time as "FR 11:19 AM 7/31/2015".

Data import > 2 Go

The screenshot shows a web browser window displaying the Galaxy/ABiMS interface. The browser's address bar shows the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The main content area features a large white box with the heading "Download data directly from web or upload files from your disk" and the instruction "You can Drag & Drop files into this box." Below this is a dashed-line box representing the drop area. A smaller dialog box is overlaid on the page, titled "Unsecured FTP connection", with the text: "Unsecured FTP connection. Password will be sent in plaintext. Please contact your web hosting service provider for assistance." The dialog has "Continue" and "Disconnect" buttons, and a "Don't show again" checkbox. At the bottom of the browser window, a taskbar shows various application icons, including Internet Explorer, Chrome, and Word. The system tray in the bottom right corner displays the date and time as "FR 11:19 AM 7/31/2015".

Data import > 2 Go

The screenshot shows a web browser window displaying the Galaxy/ABiMS interface. The browser's address bar shows the URL `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 features a "Tools" sidebar on the left and a central panel with the heading "Download data directly from web or upload files from your disk". Below this heading, there is a dashed box containing the text "You can Drag & Drop files into this box." and a message stating "Your FTP directory does not contain any files." A tooltip is visible over the "Choose FTP file" button, explaining that it allows uploading files via FTP and requires logging in at `galaxy.sb-roscoff.fr` with Galaxy credentials. At the bottom of the browser window, a notification bubble indicates "Connection opened galaxy.sb-roscoff.fr".

Overlaid on the browser window is an FTP client window titled "lecorguille@galaxy.sb-roscoff.fr - FTP". The window shows a menu bar (File, Edit, View, Go, Bookmark, Window, Help) and a toolbar with "Open Connection", "Quick Connect", "Action", "Get Info", and "Refresh" buttons. Below the toolbar is a search bar and a file list table with columns for "Filename", "Size", and "Modified". The file list is currently empty, showing "0 Files". A mouse cursor is hovering over a "Copy to /" button in the bottom left corner of the file list area.

Data import > 2 Go

The screenshot displays a web browser window with the URL `galaxy.sb-roscoff.fr`. The browser shows a 'Transfers' window for the file `left_kept_reads.bam`. The progress bar indicates that 50.6 MiB (53,018,624 bytes) of 91.6 MiB (55%) has been uploaded 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`.

In the background, the Galaxy web interface is visible, showing a 'from your disk' section with a 'Drop files into this box.' instruction. A notification bubble at the bottom right states 'Connection opened galaxy.sb-roscoff.fr'.

At the bottom of the screen, the Windows taskbar shows the system tray with the date and time: 11:19 AM 7/31/2015.

Data import > 2 Go

The screenshot shows the Galaxy/ABiMS web interface with a central dialog box titled "Download data directly from web or upload files from your disk". The dialog contains the text "You can Drag & Drop files into this box." and a large dashed box for file upload. Below the dialog, there are buttons for "local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Cancel".

An FTP client window titled "lecorguille@galaxy.sb-roscoff.fr - FTP" is open in the foreground. It shows a file list with the following details:

Filename	Size	Modified
left_kept_reads.bam	91.6 MiB	7/31/2015 9:19:00 AM

The FTP client window also displays a message: "Your FTP directory does not contain any files." and a notification bubble at the bottom right that says "Upload complete left_kept_reads.bam".

Data import > 2 Go

Galaxy / ABiMS

galaxy.sb-roscoff.fr

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 46%

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

Upload complete left kept reads bam

FR 11:19 AM 7/31/2015

Data import > 2 Go

The screenshot shows a web browser window with the URL `galaxy.sb-roscoff.fr`. The page title is "Galaxy / ABiMS". The navigation menu includes "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". A "Tools" sidebar on the left lists various categories like "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 can Drag & Drop files into this box." and a large dashed rectangular area for file upload. Below this area are two dropdown menus: "Type (set all):" with "Auto-detect" selected, and "Genome (set all):" with "unspecified (?)" selected. At the bottom of the dialog are several buttons: "Choose local file", "Choose HTTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close".

The Windows taskbar at the bottom shows the system tray with the date "7/31/2015" and time "11:19 AM". A notification bubble in the bottom right corner says "Upload complete".

Data import > 2 Go

The screenshot shows the Galaxy/ABiMS web interface. A dialog box titled "Download data directly from web or upload files from your disk" is open. Inside the dialog, 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)." Below this text is a table of available files:

<input type="checkbox"/>	Name	Size	Created
<input type="checkbox"/>	left_kept_reads.bam	96 MB	07/31/2015 11:19:45 AM

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 "Choose FTP file" button is highlighted with a dashed border. The background shows the Galaxy/ABiMS navigation menu with categories like "Tools", "Get Data", "COMMON TOOLS", "SEARCHING TOOLS", and "NGS TOOLS".

Data import > 2 Go

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

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

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

<input checked="" type="checkbox"/>	Name	Size	Created
<input checked="" type="checkbox"/>	left_kept_reads.bam	96 MB	07/31/2015 11:19:45 AM

Type (set a

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

Using 46%

Assets

History

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

FR 11:19 AM 7/31/2015

lecorguille

left_kept_r...

left_kept_r...

ABiMS

Data import > 2 Go

Galaxy / ABiMS

galaxy.sb-roscoff.fr

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 46%

Download data directly from web or upload files from your disk

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

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

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

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

lecorguille

left_kept_r...

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

ABiMS

FR 11:20 AM 7/31/2015

Data import > 2 Go

Galaxy / ABiMS

galaxy.sb-roscoff.fr

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.

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

Type (set all): Auto-detect

Genome (set all): unspecified (?)

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

Using 46%

assets

tory

ads.bam

FR 11:20 AM 7/31/2015

Data import > 2 Go

The screenshot shows a web browser window displaying the Galaxy / ABiMS interface. The browser's address bar shows the URL `galaxy.sb-roscoff.fr`. The page header includes navigation tabs for `Analyze Data`, `Workflow`, `Shared Data`, `Visualization`, `Admin`, `Help`, and `User`. A green banner at the top reads "Welcome to galaxy.sb-roscoff.fr". Below this, an information box provides contact details for support at `support.abims@sb-roscoff.fr`. The main content area features the ABiMS logo (Analyses and Bioinformatics for Marine Science) and the Station Biologique Roscoff logo. A sidebar on the left lists various tool categories such as "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". A right-hand sidebar shows a "History" section with a search bar and a list of datasets, including "left_kept_reads.bam". The Windows taskbar at the bottom displays several application icons, including Internet Explorer, Chrome, Word, Excel, PowerPoint, and a duck icon. The system tray in the bottom right corner shows the date and time as "FR 11:20 AM 7/31/2015".

Data import > 2 Go

The screenshot displays the Galaxy / ABiMS web interface. The browser 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 green banner at the top reads "Welcome to galaxy.sb-roscoff.fr". Below this, an information box provides contact details: "For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr".

The central content area features the **ABiMS** logo (Analyses and Bioinformatics for Marine Science) and the **Station Biologique Roscoff** logo. Below the logo are expandable sections for **Changelog** and **Tutorials**. A paragraph of text describes the Galaxy 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."

The left sidebar contains a **Tools** section with a search bar and a list of 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**.

The right sidebar shows the **History** section with a search bar, **Unnamed history** (1 shown, 0 bytes), and a dataset entry **left_kept_reads.bam** with a size of 1 byte.

The Windows taskbar at the bottom includes icons for Internet Explorer, Google Chrome, File Explorer, Word, Excel, PowerPoint, a calculator, a rubber duck, and a traffic cone. The system tray shows the date and time as 11:20 AM on 7/31/2015.

Data import > 2 Go

The screenshot displays the Galaxy / ABiMS web interface. The browser address bar shows 'galaxy.sb-roscoff.fr'. The main navigation bar includes 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. A 'Using 46%' indicator is visible in the top right.

The left sidebar contains a 'Tools' section with a search bar and a list of 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'.

The main content area features a green 'Welcome to galaxy.sb-roscoff.fr' message, an information box with contact details for support.abims@sb-roscoff.fr, and the ABiMS logo with the tagline 'Analyses and Bioinformatics for Marine Science'. Below the logo are links for 'Changelog' and 'Tutorials'. A paragraph of text describes Galaxy as an open, web-based platform for data-intensive biomedical research, supported by various institutions.

The right sidebar shows a 'History' section with a search bar and a list of datasets. The first entry is 'left_kept_reads.bam', which is 91.6 MB in size, in BAM format, and is identified as an 'uploaded bam file'. It includes options to 'display in IGB View' and 'Binary bam alignments file'.

The Windows taskbar at the bottom shows the system clock at 11:20 AM on 7/31/2015, along with various application icons.

~~DATA IMPORT~~

Data import

For HUGE public resources: genome, databank ...

--> Make a request to the support team

The screenshot displays the Galaxy/ABiMS interface for the NCBI BLAST+ tool. The tool title is "NCBI BLAST+ blastn Search nucleotide database with nucleotide query sequence(s) (Galaxy Version 0.1.08)". The "Nucleotide query sequence(s)" field is empty, showing "No fasta dataset available." The "Subject database/sequences" is set to "Locally installed BLAST database". A red box highlights the "Nucleotide BLAST database" section, which includes a "Select/Unselect all" checkbox and a list of databases: nt, genbank, genbank Bacterial, genbank Environmental sampling, genbank EST (expressed sequence tag), genbank GSS (genome survey sequence), genbank HTC (high throughput cDNA sequencing), and genbank HTGS (high throughput genomic sequencing). The "Set expectation value cutoff" is set to 0.001. The "Output format" is set to "Text (FASTA, BLAST, etc.)". The left sidebar shows various tool categories, and the right sidebar shows a history of datasets, including "eba 2016 sartools" and several "SARTools DESeq2" outputs.

Hands-on

DATA IMPORT





1. Fetch the file with your internet browser
(see given URL)

2. Upload this file into Galaxy
 - a. First, as you want
 - b. Consider that it is **>2 Go**

TOOLS

Tools - panel

The screenshot displays the Galaxy/ABiMS web interface. At the top, the navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and a 'Using 0%' 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'. Under 'Fasta Fastq Manipulation', there are links for 'Filter sequences by ID from a tabular file', 'FastQC Read Quality reports', and 'FASTQ Groomer'. The main content area features a green 'Welcome to galaxy3.sb-roscoff.fr' message, an information box with contact details for support.abims@sb-roscoff.fr, the ABiMS logo (Analyses and Bioinformatics for Marine Science), and links to 'Changelog' and 'Tutorials'. A paragraph of text 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, titled 'History', shows a search bar and a list of datasets: 'Trinity example' (2 shown, 3 deleted, 40.02 KB) and two files named '4: reads.left.fg' and '3: reads.right.fg'.

Tools - panel

The screenshot displays the Galaxy / ABiMS web interface. At the top, the navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right corner.

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'. Under 'Fasta Fastq Manipulation', specific tools like 'Filter sequences by ID from a tabular file', 'FastQC Read Quality reports', and 'FASTQ Groomer' are listed.

The central workspace features a green 'Welcome to galaxy3.sb-roscoff.fr' message, an information box with contact details for support.abims@sb-roscoff.fr, and the ABiMS logo (Analyses and Bioinformatics for Marine Science) with the Station Biologique Roscoff emblem.

The right sidebar, titled 'History', shows a search bar for datasets and a list of recent jobs: 'Trinity example' (2 shown, 3 deleted, 40.02 KB) and two 'reads.fq' files (4: reads.left.fg and 3: reads.right.fg).

A large white box with a black border is overlaid on the central workspace, containing the text: 'What tools are available?'

Below the box, a paragraph of text describes Galaxy: '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.'

>80 public Galaxy servers available:

<https://galaxyproject.org/public-galaxy-servers>

>80 public Galaxy servers available:

<https://galaxyproject.org/public-galaxy-servers>

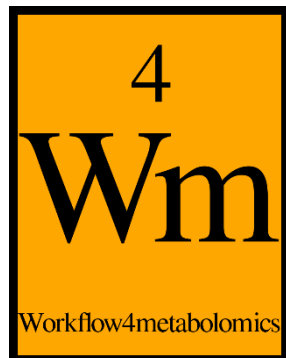


RNAseq: <http://galaxy3.sb-roscoff.fr>

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

Metagenomics: <http://galaxy4frogs.sb-roscoff.fr>

Metabolomics:



ChIP-seq:



an open and powerful Galaxy instance
for integrative Omics data analysis



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

ABiMS
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

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

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

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

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 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?

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

 2: xset.RData



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

 **2: xset.RData**

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 ☹️

2: xset.RData



Job is finished and status is OK

But warnings or errors can be hidden behind!



16: xset.RData



Job is finished but with an error status

= the program sends an error

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



16: xset.RData



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

The screenshot displays the Galaxy / ABiMS web interface. At the top, the navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and a 'Using 0%' 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), and 'SEARCHING TOOLS' (Diamond). The main content area features a green 'Welcome to galaxy3.sb-roscoff.fr' message, an information box with contact details for support.abims@sb-roscoff.fr, the ABiMS logo (Analyses and Bioinformatics for Marine Science), and links to 'Changelog' and 'Tutorials'. Below these is a paragraph describing Galaxy as an open, web-based platform for data-intensive biomedical research, supported by various institutions. The right sidebar shows a 'History' panel with a search bar and a list of datasets, including 'eba 2016 sartools' (1.59 MB) and several 'SARTools edgeR' analyses (40-48).

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

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

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

Dataset generation errors

Dataset 48: group2_count2.txt

The Galaxy framework encountered the following error while attempting to run the tool:

```
Traceback (most recent call last):  
  File "/w/galaxy/galaxy3/galaxy/lib/galaxy/jobs/runners/local.py", line  
    stdout_file.close()  
  File "/opt/python/lib/python2.7/tempfile.py", line 403, in close  
    self.unlink(self.name)  
OSError: [Errno 2] No such file or directory: '/w/galaxy/galaxy3/galaxy/database'
```

Tool execution generated the following error message:

```
failure running job
```

Report this error to the local Galaxy administrators

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.

Error Report

Your email

galaxy3.sb-roscoff.fr/dataset/errors?id=53eb0ef5c8056a28 e.gueguen@sb-roscoff.fr

History search datasets

eba 2016 sartools
28 shown, 14 deleted
1.59 MB

48: group2_count2.txt
tool error
An error occurred with this dataset:
failure running job

View or report this error

gene1	1353
gene10	72
gene100	496
gene1000	50

47: group2_count1.txt
46: group1_count2.txt
45: group1_count1.txt

Tools - Handle errors

Sent to the support team

The screenshot shows the Galaxy/ABiMS web interface. At the top, there is a navigation bar with 'Galaxy / ABiMS' and various menu items like 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right.

The main content area displays an error message: 'Tool execution generated the following error message: failure running job'. Below this, a large heading reads 'Report this error to the local Galaxy administrators'. The text explains that local administrators review errors and that providing additional information and a contact email address will help them investigate the problem.

An 'Error Report' form is shown with the following fields:

- Your email:** A text input field containing 'loraine.gueguen@sb-roscoff.fr'.
- Message:** A large text area for providing details about the error.
- Report:** A button to submit the error report.

On the right side, a 'History' panel is visible, showing a list of datasets. The top dataset is 'eba 2016 sartools' (1.59 MB). Below it, a red error entry is highlighted, labeled '48: group2_count2.txt', with the message 'tool error: An error occurred with this dataset: failure running job'. Below the error entry is a table of gene counts:

Gene	Count
1.gene0	2,1813
gene1	1353
gene10	72
gene100	496
gene1000	50

Below the error entry are two green entries: '47: group2_count1.txt' and '45: group1_count1.txt'. A red box highlights a scroll bar on the right side of the history panel.

DATASET

Dataset

Both inputs and outputs

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and a 'Using 0%' 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 green welcome banner for 'galaxy3.sb-roscoff.fr', an information box with contact details, and the ABiMS logo with the text 'Analyses and Bioinformatics for Marine Science'. Below the logo are links for 'Changelog' and 'Tutorials'. A paragraph of text describes Galaxy as an open, web-based platform for data-intensive biomedical research, supported by various institutions. On the right, the 'History' panel shows a list of datasets. The dataset '5: Trinity on data 3 and data 4: Assembled Transcripts' is highlighted with a red box, indicating it is the current dataset. Other visible datasets include '4: reads.left.fq' and '3: reads.right.fq'. The 'Trinity example' section shows '3 shown, 3 deleted' and '40.3 KB'.

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](#) @_g1_i1 len=541 path=[519:0-544
GTCTGAATTGCGATGTAATGCAGCTTTCCAGACACAAGTATGG
TCGCCATTGTGCAAAATATGTGTCTGATAGACCSCAGGCTTTCA
TGACATGAGCGTGGCACCTGAAGACAGGGTGTGGGTGAGAGGGTC
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 `3_g1_i1 len=541 path=[519:0-540]`

```
GTCTGAATTGCGATGTAATGCAGCTTTCCAGACACAAGTATGG  
TCGCCATTGTGCAAAATATGTGTCTGATAGACCSCAGGCTTTCA  
TGACATGAGCGTGGCACCTGAAGACAGGGTGTGGGTGAGAGGGTC  
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 [DN0_c0_g1_i1 len=541 path=\[519:0-544\]](#)

```
GTCTGAATTGCGATGTAATGCAGCTTTCCAGACACAAGTATGG  
TCGCCATTGTGCAAAATATGTGTCTGATAGACCCGACGGCTTCA  
TGACATGAGCGTGSCACCTGAAGACAGGGTGTGGSTGAGAGGGTC  
TGAGTTGTCTTGTATCATCAATAGATGCAAAATTAGATGTAAGAAC
```

4: reads.left.fq

3: reads.right.fq

Dataset

Re-run a job

The screenshot displays the Galaxy ABiMS interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and a 'Using 0%' indicator. The left sidebar contains a 'Tools' 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 workspace shows the configuration for a 'Trinity de novo assembly of RNA-Seq data (Galaxy Version 2.2.0.0)'. The 'Paired or Single-end data?' dropdown is set to 'Paired'. Under 'Left/Forward strand reads', the input list includes '5: Trinity on data 3 and data 4: Assembled Transcripts', '4: reads.left.fq' (highlighted), and '3: reads.right.fq'. The 'Right/Reverse strand reads' section has the same list with '3: reads.right.fq' highlighted. The 'Strand specific data' section has 'Yes' selected. The 'Strand-specific Library Type' dropdown is set to 'Reverse-Forward'. The 'Jaccard Clip options' section has 'Yes' selected. A note at the bottom states: 'You expect high gene density with UTR overlap (--jaccard_clip)'. The URL at the bottom is 'galaxy3.sb-roscoff.fr/tool_runner/rerun?id=c10ec933dc50450a'.

The right sidebar shows the 'History' panel with a search bar and a list of datasets. The top entry is 'Trinity example' (3 shown, 3 deleted, 40.3 KB). Below it is a highlighted entry: '5: Trinity on data 3 and data 4: Assembled Transcripts' (7 sequences, format: fasta, database: ?). A red box highlights a 'Run this job again' button with a tooltip showing 'len=541 path=[519:0-541]'. Below this are entries for '4: reads.left.fq' and '3: reads.right.fq'.

Dataset

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

The screenshot displays the Galaxy/ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right. 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 area shows a tabular dataset with 6 columns and 20 rows. The columns are labeled 1 through 6, and the rows are labeled with transcript IDs. The right sidebar shows a 'History' panel with a list of operations, including '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', '6: de results', '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', and '1: samples.txt'. A 'View data' tooltip is visible over the '5:' entry. The bottom status bar shows the URL: 'http://galaxy3.sb-roscoff.fr/datasets/4437d546e349a03a/display/?preview=True'.

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

Dataset

Renaming and annotation

The screenshot displays the Galaxy/ABiMS web interface. At the top, the navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right.

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 has four tabs: 'Attributes', 'Convert Format', 'Datatype', and 'Permissions'. The 'Attributes' tab is active, showing the following fields:

- Name:** matrix.counts.matrix
- Info:** uploaded tabular file
- Annotation / Notes:** This is my expression matrix.
- Database/Build:** unspecified (?)

Below these fields are 'Save' and 'Auto-detect' buttons. A note states: 'Add an annotation or notes to a dataset; annotations are available when a history is viewed. 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 operations:

- 8: 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 (41 lines, format: txt, database: ?) (uploaded tabular file) (Tags: trinity) (Annotation: This is my expression matrix.)

A tooltip 'Edit attributes' is visible over the '5:' entry in the history panel. The bottom of the page shows the URL: galaxy3.sb-roscoff.fr/datasets/4437d546e349a03a/edit.

Dataset

Renaming and annotation

The screenshot displays the Galaxy/ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right.

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 'Attributes' and includes tabs for 'Convert Format', 'Datatype', and 'Permissions'. The 'Edit Attributes' section contains the following fields:

- Name:** matrix.counts.matrix
- Info:** uploaded tabular file
- Annotation / Notes:** This is my expression matrix.
- Database/Build:** unspecified (?)

Below these fields are 'Save' and 'Auto-detect' buttons. A note states: 'Add an annotation or notes to a dataset; annotations are available when a history is viewed. 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 operations:

- 8: 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 (41 lines, format: txt, database: ?)

The '5: matrix.counts.matrix' entry is highlighted in green and includes a tooltip 'Edit attributes'. Below it, the 'Tags' section is highlighted with a red box and contains a tag 'x trinity'. The 'Annotation' field for this entry 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 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and 'Using 0%'. The left sidebar 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', 'Fasta Fastq Manipulation', 'COMMON NGS TOOLS', 'NGS:Samtools', 'NGS:Mapping', 'NGS:Bedtools', 'NGS:Picard Tools', 'SEARCHING TOOLS', and 'Diamond'. The main content area has four tabs: 'Attributes', 'Convert Format', 'Datatype' (highlighted with a red box), and 'Permissions'. Below the 'Datatype' tab, there is a section titled 'Change data type' with a 'New Type:' dropdown menu. The dropdown menu is open, showing a list of data types: 'txt', 'supermatcher', 'svg', 'swiss', 'syco', 'tabix', 'table', 'tabular' (highlighted in blue), 'tagseq', and 'tandem'. A text box next to the dropdown contains the text: 'Changing dataset but *not* modify its contents. and the type of your dataset.' The right sidebar shows a 'History' panel with a list of datasets, each with a title, a list of datasets, and icons for viewing, editing, and deleting. The datasets listed are: 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: extracted differentially expressed genes, and 8: Extract and cluster differentially expressed transcripts on data 2, data 3, and others.

1 Visualize charts

name	Bl...	Scatterplot
C1_011	HU...	...

New Chart Cancel Draw

Start Configuration 1: Data label Add Data

Provide a chart title:

New Chart

How many data points would you like to analyze?

Few (<500) Some (<10k) Many (>10k)

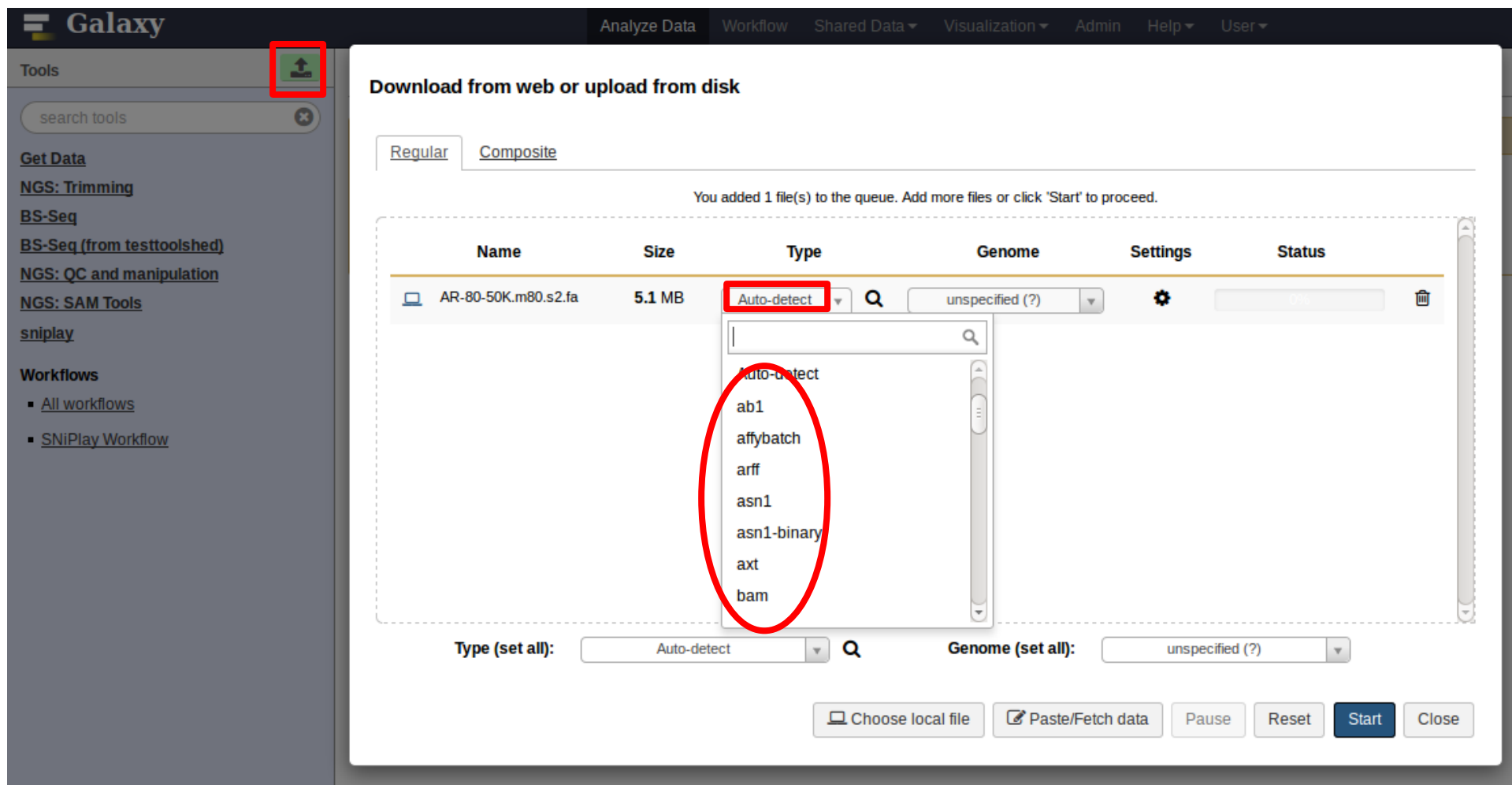
- Bar diagrams
 - Regular (NVD3)
 - Stacked (NVD3)
 - Horizontal (NVD3)
 - Stacked horizontal (NVD3)
- Others
 - QLine with focus (NVD3)
 - QLine chart (NVD3)
 - QScatter plot (NVD3)
 - QHeatmap (Custom)
- Area charts
 - QRegular (NVD3)
 - QExpanded (NVD3)
 - QStream (NVD3)
 - Pie chart (NVD3)
- Data processing (requires 'charts' tool from Toolshed)
 -
 -
 -
 -



Datatypes

DATASET

- 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

Dataset - Datatypes

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

The screenshot shows the Galaxy ABiMS interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and 'Using 0%'. 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', and 'SEARCHING TOOLS', 'Diamond'.

The main panel shows the 'Datatype' tab for a dataset. The 'Change data type' section is active, and the 'New Type:' dropdown menu is open, showing a list of datatypes: txt, supermatcrier, svg, swiss, syco, tabix, table, **tabular**, tagseq, and tandem. A red box highlights the 'tabular' option. A red box also highlights a heatmap visualization of a dataset in the history panel, with a red line connecting it to the 'tabular' option in the dropdown.

The history panel on the right shows a list of datasets, including 'matrix.counts.matrix' and 'samples.txt'. The 'matrix.counts.matrix' dataset is highlighted with a red box, and a red line connects it to the 'tabular' option in the dropdown.

- 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

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')

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

- ***html***: standard language for web pages

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

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

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

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](#)

Changelog

Tutorials

Galaxy is an open source platform for collaborative biomedical research. The Galaxy team is a mix of biologists, computer scientists, and statisticians. The Galaxy Project is supported by the French National Research Agency (ANR), the French League for the Advancement of Science, The Institute for Genome Sciences and Policy, and the Center for Computational Biology and Bioinformatics.

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

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

Filter and Sort

Filter data on any column using simple expressions

Sort data in ascending or descending order

Select lines that match an expression

GFF

Extract features from GFF data

Filter GFF data by attribute using simple expressions

Filter GFF data by feature count using simple expressions

Filter GTF data by attribute

Changelog

Tutorials

Galaxy is an open source platform for collaborative biomedical research. The Galaxy team is a multidisciplinary team in Computer Science, Biology and Mathematics and The Galaxy Project is supported in part by NHGRI, the National Center for Human Genome Research, 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

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 [support@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
atgcgtttgcgtgcatgcgtttgcgtgcatgcgtttgcgtgcatgcgtttgcgtgc
atgcgtttgcgtgc
>sequence2
tttcgtgcgatatagtttcgtgcgatatagtttcgtgcgatatag
tggcgcggt
```

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

```
@SEQ_ID
GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTTT
+
!''*(((((***+))%%%+)) (%%%) .1***-+*'' ) **55CCF>>>>>CCCCCCC65
@SEQ_ID2
GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTTT
+
!''*(((((***+))%%%+)) (%%%) .1***-+*'' ) **55CCF>>>>>CCCCCCC65
```

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

- [Filter sequences by ID from a tabular file](#)
- [FastQC Read Quality reports](#)
- [FASTQ Groomer convert between various FASTQ quality formats](#)

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
cttctgggcgtacccgattctcggagaacttgccgcaccattccgccttg
tgttcattgctgcctgcatgttcattgtctacctcggctacgtgtggcta
...
```

Cleanup

DATASET



Dataset

Delete a dataset

The screenshot shows the Galaxy / ABiMS web interface. The top navigation bar includes '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', 'COMMON NGS TOOLS', and 'SEARCHING 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 text 'Analyses and Bioinformatics for Marine Science'. Below the logo are links for 'Changelog' and 'Tutorials'. A paragraph of text describes Galaxy as an open, web-based platform for data intensive biomedical research. The right sidebar shows a 'History' panel with a search bar and a list of datasets. The dataset '62: SARTools DESeq2 R objects (.RData)' is highlighted in green, and its 'Delete' button (marked with a red 'X') is being clicked, as indicated by a tooltip.

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

javascript:void(0);

✓ 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

ABiMS
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

eba 2016 sartools
41 shown, 1 deleted
1.59 MB

62: SARTools DESeq2 R objects (.RData) Delete

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

The screenshot shows the Galaxy / ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and a 'Using 0%' 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 green welcome banner for 'galaxy3.sb-roscoff.fr', an information box with contact details, the ABiMS logo (Analyses and Bioinformatics for Marine Science), and a list of links for 'Changelog' and 'Tutorials'. The right sidebar shows a 'History' panel with a search bar and a list of datasets. The dataset 'eba 2016 sartools' is highlighted with a red box, showing '42 shown, hide deleted' and '1.59 MB'. Below it, a yellow warning box states 'This dataset has been deleted' and provides links to 'Undelete it' and 'Permanently remove it from disk'. The history list includes items like '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)', and '56: SARTools edgeR'.

Dataset



“Empty Trash” : to free up disk space

The screenshot shows the Galaxy / ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and 'Using 0%'. 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', and 'SEARCHING TOOLS' with 'Diamond' listed.

The main content area features a green welcome banner: 'Welcome to galaxy3.sb-roscoff.fr'. Below it is an information box: 'Information For any question or request for tools or account, send an email at support.abims@sb-roscoff.fr'. The center displays the 'ABiMS 4 Analyses and Bioinformatics for Marine Science' logo and the 'Station Biologique Roscoff' logo. There are links for 'Changelog' and 'Tutorials'. A paragraph of text describes Galaxy as an open, web-based platform for data intensive biomedical research, mentioning its affiliation with Penn State and Emory University, and its funding sources (NHGRI, NSF, etc.).

The right sidebar shows a 'History' section with a dropdown menu open. The menu items are: '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' (highlighted), 'DOWNLOADS', 'Export Tool Citations', 'Export History to File', and 'OTHER ACTIONS', 'Import from File'. A red box highlights the gear icon in the History header.

The browser address bar at the bottom shows: galaxy3.sb-roscoff.fr/history/purge_deleted_datasets

HISTORY

History panel

Both inputs and outputs

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and a 'Using 0%' 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), and 'SEARCHING TOOLS' (Diamond).

The main content area features a green welcome banner for 'galaxy3.sb-roscoff.fr', an information box with contact details for support.abims@sb-roscoff.fr, and the ABiMS logo (Analyses and Bioinformatics for Marine Science) with links to Changelog and Tutorials. Below the logo is a paragraph of text: '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.'

The right sidebar shows the 'History' panel with a search bar and a list of datasets. The top dataset is 'eba 2016 sartools' (28 shown, 14 deleted, 1.59 MB). Below it is a red dataset '48: group2_count2.txt'. The following datasets are green: '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'. Each dataset entry includes icons for viewing, editing, and deleting.

History panel

Both inputs and outputs

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and a 'Using 0%' 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 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'. The right sidebar shows the 'History' panel with a search bar and a list of datasets. A red box highlights the top entries in the history list:

- 48: group2_count2.txt (marked with a red X)
- 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 panel renaming and annotation

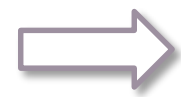
History

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:
 RNAseq
 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 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right. 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 green welcome message for 'galaxy3.sb-roscoff.fr', an information box with contact details, and the 'ABiMS 4' logo for 'Analyses and Bioinformatics for Marine Science' at 'Station Biologique Roscoff'. Below the logo are links for 'Changelog' and 'Tutorials', followed by a paragraph of text describing the Galaxy platform. On the right, the 'History' panel is open, showing a menu with categories: '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). A red box highlights the gear icon in the History panel header.

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 dese2	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

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

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

History panel

Saved histories: Switch histories

The screenshot displays the Galaxy ABiMS interface. The top navigation bar includes 'Galaxy / ABiMS' and various menu items like 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. The left sidebar contains a 'Tools' section with a search bar and several 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' with 'Diamond' listed.

The main area shows the 'History' panel with a search bar and an 'Advanced Search' link. Below is a table of saved histories:

<input type="checkbox"/>	Name	Datasets	Tags	Sharing	Size on Disk	Created	Last
<input type="checkbox"/>	eba2016 deseq2	8	0 Tags		829.2 KB	Nov 14, 2016	Nov 14, 2016
<input type="checkbox"/>	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
<input type="checkbox"/>	eba 2016 macs2	3 2	0 Tags		602.8 KB	Nov 09, 2016	Nov 09, 2016
<input type="checkbox"/>	eba 2016 sickle	10	0 Tags		4.2 MB	Oct 04, 2016	Nov 04, 2016

Below the table, there is a pagination control: 'Page: 1 2 | Show All'. At the bottom of the history list, it says 'For 0 selected histories: Rename Delete Delete Permanently Undelete'.

A dropdown menu is open for the 'eba2016 deseq2' history, showing 'HISTORY LISTS' with 'Saved Histories' selected, and 'HISTORY ACTIONS' with options like 'Create New', 'Copy History', 'Share or Publish', 'Show Structure', 'Extract Workflow', 'Delete', 'Delete Permanently', and 'DATASET ACTIONS' with options like 'Copy Datasets', 'Dataset Security', 'Resume Paused Jobs', 'Collapse Expanded Datasets', 'Unhide Hidden Datasets', 'Delete Hidden Datasets', 'Purge Deleted Datasets', 'DOWNLOADS' with 'Export Tool Citations', 'Export History to File', and 'OTHER ACTIONS' with 'Import from File'.

At the bottom of the history panel, a note states: 'Histories that have been deleted for more than a time period specified by the Galaxy administrator(s) may be permanently deleted.'

The browser address bar at the bottom left shows 'galaxy3.sb-roscoff.fr/history/list'.

History panel

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

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right. 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 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', followed by a paragraph of text describing the Galaxy platform and its affiliations. 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 also highlighted with a red box. Other actions include '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'.

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

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

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and a 'Using 0%' 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 green 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 of text describes Galaxy as an open, web-based platform for data-intensive biomedical research, supported by various institutions. The right sidebar is the 'History' panel, which lists datasets with their names, sizes, and actions. The top of the history panel has a search bar and icons for refresh, settings, and a red-bordered icon. The history list includes datasets like 'eba 2016 sartools', 'group2_count2.txt', 'group2_count1.txt', 'group1_count2.txt', 'group1_count1.txt', 'SARTools edgeR R objects (.RData)', 'SARTools edgeR R log', 'SARTools edgeR figures', 'SARTools edgeR tables', and '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

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

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.fg](#)
- 3: [reads.right.fg](#)

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 da

- 16: [Extra differentially transcripts of RData file](#)
- 15: [Extract expressed to data 3, and depleted cat a list of datase](#)
- 14: [Extract expressed to data 3, and a list of datase](#)
- 13: [Extra differentially transcripts of](#)
- 12: [Extra differentially](#)

History panel

The screenshot shows the Galaxy / ABiMS interface with a history panel. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and 'Using 0%'. Below the navigation bar, there are search bars for 'search histories' and 'search all datasets'. The main area displays a 'Current History' panel with a 'Switch to' button highlighted by a red box. The history panel is divided into several columns, each representing a different history item. The first column is titled 'eba 2016 sartools' (1.59 MB) and contains items 40 through 48. The second column is titled 'Trinity example' (40.3 KB) and contains items 3, 4, and 5. The third column is titled 'trinity_contig_exn50_statistic' (47.01 KB) and contains items 8 through 14. The fourth column is titled 'eba 2016 tr' (21.92 KB) and contains items 12 through 16. Each item in the history panel has a title, a description, and a set of icons (eye, pencil, and X) for viewing, editing, and deleting the item.

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 Switch to 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.fg
- 3: reads.right.fg

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 da

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

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.fg
- 3: reads.right.fg

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 da

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

History panel

The screenshot displays the Galaxy / ABiMS interface. At the top, there is a navigation bar with the following items: Galaxy / ABiMS, Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, User, and a grid icon. On the right side of the navigation bar, it says "Using 0%" and "Create new". Below the navigation bar, there is a search bar with two input fields: "search histories" and "search all datasets". A red box highlights these two search fields. Below the search bar, there are three panels, each with a "Switch to" dropdown menu. The first panel is titled "Current History" and shows a list of datasets. The second panel is titled "Trinity example" and shows a list of datasets. The third panel is titled "trinity_contig_exn50_statistic" and shows a list of datasets. The fourth panel is partially visible on the right and is titled "eba 2016 tr". Each panel has a search bar and a list of datasets with icons for viewing, editing, and deleting. The datasets in the "Current History" panel include: 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, and 40: SARTools edgeR report. The datasets in the "Trinity example" panel include: 5: Trinity on data 3 and data 4: Assembled Transcripts, 4: reads.left.fg, and 3: reads.right.fg. The datasets in the "trinity_contig_exn50_statistic" panel include: 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), and 8: Build expression matrix on data 7 and data 6: estimated.

History panel

The screenshot displays the Galaxy / ABiMS interface. At the top, there is a navigation bar with the following elements: the Galaxy / ABiMS logo, a search bar for "search histories", another search bar for "search all datasets", and a "Create new" button highlighted with a red box. Below the navigation bar, the "Current History" section is visible, showing a list of workflow history items. Each item is represented by a card with a title, a description, a size, and a search bar. The items are organized into three columns, each with a "Switch to" button. The first column shows items 40 through 48, including "SARTools edgeR report" and "group2_count1.txt". The second column shows items 3 through 5, including "reads.left.fg" and "Trinity on data 3 and data 4: Assembled Transcripts". The third column shows items 8 through 14, including "Build expression matrix on data 7 and data 6: estimated RNA-Seq fragment counts (raw counts)".

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 Switch to 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.fg

3: reads.right.fg

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 da

16: Extra differentially transcripts of RData file

15: Extract expressed to data 3, and depleted cat a list of datase

14: Extract expressed to data 3, and a list of datase

13: Extra differentially transcripts of

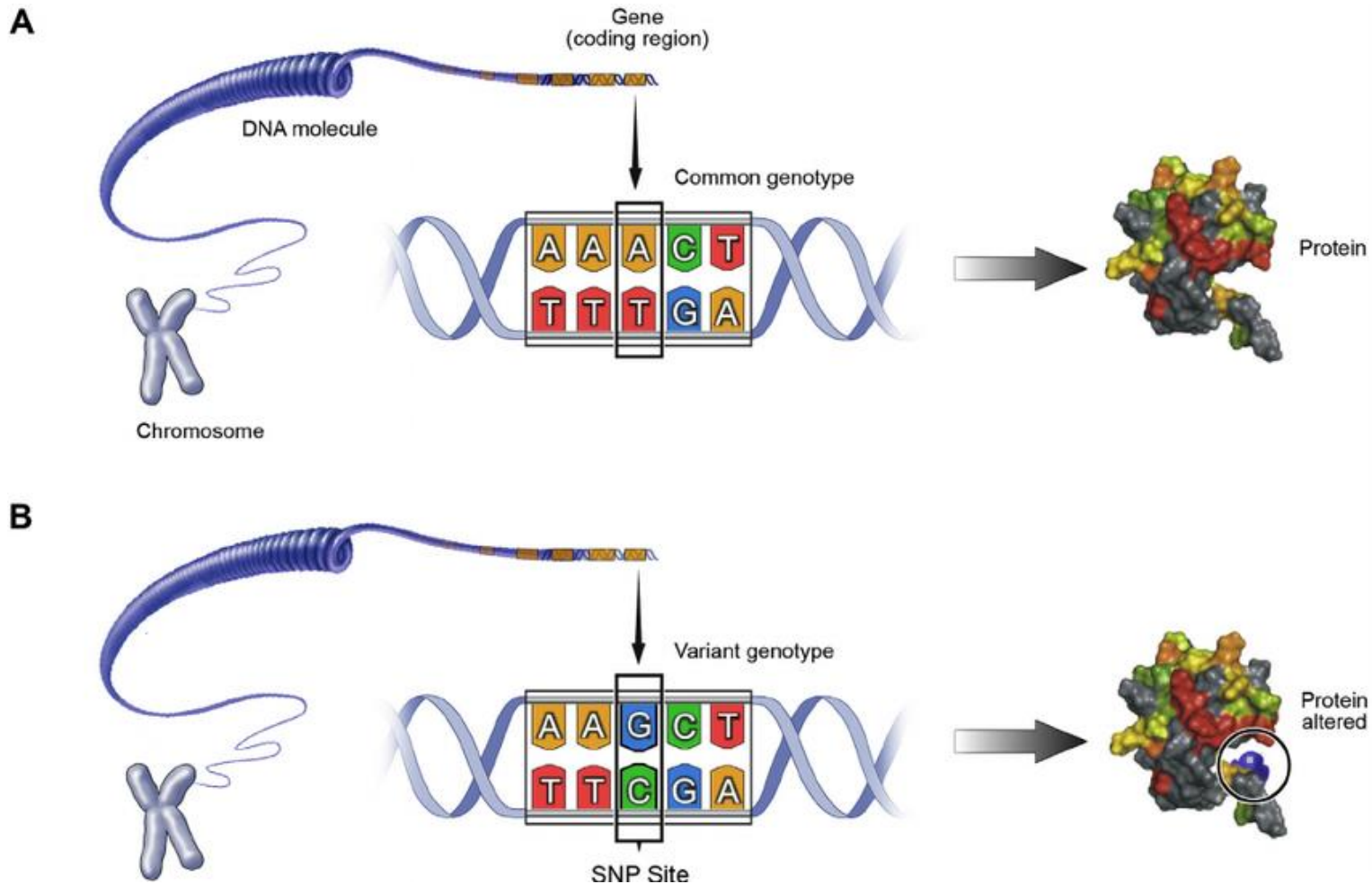
12: Extra differentially

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?

1. Get the data in a new history
2. Join exons with SNPs
3. Count the number of SNPs per exon
4. Sort exons by SNP count
5. Select top five
6. Build a bar diagram
7. Recover exon info and display data in genome browsers



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

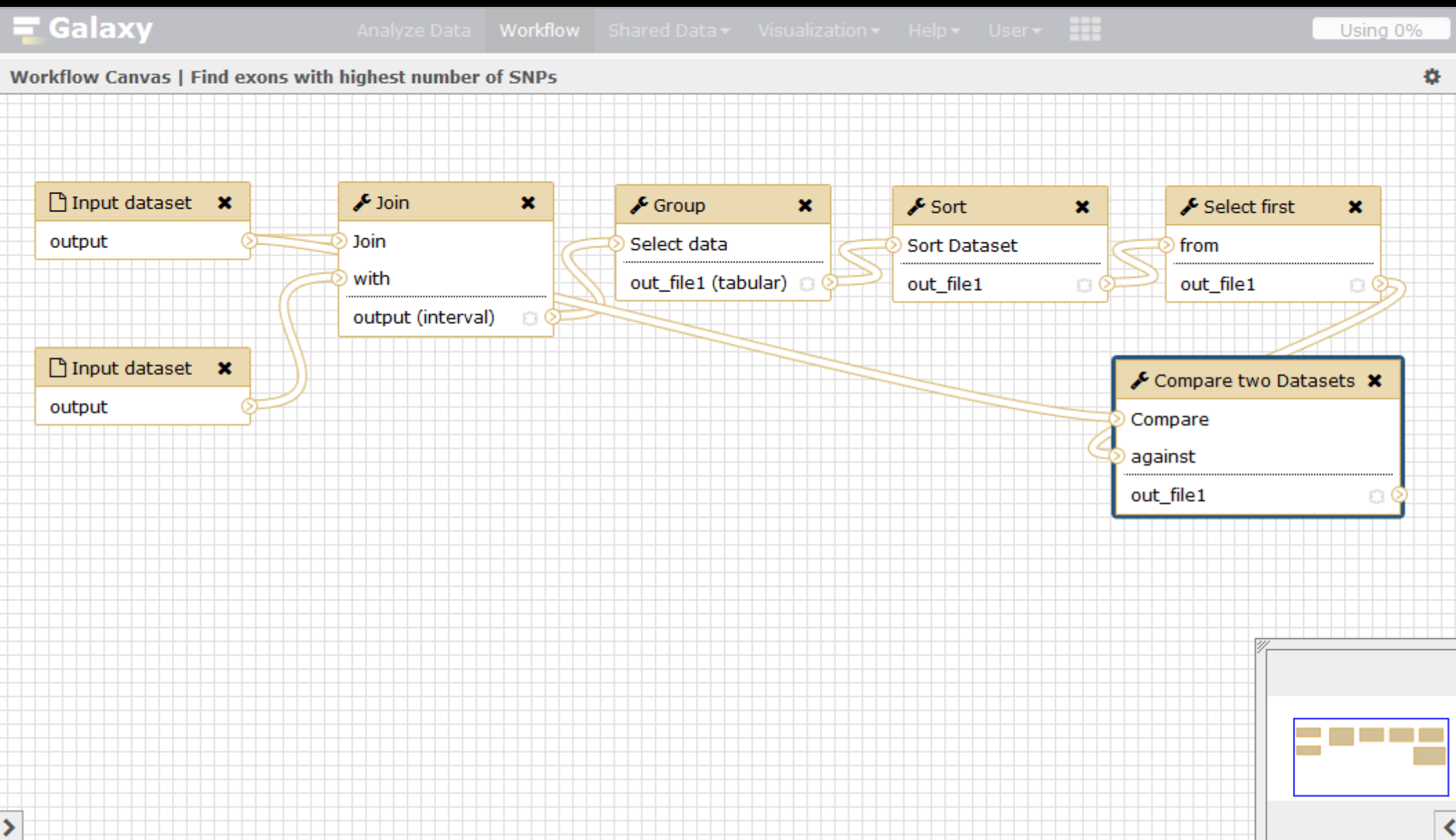
1. Get the data in a new history
2. Join exons with SNPs *Operate on Genomics Intervals -> Join*
3. Count the number of SNPs per exon *Join, Subtract, and Group -> Group*
4. Sort exons by SNP count *Filter and Sort -> Sort*
5. Select top five *Text Manipulation -> Select First*
6. Build a bar diagram *Visualize -> Charts*
7. Recover exon info and display data in genome browsers *Join, Subtract and Group -> Compare two Datasets*

WORKFLOW

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

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 <input checked="" type="checkbox"/> Treat as input dataset Exons
UCSC Main <i>This tool cannot be used in workflows</i>	2 SNPs <input checked="" type="checkbox"/> Treat as input dataset SNPs
Join <input checked="" type="checkbox"/> Include "Join" in workflow	3 Join on data 2 and data 1
Group <input checked="" type="checkbox"/> Include "Group" in workflow	5 Group on data 3
Sort <input checked="" type="checkbox"/> Include "Sort" in workflow	6 Sort on data 5

https://usegalaxy.org/workflow/build_from_current_history

History [Refresh] [Close]

- 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

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 <i>This tool cannot be used in workflows</i>	1 Exons <input checked="" type="checkbox"/> Treat as input dataset Exons
UCSC Main <i>This tool cannot be used in workflows</i>	2 SNPs <input checked="" type="checkbox"/> Treat as input dataset SNPs
Join <input checked="" type="checkbox"/> Include "Join" in workflow	3 Join on data 2 and data 1
Group <input checked="" type="checkbox"/> Include "Group" in workflow	5 Group on data 3
Sort <input checked="" type="checkbox"/> 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

Workflow

From history

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 <i>This tool cannot be used in workflows</i>	1 Exons <input checked="" type="checkbox"/> Treat as input dataset Exons
UCSC Main <i>This tool cannot be used in workflows</i>	2 SNPs <input checked="" type="checkbox"/> Treat as input dataset SNPs
Join <input checked="" type="checkbox"/> Include "Join" in workflow	3 Join on data 2 and data 1
Group <input checked="" type="checkbox"/> Include "Group" in workflow	5 Group on data 3
Sort <input checked="" type="checkbox"/> 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

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](#)

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

Workflow

Edit a workflow: add tags and annotation

The screenshot displays the Galaxy workflow editor interface. At the top, the navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right corner. The main area is divided into three sections: 'Tools', 'Workflow Canvas', and 'Details'.

Tools: A search bar labeled 'search tools' is at the top. Below it, a list of tool categories is provided, including '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', and 'NGS: Du Novo'.

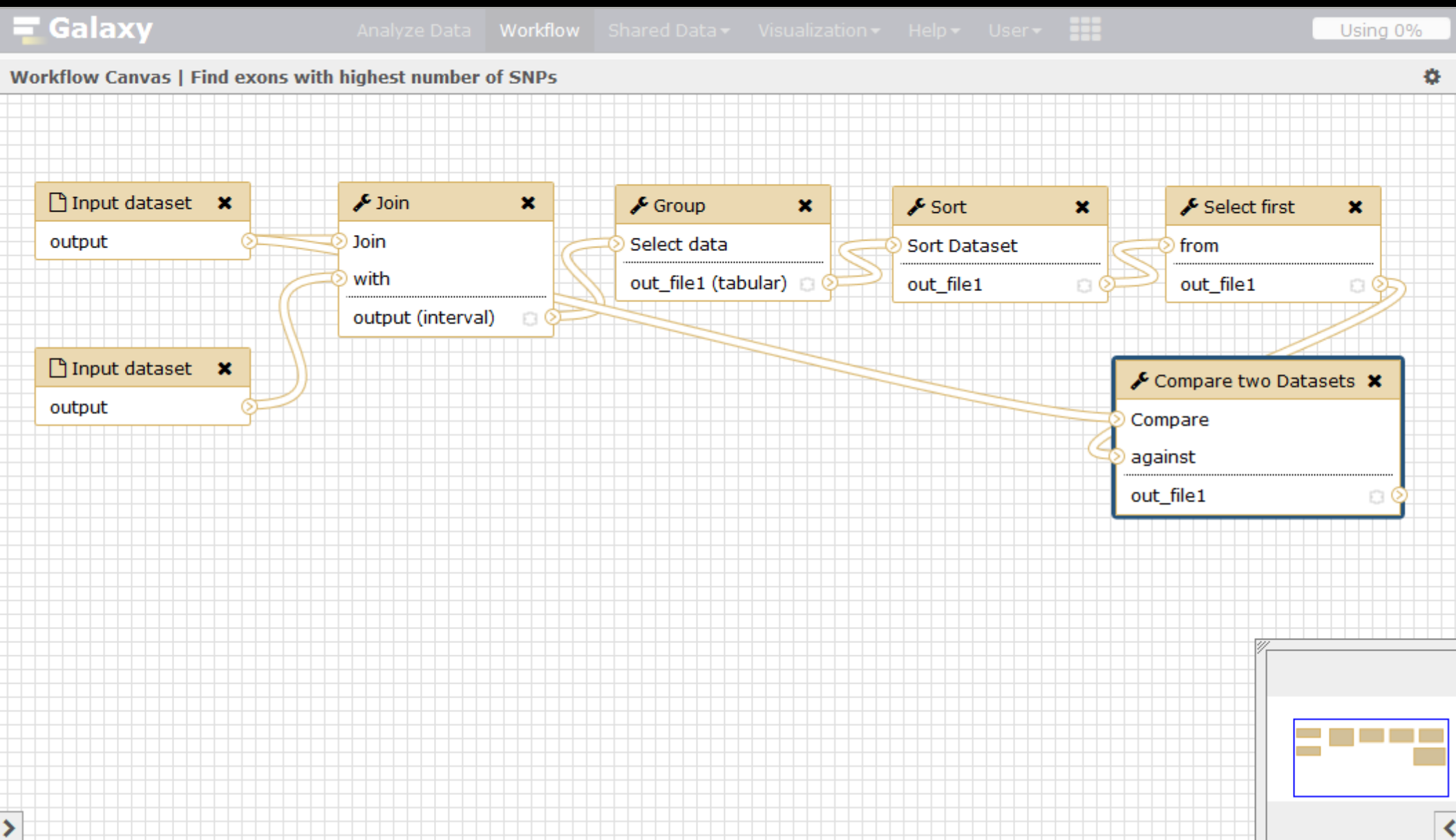
Workflow Canvas: The canvas is titled 'Workflow Canvas | Find exons with highest number of SNPs'. It features a grid background with three main workflow steps: two 'Input dataset' blocks on the left, a central 'Join' block, and a 'Group' block on the right. The 'Join' block is configured with 'with output (interval)'. The 'Group' block is configured with 'out_file1 (tab)'. Arrows indicate the flow of data from the input datasets through the join step to the group step.

Details: The 'Details' panel on the right is titled 'Edit Workflow Attributes' and is highlighted with a red border. It contains the following information:

- Name:** Find exons with highest number of SNPs
- Tags:** A section with a plus icon and a text box containing 'out_file1 (tab)'. Below this, a description reads: 'Apply tags to make it easy to search for and find items with the same tag.'
- Annotation / Notes:** A section with a text box containing the text: '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

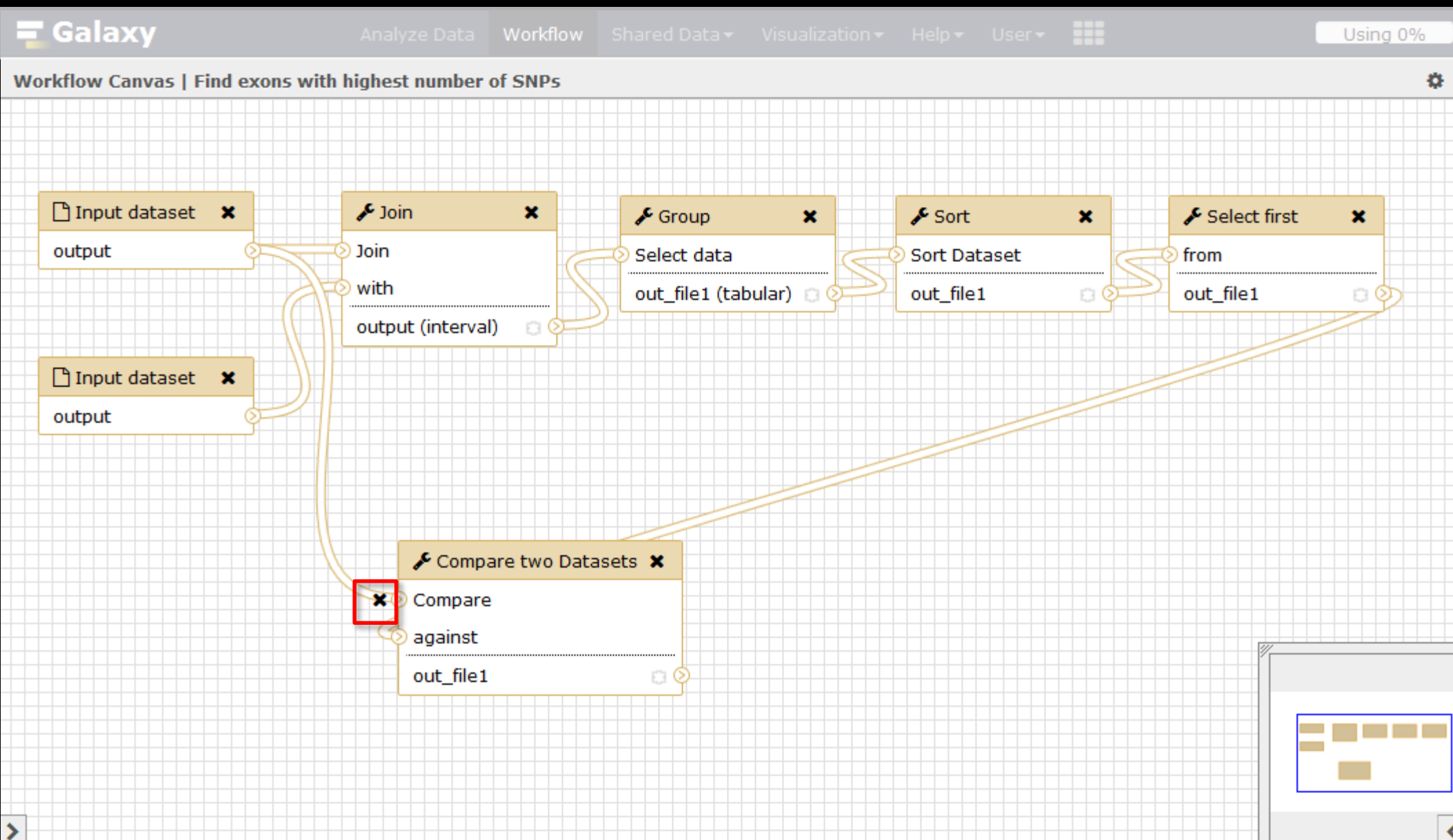
The screenshot displays the Galaxy Workflow Canvas interface. At the top, the navigation bar includes the Galaxy logo, menu items for 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User', and a 'Using 0%' indicator. The main title of the canvas is 'Workflow Canvas | Find exons with highest number of SNPs'. The workflow consists of the following steps:

- Input dataset**: Two separate input boxes, each labeled 'Input dataset' with a close button (x) and an 'output' field.
- Join**: A 'Join' tool box with a close button (x). It contains the text 'Join with output (interval)' and a refresh icon.
- Group**: A 'Group' tool box with a close button (x). It contains the text 'Select data out_file1 (tabular)' and a refresh icon.
- Sort**: A 'Sort' tool box with a close button (x). It contains the text 'Sort Dataset out_file1' and a refresh icon.
- Select first**: A 'Select first' tool box with a close button (x). It contains the text 'from out_file1' and a refresh icon.
- Compare two Datasets**: A 'Compare two Datasets' tool box with a close button (x). It contains the text 'Compare against out_file1' and a refresh icon.

Orange lines represent the workflow connections. The two 'Input dataset' boxes connect to the 'Join' tool. The 'Join' tool connects to the 'Group' tool. The 'Group' tool connects to the 'Sort' tool. The 'Sort' tool connects to the 'Select first' tool. The 'Select first' tool connects to the 'Compare two Datasets' tool. Additionally, a long orange line connects the 'Compare two Datasets' tool back to the 'Join' tool, indicating a feedback loop or comparison of the workflow's output against its input.

Workflow

Edit a workflow: delete a noodle



Workflow

Edit a workflow: delete a noodle

The screenshot shows the Galaxy Workflow Canvas interface. The title bar reads "Workflow Canvas | Find exons with highest number of SNPs". The main workspace contains a workflow with the following steps:

- Two "Input dataset" nodes, each with an "output" field.
- A "Join" step with fields "Join", "with", and "output (interval)".
- A "Group" step with fields "Select data" and "out_file1 (tabular)".
- A "Sort" step with fields "Sort Dataset" and "out_file1".
- A "Select first" step with fields "from" and "out_file1".
- A "Compare two Datasets" step with fields "Compare", "against", and "out_file1".

Connections between steps are shown as orange lines. A red box highlights the "Compare two Datasets" step, indicating it is the target for deletion. The interface includes a top navigation bar with "Galaxy" and various menu items, and a bottom status bar showing "Using 0%".

Workflow

Edit a workflow: add a tool

The screenshot displays the Galaxy workflow editor interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right. The main area is titled 'Workflow Canvas | Find exons with highest number of SNPs'. On the left, a 'Tools' panel lists various categories: 'Get Data', 'Send Data', 'Lift-Over', 'Text Manipulation', and 'Merge Columns together' (highlighted with a red box). The workflow canvas shows a sequence of tools: two 'Input dataset' tools feeding into a 'Join' tool, which then feeds into a 'Group' tool, followed by a 'Sort' tool. A 'Merge Columns' tool (highlighted with a red box) is positioned below the 'Join' tool, and a 'Compare two Datasets' tool is positioned below the 'Group' tool. A small preview window in the bottom right corner shows a grid of data points.

Workflow

Edit a workflow: add a noodle

The screenshot shows the Galaxy workflow editor interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right. The left sidebar contains a 'Tools' panel with a search bar and a list of tool categories: 'Get Data', 'Send Data', 'Lift-Over', and 'Text Manipulation'. The 'Text Manipulation' section is expanded, showing tools like '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', and 'Remove beginning of a file'. The main 'Workflow Canvas' is titled 'Find exons with highest number of SNPs'. It contains several tools: two 'Input dataset' tools, a 'Join' tool, a 'Merge Columns' tool, a 'Compare two Datasets' tool, a 'Group' tool, and a 'Sort' tool. The 'Join' tool is highlighted with a red box, and a red arrow points to its output port. A green line connects the output of the 'Join' tool to the input of the 'Merge Columns' tool. The 'Merge Columns' tool is highlighted with a blue box. The 'Compare two Datasets' tool is also highlighted with a blue box. The 'Group' tool has a 'Select data' section with 'out_file1 (tabular)'. The 'Sort' tool has a 'Sort Dataset' section with 'out_file1'. A small preview window in the bottom right corner shows a grid of data.

Workflow

Edit a workflow: hide intermediate steps

The screenshot shows the Galaxy Workflow Canvas interface. The title bar reads "Workflow Canvas | Find exons with highest number of SNPs". The navigation menu includes "Analyze Data", "Workflow", "Shared Data", "Visualization", "Help", and "User". A "Using 0%" indicator is in the top right.

The workflow consists of the following steps:

- Join**: Takes two "Input dataset" steps as input. The output is labeled "output (interval)".
- Group**: Takes the output from the "Join" step as input. The output is labeled "out_file1 (tabular)".
- Sort**: Takes the output from the "Group" step as input. The output is labeled "out_file1".
- Select first**: Takes the output from the "Sort" step as input. The output is labeled "out_file1".
- Compare two Datasets**: Takes the output from the "Select first" step as input. The output is labeled "out_file1".

A tooltip is displayed over the "Compare two Datasets" step, containing the following text:

Mark dataset as a workflow output. All unmarked datasets will be hidden.

Workflow

Edit a workflow: set or release a parameter

The screenshot shows the Galaxy Workflow Canvas interface. The workflow is titled "Find exons with highest number of SNPs" and consists of the following steps:

- Group**: Select data from `out_file1 (tabular)`.
- Sort**: Sort Dataset from `out_file1`.
- Select first**: Select first 5 lines from `out_file1`. This tool is highlighted with a red box.
- Compare two Datasets**: Compare against `out_file1`.

The configuration panel for the **Select first** tool is open on the right, showing the following details:

- Select first lines from a dataset (Galaxy Version 1.0.0)**
- Select first**: lines
- from**: Data input 'input' (txt)
- Annotation / Notes**:
- Email notification**: An email notification will be sent when the job has completed.
- Output cleanup**:

Workflow

Edit a workflow: set or release a parameter

The screenshot displays the Galaxy Workflow Canvas for a workflow titled "Find exons with highest number of SNPs". The workflow is composed of the following steps:

- Group**: Select data (out_file1 (tabular))
- Sort**: Sort Dataset (out_file1)
- Select first**: from (out_file1) - This step is highlighted with a blue border.
- Compare two Datasets**: Compare against (out_file1)
- Select first**: (out_file1)

The configuration panel for the selected "Select first" step is visible on the right, showing the following details:

- Select first lines from a dataset (Galaxy Version 1.0.0)**
- Select first** (highlighted with a red box)
- Set at Runtime** (dropdown menu)
- lines**
- from**: Data input 'input' (txt)
- Annotation / Notes**: (Empty text area)
- Email notification**: Yes No (radio buttons)
- Output cleanup**: Yes No (radio buttons)

Workflow

Edit a workflow: rename the outputs

The screenshot shows the Galaxy workflow editor interface. The workflow canvas displays four steps connected by arrows:

- Group**: Select data, out_file1 (tabular)
- Sort**: Sort Dataset, out_file1
- Select first**: from, out_file1
- Compare two Datasets**: Compare, against, out_file1

The 'Compare two Datasets' step is highlighted with a red box. Its configuration panel is open, showing the 'Rename dataset' section with a text input field containing 'Top exons'.

Details

step. It will be shown with the workflow.

Email notification

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

Output cleanup

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

Group Select data out_file1 (tabular)

Sort Sort Dataset out_file1

Select first from out_file1

Compare two Datasets Compare against out_file1

Save
Save As
Run
Edit Attributes
Auto Re-layout
Close

Details

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

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow' (highlighted with a red box), 'Shared Data', 'Visualization', 'Help', 'User', and a 'Using 0%' indicator. Below the navigation bar, the 'Your workflows' section contains two buttons: 'Create new workflow' and 'Upload or import workflow'. A table lists workflows with columns for 'Name' and '# of Steps'. The first workflow, 'Find exons with highest number of SNPs', is highlighted with a red box. A context menu is open over this workflow, listing actions: 'Edit', 'Run' (highlighted with a red box), 'Share or Download', 'Copy', 'Rename', 'View', and 'Delete'. Below the table, there are sections for 'Workflows shared by others' and 'Other options'.

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

The screenshot displays the Galaxy workflow editor interface. The main workspace shows a workflow titled "Workflow: Find exons with highest number of SNPs" with a "Run workflow" button. The workflow steps are:

- History Options**: Send results to a new history (Yes/No).
- 1: Input dataset**: A dropdown menu showing "2: Exons" (highlighted with a red box).
- 2: Input dataset**: A dropdown menu showing "1: Repeats" (highlighted with a red box).
- 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)**: A sub-section with a "Select first" label and a text input field containing "20" (highlighted with a red box), with "lines" written below it.

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, 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 right sidebar shows the "History" section with a search bar and a list of datasets: "Galaxy initiation - workflow" (2 shown, 2.77 MB) and "2: Exons" (highlighted in green). Below it, "1: Repeats" is also highlighted in green.

Workflow

Run a workflow

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

Tools search tools

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%

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 2 shown
Galaxy initiation - workflow 2.77 MB
2: Exons
1: Repeats

Workflow

Run a workflow

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 visible in the top right corner. On the left, a 'Tools' sidebar lists various categories such as 'Get Data', 'Send Data', 'Text Manipulation', and 'NGS: QC and manipulation'. The main workspace 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.' On the right, the 'History' pane shows a search bar and a list of workflow steps. Steps 1 through 7 are listed, with steps 3 through 7 highlighted in a red box. Step 1 is '1: Repeats', step 2 is '2: Exons', step 3 is '3: Join on data 1 and data 2', step 4 is '4: Group on data 3', step 5 is '5: Sort on data 4', step 6 is '6: Top exons', and step 7 is '7: Top exon genetic location'. Each step includes icons for viewing, editing, and deleting.

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.**
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 - 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. At the top, the navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right corner. On the left, a 'Tools' sidebar lists various categories such as 'Get Data', 'Send Data', 'Text Manipulation', and 'NGS: QC and manipulation'. The main workspace contains 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.' On the right, the 'History' panel shows a search bar and a list of workflow steps. The top step is 'Galaxy initiation - workflow' with 4 shown and 3 hidden. Below it are steps '7: Top exon genetic location', '6: Top exons', '2: Exons', and '1: Repeats', each with view, edit, and delete icons. A red box highlights the '7: Top exon genetic location' and '6: Top exons' steps.

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.**
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 - workflow
4 shown 3 hidden

2.92 MB

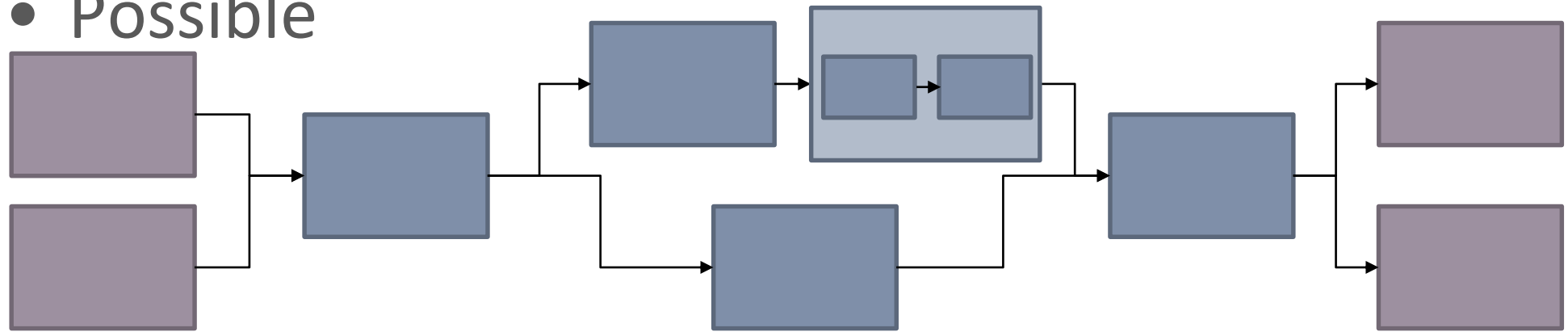
7: Top exon genetic location

6: Top exons

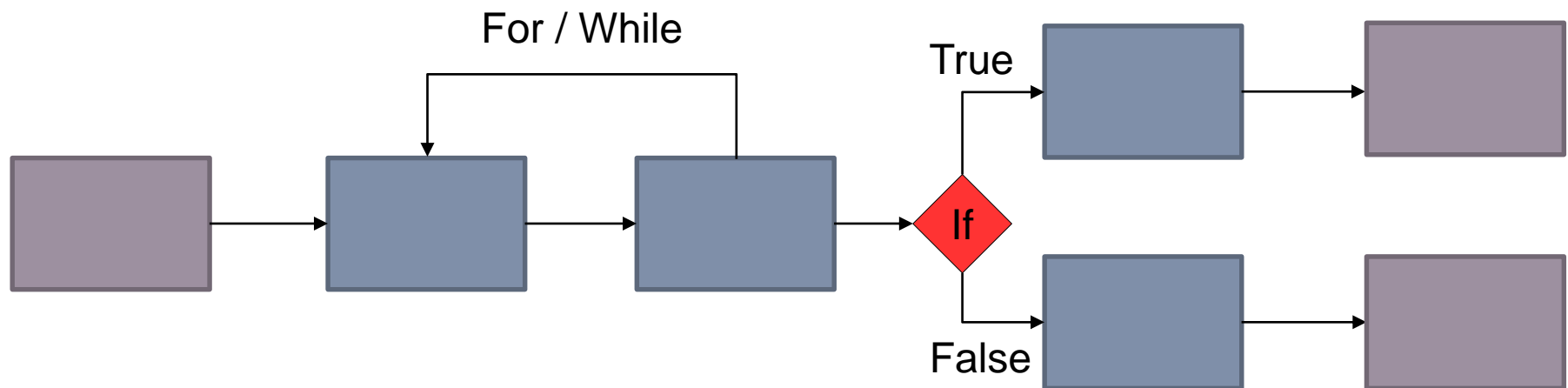
2: Exons

1: Repeats

- Possible



- Impossible (until now)



SHARE

biologist ↔ 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 displays the Galaxy web interface. The main content area is titled "Saved Histories" and contains a table of saved history items. A context menu is open over the first item, "Galaxy initiation - workflow", with the "Share or Publish" option highlighted. The right sidebar shows a "History" panel with a search bar and a list of datasets.

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

Saved Histories

search history names and tags

Advanced Search

Name	Datasets	Tags	Sharing	Size on Disk	Created	Last Upd
Galaxy initiation - workflow	4	0 Tags		2.9 MB	~5 hours ago	~4 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 search datasets

Galaxy initiation - workflow
4 shown, 3 hidden
2.92 MB

7: Top exon genetic location
6: Top exons
2: Exons
1: Repeats

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

Share

Workflow



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

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

Workflows shared with others

No workflows have been shared with others

Other options

Configure your workflow menu

Share

Mode

Galaxy Analyze Data Workflow Shared Data Visualization Help User 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.

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

→ **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.

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

Export

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.

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

<input type="checkbox"/>	Name	Datasets	Created	Last Updated↑	Shared by
<input type="checkbox"/>	mmonsoor	6	Apr 28, 2014	~2 days ago	mmonsoor@sb-roscoff.fr

For 0 selected histories: Copy Unshare

History ↻ ⚙️

- HISTORY LISTS
- in Saved Histories
- 6: **Histories Shared with Me**
- 24: CURRENT HISTORY
- xs: Create New
- 5: Copy History
- 2: 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 Q

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](#)

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

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	

Import history

About this History

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

Collection

DATASET

- 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%

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 'Using 0%' indicator is visible in the top right. On the left, a 'Tools' sidebar lists various categories such as 'Get Data', 'Send Data', 'Text Manipulation', and 'NGS: DeepTools'. The main workspace contains 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' (11.68 MB). Below it, several datasets are listed, with a red box highlighting the following four: '14: Top exon genetic location', '13: Top exons', '12: Top exon genetic location', and '10: Top 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

- 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 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. A 'Using 0%' indicator is visible in the top right. 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 green welcome message for 'galaxy3.sb-roscoff.fr', an information box with contact details, and the ABiMS logo (Analyses and Bioinformatics for Marine Science) with links to 'Changelog' and 'Tutorials'. Below the logo is a paragraph of text describing Galaxy as an open, web-based platform for data-intensive biomedical research, supported by various institutions. The right sidebar shows a 'History' section with a search bar and a list of datasets under the heading 'Galaxy initiation - collection'. The list includes '3: Repeats', '2: SNPs', and '1: Exons'. A red box highlights a checkmark icon next to the '3: Repeats' entry, with a tooltip that reads 'Operations on multiple datasets'. The bottom of the interface shows a JavaScript console with 'javascriptvoid(0);'.

Dataset collection

Create a dataset collection

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and a 'Using 0%' indicator. The left sidebar lists various tool categories such as 'Get Data', 'Send Data', 'Collection Operations', and 'COMMON TOOLS'. 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'. Below the logo are links for 'Changelog' and 'Tutorials', followed by a paragraph of text describing the Galaxy platform and its affiliations.

On the right side, the 'History' panel shows a search bar and a list of datasets under the heading 'Galaxy initiation - collection'. A context menu is open over the dataset list, with the following options:

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

The 'Build Dataset List' option is highlighted with a red border, indicating the current step in the process.

Dataset collection

Create a dataset collection

The screenshot shows the Galaxy ABiMS interface with a dialog box titled "Create a collection from a list of datasets". The dialog box contains the following elements:

- Header:** "Create a collection from a list of datasets"
- Description:** "Collections of datasets are permanent, ordered lists of datasets that can be passed to tools and workfl... [More help](#)"
- Start over:** A link to "Start over"
- Dataset Selection:** Two input fields with "Repeats" and "SNPs" selected. Each field has a "Discard" button to its right.
- Name:** A text input field containing "Collection of different features", which is highlighted with a red border.
- Buttons:** "Cancel" and "Create list" (highlighted with a blue border).

The background shows the Galaxy ABiMS navigation menu with categories like "Tools", "COMMON TOOL", "COMMON NGS TOOLS", and "SEARCHING TOOLS". A "Using 0%" indicator is visible in the top right corner.

Dataset collection

Create a dataset collection

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and a 'Using 0%' 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), and 'SEARCHING TOOLS' (Diamond).

The main content area features a green welcome banner for 'galaxy3.sb-roscoff.fr' and an information box with contact details for support.abims@sb-roscoff.fr. Below this is the ABiMS logo and the text 'Analyses and Bioinformatics for Marine Science'. A 'History' panel on the right shows a search bar and a list of dataset collections. The collection '4: Collection of different features' is highlighted with a red border and contains a list of datasets. Other collections listed are '3: Repeats', '2: SNPs', and '1: Exons'. The 'History' panel also shows a total size of 11.32 MB and selection options: 'All', 'None', and 'For all selected...'. The bottom of the interface has navigation arrows.

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%' indicator. The left sidebar lists various tool categories such as '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 to galaxy3.sb-roscoff.fr' message, an 'Information' box with contact details, the ABiMS logo (Analyses and Bioinformatics for Marine Science), and links to 'Changelog' and 'Tutorials'. A paragraph of text describes Galaxy as an open, web-based platform for data-intensive biomedical research, supported by various institutions. The right-hand panel shows a 'History' section with a red-bordered box highlighting a collection titled 'Collection of different features' (a list of datasets), which includes sub-sections for 'Repeats' and 'SNPs', each with view and edit icons.

Dataset collection

Tools for collection operations

The screenshot displays the Galaxy / ABiMS web interface. The top navigation bar includes 'Galaxy / ABiMS', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', 'User', and a 'Using 0%' indicator. The left sidebar contains a 'Tools' section with a search bar and several tool categories: 'Get Data', 'Send Data', 'Collection Operations' (highlighted with a red box), 'COMMON TOOLS', and 'COMMON NGS TOOLS'. The 'Collection Operations' category lists: 'Unzip Collection', 'Zip Collection', 'Filter failed datasets from a list', and 'Flatten Collection into a flat list of datasets'. The main content area features a green welcome banner for 'galaxy3.sb-roscoff.fr', an information box with contact details for support.abims@sb-roscoff.fr, the ABiMS logo (Analyses and Bioinformatics for Marine Science), and links for 'Changelog' and 'Tutorials'. A paragraph of text describes Galaxy as an open, web-based platform for data-intensive biomedical research, mentioning affiliations with Penn State and Emory University. The right sidebar shows a 'History' section with a search bar, a collection named 'Galaxy initiation - collection' (11.32 MB), and a list of features: '4: Collection of different features', '3: Repeats', '2: SNPs', and '1: Exons'.

Dataset collection

Use a collection as input

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
```

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 shows the Galaxy web interface. At the top, there is a navigation bar with 'Galaxy' logo and menu items: 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A 'Using 0%' indicator is in the top right. On the left, a 'tools' sidebar contains a search bar and a list of tool categories: 'Get Data', 'Send Data', 'Lift-Over', 'Text Manipulation', 'Datamash', and 'Convert Formats'. The main workspace displays a green notification box with a checkmark icon, stating: '2 jobs have been successfully added to the queue - resulting in the following datasets: 16: BED-to-GFF on data 4, 17: BED-to-GFF on data 2. 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.' On the right, the 'History' panel shows a search bar and a list of jobs. A red box highlights the job '18: BED-to-GFF on collection 5', which is described as 'a list of datasets'. Below it are jobs '17: BED-to-GFF on data 2' and '16: BED-to-GFF on data 4'. Other jobs in the history include '5: Collection of different features', '4: Repeats', '2: SNPs', and '1: Exons'.

Hands-on
COLLECTION





1. Create a dataset list with SNPs, repeats and exons
2. Run tool “Convert Formats -> BED-to-GFF converter” on the dataset list

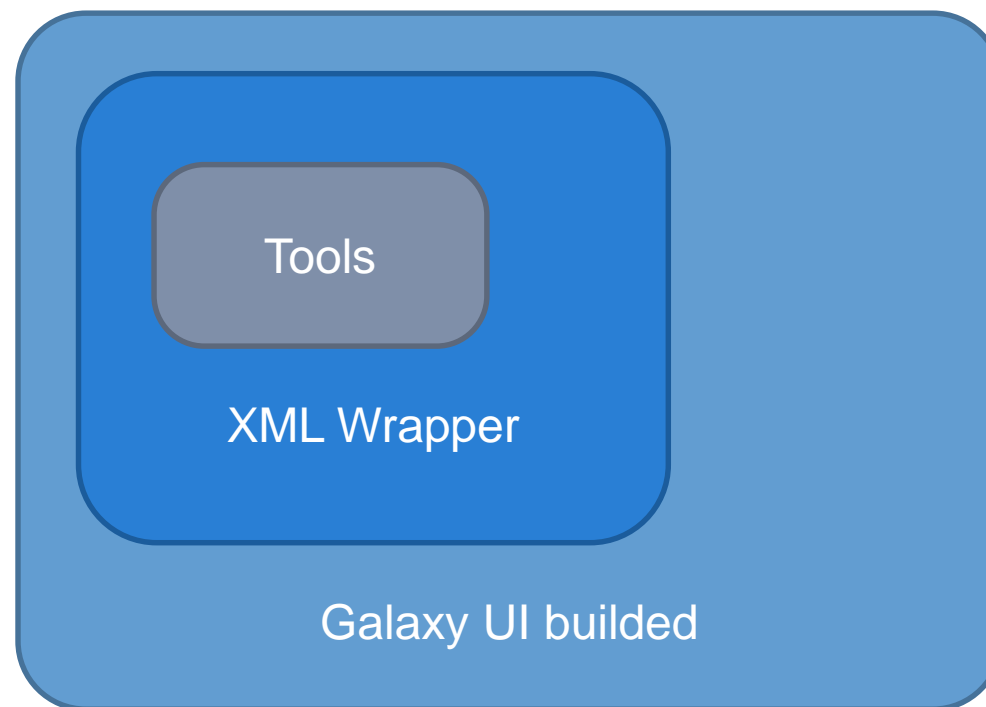
END

BONUS

How are tools born?

BONUS

- How to import a tool in Galaxy?

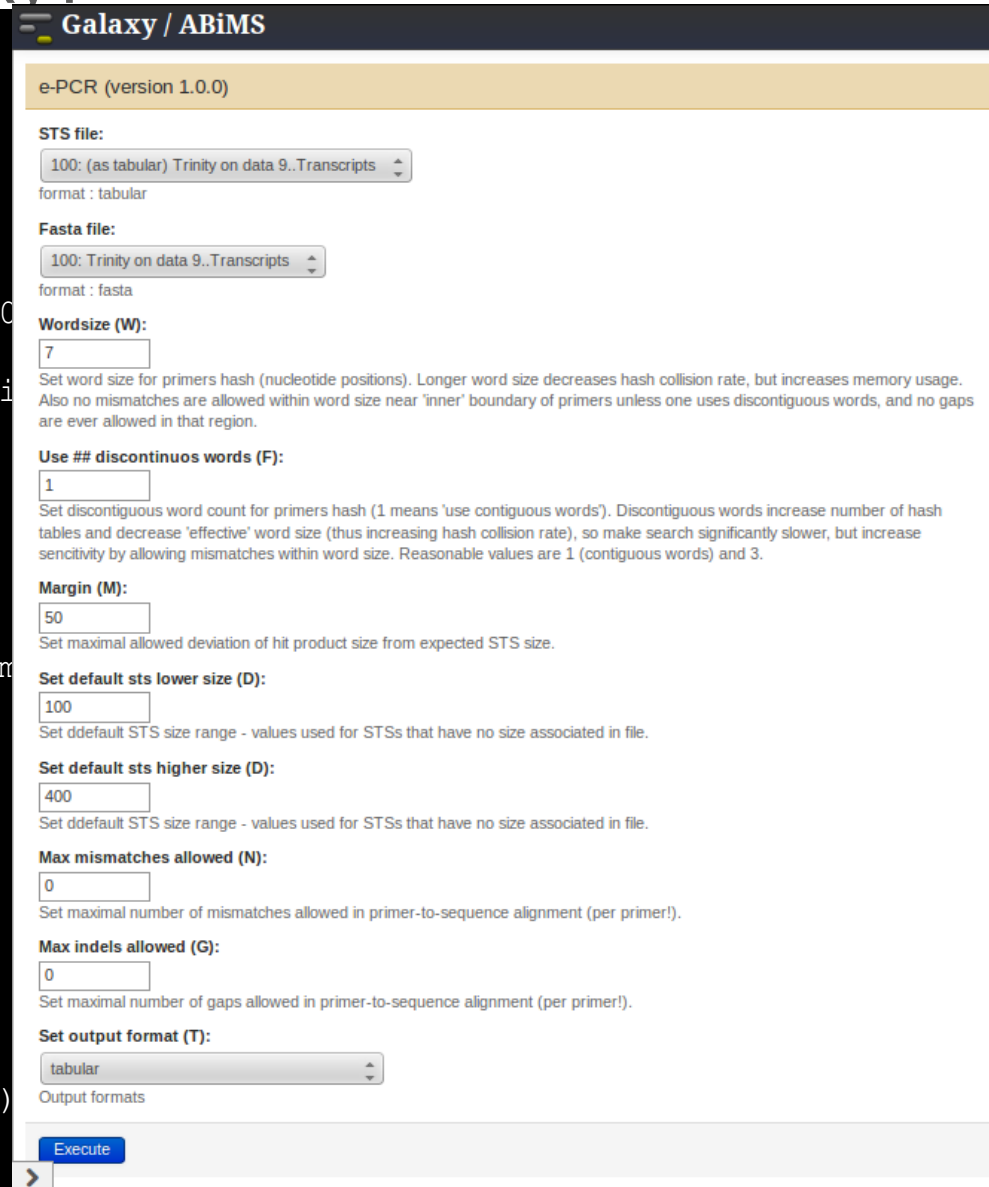


• How to import a tool in Galaxy?

```

[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 ## discontinuos words, slow i
                ##>1
  -o ##      Set output file
  -t ##      Set output format:
                1 - classic, range (pos1..pos2)
                2 - classic, midpoint
                3 - tabular
                4 - tabular with alignment in com
                    (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?

```

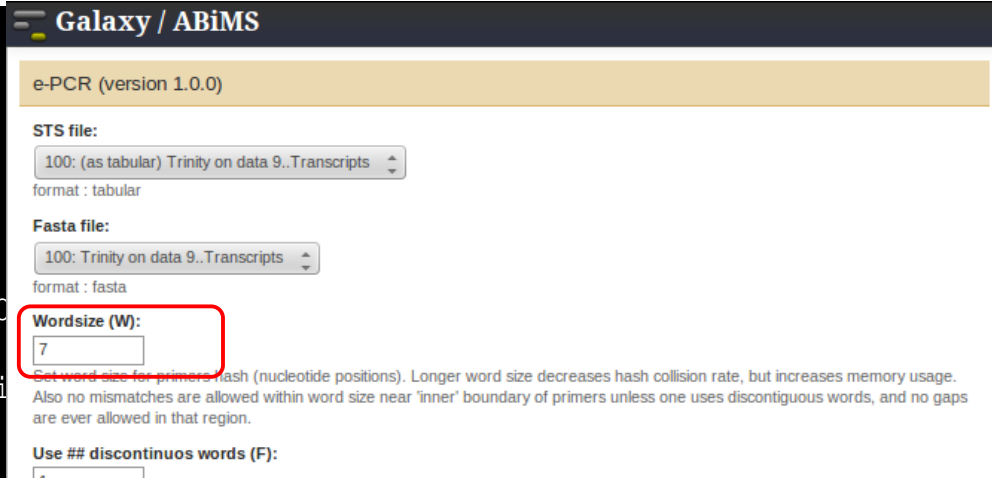
<tool id="abims_epcr" name="e-PCR">
>   <!-- author : lecorguille@sb-roscoff.fr -->
>   <!-- date : 11-05-12 -->
>   <description>e-PCR parses stsfile in unists format, then reads nucleotide sequence data in FASTA format from files listed in commandline if
any, or from stdin otherwise. For input sequences 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 i
                ##>1
    -o ##      Set output file
    
```



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):

```

<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
  
```