



Biogenouest
BIOGÉOINFORMATIQUE



OCEANOMICS



AB⁴BIMS

30/05/2016

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.

Initiation

Misharl Monsoor

Loraine Guéguen

Credit to Gildas Le Corguillé - V1.07

UPMC
SORBONNE UNIVERSITÉS



INTRODUCTION / PROBLEMATIC

- Setup TP
 - <http://galaxy.sb-roscoff.fr>
- Account
 - login@sb-roscoff.fr
 - *****

```
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-ro
#$ -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!

l.bam

prep_reads.info tmp

Introduction



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



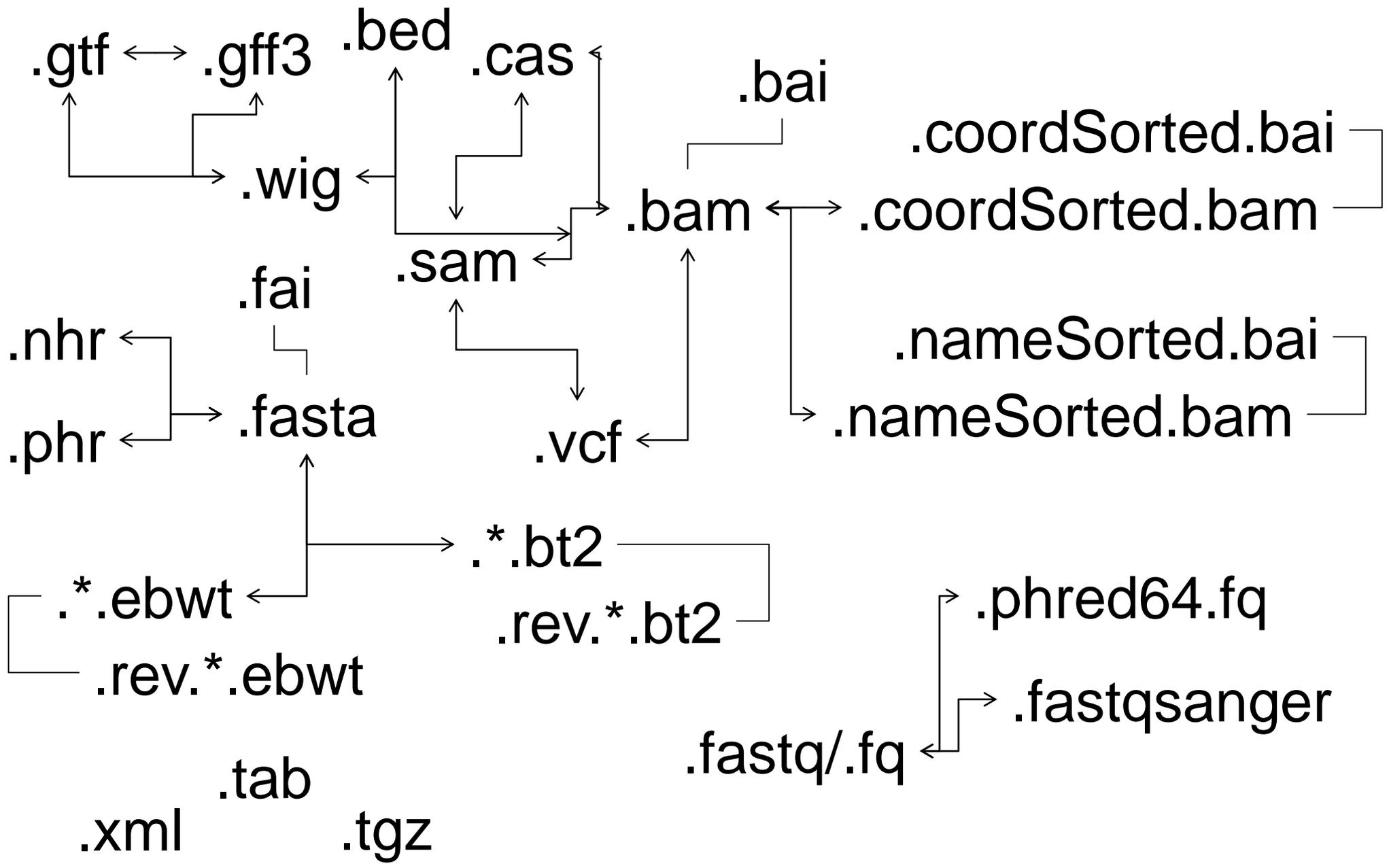
« I have a bunch of cool tools!
But I'm the only one who can launch them.

Comments? »

Introduction

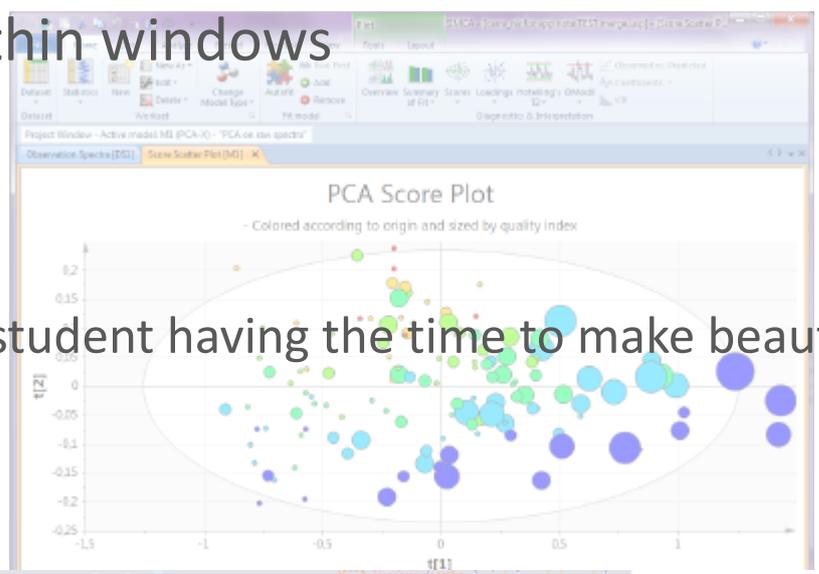
.gtf .gff3 .bed .cas .bai
.wig .bam .coordSorted.bai
.fai .sam .coordSorted.bam
.nhr .nameSorted.bai
.phr .fasta .vcf .nameSorted.bam
.*.ebwt *.bt2 .phred64.fq
.rev.*.ebwt .rev.*.bt2 .fastqsanger
.tab .fastq/.fq
.xml .tgz

ABIMS Introduction



Introduction

- Graphical interface click-button tools within windows
 - + very ergonomic
 - too ergonomic → lack of flexibility
 - don't count on it ! Have you ever seen a PhD student having the time to make beautiful green buttons?
 - paying for it!



- Tools available on the internet
 - + very ergonomic
 - too ergonomic → lack of flexibility
 - A small part of the available tools
 - distributed on different universities locations
 - the submission size is often limited
 - must not be paranoid

Parameter	Value
maximum peak width	20
signal-to-noise threshold	3.0
minimum peak width	1
peak-to-peak	5
peak-to-peak	1000

```
library(xcms)
loaddata()
polar<- "Po"
```

• Command line tools

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

+ represent almost the majority of scientific tools

```
xset2<-retco(xset,method="PLW")
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)
```

+ good parameters completeness

```
# rapport final avec statistiques de différences entre les deux classes
```

```
reporttab <- diffreport(xset5, filebase =paste(pathResult,"/Rapport_",expe,"_",polar, sep=""), mzdec=4, eicmax=5000, metlin = metlin, classeic=levels(xset5@phenoData))
```

+ can be executed on high performance computers

```
#écriture du fichier Excel
```

```
dir.create(paste(pathResult,"/Rapport_",expe,"_",polar,"_diffreport/", sep=""), showWarnings = FALSE)
```

```
write.table(reporttab,paste(pathResult,"/Rapport_",expe,"_",polar,"_diffreport/resultat",expe,"_",polar, ".xls", sep=""),sep="\t")
```

+ g33ks love it, since automatable, workflowsable, ...

```
library(CAMERA)
```

```
#annotation version rapide?
```

```
an<-annotate(xset,method="CAMERA", 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)
```

- minimum linux knowledge is required

```
diffreport1<-getPeaklist(an)
```

- crucial lack of ergonomics

```
#diffreport <- annotateDiffreport(xsg,pval_th=0.05,fc=0.1, nSlaves=8, calcIso=TRUE, calcCaS=FALSE, maxcharge=3, maxiso=4, minfrac=0.5,
```

```
# ppm=15, mzabs=0.015, quick=FALSE, psg_list=NULL, rules=NULL,
```

```
# polarity=polarity, sortpval=FALSE)
```

```
diffreport<-cbind(reporttab,diffreport1[,c("isotopes", "adduct", "pcgroup")])
```

```
write.table(diffreport, file=paste(pathResult,"/result_",expe,"_",polar,"_CAMERA_diffreport-fast.xls",sep=""), row.names=FALSE, sep="\t")
```

```
library(FactoMineR)
```

```
pca3<-PCA(t(matacp), axes=c(1,2))
```

```
pca3<-PCA(t(matacp), axes=c(1,3))
```

```
pca3<-PCA(t(matacp), axes=c(2,3))
```

```
pca4<-PCA(t(matacplog2))
```

```
# -- output png --
```

```
# Percentage of variance
```

```
png("percentage_of_variance.png", width =800, height = 400);
```

```
barplot(resPCA$eig$per,xlab="Components",ylab="percentage of variance");
```

```
dev.off()
```

```
png("eigenvalue.png", width =800, height = 400);
```

```
barplot(resPCA$eig$eig,xlab="Components",ylab="eigenvalue");
```

```
dev.off()
```

```
library(ctc)
```

```
# -- Normalization: logratio --
```

```
if (normalization) {
```

```
data=t(scale(t(data)))
```

INTRODUCTION / GALAXY



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:**

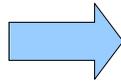
- Galaxy it's ...
 - No need to execute a command line through a terminal
 - Programming or scripting skills are not required
 - Submission of jobs is transparent through a high performance computer cluster
 - Secure histories and data manager
 - A data and protocols sharing system
 - Tool-boxes of several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



Why Galaxy ?

- Accessibility
- Reproductibility
- Transparency

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
```




```
[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

Execute

`xcmsSet.matchedFilter(object, fwhm = 30, sigma = f`

Arguments

object	xcmsRaw object
fwhm	full width at half maximum of matched filtration gaussian model peak used to calculate the actual sigma, see below.
sigma	standard deviation (width) of matched filtration model peak
max	maximum number of peaks per extracted ion chromatogram
snthresh	signal to noise ratio cutoff
step	step size to use for profile generation
steps	number of steps to merge prior to filtration
mzdiff	minimum difference in m/z for peaks with overlapping retention times
index	return indices instead of values for m/z and retention times
sleep	number of seconds to pause between plotting peak finding cycles
scanrange	scan range to process

Galaxy / 4 / Metabolomics Analyze Data Workflow

xcms.xcmsSet Filtration and Peak Identification using xcmsSet function from xcms R package to p for relative quantification and statistical analysis (Galaxy Tool Version 2.0.2)

Choose your inputs method

Zip file from your history containing your chromatograms

Zip file

No no_unzip.zip dataset available.

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 f

Full width at half maximum of matched filtration gaussian model peak

30
[fwhm] Only used to calculate the actual sigma

Advanced options

show

Maximum number of peaks per extracted ion chromatogram

5
[max]

Signal to noise ratio cutoff

10
[snthresh]

Number of steps to merge prior to filtration

2
[steps] The peak identification algorithm combines a given number of EIBPCs prior to filtration and steps argument

Galaxy interface

Galaxy / 4 / Metabolomics Analyze Data Workflow Shared Data Visualization Admin Help User Using -993344424 b

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

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 options like 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. The main content area is divided into three sections:

- Tools Panel (Left):** A sidebar with 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 (Center):** The 'Batch_correction (version 2.0.0)' tool is selected. It features several input fields and dropdown menus:
 - 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 the configuration area. Below the configuration, there is a section for 'Authors' (Jean-Francois Martin) and 'Contributors' (Melanie Petera, Etienne Thevenot).
- History Panel (Right):** A sidebar showing a list of datasets. The top entry is 'Sacuri Zip' (289.7 MB). Below it, a list of datasets is shown, each with a number (19 down to 11) and a file name, such as 'xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv (Xdiffreport)'. Each entry has icons for viewing, editing, and deleting.

Galaxy interface

Web forms / 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 : 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:

GET HELP

Galaxy / ABiMS
Using 12%

Analyze Data
Workflow
Shared Data
Visualization
Admin
Help
User

Tools

Get Data

COMMON TOOLS

Convert Formats

FASTA manipulation

Filter and Sort

Join, Subtract and Group

Text Manipulation

Graphics

Statistics

EMBOSS 5 Suite

SEARCHING TOOLS

NCBI BLAST+

Diamond

Primer/Microsatellite

NGS TOOLS

NGS: BedTools

NGS: Mapping

NGS: Picard

NGS: QC and manipulation

NGS: SAM Tools

GATK Tools

GATK2 Tools

DNA-SEQ TOOLS

DNA-Seq tools

Welcome to galaxy.sb-roscoff.fr

Warning
 17-05-16: For performance and maintenance reasons, we will disable the tool Export2Dir. You can always export your whole history in a file or download one file using the old fashion "floppy disk" icon. Export2Dir could return on day after a little refactoring?

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

sartools deseq2
7 shown

353.2 KB

7: SARTools DESeq2 R objects (.RData)

6: SARTools DESeq2 R log

5: SARTools DESeq2 figures

4: SARTools DESeq2 tables

3: SARTools DESeq2 report

2: targetT048.txt

1: t048.zip

Tools

search tools

Get Data

COMMON TOOLS

Convert Formats

FASTA manipulation

Filter and Sort

Join, Subtract and Group

Text Manipulation

Graphics

Statistics

EMBOSS 5 Suite

SEARCHING TOOLS

NCBI BLAST+

Diamond

Primer/Microsatellite

NGS TOOLS

NGS: BedTools

NGS: Mapping

NGS: Picard

NGS: QC and manipulation

NGS: SAM Tools

GATK Tools

GATK2 Tools

DNA-SEQ TOOLS

DNA-Seq tools

Welcome to galaxy.sb-roscoff.fr

Warning
17-05-16: For performance and maintenance reasons, we will disable the tool Export2Dir. You can always export your whole history in a file or download one file using the old fashion "floppy disk" icon. Export2Dir could return on day after a little refactoring?

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



- ▶ 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

sartools deseq2
7 shown
353.2 KB

7: SARTools DESeq2 R objects (.RData)

6: SARTools DESeq2 R log

5: SARTools DESeq2 figures

4: SARTools DESeq2 tables

3: SARTools DESeq2 report

2: targetT048.txt

1: t048.zip

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)

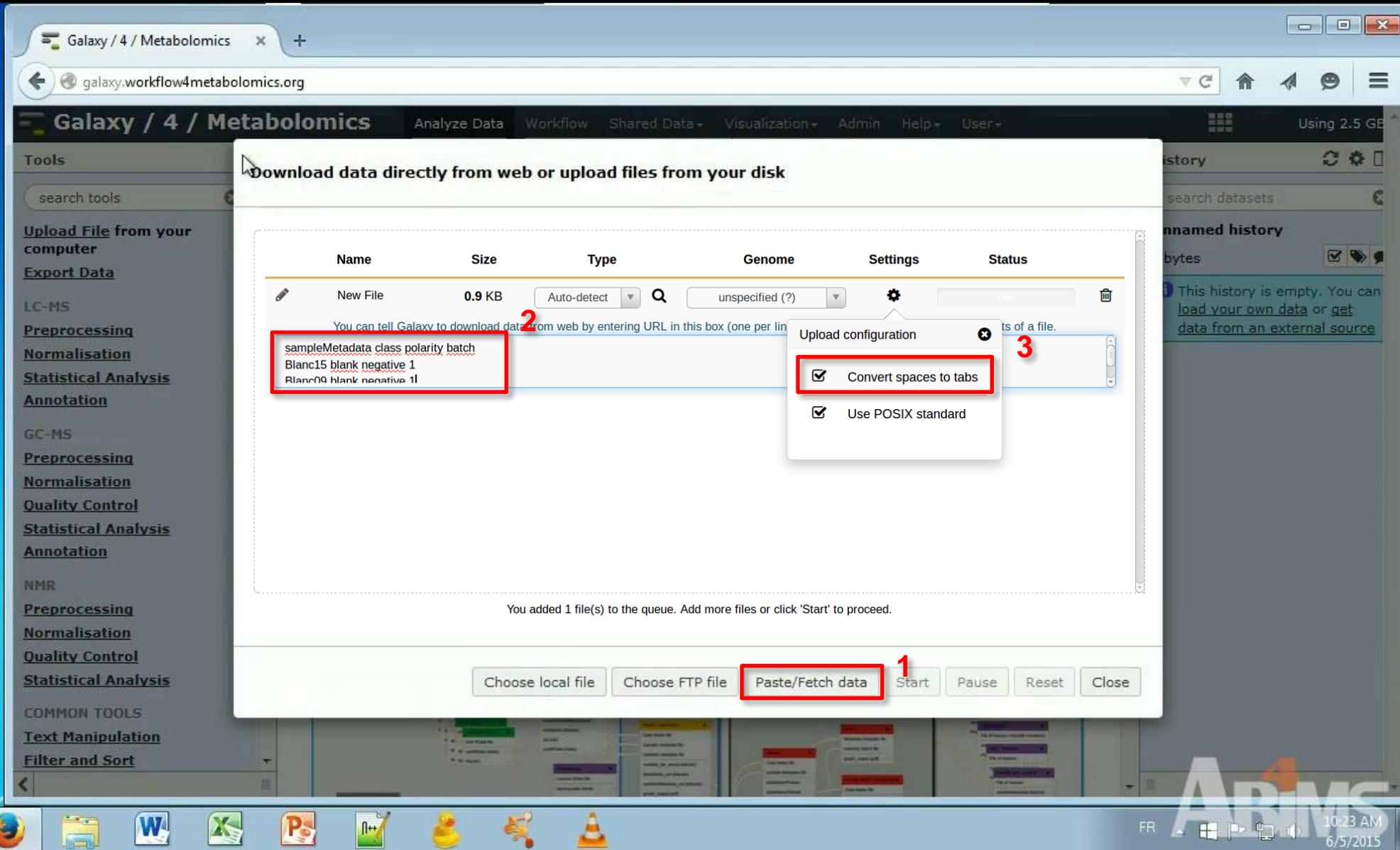
10:23 AM 6/5/2015

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 on the left containing 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 1l					

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-line box for file upload, with the text "You can Drag & Drop files into this box." below it. At the bottom of the dialog, there are 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 top navigation bar of the Galaxy application are circled in red. The Galaxy application header includes the text "Galaxy / 4 / Metabolomics" and a navigation menu with items: "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User".

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 upload. A yellow folder icon is being dragged into this box, with a "Move" tooltip visible. Below the dashed box, the text reads "You can Drag & Drop files into this box." 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 of the Galaxy interface shows a sidebar with tool categories: "Tools", "Upload File from your computer", "Export Data", "LC-MS", "Preprocessing", "Normalisation", "Statistical Analysis", "Annotation", "GC-MS", "Preprocessing", "Normalisation", "Quality Control", "Statistical Analysis", "Annotation", "NMR", "Preprocessing", "Normalisation", "Quality Control", "Statistical Analysis", and "COMMON TOOLS". The "COMMON TOOLS" section includes "Text Manipulation" and "Filter and Sort".

The Windows taskbar at the bottom shows various application icons, including Internet Explorer, File Explorer, Word, Excel, PowerPoint, and a yellow duck. The system tray on the right shows the date and time: "10:23 AM 6/5/2015".

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

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

ARMS

FR 10:23 AM 6/5/2015

Data import < 2 Go

From local files

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

Name	Size	Type	Genome	Settings	Status
sacuri.zip	0.2 GB	Auto-det...	unspecified (?)	⚙️	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 interface shows a sidebar with tool categories: 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, and COMMON TOOLS (Text Manipulation, Filter and Sort). The top navigation bar includes "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". The system tray at the bottom shows the date and time: 10:23 AM 6/5/2015.

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

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 workflow4metabolomics.org v2.0 web interface. The browser address bar shows the URL galaxy.workflow4metabolomics.org. The main navigation bar includes 'Galaxy / 4 / Metabolomics' and various menu items like 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. The interface is divided into several sections:

- Tools Panel (Left):** Contains a search bar and a list of tool categories: LC-MS, NMR, and COMMON TOOLS. Under LC-MS, there are sub-categories for Preprocessing, Normalisation, Statistical Analysis, and Annotation. Under NMR, there are sub-categories for Preprocessing, Normalisation, Quality Control, and Statistical Analysis. Under COMMON TOOLS, there are sub-categories for Text Manipulation and Filter and Sort.
- Welcome Banner (Top Center):** A green banner with a checkmark icon, titled 'Welcome to workflow4metabolomics.org v2.0'. It includes a publication reference: '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'. It also provides a help and support email: support@workflow4metabolomics.org.
- Latest News (Middle):** A light blue section with an information icon, titled 'Latest news'. It lists three news items:
 - 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
- History Panel (Right):** A panel titled 'History' with a search bar for datasets. It shows an 'Unnamed history' with 0 bytes. Below it, a job named '1: sacuri.zip' is highlighted in yellow, with a status of 'This job is currently running'. The job has icons for refresh, edit, and delete.
- Footer (Bottom):** A navigation bar with icons for 'LC/MS', 'MS', and 'Common', each with a corresponding workflow diagram. The system tray at the bottom right shows the date and time: 10:23 AM 6/5/2015.

Step 1: Choose a FTP Client

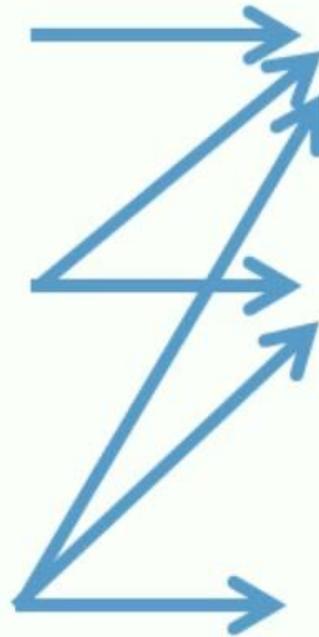
DATA IMPORT

> 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



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 main navigation bar includes "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". A "Using 46%" indicator is visible in the top right corner.

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" banner. Below it is an "Information" box with an "i" icon and 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 (with a large "4" above the "i") and the text "Analyses and Bioinformatics for Marine Science". To the right of the logo is the logo for "CNRS UPRC Station Biologique Roscoff".

Below the logo are two expandable sections: "Changelog" and "Tutorials".

At the bottom of the main content area, there 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 a "History" section with a search bar for "search datasets" and a message: "Unnamed history 0 bytes. 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, File Explorer, Word, Excel, PowerPoint, and a duck icon. The system tray in the bottom right corner shows 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. The Windows taskbar at the bottom shows various application icons and the system clock displaying "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 is visible on the left, listing 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-line box for file upload. Below the box, there are two dropdown menus: "Type (set all):" with "Auto-detect" selected, and "Genome (set all):" with "unspecified (?)" selected. At the bottom of the dialog, there are several buttons: "Choose local file", "Choose FTP file" (which is being clicked by the mouse), "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". The taskbar also contains icons for various applications like Internet Explorer, Chrome, Word, Excel, PowerPoint, and a duck icon.

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

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

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

ory

ry 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

lecorguille

left_kept_r...

ABiMS

11:19 AM
7/31/2015

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 web page has a dark header with navigation links: Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User. Below the header, there's a sidebar with 'Tools' and 'Get Data' sections. The main content area features a large white box with the heading 'Download data directly from web or upload files from your disk' and a dashed border indicating a drag-and-drop zone. A small tooltip is visible over this area, explaining that files can be dragged and dropped into the box. In the foreground, a Cyberduck file manager 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 11:19 AM on 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 main content area of the browser has a title "Download data directly from web or upload files from your disk" and a large dashed box for file uploads. A message in the background says "You can Drag & Drop files into this box." and another message below it says "Your FTP directory does not contain any files." At the bottom of the browser window, there 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 says "Cyberduck" and "Unregistered". The menu bar includes "File", "Edit", "View", "Go", "Bookmark", "Window", and "Help". The main area of Cyberduck shows an "Open connection" dialog with the following fields and options:

- Protocol: FTP (File Transfer Protocol)
- Server: [Empty text box]
- Port: 21
- URL: [Empty text box]
- Username: [Empty text box]
- Password: [Empty text box]
- Anonymous Login
- Save Password
- Buttons: "Connect" and "Cancel"
- More Options: [Expanded arrow]

The Windows taskbar at the bottom shows the system tray with the date and time: "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 main content area of the browser has a header "Galaxy / ABiMS" and a navigation menu with items like "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". Below the header, there is a "Tools" section with a search bar and a "Get Data" section. A large white dialog box is overlaid on the browser, 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 this, there is a smaller dialog box with the text "allows you to upload files via FTP. To upload some files, log at **galaxy.sb-roscoff.fr** using your Galaxy credentials (password)." and a yellow warning box that says "Your FTP directory does not contain any files." At the bottom of the browser window, 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" dialog. The dialog is titled "Open Connection" and has a menu bar with "File", "Edit", "View", "Go", "Bookmark", "Window", and "Help". The "FTP (File Transfer Protocol)" option is selected in the dropdown menu. The "Server" field contains `galaxy.sb-roscoff.fr` and the "Port" field contains `21`. The "URL" field contains `ftp://lecorguille@galaxy.sb-roscoff.fr:21/`. The "Username" field contains `lecorguille` and the "Password" field is filled with dots. There are checkboxes for "Anonymous Login" and "Save Password", both of which are unchecked. At the bottom of the dialog, there are "Connect" and "Cancel" buttons. A mouse cursor is pointing at the "Connect" button.

The Windows taskbar at the bottom of the screen shows several icons, including the Start button, Internet Explorer, Google Chrome, File Explorer, Microsoft Word, Microsoft Excel, Microsoft PowerPoint, a green application icon, a yellow duck icon, and a traffic cone icon. The system tray on the right shows the time as 11:19 AM on 7/31/2015 and the language as FR.

Data import > 2 Go

The screenshot displays the Galaxy/ABiMS web interface in a browser window. The main heading is "Download data directly from web or upload files from your disk". Below this, there is a large dashed box with the text "You can Drag & Drop files into this box." To the left, a sidebar contains navigation options like "Tools", "Get Data", and "COMMON TOOLS".

Overlaid on the interface are two security warning dialogs. The first, titled "Unsecured FTP connection", contains the text: "Unsecured FTP connection. Password will be sent in plaintext. Please contact your web hosting service provider for assistance." It features "Continue" and "Disconnect" buttons, along with a "Don't show again" checkbox and a "Help" link. The second dialog, partially obscured, contains the text: "allows you to upload files via FTP. To upload some files, log in at galaxy.sb-roscoff.fr using your Galaxy credentials (password)." Below this, it states "Your FTP directory does not contain any files." and includes buttons for "local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close".

The desktop environment is visible at the bottom, showing the Windows taskbar with various application icons and the system tray displaying the date and time as 11:19 AM on 7/31/2015.

Data import > 2 Go

The screenshot shows a Windows desktop environment with a web browser window open to the Galaxy/ABiMS website. The browser address bar shows `galaxy.sb-roscoff.fr`. The website header includes navigation links: `Analyze Data`, `Workflow`, `Shared Data`, `Visualization`, `Admin`, `Help`, and `User`. A central panel displays the text: **Download data directly from web or upload files from your disk**. Below this, it says "You can Drag & Drop files into this box." and shows a large dashed-line box for file uploads. A tooltip explains that files can be uploaded via FTP, providing instructions to log in at `galaxy.sb-roscoff.fr`. A yellow message box states: "Your FTP directory does not contain any files." At the bottom of the page, 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, showing a menu bar, a toolbar with `Open Connection`, `Quick Connect`, `Action`, `Get Info`, and `Refresh`, and a file list table with columns `Filename`, `Size`, and `Modified`. The file list is currently empty, showing a "0 Files" status. A system tray notification in the bottom right corner reads: **Connection opened** galaxy.sb-roscoff.fr. The system clock shows 11:19 AM on 7/31/2015.

Data import > 2 Go

The screenshot displays a Windows desktop environment with a web browser window open to the Galaxy web interface at `galaxy.sb-roscoff.fr`. The browser window shows a file upload progress window for `left_kept_reads.bam`. The progress bar indicates that 50.6 MiB (53,018,624 bytes) of 91.6 MiB (55%) has been uploaded, with a speed of 70.1 MB/sec and 1 second remaining. The local file path is `C:\Users\lecorguille\Desktop\left_kept_reads.bam` and the URL is `ftp://galaxy.sb-roscoff.fr/left_kept_reads.bam`. A notification bubble in the bottom right corner states "Connection opened galaxy.sb-roscoff.fr".

Galaxy / ABiMS

galaxy.sb-roscoff.fr

Transfers

Resume Reload Stop Remove Open Show

left_kept_reads.bam

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

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

Connection opened galaxy.sb-roscoff.fr

11:19 AM 7/31/2015

Data import > 2 Go

The screenshot shows the Galaxy/ABiMS web interface in a browser window. The browser address bar shows 'galaxy.sb-roscoff.fr'. The page title is 'Galaxy / ABiMS'. The main content area has a heading 'Download data directly from web or upload files from your disk' and a sub-heading 'You can Drag & Drop files into this box.' Below this is a large dashed box for file upload. A yellow message box states 'Your FTP directory does not contain any files.' At the bottom of the page are buttons: 'local file', 'Choose FTP file', 'Paste/Fetch data', 'Start', 'Pause', 'Reset', and 'Cancel'. An FTP client window is open in the foreground, titled 'lecorguille@galaxy.sb-roscoff.fr - FTP'. It shows a file list with one file: 'left_kept_reads.bam' (91.6 MiB, 7/31/2015 9:19:00 AM). A notification bubble in the bottom right corner says 'Upload complete left_kept_reads.bam'. The Windows taskbar at the bottom shows various application icons and the system tray with the date '7/31/2015' and time '11:19 AM'.

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

Get Data

COMMON TOOLS

Convert Formats

Download data directly from web or upload files from your disk

You can Drag & Drop files into this box.

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.

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

lecorguille@galaxy.sb-roscoff.fr - FTP

Unregistered

File Edit View Go Bookmark Window Help

Quick Connect Action Get Info Refresh

Open Connection

Filename Size Modified

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

1 Files

Upload complete left_kept_reads.bam

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

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

Using 46%

Assets

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

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 is visible on the left, listing 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 instructions on how to upload files via FTP. A table lists available files, including "left_kept_reads.bam" which is 96 MB and was created on 07/31/2015 at 11:19:45 AM. The dialog also includes buttons for "Choose local file", "Choose FTP file", "Paste/Fetch data", "Start", "Pause", "Reset", and "Close".

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

search tools

Get Data

COMMON TOOLS

Convert Formats

FASTA manipulation

Filter and Sort

Join, Subtract and

Text Manipulation

Graphics

Statistics

EMBOSS 5 Suite

SEARCHING TOOLS

NCBI BLAST+

Diamond

Primer/Microsatelli

NGS TOOLS

NGS: BedTools

NGS: Mapping

NGS: Picard

NGS: QC and manip

Download data directly from web or upload files from your disk

You can Drag & Drop files into this box.

FTP files

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

Available files: 1 files 96 MB

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

Type (set a

specified (?)

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

lecorguille

left_kept_r...

Using 46%

Assets

History

is empty. You can

own data or get

an external source

ABiMS

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

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 (?)		

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) 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

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 11:19 AM 7/31/2015

Data import > 2 Go

Galaxy / ABiMS

galaxy.sb-roscoff.fr

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 46%

Tools

search tools

Get Data

COMMON TOOLS

Convert Formats

FASTA manipulation

Filter and Sort

Join, Subtract and

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 (?)	⚙	🗑

Type (set all): Auto-detect

Genome (set all): unspecified (?)

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

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

actory

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. The main content area features the ABiMS logo and the text "Analyses and Bioinformatics for Marine Science", along with links to "Changelog" and "Tutorials". A paragraph of text describes the Galaxy platform and its affiliations. On the left, a sidebar lists various tool categories such as "Get Data", "COMMON TOOLS", "FASTA manipulation", and "NGS TOOLS". On the right, a "History" panel shows a dataset named "left_kept_reads.bam". The Windows taskbar at the bottom includes icons for Internet Explorer, Chrome, File Explorer, Word, Excel, PowerPoint, and other applications. 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 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 it, 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 links for **Changelog** and **Tutorials**. A paragraph of text describes the platform: "Galaxy is an open, web-based platform for data intensive biomedical research. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Emory University."

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

The right sidebar shows the **History** panel with a search box, "Unnamed history" (1 shown, 0 bytes), and a dataset entry: **left_kept_reads.bam**.

The Windows taskbar at the bottom shows various application icons and the system tray with the date **11:20 AM 7/31/2015**.

Data import > 2 Go

The screenshot shows the Galaxy / ABiMS web interface. The browser address bar displays 'galaxy.sb-roscoff.fr'. The main navigation bar includes '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. The central area features the 'ABiMS 4' logo and the text 'Analyses and Bioinformatics for Marine Science'. A sidebar on the left lists various tools under categories like 'Get Data', 'COMMON TOOLS', 'SEARCHING TOOLS', and 'NGS TOOLS'. On the right, the 'History' panel shows a dataset named 'left_kept_reads.bam' with a size of 91.6 MB, format 'bam', and database '?'. The taskbar at the bottom shows various application icons and the system clock indicating 11:20 AM on 7/31/2015.

Exercise

DATA IMPORT





- Exercise
 - Fetch this file

<http://tinyurl.com/GI-input0>

1. First, as you want
2. Then, consider that it is **>2 Go**

~~DATA IMPORT~~

Data import

For HUGE public resources: genome, databank ...

--> Make a request to the support team

Galaxy / ABiMS Analyze Data Workflow Shared Data Visualization Help User Using 42%

Tools

search tools

Get Data

- Upload File from your computer

ABiMS WORKFLOWS

- Workflow RNA-seq de novo by ABiMS
- Workflow RNA-seq with reference by ABiMS
- Workflow 4 Metabolomics

ABiMS TOOLS

- Primer
- RNASeq
- InterEsil
- Statistics
- Utils
- Phylogenetics
- Debug

COMMON TOOLS

- Text Manipulation
- FASTA manipulation
- Join, Subtract and Group
- Filter and Sort

NCBI BLAST+

- NCBI BLAST+ `blastn` Search nucleotide database with nucleotide query sequence(s)
- NCBI BLAST+ `blastp` Search protein database with protein query sequence(s)
- NCBI BLAST+ `blastx` Search protein database with translated nucleotide query sequence(s)
- NCBI BLAST+ `tblastn` Search translated nucleotide database

NCBI BLAST+ `blastn` (version 0.0.17)

Nucleotide query sequence(s):

Subject database/sequences: BLAST Database

Nucleotide BLAST database:

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

Execute

Note. Database searches may take a substantial amount of time. For large input datasets it is advisable to allow overnight processing.

What it does

Search a *nucleotide database* using a *nucleotide query*, using the NCBI BLAST+ `blastn` command line tool. Algorithms include `blastn`, `megablast`, and `discontiguous megablast`.

Note. You can also search against a FASTA file of subject nucleotide sequences. This is *not* advised because it is slower (only one CPU is used), but more importantly gives e-values for pairwise searches (very small e-values which will look overly significant). In most cases you should instead turn the other FASTA file into a database first using `makeblastdb` and search against that.

History

Unnamed history

0 bytes

Your history is empty. Click 'Get Data' on the left pane to start

Output format

TOOLS

Tools - panel

Galaxy / ABiMS Analyze Data Workflow Shared Data Visualization Help User Using 41%

Tools

search tools

Get Data

- Upload File from your computer

ABiMS WORKFLOWS

- Workflow RNA-seq de novo by ABiMS
- Workflow RNA-seq with reference by ABiMS
- Workflow 4 Metabolomics

ABiMS TOOLS

- Primer
- RNASeq
- InterEsil
- Statistics
- Utils
- Phylogenetics
- Debug

COMMON TOOLS

- Text Manipulation
- FASTA manipulation
- Join, Subtract and Group
- Filter and Sort
- NCBI BLAST+**

- NCBI BLAST+ **blastn** Search nucleotide database with nucleotide query sequence(s)
- NCBI BLAST+ **blastp** Search protein database with protein query sequence(s)
- NCBI BLAST+ **blastx** Search protein database with translated nucleotide query sequence(s)
- NCBI BLAST+ **tblastn** Search translated nucleotide database

NCBI BLAST+ blastx (version 0.0.17)

Nucleotide query sequence(s):

1: human_protein.fas

Subject database/sequences:

FASTA file from your history (see warning note below)

Protein FASTA file to use as database:

1: human_protein.fas

Query genetic code:

1. Standard

Set expectation value cutoff:

0.001

Output format:

Tabular (extended 24 columns)

Advanced Options:

Hide Advanced Options

Execute

Note. Database searches may take a substantial amount of time. For large input datasets it is advisable to allow overnight processing.

What it does

Search a *protein database* using a *translated nucleotide query*, using the NCBI BLAST+ blastx command line tool.

Note You can also search against a FASTA file of subject protein sequences. This is *not* advised because it is slower (only one CPU is used), but more importantly gives e-values for pairwise searches (very small e-values which will look overly significant). In most cases you should instead turn the other FASTA file into a database first using *makeblastdb* and search against that.

Output format

Because Galaxy focuses on processing tabular data, the default output of this tool is tabular. The standard BLAST+ tabular output contains 12 columns:

Column	NCBI name	Description
--------	-----------	-------------

History

- Human protein study
5.3 MB
- 2: chr22_check.gff3
- 1: human_protein.fas

76 /

Tools - panel

Galaxy / ABiMS Analyze Data Workflow Shared Data Visualization Help User Using 41%

Tools

blast

NCBI BLAST+

- NCBI BLAST+ [blastn](#) Search nucleotide database with nucleotide query sequence(s)
- NCBI BLAST+ [blastp](#) Search protein database with protein query sequence(s)
- NCBI BLAST+ [blastx](#) Search protein database with translated nucleotide query sequence(s)
- NCBI BLAST+ [tblastn](#) Search translated nucleotide database with protein query sequence(s)
- NCBI BLAST+ [tblastx](#) Search translated nucleotide database with translated nucleotide query sequence(s)
- BLAST XML to tabular Convert BLAST XML output to tabular

Workflows

- All workflows

Online

- 07-06-13: Metabolomic : Workflow 4 Metabolomics, updated to version 2.1.0 (2013_06_07) ⓘ
- 30-04-13: RNASeq : DESeq is now available for RNASeq expression data with reference (with gtf input).
- 26-04-13: RNASeq : DESeq is now available for denovo RNASeq expression data (without gtf input).
- 26-04-13: RNASeq : sam2counts is now available to count the reads coverage by transcrit. It's also a requirement for DESeq denovo.
- 26-04-13: Metabolomic : Workflow Metabolomic by ABiMS, updated to version 2.0.0 (2013_04_18) ⓘ

Abi4
AbiMS

Analyses and Bioinformatics for Marine Science

 CNRS UPMC
Station Biologique Roscoff

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

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

Human protein study
5.3 MB

2: chr22_check.gff3

1: human_protein.fas

77 /

Tools - form

Galaxy / 4 / Metabolomics Analyze Data Workflow Shared Data Visualization Admin Help User Using -993344424 h

Tools search tools

Upload File from your computer
Export Data

LC-MS
Format Conversion
Preprocessing

[xcms.xcmsSet](#) Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis

[xcms.group](#) Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.

[xcms.retcor](#) Retention Time Correction using retcor function from xcms R package

[xcms.fillPeaks](#) Integrate the signal in the region of that peak group not represented and create a new peak

[CAMERA.annotate](#) CAMERA annotate function. Returns annotation results (isotope peaks, adducts and fragments) and a diffreport if more than one condition.

[CAMERA.combinexsAnnos](#) Wrapper function for the combinexsAnnos CAMERA function.

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.

History search datasets

Sacuri Zip
1 shown
191.3 MB ✓ 📄 💬

1: sacuri.zip 👁 ✎ ✕

Tools - form

Galaxy / 4 / Metabolomics Analyze Data Workflow Shared Data Visualization Admin Help User Using -993344424 h

Tools search tools

Upload File from your computer
Export Data

LC-MS
Format Conversion
Preprocessing

[xcms.xcmsSet](#) Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis

[xcms.group](#) Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.

[xcms.retcor](#) Retention Time Correction using retcor function from xcms R package

[xcms.fillPeaks](#) Integrate the signal in the region of that peak group not represented and create a new peak

[CAMERA.annotate](#) CAMERA annotate function. Returns annotation results (isotope peaks, adducts and fragments) and a diffreport if more than one condition.

[CAMERA.combinexsAnnos](#) Wrapper function for the combinexsAnnos CAMERA function.

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.

History search datasets

Sacuri Zip
1 shown
191.3 MB

1: sacuri.zip

Tools - form

Galaxy / 4 / Metabolomics Analyze Data Workflow Shared Data Visualization Admin Help User Using -993344424 h

Tools search tools

Upload File from your computer
Export Data

LC-MS
Format Conversion
Preprocessing

[xcms.xcmsSet](#) Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis

[xcms.group](#) Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.

[xcms.retcor](#) Retention Time Correction using retcor function from xcms R package

[xcms.fillPeaks](#) Integrate the signal in the region of that peak group not represented and create a new peak

[CAMERA.annotate](#) CAMERA annotate function. Returns annotation results (isotope peaks, adducts and fragments) and a diffreport if more than one condition.

[CAMERA.combinexsAnnos](#) Wrapper function for the combinexsAnnos CAMERA function.

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.

History search datasets

Sacuri Zip
1 shown
191.3 MB 📄 🗑️ 🗨️

1: sacuri.zip 👁️ ✎️ ✕️

Tools - form

Tools can have some advanced options

The screenshot displays the Galaxy web interface for the `xcms.xcmsSet` tool (version 2.0.1). The interface is divided into several 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](#), and [Preprocessing](#). Under [Preprocessing](#), the `xcms.xcmsSet` tool is listed with a description: "Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis".
- Tool Configuration Panel (Center):** Shows the configuration for `xcms.xcmsSet`. It includes:
 - Choose your inputs method:** A dropdown menu set to "Zip file from your history containing your chromatograms".
 - Zip file:** A dropdown menu set to "1: sacuri.zip".
 - Extraction method for peaks detection:** A dropdown menu set to "matchedFilter".
 - Step size to use for profile generation:** A text input field set to "0.01".
 - Full width at half maximum of matched filtration gaussian model peak:** A text input field set to "30".
 - Advanced options:** A section highlighted with a red box, containing:
 - A "show" dropdown menu.
 - Maximum number of peaks per extracted ion chromatogram:** A text input field set to "5".
 - Signal to noise ratio cutoff:** A text input field set to "10".
 - Number of steps to merge prior to filtration:** A text input field set to "2".
- History Panel (Right):** Shows a list of datasets. The top entry is "Sacuri Zip" (191.3 MB). Below it, the dataset "1: sacuri.zip" is highlighted in green.

At the bottom of the tool configuration panel, there is an "Execute" button.

Tools - form

Galaxy / 4 / Metabolomics Analyze Data Workflow Shared Data Visualization Admin Help User Using -993344424 h

Tools search tools

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

LC-MS
[Format Conversion](#)
[Preprocessing](#)

[xcms.xcmsSet](#) Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis

[xcms.group](#) Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.

[xcms.retcor](#) Retention Time Correction using retcor function from xcms R package

[xcms.fillPeaks](#) Integrate the signal in the region of that peak group not represented and create a new peak

[CAMERA.annotate](#) CAMERA annotate function. Returns annotation results (isotope peaks, adducts and fragments) and a diffreport if more than one condition.

[CAMERA.combinexsAnnos](#) Wrapper function for the combinexsAnnos CAMERA function.

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:
show

Maximum number of peaks per extracted ion chromatogram:
5
[max]

Signal to noise ratio cutoff:
10
[snthresh]

Number of steps to merge prior to filtration:
2
[steps] The peak identification algorithm combines a given number of EIBPCs prior to filtration and peak detection, as defined by the steps argument

Execute

History ↺ ⚙️ 🗑️

search datasets

Sacuri Zip
1 shown

191.3 MB 📄 🗑️ 🗨️

1: sacuri.zip 👁️ ✎️ 🗑️

Tools - form

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Using -993344424 h

Tools

search tools

Upload File from your computer

Export Data

LC-MS

Format Conversion

Preprocessing

[xcms.xcmsSet](#) Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis

[xcms.group](#) Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.

[xcms.retcOR](#) Retention Time Correction using retcor function from xcms R package

[xcms.fillPeaks](#) Integrate the signal in the region of that peak group not represented and create a new peak

[CAMERA.annotate](#) CAMERA annotate function. Returns annotation results (isotope peaks, adducts and fragments) and a diffreport if more than one condition.

[CAMERA.combinexsAnnos](#) Wrapper function for the combinexsAnnos CAMERA function.

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

- 2: xset.RData
- 3: sampleMetadata.tsv
- 4: xset.TICs_raw.pdf
- 5: xset.BPCs_raw.pdf
- 6: xset.log.txt

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

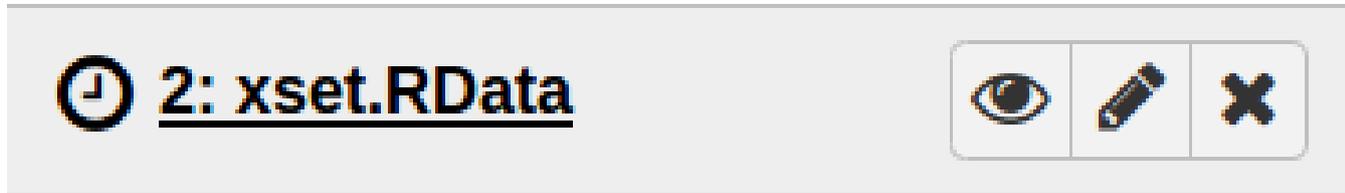
Sacuri Zip

1 shown

191.3 MB

- 6: xset.log.txt
- 5: xset.BPCs_raw.pdf
- 4: xset.TICs_raw.pdf
- 3: sampleMetadata.tsv
- 2: xset.RData
- 1: sacuri.zip

- Status



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 or on the requested number of processors

- Status



Job is currently running

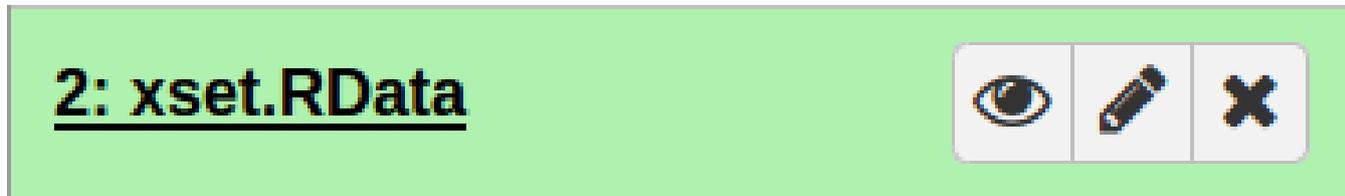
= the job is being executed on the computing cluster

Duration time of this status depends completely 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☹

- Status

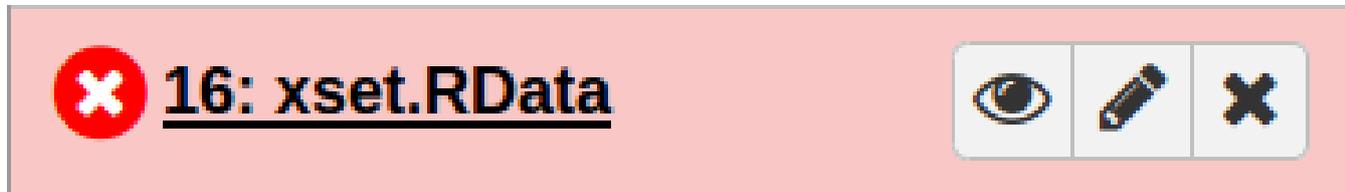


Job is finished

It's status is OK

but warnings or errors can be hidden behind. Ah hum !

- Status

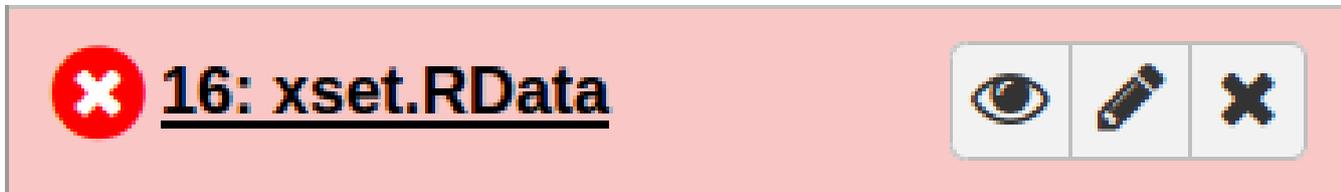


Job is finished but with an error status

= the program sends an error

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

- Status



Job is finished but with an error status

Error causes :

- ~~The user~~ :P
- Bad usage : input file, format or option
- Wrong porting of the program through Galaxy ... sorry :/
- Non anticipated crash of the program

Exercise
TOOLS

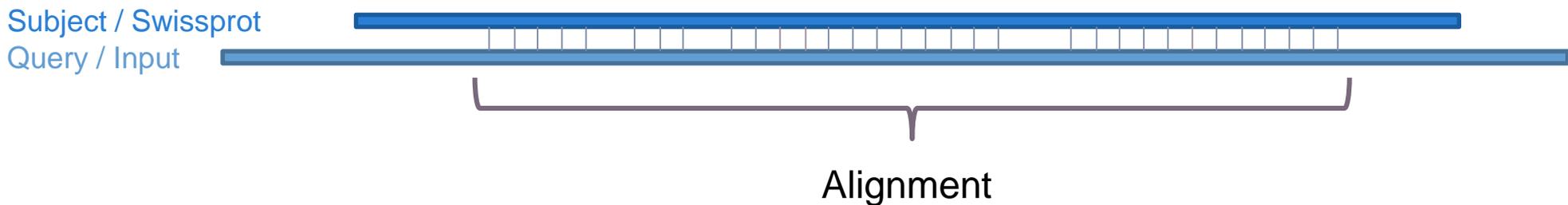




- Aim of this Exercise:
 - Get blast hits with :
 - Identity > 75%
 - Alignment coverage > 75%
 - Execute and chain example of little Galaxy friendly tools together.



- Aim of this Exercise:
 - Get blast hits with :
 - Identity > 75%
 - Alignment coverage > 75%



$$\text{Coverage} = \text{alignment length} / \text{query length}$$



- Fetch these two files (< 2Go)
 - Link1:
<http://tinyurl.com/GI-Roscoff1>
 - Link2:
<http://tinyurl.com/GI-Roscoff2>
- Check their contents and datatypes through Galaxy.



Get blast hits with :

- Identity > 75%
- Alignment coverage > 75%



Get blast hits with :

- Identity > 75%
 - Alignment coverage > 75%
-
- Step 1: use the tool « Compute sequence length »



Get blast hits with :

- Identity > 75%
 - Alignment coverage > 75%
-
- Step 1: use the tool « Compute sequence length »
 - Step 2: use the tool « Join Two Datasets »



Get blast hits with :

- Identity > 75%
 - Alignment coverage > 75%
-
- Step 1: use the tool « Compute sequence length »
 - Step 2: use the tool « Join Two Datasets »
 - help : join with column 1 and column 1



Get blast hits with :

- Identity > 75%
 - Alignment coverage > 75%
-
- Step 1: use the tool « Compute sequence length »
 - Step 2: use the tool « Join Two Datasets »
 - help : join with column 1 and column 1
 - Step 3: use the tool « Compute an expression... »



Get blast hits with :

- Identity > 75%
 - Alignment coverage > 75%
-
- Step 1: use the tool « Compute sequence length »
 - Step 2: use the tool « Join Two Datasets »
 - help : join with column 1 and column 1
 - Step 3: use the tool « Compute an expression... »
 - help : $((c8-c7+1)/c14)*100$



Get blast hits with :

- Identity > 75%
- Alignment coverage > 75%

- Step 1: use the tool « Compute sequence length »
- Step 2: use the tool « Join Two Datasets »
 - help : join with column 1 and column 1
- Step 3: use the tool « Compute an expression... »
 - help : $((c8-c7+1)/c14)*100$
- Step 4: use the tool « Filter »



Get blast hits with :

- Identity > 75%
 - Alignment coverage > 75%
-
- Step 1: use the tool « Compute sequence length »
 - Step 2: use the tool « Join Two Datasets »
 - help : join with column 1 and column 1
 - Step 3: use the tool « Compute an expression... »
 - help : $((c8-c7+1)/c14)*100$
 - Step 4: use the tool « Filter »
 - help : $c3 \geq 75$ and $c15 \geq 75$

Part II

TOOLS

Tools – Handle errors

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User Using 19.1 GB

Tools

search tools

Get Data

ABiMS WORKFLOWS

- [Workflow RNA-seq de novo by ABiMS](#)
- [Workflow RNA-seq with reference by ABiMS](#)
- [Workflow 4 Metabolomics](#)

ABiMS TOOLS

- [Primer](#)
- [RNASeq](#)
- [InterEsil](#)
- [Statistics](#)
- [Utils](#)
- [Phylogenetics](#)
- [Debug](#)

COMMON TOOLS

- [Text Manipulation](#)
- [FASTA manipulation](#)
- [Join, Subtract and Group](#)
- [Filter and Sort](#)
- [Graphics](#)
- [NCBI BLAST+](#)
- [NGS: QC and manipulation](#)
- [NGS: RNA Analysis](#)
- [NGS: Mapping](#)
- [NGS: Picard \(beta\)](#)
- [NGS: SAM Tools](#)
- [NGS: GATK Tools \(beta\)](#)
- [SVDetect](#)
- [VarScan](#)
- [Searching sequence tools](#)

Online

Information

- 07-06-13: Metabolomic : Workflow 4 Metabolomics, updated to version 2.1.0 (2013_06_07) [↓](#)
- 30-04-13: RNASeq : DESeq is now available for RNASeq expression data with reference (with gtf input).
- 26-04-13: RNASeq : DESeq is now available for denovo RNASeq expression data (without gtf input).
- 26-04-13: RNASeq : sam2counts is now available to count the reads coverage by transcrit. It's also a requirement for DESeq denovo.
- 26-04-13: Metabolomic : Workflow Metabolomic by ABiMS, updated to version 2.0.0 (2013_04_18) [↓](#)



Analyses and Bioinformatics for Marine Science



CNRS UPMC
Station Biologique
Roscoff

Information

For any question or request for tools or account, send an email at [support.abims 'AT' sb-roscoff.fr](mailto:support.abims@at.sb-roscoff.fr)

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

Copper Stress v3
133.9 MB

XZ: [mzXML_copper_stress.group.retcor.gro](#)
[up.fillPeaks.annotateDiffreport.data](#)
[ma](#)
[trix.tsv_anova_pvalue.tabular](#)

6: [mzXML_copper_stress.group.retcor.gro](#)
[up.fillPeaks.annotateDiffreport.Rdata](#)

5: [mzXML_copper_stress.group.retcor.gro](#)
[up.fillPeaks.RData](#)

4: [mzXML_copper_stress.group.retcor.gro](#)
[up.RData](#)

3: [mzXML_copper_stress.RData](#)

2: [sampleInfo.tab](#)

1: [mzXML_copper_stress.ms.zip](#)

HISTORY

History panel

Both inputs and outputs

The screenshot displays the Galaxy web interface. The top navigation bar includes 'Galaxy / 4 / Metabolomics', '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: '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'. The main workspace shows the 'Batch_correction (version 2.0.0)' tool configuration. The 'Data Matrix file' is set to '17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv', 'Sample metadata file' to '3: sampleMetadata.tsv', and 'Variable metadata file' to '16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv'. The 'Type of regression model' is 'linear', and the 'Factor of interest' is 'batch'. The 'Level of details for plots' is 'basic'. An 'Execute' button is visible. Below the configuration, the 'Authors' and 'Contributors' sections are shown. The right sidebar features a 'History' panel with a search bar and a list of datasets. The datasets listed include 'Sacuri Zip' (289.7 MB) and a series of numbered entries (19 down to 11) with their respective file names and icons for viewing, editing, and deleting.

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:

History panel

Both inputs and outputs

The screenshot displays the Galaxy web interface for the 'Batch_correction' tool (version 2.0.0). The tool configuration includes the following fields:

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

The 'Execute' button is visible at the bottom of the tool configuration. Below the configuration, the 'Authors' and 'Contributors' sections are displayed.

The 'History' panel on the right shows a list of datasets generated by the tool, with steps 11 through 19 highlighted in green. The datasets listed are:

- 19: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv (Xdiffreport)
- 18: xset.group.retcor.group.fillPeaks.annotate.negative.Rdata
- 17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv
- 16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv
- 15: xset.group.retcor.group.fillPeaks.RData
- 14: xset.group.retcor.group.Rplots.pdf
- 13: xset.group.retcor.group.RData
- 12: xset.group.retcor.BPCs_corrected.pdf
- 11: (partially visible)

A large red watermark with the word 'History' is overlaid on the tool configuration area.

History panel renaming and annotation

History

search datasets

Unnamed history
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:



History

search datasets

Sacuri
19 shown
289.7 MB

Click to rename history

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:



History

search datasets

Sacuri
19 shown
289.7 MB

Tags:
human LC-MS

Annotation:
Annotation of the Human Adult Urinary Metabolome and Metabolite Identification Using Ultra High Performance Liquid Chromatography Coupled to a Linear Quadrupole Ion Trap-Orbitrap Mass Spectrometer.

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

History panel

Saved histories: Rename, Delete, **Delete Permanently**

Galaxy / 4 / Metabolomics Analyze Data Workflow Shared Data Visualization Admin Help User Using -99334424 b

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](#)

Saved Histories

Q

Advanced Search

<input type="checkbox"/>	Name	Datasets	Tags	Sharing	Size on Disk	Created	Last Updated	Status
<input type="checkbox"/>	Sacuri	19	2 Tags		289.7 MB	Sep 02, 2015	~3 days ago	current history
<input type="checkbox"/>	Sacuri Lib	30	0 Tags		17.3 MB	May 14, 2014	Sep 02, 2015	
<input type="checkbox"/>	Cooper Stress Lib	19	0 Tags		7.8 MB	May 13, 2014	Sep 02, 2015	

For 0 selected histories:

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

History

- HISTORY LISTS
- Saved Histories**
- Histories Shared with Me
- CURRENT HISTORY
- Create New
- Copy History
- Copy Datasets
- Share or Publish
- Extract Workflow
- Dataset Security
- Resume Paused Jobs
- Collapse Expanded Datasets
- Unhide Hidden Datasets
- Delete Hidden Datasets
- Purge Deleted Datasets
- Show Structure
- Export Citations
- Export to File
- Delete
- Delete Permanently
- OTHER ACTIONS
- Import from File

19: xset.gr... tate.var...
18: xset.gr... tate.ne...
17: xset.gr... tate.da...
16: xset.gr... tate.va...
15: xset.gr...
14: xset.gr...
13: xset.group.retcor.group.RData
12: xset.group.retcor.BPCs_corrected.pdf
11:

History panel

Saved histories: Switch histories

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Using -99334424 b

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

Saved Histories

search history names and tags

Advanced Search

<input type="checkbox"/>	Name	Datasets	Tags	Sharing	Size on Disk	Created	Last Updated↑	Status
<input type="checkbox"/>	Sacuri	19	2 Tags		289.7 MB	Sep 02, 2015	~3 days ago	current history
<input type="checkbox"/>	Sacuri Lib	30	0 Tags		17.3 MB	May 14, 2014	Sep 02, 2015	
<input type="checkbox"/>	Cooper Stress Lib	19	0 Tags		7.8 MB	May 13, 2014	Sep 02, 2015	

For 0 selected histories:

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

History

search

HISTORY LISTS

- Saved Histories
- Histories Shared with Me

CURRENT HISTORY

- Create New
- Copy History
- Copy Datasets
- Share or Publish
- Extract Workflow
- Dataset Security
- Resume Paused Jobs
- Collapse Expanded Datasets
- Unhide Hidden Datasets
- Delete Hidden Datasets
- Purge Deleted Datasets
- Show Structure
- Export Citations
- Export to File
- Delete
- Delete Permanently

OTHER ACTIONS

- Import from File

13: xset.group.retcor.group.RData

12: xset.group.retcor.BPCs_corrected.pdf

11:

DATASET

Dataset

Both inputs and outputs

The screenshot displays the Galaxy web interface. The top navigation bar includes 'Galaxy / 4 / Metabolomics', '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: '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'. The main workspace shows the 'Batch_correction (version 2.0.0)' tool configuration. The 'Data Matrix file' is set to '17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv', the 'Sample metadata file' is '3: sampleMetadata.tsv', and the 'Variable metadata file' is '16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv'. The 'Type of regression model' is 'linear', the 'Factor of interest' is 'batch', and the 'Level of details for plots' is 'basic'. An 'Execute' button is visible. Below the tool configuration, there are sections for 'Authors' (Jean-Francois Martin) and 'Contributors' (Melanie Petera, Etienne Thevenot). On the right, the 'History' panel shows a list of datasets. The dataset '19: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv (Xdiffreport)' is highlighted with a red box. Other datasets in the history include '18: xset.group.retcor.group.fillPeaks.annotate.negative.Rdata', '17: xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv', '16: xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv', '15: xset.group.retcor.group.fillPeaks.RData', '14: xset.group.retcor.group.Rplots.pdf', '13: xset.group.retcor.group.RData', '12: xset.group.retcor.BPCs_corrected.pdf', and '11:'. A large red text 'Dataset' is overlaid on the right side of the main workspace.

Dataset

Informations

Galaxy / 4 / Metabolomics Analyze Data Workflow Shared Data Visualization Admin Help User Using -99334424

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 Statistics Graph/Display Data Deprecated Tools Multiple regression

Workflows All workflows

Tool: xcms.group

Name: xset.group.retcor.group.RData
Created: Wed Sep 2 09:11:46 2015 (UTC)
Filesize: 5.2 MB
Dbkey: ?
Format: rdata.xcms.group
Galaxy Tool ID: toolshed.france-bioinformatique.fr/repos/lecorguille/xcms_group/abims_xcms_group/2.0.1
Galaxy Tool Version: 2.0.1
Tool Version:
Tool Standard Output: stdout
Tool Standard Error: stderr
Tool Exit Code: 0
API ID: 9265a1b3d61fdbeb
History ID: fd7c05917f9701f7
UUID: 72a13a4b-6e2e-47a6-b152-27d5187df767

Input Parameter	Value	Note for rerun
xset RData file	9: xset.group.retcor.RData	
Method to use for grouping	density	
Bandwidth	5	
Minimum fraction of samples necessary	0.3	
Width of overlapping m/z slices	0.01	
Advanced options	show	
Maximum number of groups to identify in a single m/z slice	50	

Inheritance Chain

xset.group.retcor.group.RData

History

- 13: xset.group.retcor.group.RData (5.2 MB, format: rdata.xcms.group, database: 2)
PACKAGE INFO
parallel 3.1.1
BiocGenerics 0.12.1
Biobase 2.26.0
Rcpp 0.11.5
mzR 2.0.0
igraph 0.7.1
xcms 1.42.0
snow 0.3.13
batch 1.1.4
ARGUMENTS INFO
xfunction group
image /w/galaxy/galaxy4metabolomics/galaxy-dist/database/files/045/dataset_45527.dat
m
- 14: xset.group.retcor.group.Rplots.pdf
- 12: xset.group.retcor.BPCs_corrected.pdf
- 11: xset.group.retcor.TICs_corrected.pdf
- 10: xset.group.retcor.Rplots.pdf
- 9: xset.group.retcor.RData
- 8: xset.group.Rplots.pdf

galaxy4metabolomics.sb-roscoff.fr/datasets/9265a1b3d61fdbeb/show_params

Dataset

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

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

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

name	Blanc15	Blanc09	Blanc12	Blanc06	Blanc17	
M100T293	0	0	0	0	0	
M100T313	14737.3434458556	70497.1552979614	29398.2144370894	74121.4139604603	8413.61560029072	61416.05
M100T318	1396.8756293629	6403.15709537553	3483.72951135339	6365.40360216962	1256.93101995603	9009.046
M100T415	4103.24769663852	6007.46669614238	5686.56418559869	18169.7184114364	1394.93820578164	11794.10
M101T308	0	0	1354.41420127279	0	0	
M101T63	927.737622943425	0	0	0	0	
M102T348	649.422177375992	0	0	0	0	
M102T379	7957.3450177005	17013.5105876073	11971.525671295	24170.4556255728	5403.87982724869	18730.33
M102T59	0	0	0	0	480.552181164762	
M103T1003	0	0	0	0	0	
M103T1012	0	1776.41171822315	0	0	0	
M103T152	0	0	0	0	0	
M103T45	26872.0684617598	51859.8370837991	1335.11123434995	15915.782182274	13919.992158649	117195.7
M103T50	26002.5859959576	100564.745002913	6153.36022361006	82775.7546772299	17210.0193279638	109490.6
M103T63	26143.9616699194	68658.1859951143	45073.1929194364	65013.3829999986	15636.7248046881	88253.90
M105T50	37864.7519906614	172016.779677334	104572.584541783	186717.191434361	0	225744.8
M105T57	22972.800179443	175038.512133167	125815.056393625	232961.533116152	0	265084.4
M107T348	39111.6763561207	111455.640695432	63522.7138126517	94976.4981542975	13711.7442593179	117586.6
M107T379	0	66740.2961367195	109739.70423633	113615.958986816	2009.03136237782	12746.20
M108T336	2875.53106221982	65289.224125976	59553.329443204	162505.874087406	0	13579.44
M108T379	507.943477920085	52885.4503151831	6633.9879785059	140119.471455469	519.076044539372	55462.93
M109T294	860.602751283829	84109.0329002055	30997.4207106858	137467.156183397	506.708184882696	6629.054
M109T51	3163.71089178629	14569.0030136036	23613.854648709	264370.767449398	830.560495430845	24031.59
M110T294	1324.57426677155	15356.52582916	55063.9665625007	86087.9042496742	0	24399.01
M110T313	0	11985.727034715	72124.8941731575	71361.0011601552	2461.35172321908	36048.75
M110T55	2572.37712822047	13158.4785864824	2853.71333606008	16036.855922733	0	7393.124
M111T273	16799.0249907129	130599.823855273	42009.1613629325	78842.8338411576	3967.62968023854	29231.52
M111T338	1946.90464648183	47811.3018066406	7387.00068640994	25405.6901166716	466.69084317689	3325.518
M111T51	3395.04200209943	5670.38219377001	9834.69273551838	2592.88676690728	7642.34685296875	
M111T58	0	0	0	2331.99952486997	0	

History

search datasets

Sacuri

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](#) View data

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: [xset.group.retcor.BPCs_corrected.pdf](#)

galaxy4metabolomics.sb-roscoff.fr/datasets/b787ed13b1b7c1bf/display/?preview=True

Dataset

Renaming and annotation

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Using -993344424 b

Attributes Convert Format Datatype Permissions

Edit Attributes

Name:
xset.group.retcor.group.fillPeaks.annotate

Info:
adding: bio_vs_blank.annot.tsv (deflated 52%)
adding: bio_vs_blank_box/ (stored 0%)

Annotation / Notes:

Add an annotation or notes to a dataset; annotations are available when a history is viewed.

Database/Build:
unspecified (?)

Save

Auto-detect

This will inspect the dataset and attempt to correct the above column values if they are not accurate.

History

search datasets

Sacuri
19 shown
289.7 MB

19:
[xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv \(Xdiffreport\)](#)

18:
[xset.group.retcor.group.fillPeaks.annotate.negative.Rdata](#) **Edit attributes**

7.5 MB
format: **rdata.camera.negative**,
database: **?**

adding: bio_vs_blank.annot.tsv (deflated 52%)
adding: bio_vs_blank_box/ (stored 0%)
adding: bio_vs_blank_box/001.png (deflated 23%)
adding: bio_vs_blank_box/002.png (deflated 23%)
adding: bio_vs_blank_box/003.png (deflated 23%)
adding: bio_vs_b

binary data

17:
[xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv](#)

Dataset

Change the Datatype of the Dataset

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Attributes Convert Format **Datatype** Permissions

Change data type

New Type:

- txt
- rgb
- sam
- scf
- sff
- sif
- svg
- tabix
- tabular**

dataset but *not* modify its contents. Use this if Galaxy has incorrectly guessed the type of your dataset.

History

- 14: [xset.group.retcor.group.fillPeaks.RData](#)
- 13: [xset.group.retcor.group.Rplots.pdf](#)
- 12: [xset.group.retcor.BPCs_corrected.pdf](#)
- 11: [xset.group.retcor.TICs_corrected.pdf](#)
- 10: [xset.group.retcor.Rplots.pdf](#)
- 9: [xset.group.retcor.RData](#)
- 8: [xset.group.Rplots.pdf](#)
- 7: [xset.group.RData](#)
- 6: [xset.log.txt](#)**
- 5: [xset.BPCs_raw.pdf](#)
- 4: [xset.TICs_raw.pdf](#)
- 3: [sampleMetadata.tsv](#)
- 2: [xset.RData](#)
- 1: [sacuri.zip](#)

1 Visualize charts

name	Bl...	Scatterplot
C1_011	HU...	

New Chart

Start Configuration 1: Data label Add Data

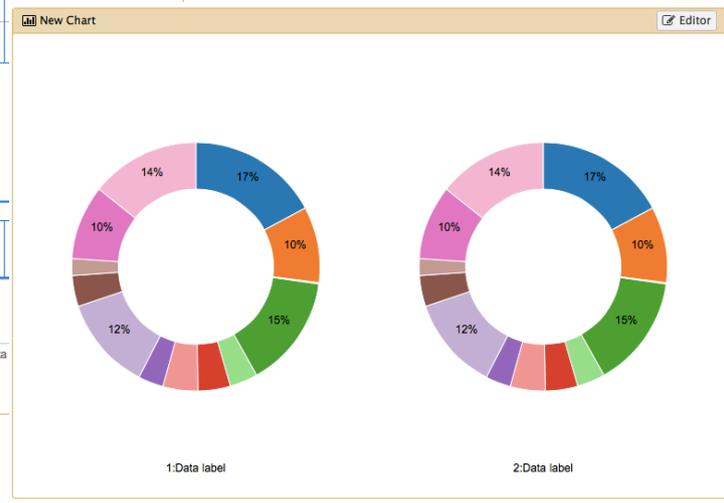
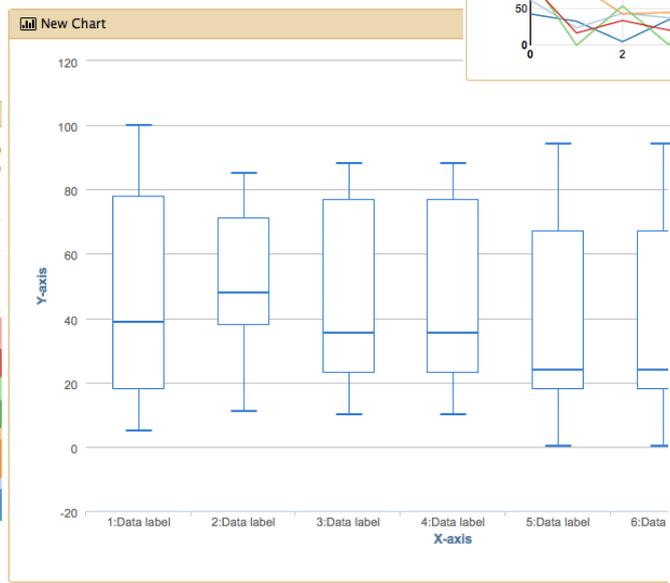
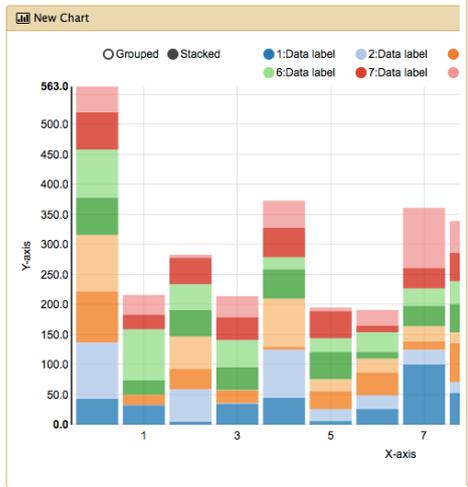
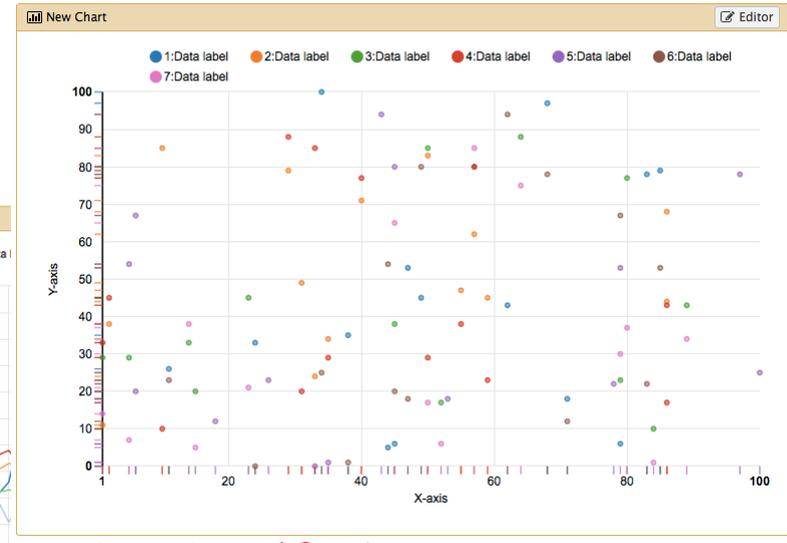
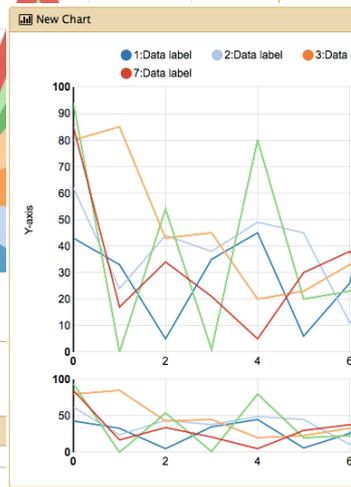
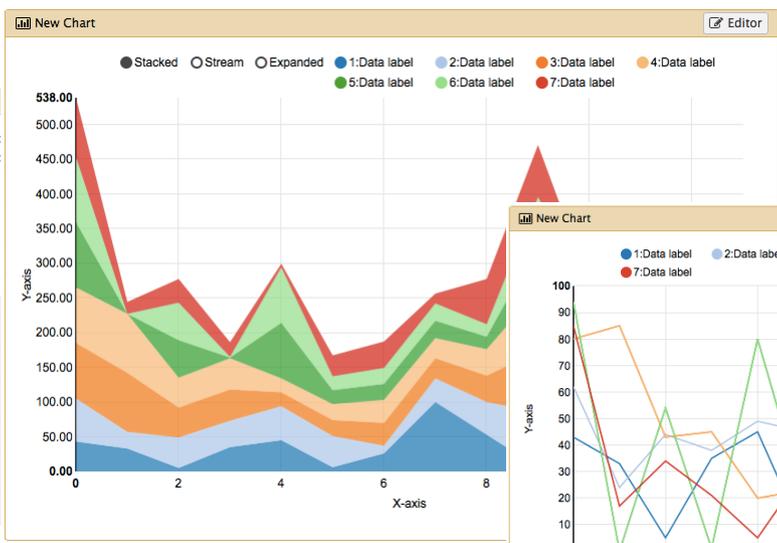
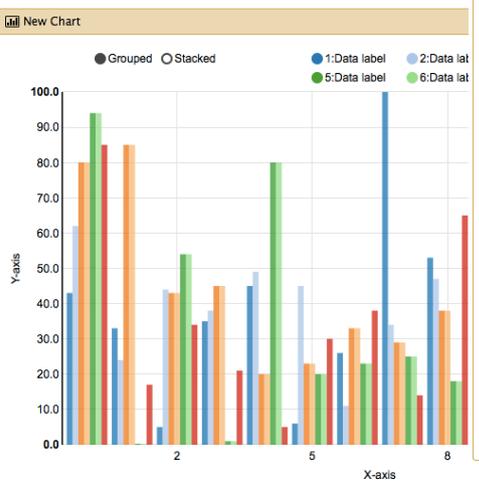
Provide a chart title:

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)



Dataset

Re-run a job

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Using -99334424

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

CAMERA.annotate (version 2.0.0)

RData file:
15: xset.group.retcor.group.fillPeaks.RData
output file from another function xcms (fillPeaks)

Convert retention time (seconds) into minutes:

Convert the columns rtmed, rtmin and rtmax into minutes

num_digits:

Number of decimal places for mass values reported in ions identifiers

groupFWHM: multiplier of the standard deviation:

[sigma]

groupFWHM: percentage of FWHM width:

[perfwHM]

findIsotopes: max. ion charge:

[maxcharge]

findIsotopes: max. number of expected isotopes:

[maxiso]

findIsotopes: The percentage number of samples, which must satisfy the C12/C13 rule for isotope annotation:

[minfrac]

General ppm error:

[ppm]

History

289.7 MB

19: [xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv \(Xdiffreport\)](#)

18: [xset.group.retcor.group.fillPeaks.annotate.negative.Rdata](#)
7.5 MB
format: **rdata.camera.negative**,
database: ?
adding: bio_vs_blank.annot.tsv (deflated 52%)
adding: bio_vs_blank_box/ (stored 0%)
adding: bio_vs_blank_box/001.png (deflated 23%)
adding: bio_vs_blank_box/002.png (deflated 23%)
adding: bio_vs_blank_box/003.png (deflated 23%)
adding: bio_vs_b

Run this job again

17: [xset.group.retcor.group.fillPeaks.annotate.dataMatrix.tsv](#)

16: [xset.group.retcor.group.fillPeaks.annotate.variableMetadata.tsv](#)

15:

galaxy4metabolomics.sb-roscoff.fr/tool_runner/rerun?id=89e90f36310e084c

Cleanup

DATASET



Dataset / Cleanup

Delete a dataset

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Using -993344424

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

CAMERA.annotate (version 2.0.0)

RData file:
15: xset.group.retcor.group.fillPeaks.RData
output file from another function xcms (fillPeaks)

Convert retention time (seconds) into minutes:

Convert the columns rtmed, rtmin and rtmax into minutes

num_digits:

Number of decimal places for mass values reported in ions identifiers

groupFWHM: multiplier of the standard deviation:

[sigma]

groupFWHM: percentage of FWHM width:

[perfwHM]

findIsotopes: max. ion charge:

[maxcharge]

findIsotopes: max. number of expected isotopes:

[maxiso]

findIsotopes: The percentage number of samples, which must satisfy the C12/C13 rule for isotope annotation:

[minfrac]

General ppm error:

[ppm]

History

search datasets

Sacuri
19 shown
289.7 MB

19: xset.group.retcor.group.fillPeaks.tate.variableMetadata.tsv (Xdiffre) 

18: xset.group.retcor.group.fillPeaks.anno.tate.negative.Rdata 

17: xset.group.retcor.group.fillPeaks.anno.tate.dataMatrix.tsv 

16: xset.group.retcor.group.fillPeaks.anno.tate.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: 

Dataset / Cleanup



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

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User Using -99334424

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

CAMERA.annotate (version 2.0.0)

RData file:
15: xset.group.retcor.group.fillPeaks.RData
output file from another function xcms (fillPeaks)

Convert retention time (seconds) into minutes:

Convert the columns rtmed, rtmin and rtmax into minutes

num_digits:
0
Number of decimal places for mass values reported in ions identifiers

groupFWHM: multiplier of the standard deviation:
6
[sigma]

groupFWHM: percentage of FWHM width:
0.6
[perfwHM]

findIsotopes: max. ion charge:
2
[maxcharge]

findIsotopes: max. number of expected isotopes:
2
[maxiso]

findIsotopes: The percentage number of samples, which must satisfy the C12/C13 rule for isotope annotation:
0.5
[minfrac]

General ppm error:
5
[ppm]

History

search datasets

Sacuri
19 shown [hide deleted](#)
289.7 MB

This dataset has been deleted
Undelete it
Permanently remove it from disk

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](#)

Dataset / Cleanup



“Empty Trash” : to free up disk space

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User Using -99334424 b

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

CAMERA.annotate (version 2.0.0)

RData file:
15: xset.group.retcor.group.fillPeaks.RData
output file from another function xcms (fillPeaks)

Convert retention time (seconds) into minutes:

Convert the columns rtmed, rtmin and rtmax into minutes

num_digits:
0
Number of decimal places for mass values reported in ions identifiers

groupFWHM: multiplier of the standard deviation:
6
[sigma]

groupFWHM: percentage of FWHM width:
0.6
[perfwHM]

findIsotopes: max. ion charge:
2
[maxcharge]

findIsotopes: max. number of expected isotopes:
2
[maxiso]

findIsotopes: The percentage number of samples, which must satisfy the C12/C13 rule for isotope annotation:
0.5
[minfrac]

General ppm error:
5
[ppm]

History

search

Sacuri
19 show

289.7 M

! This
Un
Pe

19:
xset.gr
tate.var

18:
xset.gr
tate.ne

17:
xset.gr
tate.dat

16:
xset.gr
tate.var

15:
xset.gr
a

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

13:
xset.group.retcor.group.RData

HISTORY LISTS
Saved Histories
Histories Shared with Me

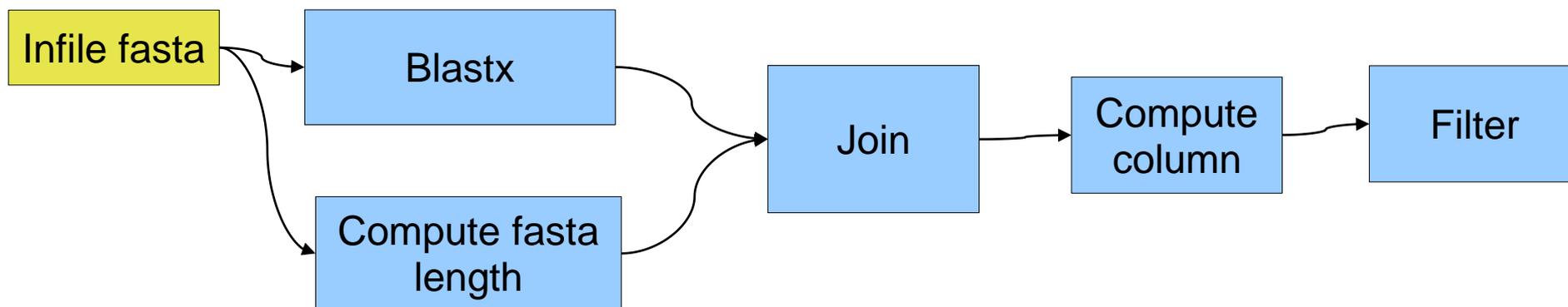
CURRENT HISTORY
Create New
Copy History
Copy Datasets
Share or Publish
Extract Workflow
Dataset Security
Resume Paused Jobs
Collapse Expanded Datasets
Unhide Hidden Datasets
Delete Hidden Datasets
Purge Deleted Datasets
Show Structure
Export Citations
Export to File
Delete
Delete Permanently

OTHER ACTIONS
Import from File

WORKFLOW

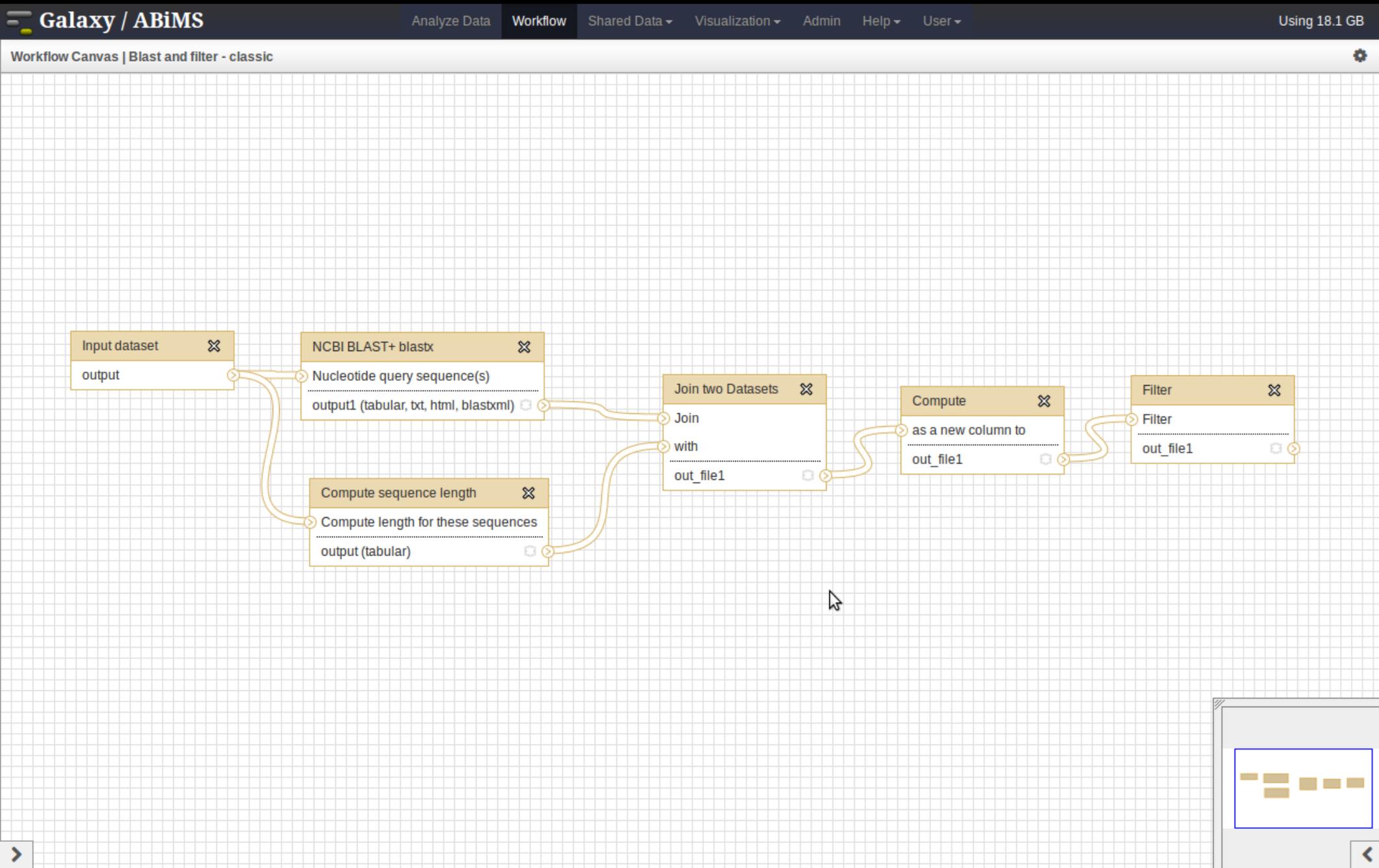
- A workflow is a sequence of tool operations and parameters
- Can match the experiment protocol
- A workflow is built to be replayed (more or less strict)

- Our workflow



Workflow

Our workflow with Galaxy



Workflow

From a history

The screenshot shows the Galaxy/ABiMS web interface. The top navigation bar includes 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. The right corner indicates 'Using 18.1 GB'. The left sidebar contains a 'Tools' section with a search bar and various tool categories like 'Get Data', 'ABiMS WORKFLOWS', 'ABiMS TOOLS', and 'COMMON TOOLS'. The main content area features a green 'Online' status bar, an information box with a list of workflow updates, the 'Ab4 AbiMS' logo, the text 'Analyses and Bioinformatics for Marine Science', and the 'Station Biologique Roscoff' logo. A 'History' panel on the right shows a list of workflow entries, and a context menu is open over the entry '7: blastx', listing actions such as 'Extract Workflow', 'Copy History', and 'Delete'.

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User Using 18.1 GB

Tools

search tools

Get Data

ABiMS WORKFLOWS

Workflow RNA-seq de novo by ABiMS

Workflow RNA-seq with reference by ABiMS

Workflow 4 Metabolomics

ABiMS TOOLS

Primer

RNASeq

InterEsil

Statistics

Utils

Phylogenetics

Debug

COMMON TOOLS

Text Manipulation

FASTA manipulation

Join, Subtract and Group

Filter and Sort

Graphics

NCBI BLAST+

NGS: QC and manipulation

NGS: RNA Analysis

NGS: Mapping

NGS: Picard (beta)

NGS: SAM Tools

NGS: GATK Tools (beta)

SVDetect

VarScan

Muscle

RAxML

Online

07-06-13: Metabolomic : Workflow 4 Metabolomics, updated to version 2.1.0 (2013_06_07)

30-04-13: RNASeq : DESeq is now available for RNASeq expression data with reference (with gtf input).

26-04-13: RNASeq : DESeq is now available for denovo RNASeq expression data (without gtf input).

26-04-13: RNASeq : sam2counts is now available to count the reads coverage by transcrit. It's also a requirement for DESeq denovo.

26-04-13: Metabolomic : Workflow Metabolomic by ABiMS, updated to version 2.0.0 (2013_04_18)

Ab4 AbiMS

Analyses and Bioinformatics for Marine Science

CNRS UPMC

Station Biologique Roscoff

Information

For any question or request for tools or account, send an email at support.abiMS 'AT' sb-roscoff.fr

History

HISTORY LISTS

Saved Histories

Histories Shared with Me

CURRENT HISTORY

Create New

Copy History

Copy Datasets

Share or Publish

Extract Workflow

Dataset Security

Resume Paused Jobs

Collapse Expanded Datasets

Include Deleted Datasets

Include Hidden Datasets

Unhide Hidden Datasets

Delete Hidden Datasets

Purge Deleted Datasets

Show Structure

Export to File

Delete

Delete Permanently

OTHER ACTIONS

Import from File

129 /

Workflow

From a history

The screenshot displays the Galaxy/ABiMS web interface. The top navigation bar includes 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. The right corner shows 'Using 18.1 GB'. On the left, a 'Tools' sidebar lists various categories like 'Get Data', 'ABiMS WORKFLOWS', 'ABiMS TOOLS', and 'COMMON TOOLS'. The main content area features a 'History' panel on the right with a gear icon. A context menu is open over the history, listing actions such as 'Extract Workflow', 'Copy History', and 'Delete Permanently'. The central workspace contains a workflow history entry with a large black icon of two figures (one sitting, one standing) and the text '4' and 'Science'.

Galaxy / ABiMS

Analyze Data Workflow Shared Data Visualization Admin Help User Using 18.1 GB

Tools

search tools

Get Data

ABiMS WORKFLOWS

Workflow RNA-seq de novo by ABiMS

Workflow RNA-seq with reference by ABiMS

Workflow 4 Metabolomics

ABiMS TOOLS

Primer

RNASeq

InterEsil

Statistics

Utils

Phylogenetics

Debug

COMMON TOOLS

Text Manipulation

FASTA manipulation

Join, Subtract and Group

Filter and Sort

Graphics

NCBI BLAST+

NGS: QC and manipulation

NGS: RNA Analysis

NGS: Mapping

NGS: Picard (beta)

NGS: SAM Tools

NGS: GATK Tools (beta)

SVDetect

VarScan

Muscle

RAxML

Online

07-13: Metabolomic : Workflow 4 Metabolomics, updated to version 2.1.0 (2013_06_07)

30-13: RNASeq : DESeq is now available for RNA-seq expression data analysis (with gtf files)

26-13: RNASeq : DESeq is now available for denovo transcriptome expression (without gtf files) (requirement)

26-13: RNASeq : sam2counts is now available to count reads from bam files (requirement)

26-13: Metabolomic : Workflow Metabolomic by ABiMS, updated to version 2.1.0 (2013_06_07)

4

Science

History

HISTORY LISTS

Saved Histories

Histories Shared with Me

CURRENT HISTORY

Create New

Copy History

Copy Datasets

Share or Publish

Extract Workflow

Dataset Security

Resume Paused Jobs

Collapse Expanded Datasets

Include Deleted Datasets

Include Hidden Datasets

Unhide Hidden Datasets

Delete Hidden Datasets

Purge Deleted Datasets

Show Structure

Export to File

Delete

Delete Permanently

OTHER ACTIONS

Import from File

Information

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

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.

130 /

Workflow

From a history

Galaxy / 4 / Metabolomics

Analyze Data Workflow Shared Data Visualization Admin Help User

Using -99334424 b

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

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
LS-MS

Create Workflow Check all Uncheck all

Tool	History items created
Upload File <i>This tool cannot be used in workflows</i>	1: sacuri.zip <input checked="" type="checkbox"/> Treat as input dataset
xcms.xcmsSet <input checked="" type="checkbox"/> Include "xcms.xcmsSet" in workflow	2: xset.RData 3: sampleMetadata.tsv 4: xset.TICs_raw.pdf 5: xset.BPCs_raw.pdf 6: xset.log.txt
xcms.group <input checked="" type="checkbox"/> Include "xcms.group" in workflow	7: xset.group.RData 8: xset.group.Rplots.pdf 9: xset.group.retcor.RData 10: xset.group.retcor.Rplots.pdf 11: xset.group.retcor.TICs_corrected.pdf 12: xset.group.retcor.BPCs_corrected.pdf
xcms.retcor <input checked="" type="checkbox"/> Include "xcms.retcor" in workflow	13: xset.group.retcor.group.RData

History

search datasets

Sacuri
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:

Workflow

Workflow manager

Your workflows

1

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

Name	# of Steps
LS-MS ▾	7
Copy of 'gigaXml' shared by 'ethevenot@sb-roscoff.fr' ▾	13
Workflow LC/MS ▾	6
Community ▾	10
Full_workflow ▾	19
Workflow XCMS ▾	8

Workflows shared with you by others

Name	Owner	# of Steps
demo_workflow_06_annotation ▾	mlandi@sb-roscoff.fr	6
cohort ▾	ethevenot@sb-roscoff.fr	15
gigaRaw-convert ▾	ethevenot@sb-roscoff.fr	1

Other options

[Configure your workflow menu](#)

Workflow

Edit a workflow

Your workflows

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

Name	# of Steps
I.S.MS 1	7
by 'ethevenot@sb-roscoff.fr'	13
	6
	10
	19
	8

Workflows shared with you by others

Name	Owner	# of Steps
demo_workflow_06_annotation	mlandi@sb-roscoff.fr	6
cohort	ethevenot@sb-roscoff.fr	15
gigaRaw-convert	ethevenot@sb-roscoff.fr	1

Other options

[Configure your workflow menu](#)

Workflow

Edit a workflow : drag and drop

Galaxy / METABO

Analyze Data Workflow Shared Data Visualization Admin Help User Using 7.8 MB

Tools

search tools

Get Data

WORKFLOW 4 METABOLOMICS

- [2-Preprocessing](#)
- [3-Normalisation](#)
- [4-Quality Control](#)
- [5-Statistical Analysis](#)
- [6-Annotation](#)

COMMON TOOLS

- [Text Manipulation](#)
- [Filter and Sort](#)
- [Join, Subtract and Group](#)
- [Statistics](#)
- [Graph/Display Data](#)
- [Multiple regression](#)

Workflow control

Inputs

Workflow Canvas | Workflow XCMS

```
graph LR; T1[xcms.xcmsSet] -- "output (rdata)" --> T2[xcms.group]; T2 -- "output (rdata)" --> T3[xcms.refine]; T3 -- "output (rdata)" --> T4[xcms.group]; T4 -- "output (rdata)" --> T5[xcms.fillPeaks];
```

Tool	Input	Output
xcms.xcmsSet		output (rdata)
xcms.group	RData file	output (rdata), rplots (pdf), log (txt)
xcms.refine	RData file	output (rdata), rplots (pdf), tics_cor (pdf), log (txt)
xcms.group	RData file	output (rdata), rplots (pdf), log (txt)
xcms.fillPeaks	RData file	output (rdata), log (txt)

Workflow

Edit a workflow : drag and drop

The screenshot displays the Galaxy METABO interface. The top navigation bar includes 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. The left sidebar lists tool categories: 'Get Data', 'WORKFLOW 4 METABOLOMICS' (with sub-items 2-Preprocessing, 3-Normalisation, 4-Quality Control, 5-Statistical Analysis, 6-Annotation), 'COMMON TOOLS', 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Statistics', 'Graph/Display Data', 'Multiple regression', 'Workflow control', and 'Inputs'. The main 'Workflow Canvas | Workflow XCMS' area shows a sequence of tool nodes: 'xcms.xcmsSet' (output: rdata, output_info, tics_raw, log), 'xcms.group' (output: rdata, rplots, log), 'xcms.rawCor' (output: rdata, rplots, tics_cor, log), another 'xcms.group' (output: rdata, rplots, log), and 'xcms.fillPeaks' (output: rdata, log). A blue box highlights the 'xcms.rawCor' node, and a yellow arrow points to its 'RData file' input, which is connected to the 'output (rdata)' of the second 'xcms.group' node.

Workflow

Edit a workflow : delete a noodle

Galaxy / METABO

Analyze Data Workflow Shared Data Visualization Admin Help User Using 7.8 MB

Tools

search tools

Get Data

WORKFLOW 4 METABOLOMICS

[2-Preprocessing](#)

[3-Normalisation](#)

[4-Quality Control](#)

[5-Statistical Analysis](#)

[6-Annotation](#)

COMMON TOOLS

[Text Manipulation](#)

[Filter and Sort](#)

[Join, Subtract and Group](#)

[Statistics](#)

[Graph/Display Data](#)

[Multiple regression](#)

[Workflow control](#)

[Inputs](#)

Workflow Canvas | Workflow XCMS

```
graph LR; A[xcms.xcmsSet] --> B[xcms.group]; B --> C[xcms.retcor]; C --> D[xcms.group]; D --> E[xcms.fillPeaks];
```

The workflow consists of the following tools and their connections:

- xcms.xcmsSet** (leftmost tool) connects to the **xcms.group** tool below it.
- The **xcms.group** tool below it connects to the **xcms.retcor** tool above it.
- The **xcms.retcor** tool connects to the **xcms.group** tool below it.
- The **xcms.group** tool below it connects to the **xcms.fillPeaks** tool on the far right.

The **xcms.group** tool below it is highlighted with a blue border. The connection point between the **xcms.group** tool below it and the **xcms.fillPeaks** tool is highlighted with a yellow circle and a mouse cursor.

Workflow

Edit a workflow : add a tool

Galaxy / METABO

Analyze Data Workflow Shared Data Visualization Admin Help User Using 7.8 MB

Tools

search tools

Get Data

WORKFLOW 4 METABOLOMICS

2-Preprocessing

- [xcms.xcmsSet](#) Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis
- [xcms.group](#) Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.
- [xcms.retcor](#) Retention Time Correction using retcor function from xcms R package
- [xcms.fillPeaks](#) Integrate the signal in the region of that peak group not represented and create a new peak
- [xcms.diffreport](#) A report showing the most statistically significant differences in analyte intensities
- [CAMERA.annotateDiffreport](#) Wrapper function for the

Workflow Canvas | Workflow XCMS

The workflow canvas shows a sequence of tools: **xcms.xcmsSet** (output: rdata, output_info (tabular), tics_raw (pdf), log (txt)) feeds into **xcms.group** (output: rdata, rplots (pdf), log (txt)). This **xcms.group** tool feeds into another **xcms.group** tool, which has a sub-tool **xcms.retcor** (output: rplots (pdf), log (txt)) highlighted with a blue box. The **xcms.retcor** tool also feeds into a final **xcms.fillPeaks** tool (output: rdata, log (txt)).

galaxy4metabolomics.sb-roscoff.fr/workflow/editor?id=c1992d25d70f8c1f#

Workflow

Edit a workflow : add a noodle

Galaxy / METABO

Analyze Data Workflow Shared Data Visualization Admin Help User Using 7.8 MB

Tools

search tools

Get Data

WORKFLOW 4 METABOLOMICS

2-Preprocessing

- [xcms.xcmsSet](#) Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis
- [xcms.group](#) Group peaks together across samples using overlapping m/z bins and calculation of smoothed peak distributions in chromatographic time.
- [xcms.retcor](#) Retention Time Correction using retcor function from xcms R package
- [xcms.fillPeaks](#) Integrate the signal in the region of that peak group not represented and create a new peak
- [xcms.diffreport](#) A report showing the most statistically significant differences in analyte intensities
- [CAMERA.annotateDiffreport](#) Wrapper function for the

Workflow Canvas | Workflow XCMS

The workflow canvas displays the following tools and their connections:

- xcms.xcmsSet** (leftmost) connects to the first **xcms.group**.
- The first **xcms.group** connects to the first **xcms.retcor**.
- The first **xcms.retcor** connects to the second **xcms.group**.
- The second **xcms.group** connects to the second **xcms.retcor**.
- The second **xcms.retcor** connects to the third **xcms.group**.
- The third **xcms.group** connects to the **xcms.fillPeaks** tool.

A green noodle is being added to the workflow, connecting the output of the second **xcms.group** tool to the input of the second **xcms.retcor** tool. A red box highlights the connection point on the second **xcms.group** tool.

Workflow

Edit a workflow : set or release a parameter

Galaxy / METABO Analyze Data Workflow Shared Data Visualization Admin Help User Using 7.8 MB

Workflow Canvas | Workflow XCMS

Details

xcms.xcmsSet Filtration and Peak Identification using xcmsSet function from xcms R package to preprocess LC/MS data for relative quantification and statistical analysis (Galaxy Tool Version 2.0.1)

Choose your inputs method

Zip file from your history c...

Zip file

Data input 'zip_file' (no_unzip.zip)

Extraction method for peaks detection

matchedFilter

[method] See the help section below

Step size to use for profile

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

4

[fwhm] Only used to calculate the actual sigma

Workflow

Run a workflow

Galaxy / METABO

Analyze Data Workflow Shared Data Visualization Admin Help User

Using 7.8 MB

Tools

search tools

Get Data

WORKFLOW 4 METABOLOMICS

2-Preprocessing

3-Normalisation

4-Quality Control

5-Statistical Analysis

6-Annotation

COMMON TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Statistics

Graph/Display Data

Multiple regression

Workflows

All workflows

Running workflow "LS-MS"

Expand All Collapse

Step 1: [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

Step size to use for profile generation

0.01

Full width at half maximum of matched filtration gaussian model peak

4

Advanced options

show

Maximum number of peaks per extracted ion chromatogram

50

Signal to noise ratio cutoff

3

Number of steps to merge prior to filtration

2

Step 2: [xcms.group](#) (version 2.0.1)

xset RData file

Output dataset 'xsetRData' from step 1

Method to use for grouping

density

Bandwidth

30

Minimum fraction of samples necessary

0.3

History

sacuri

0 bytes

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

Workflow

Run a workflow : HOP!

Galaxy / METABO Analyze Data Workflow Shared Data Visualization Admin Help User Using 7.8 MB

Tools search tools

Get Data

WORKFLOW 4 METABOLOMICS

- [2-Preprocessing](#)
- [3-Normalisation](#)
- [4-Quality Control](#)
- [5-Statistical Analysis](#)
- [6-Annotation](#)

COMMON TOOLS

- [Text Manipulation](#)
- [Filter and Sort](#)
- [Join, Subtract and Group](#)
- [Statistics](#)
- [Graph/Display Data](#)
- [Multiple regression](#)

Workflows

- [All workflows](#)

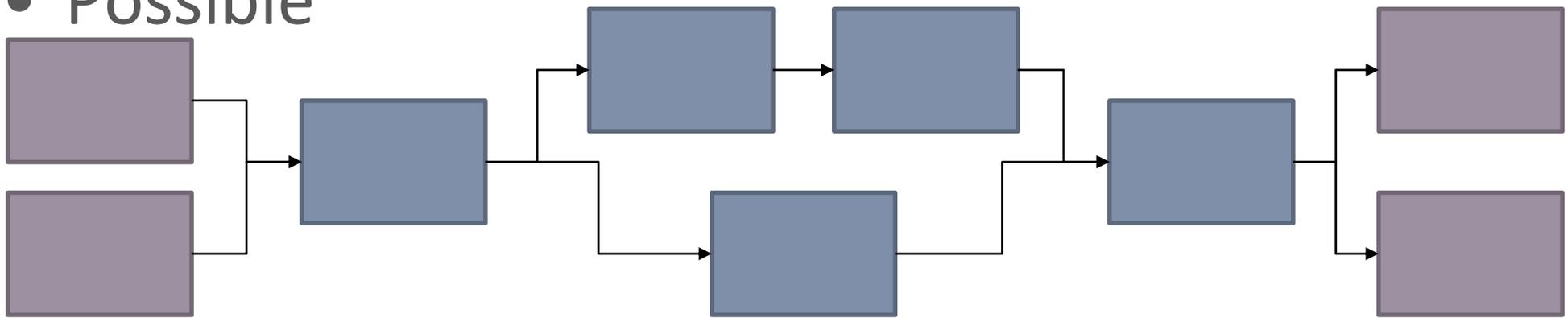
History sacuri 0 bytes

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

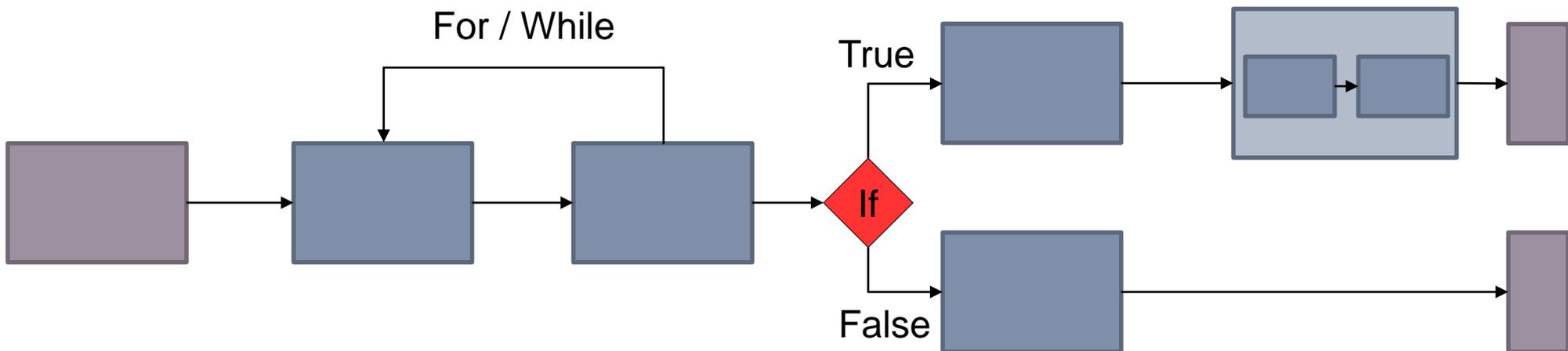
Successfully ran workflow "Workflow XCMS". The following datasets have been added to the queue:

- 1: xset.RData
- 2: sampleMetadata.tsv
- 3: xset.TICs_raw.pdf
- 4: xset.log.txt
- 5: xset.group.RData
- 6: xset.group.Rplots.pdf
- 7: xset.group.log.txt
- 8: xset.group.retcor.RData
- 9: xset.group.retcor.TICs_corrected.pdf
- 10: xset.group.retcor.log.txt
- 11: xset.group.retcor.group.RData
- 12: xset.group.retcor.group.Rplots.pdf
- 13: xset.group.retcor.group.log.txt
- 14: xset.group.retcor.group.retcor.RData
- 15: xset.group.retcor.group.retcor.TICs_corrected.pdf
- 16: xset.group.retcor.group.retcor.log.txt
- 17: xset.group.retcor.group.retcor.group.RData
- 18: xset.group.retcor.group.retcor.group.Rplots.pdf
- 19: xset.group.retcor.group.retcor.group.log.txt
- 20: xset.group.retcor.group.retcor.group.fillPeaks.RData
- 21: xset.group.retcor.group.retcor.group.fillpeaks.log.txt
- 22: xset.group.retcor.group.retcor.group.fillPeaks.annotateDiffreport.variableMetadata.tsv
- 23: xset.group.retcor.group.retcor.group.fillPeaks.annotateDiffreport.dataMatrix.tsv
- 24: xset.group.retcor.group.retcor.group.fillPeaks.annotateDiffreport.zip
- 25: xset.group.retcor.group.retcor.group.fillPeaks.annotateDiffreport.log.txt

- 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

Datasets

Galaxy / METABO Analyze Data Workflow Shared Data Visualization Help User Using 47.0 MB

Tools ↑

search tools ×

Get Data

WORKFLOW 4 METABOLOMICS

- 2-Preprocessing
- 3-Normalisation
- 4-Quality Control
- 5-Statistical Analysis
- 6-Annotation

COMMON TOOLS

- Text Manipulation
- Filter and Sort
- Join, Subtract and Group
- Statistics
- Graph/Display Data
- Multiple regression

Workflows

- All workflows

Saved Histories

search history names and tags Q

[Advanced Search](#)

<input type="checkbox"/>	Name	Datasets	Tags	Sharing	Size on Disk	Created	Last Updated	Status
<input type="checkbox"/>	Preprocessing ▼	8	1	0 Tags	45.6 MB	~18 hours ago	~less than ago	current history
<input type="checkbox"/>	PRC ▼			0 Tags	0 bytes	~2 days ago	~2 minutes ago	
<input type="checkbox"/>	test ▼	1		0 Tags	4.0 KB	Apr 28, 2014	~4 minutes ago	
<input type="checkbox"/>	After_Preprocessing ▼	3		0 Tags	1.4 MB	~37 minutes ago	~7 minutes ago	
<input type="checkbox"/>	Unnamed history ▼			0 Tags	0 bytes	Apr 28, 2014	Apr 28, 2014	

For 0 selected histories: Rename Delete Delete Permanently Undelete

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

Switch

View

Share or Publish

Rename

Delete

Delete Permanently

History ↺ ⚙

Preprocessing 45.6 MB Q ✓ 🗑 💬

8: 👁 ✎ ✕
[xset.group.retcor.group.fillPeaks.diffreport.tsv.tabular](#)

7: 👁 ✎ ✕
[xset.group.retcor.group.fillPeaks.diffreport.RData.rdata](#)

6: 👁 ✎ ✕
[xset.group.retcor.group.fillPeaks.diffreport.data_matrix.tsv.tabular](#)

5: 👁 ✎ ✕
[bio_vs_blank_box/050.png](#)

4: 👁 ✎ ✕
[xset.group.retcor.group.fillPeaks.annotateDiffreport.tsv.tabular](#)

3: 👁 ✎ ✕
[xset.group.retcor.group.fillPeaks.annotateDiffreport.Rdata.rdata](#)

2: 👁 ✎ ✕
[xset.group.retcor.group.fillPeaks.annotateDiffreport.data_matrix.tsv.tabular](#)

1: 👁 ✎ ✕
[xset.group.RData.rdata](#)

Share

Workflow

Your workflows

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

Name	# of Steps
complete_workflow_RFM	17

Workflow you by others

No workflows ha

Other opt

Configure you

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

Share

Mode

Share or Publish Workflow 'complete_workflow_RFMF'

Make Workflow Accessible via Link and Publish It

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

Make Workflow Accessible via Link



Restricted community

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

Make Workflow Accessible and Publish



All the Galaxy server users

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

Share Workflow with Individual Users

You have not shared this workflow with any users.

Share with a user



Designated community (login@sb-roscoff.fr)

[Back to Workflows List](#)

- 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 **Import history** About this History

TP1 xcms sacuri

65.4 MB

search datasets

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

Author
mmonsoor

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

Rating
Community (0 ratings, 0.0 average)
Yours

Tags
Community: none

Histories

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

Your workflows

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

Name	# of Steps
complete_workflow_RFMF	17

Workflows shared with you by others

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

Other

Configure View Run **Copy** Remove



Level 5

- Share of tools and descriptions in the ToolShed

Level 4



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



Level 3

- Launch autonomously tools
- Use workflow more or less presetted



Level 2

- Use presetted workflow



Level 1

- Share his data to colleagues Level 2-5

END

BONUS

How are tools born?

BONUS

- How to import a tool in Galaxy?

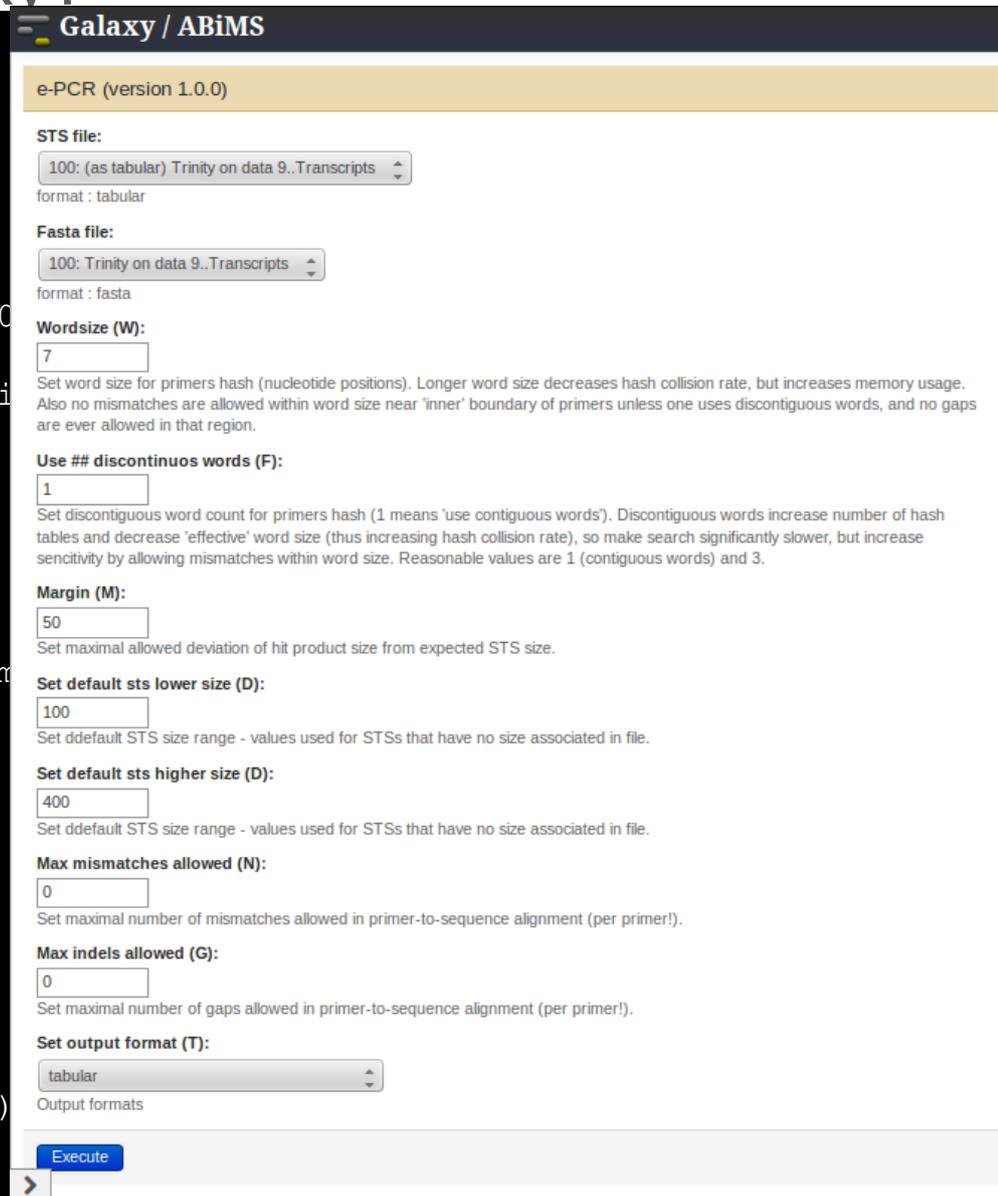


• How to import a tool in Galaxy?

```

[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?

```

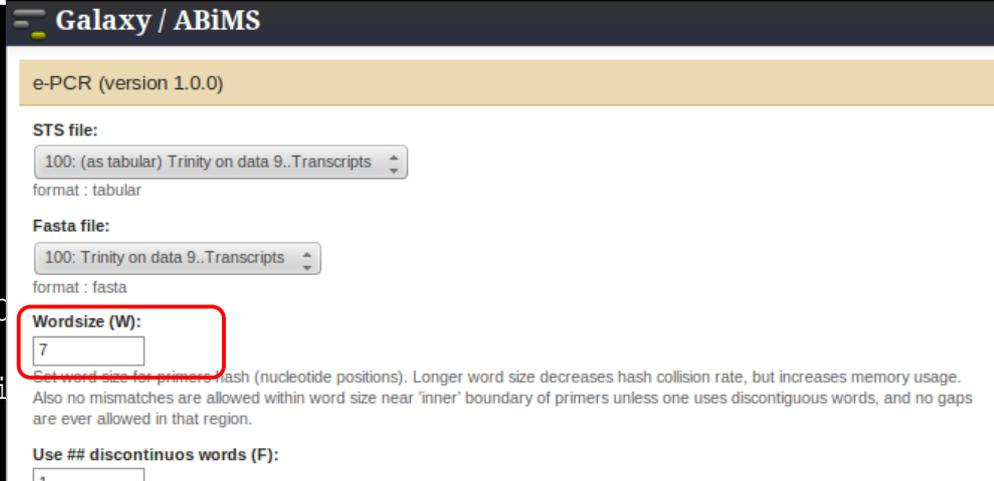
<tool id="abims_epcr" name="e-PCR">
>   <!-- author : lecorguille@sb-roscoff.fr -->
>   <!-- date : 11-05-12 -->
>   <description>e-PCR parses stsfile in unists format, then reads nucleotide sequence data in FASTA format from files listed in commandline if
any, or from stdin otherwise. For input sequences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