INTRODUCTION / PROBLEMATIC
library(xcms)
loaddata()
polar<-"Pos"

noise=250000
xset <- xcmsSet(cdffiles, ppm=ppm, mzdif=mzWid, peakwidth=peakwidth, noise=noise, snthresh=sn, method="centWave", fitgauss=TRUE, nSlaves=8)
xset2<-retrcor(xset, method="obiwarp", plottype="deviation")
dev.copy2pdf(device = 2, file = paste(pathResult, "/Ret_Cor-Graph",expe,"",
polar,"",
pdf",sep=""), paper="a4", height=9, width=14)
xset3<-group(xset2, minfrac = 0.2, bw = bw, minsamp = 1, mzWid = mzWid, max = 50, sleep = 0)
xset5<-fillPeaks(xset3)

# rapport final avec statistiques de différences entre les deux classes
reporttab <- diffreport(xset5, filebase =paste(pathResult,"/Rapport_",expe,"",
polar, sep=""), mzdec=4, eicmax=5000, metlin = metlin, classeic=1)

#écriture du fichier Excel
dir.create(paste(pathResult,"/Rapport_",expe,"",
polar,"_diffreport/", sep=""), showWarnings = FALSE)
write.table(reporttab,paste(pathResult,"/Rapport_",expe,"",
polar,"_diffreport/resultat_",expe,"",
polar,".xls", sep=""), sep="\t")

library(CAMERA)

annotate(xsg,pval=0.05, 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)
diffreport<-getPeakList(an)

diffreport <- annotateDiffreport(xsg,pval=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(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)

pc3<-PCA(t(matapc), axes=c(1,2))

pc3<-PCA(t(matapc), axes=c(1,3))

pc3<-PCA(t(matapc), axes=c(2,3))

pc4<-PCA(t(matapclog2))

# -- 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(ctt)
# -- Normalization: logratio --
if (normalization) {

data=t(scale(t(data)))
}
library(xcms)
loaddata()
polar="<"Pos"

noise=250000
xset <- xcmsSet(cdfiles, ppm=ppm, mzdif=mzwid, peakwidth=peakwidth, noise=noise, snthresh=snth, method="centWave", fitgauss=TRUE, nSlaves=8)
xset2<-retcor(xset, method="obiwarp"
dev.copy2pdf(device = 2, file = paste(paste(unique(cdfiles), c("_")), "Rapport_final.pdf"), width=14)
xset3<-group(xset2, minfrac = 0.2, bkgd=TRUE)
xset5<-fillPeaks(xset3)

# rapport final avec statistiques de diffreport
reporttab <- diffreport(xset5, filebase="Rapport_final"

# écriture du fichier Excel
dir.create(paste(pathResult,"/Rapport_FINAL.xlsx",sep=""), recursive=TRUE)
write.table(reporttab, paste(pathResult,"/Rapport_FINAL.xlsx",sep=""), row.names=FALSE, sep="\t")

library(CAMERA)

# annotation version rapide?
an<-annotate(xsg,pval=0.05, nSlaves=8, ppm=15, mzabs=0.015, quick=FALSE, polarity=polarity)
diffreport<-getPeaklist(an)

diffreport <- annotateDiffreport(xset5, diffreport, withpval=FALSE, withzscore=FALSE)
write.table(cbind(reporttab,diffreport), paste(pathResult,"/Rapport_FINAL.xlsx",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 =500)
braplot(resPCA$eig$per, xlab="Component", ylab="Percentage of variance")
dev.off()

png("eigenvalue.png", width =800, height =500)
braplot(resPCA$eig$eig, xlab="Component", ylab="Eigenvalues")
dev.off()

library(cct)

# -- Normalization: logratio --
if (normalization) {
  data=t(scale(t(data)))
}
Select your level:
Level 1

« I want to know which metabolites are different between my two conditions »
Level 2

« I want to know which metabolites are different between my two conditions »
« I know that I need to import files then annotate my peaks and finally draw a PCA »
Level 3

« I want to launch XCMS tools to preprocess and align
Then use CAMERA to annot my adducts.
And finally use a correlation tool to reduce my dataset before the PCA.
« I want 1TB for my project. I will launch some R scripts through SSH on the cluster using XCMS and CAMERA in multi-thread mode.
Next I want to query HMDB.
Except that, I will manage :P”
Level 5

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

Comments? »
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
Introduction

- Command line tools
  - represent almost the majority of scientific tools
  - good parameters completeness
  - can be executed on high performance computers
  - g33ks love it, since automatable, workflowsable, ...
- minimum linux knowledge is required
- cruel lack of ergonomics
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:
30

Advanced options:
hide

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


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

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

RNA-Seq Analysis Tools

COST

DIFFICULTY OF USE/LEARNING CURVE

Size of dot indicates flexibility/power
Why Galaxy?

- Accessibility
- Reproductibility
- Transparency
Introduction / Galaxy

```
[login@n0 ~] $ cdprojet
[login@n0 login]$ cd 13-07-29-panda/tmp/mapping
[login@n0 mapping]$ cat tophat.qsub
#!/bin/bash
#$ -S /bin/bash
#$ -M login@sb-roscoff.fr
#$ -m bea
#$ -V
#$ -cwd
#$ -o qsub.out
#$ -e qsub.err
tophat2 panda_v121029 ../input/IllR1-1.fq ../input/IllR1-2.fq
  ../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 533869 ("tophat.qsub") has been submitted
```
[login@n0 :-]$ cdprojet
[login@n0 login]$ cd 07-29-panda/tmp/mapping
[login@n0 mapping]$ cat tophat.qsub
#!/bin/bash
#$ -S /bin/bash
#$ -M login@sb-roscoff.fr
#$ -m bce
#$ -V
#$ -cwd
#$ -o qsub.out
#$ -e qsub.err
tophat2 panda_v121029 ../input/IllR1-1.fq ../input/IllR1-2.fq
-GTF ../input/panda_v121029.gtf --b2-sensitive --r 100
--num-threads 8
[login@n0 mapping]$ qsub -q long.q -pe thread 8 tophat.qsub
Your job 533869 ("tophat.qsub") has been submitted
Introduction / Galaxy

[lecorguille@n0 ~]$ e-PCR --help

e-PCR: invalid option --
usage: [-hV] [posix-options] stsfile [fasta ...]
[compat-options]
where posix-options are:

-m ## Margin (default 50)
-w ## Wordsize (default 7)
-n ## Max mismatches allowed (default 0)
-g ## Max indels allowed (default 0)
-f ## Use ## discontiguos words, slow if
     ##>1
-o ## Set output file
-t ## Set output format:
     1 - classic, range (pos1..pos2)
     2 - classic, midpoint
     3 - tabular
     4 - tabular with alignment in
     comments (slow)
-d##-## Set default size range
     (default 100-350)
-p +- Turn hits postprocess on/off
-v ## Verbosity flags
-a a|f Use presize alignmens (only if
     gaps>0), slow
     a - Allways or f - as Fallback
-x +- Use 5'-end lowercase masking of
     primers (default -)
-u +- Uppercase all primers (default -)

[...]
Introduction / Galaxy

```
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 indicies 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

**Choose your inputs method**
- Zip file from your history containing your chromatograms

**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 fine grid

**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 uses this argument
Introduction / Connection

http://workflow4metabolomics.org
Galaxy interface

Batch_correction (version 2.0.0)

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

Sample metadata file:
3: sampleMetadata.tsv
must contain at least the following three columns: 'batch' + 'InjectionOrder' + 'sampleType'

Variable metadata file:
16: xset.group.retro.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 Kloot algorithm; when using loess, you can choose to use pools or samples to model batch effect.

Factor of interest:
batch

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

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

Execute

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

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

Galaxy interface

Batch_correction (version 2.0.0)

Data Matrix file:
17: xset.group.rector.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.rector.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 Koot 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.

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

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

**Tool list**

- **Batch_correction** (version 2.0.0)
  - Data Matrix file: ...
  - Sample metadata file: ...
  - Variable metadata file: ...

**Type of regression model:**
- Linear

**Factor of interest:**
- Batch

**Level of details for plots:**
- Basic

Authors:
- Jean-François Martin - PF MetaToul-AXIOM; INRA; MetaboHUB (for original version of this tool and overall development of the R script)
- Elodie Thenent - LIST/LADIS; CEA; MetaboHUB (for R script and wrapper concerning "all loess pool" and "all loess sample" methods)

Contributors:
- Melanie Petere - PFEM: INRA; MetaboHUB (for R wrapper and R script improvement)
- Etienne Thenent - LIST/LADIS: CEA; MetaboHUB (for R script and wrapper concerning "all loess pool" and "all loess sample" methods)
Galaxy interface

History

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 following three 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 Koot 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.

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

Contributors
Melanie Petera - PFEM ; INRA ; MetaboHUB (for R wrapper and R script improvement)
Etienne Thevenot - LIST/LADIS ; CEA ; MetaboHUB (for R script and wrapper concerning "all loess pool" and "all loess sample" methods)
GET HELP
Workflow4Metabolomics

version 2.1


Help and support: support@workflow4metabolomics.org
version 2.1


Help and support: support@workflow4metabolomics.org

Changelog
Tutorials
Past events
Get help

Workflow4metabolomics

W4M HowTo

<table>
<thead>
<tr>
<th>Description</th>
<th>PDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Import datasets &lt; 2Gb</td>
<td>Download</td>
</tr>
<tr>
<td>Import datasets &gt; 2Gb</td>
<td>Download</td>
</tr>
<tr>
<td>Build And Configure A Workflow</td>
<td>Download</td>
</tr>
<tr>
<td>Share Histories And Workflows</td>
<td>Download</td>
</tr>
<tr>
<td>Format Data For Postprocessing</td>
<td>Download</td>
</tr>
<tr>
<td>Perform Xcms Preprocessing</td>
<td>Download</td>
</tr>
<tr>
<td>Perform Drift And Batch Correction</td>
<td>Download</td>
</tr>
<tr>
<td>Perform Univariate Analyzes</td>
<td>Download</td>
</tr>
<tr>
<td>Perform Multivariate Analyzes</td>
<td>Download</td>
</tr>
<tr>
<td>Perform LCMS Annotations</td>
<td>Download</td>
</tr>
<tr>
<td>Use NIST</td>
<td>Download</td>
</tr>
</tbody>
</table>

For requests, please fill in the webform here

workflow4metabolimcs.org
DATA IMPORT
DATA IMPORT

< 2 GO AND > 2 GO
The "raw" data within Galaxy in "Dataset Collection"

DATA IMPORT
Data import

The "raw" data within Galaxy in "Dataset Collection"
Data import

The "raw" data within Galaxy in "Dataset Collection"
Data import

The "raw" data within Galaxy in "Dataset Collection"
Data import

The "raw" data within Galaxy in "Dataset Collection"
Data import

The "raw" data within Galaxy in "Dataset Collection"
The "raw" data within Galaxy in "Dataset Collection"
Data import

The "raw" data within Galaxy in "Dataset Collection"
Data import

The "raw" data within Galaxy in "Dataset Collection"
Data import

The "raw" data within Galaxy in "Dataset Collection"
Other files without a Dataset Collection

DATA IMPORT
Data import

Other files without a Dataset Collection
Data import

Other files without a Dataset Collection
Data import

Other files without a Dataset Collection
Data import

Other files without a Dataset Collection
Paste or small text inputs

DATA IMPORT
Data import

Paste or small text inputs
Data import

Paste or small text inputs
Data import

Paste or small text inputs
Data import

Paste or small text inputs
Data import

Paste or small text inputs
Exercise
DATA IMPORT
Data import

• Exercise

1. Fetch those files
   
   goo.gl/KWIEjL
   
   goo.gl/Yyqmnh

2. Unzip the zip file sacuri.zip

3. Upload the individual files within Galaxy
   
   - As dataset collection for the samples
   
   - As standard dataset for the sampleMetadata file
TOOLS
Tools - form

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
[w/hm] Only used to calculate the actual sigma

Advanced options:
hide:

Execute

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

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.
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:
- 30

Advanced options:
- hide

Authors
Colin A. Smith (csmith@scripps.edu), Rolf Tautenhahn (rtautenh@gmail.com), Steffen Neumann (sneumann@ipb-halle.de), Paul Benton (paul.benton08@imperial.ac.uk) and Christopher Conley (cconley@ucdavis.edu)


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.
Tools - form

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
[w/h] Only used to calculate the actual sigma

Advanced options:
hide

Authors Colin A. Smith csmith@scripps.edu, Rolf Tautenhahn rtautenh@gmail.com, Steffen Neumann sneumann@ipb-halle.de, Paul Benton p.ba.benton08@imperial.ac.uk and Christopher Conley c.conley@ucdavis.edu

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.
Tools can have some advanced options

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

[w/hm] 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
Tools can have some advanced options.

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

[wmm] 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
Tools can have some advanced options

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.
Tools - Job status

- 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
Tools - Job status

- 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 😞
Tools - Job status

- Status

Job is finished

It's status is OK

but warnings or errors can be hidden behind. Ah hum!
Tools - Job status

- 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.
Tools - Job status

- 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
Tools - Exercise

• Aim of this Exercise
  – Import data into Galaxy into a new history
  – Execute and chain example of little Galaxy friendly tools together.
Tools - Exercise

• Create a New history

• Fetch these two tabular files
  – Link1: http://tinyurl.com/w4mddata2
  – Link2: http://tinyurl.com/w4mddata3
  – Tabular files (data separated by tab delimiters)
    • VariableMetadata.tsv
    • DataMatrix.tsv

• Check their contents and datatypes through Galaxy.
Tools - Exercise

• First tool:
  – Search for the tool « Compute an expression on every row » in the toolbar)
  – Calculate the average for each metabolite by sample
Tools - Exercise

• First tool:
  – Search for the tool « Compute an expression on every row » in the toolbar)
  – Calculate the average for each metabolite by sample:

• Set the parameters
  – Add expression: \((c2+c3+c4+c5+c6+c7)/6\)
  – as a new column to: Choose the **DataMatrix.tsv**
  – Skip a header line: **Yes**
  – The new column name: **average**
Tools - Exercise

• Second tool:
  – Search for the tool « Cut columns from a table » in the toolbar)
  – Keep only columns 1 and 8:
Tools - Exercise

• Second tool:
  – Search for the tool « Cut columns from a table » in the toolbar)
  – Keep only columns 1 and 8:

• Set the parameters
  – Cut columns: \texttt{c1,c8}
  – Delimited by: \texttt{Tab}
  – From?: \texttt{Compute on Data 1}
• Third tool:
  – Search for the tool « Join two Datasets side by side on a specified field » in the toolbar)
  – Join the two tab files by the metabolite name:
Tools - Exercise

• Third tool:
  – Search for the tool « Join two Datasets side by side on a specified field » in the toolbar)
  – Join the two tab files by the metabolite name:
    • Set the parameters
      – Join: `variableMetadata.tsv`
      – Using column: **column 1**
      – with: **Cut on Data 3**
      – And column: **column 1**
      – Keep the header lines: **Yes**
Part II

TOOLS
Tools – Handle errors

Abims: Analyses and Bioinformatics for Marine Science

Station Biologique Roscoff

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.
Tools – Handle errors

- 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 de novo RNASeq expression data (without gtf input).
- 26-04-13: RNASeq: sam2counts is now available to count the reads coverage by transcript. It's also a requirement for DESeq de novo.
- 26-04-13: Metabolomic: Workflow Metabolomics by ABiMS, updated to version 2.0.0 (2013_04_16)

AbiMS
Analyses and Bioinformatics for Marine Science

Station Biologique Roscoff

For any question or request for tools or account, send an email at support@abiims AT st-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.
Tools – Handle errors

Dataset generation errors

Dataset: mzXML_copper_stress.group.recor.group.fillPeaks.annotateDiffreport.data_matrix.tsv.anova.pvalue.tabular

Tool execution generated the following error message:

Fatal error: Exit code 10 ()
ERROR: There is a problem with the group of condition (presence of NA). You may need to use change the mode (column/row) Current groups: NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA

View or report this error

Sent to the support team
HISTORY
Both inputs and outputs

History panel

Batch_correction (version 2.0.0)

Data Matrix file:
17: xset.group.retor.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.retor.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 Koot algorithm; when using loess, you can choose to use pools or samples to model batch effect.

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

Level of details for plots:
basic

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

Execute

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

Contributors
Melanie Patera - 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

19: xset.group.retor.group.fillPeaks.annotate.variableMetadata.tsv (diffreport)
18: xset.group.retor.group.fillPeaks.annotate.negative.Rdata
17: xset.group.retor.group.fillPeaks.annotate.dataMatrix.tsv
16: xset.group.retor.group.fillPeaks.annotate.variableMetadata.tsv
15: xset.group.retor.group.fillPeaks.RData
14: xset.group.retor.group.Rplots.pdf
13: xset.group.retor.group.RData
12: xset.group.retor.BPCs_corrected.pdf
11:
History panel

renaming and annotation
History panel

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

- Saved Histories
  - Name: Sacuri
    - Datasets: 19
    - Tags: 2
    - Sharing: current history
    - Size on Disk: 289.7 MB
    - Created: Sep 02, 2015
    - Last Updated: ~3 days ago
  - Name: Sacuri Lib
    - Datasets: 30
    - Tags: 0
    - Sharing: current history
    - Size on Disk: 17.3 MB
    - Created: May 14, 2014
    - Last Updated: Sep 02, 2015
  - Name: Cooper Stress Lib
    - Datasets: 19
    - Tags: 0
    - Sharing: current history
    - Size on Disk: 7.8 MB
    - Created: May 13, 2014
    - Last Updated: Sep 02, 2015

For 0 selected histories: Rename, Delete, Delete Permanently, Undelete

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

Saved histories: Switch histories

Saved Histories

<table>
<thead>
<tr>
<th>Name</th>
<th>Datasets</th>
<th>Tags</th>
<th>Sharing</th>
<th>Size on Disk</th>
<th>Created</th>
<th>Last Updated</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sacuri</td>
<td>19</td>
<td>2</td>
<td></td>
<td>289.7 MB</td>
<td>Sep 02, 2015</td>
<td>~3 days ago</td>
<td>current history</td>
</tr>
<tr>
<td>Sacuri Lib</td>
<td>30</td>
<td>0</td>
<td></td>
<td>17.3 MB</td>
<td>May 14, 2014</td>
<td>Sep 02, 2015</td>
<td></td>
</tr>
<tr>
<td>Cooper Stress Lib</td>
<td>19</td>
<td>0</td>
<td></td>
<td>7.8 MB</td>
<td>May 13, 2014</td>
<td>Sep 02, 2015</td>
<td></td>
</tr>
</tbody>
</table>

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.
Both inputs and outputs

Dataset

Batch_correction (version 2.0.0)

Data Matrix file:
17: xset.group.rector.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.rector.group.fillPeaks.annotate.variableMetadata.tsv

Type of regression model:
linear

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

Factor of interest:
batch

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

Level of details for plots:

basic

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

Execute

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

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

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

---

<table>
<thead>
<tr>
<th>name</th>
<th>Blanc15</th>
<th>Blanc09</th>
<th>Blanc12</th>
<th>Blanc06</th>
<th>Blanc17</th>
</tr>
</thead>
<tbody>
<tr>
<td>M100T293</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M100T313</td>
<td>14737.3434458556</td>
<td>70497.1552979614</td>
<td>29398.2144370894</td>
<td>67121.4139064063</td>
<td>4813.6156029072</td>
</tr>
<tr>
<td>M100T318</td>
<td>1396.8756293629</td>
<td>6403.15709537553</td>
<td>3483.72951135393</td>
<td>6365.40360216962</td>
<td>1256.93109196603</td>
</tr>
<tr>
<td>M100T415</td>
<td>4103.24769663852</td>
<td>6007.46666614238</td>
<td>5866.56418559669</td>
<td>1817.6141143643</td>
<td>1394.93820578164</td>
</tr>
<tr>
<td>M101T308</td>
<td>0</td>
<td>0</td>
<td>1354.41420127279</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M101T63</td>
<td>927.737622943452</td>
<td>46094.2217735992</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M102T348</td>
<td>7957.3450177005</td>
<td>17013.5105876707</td>
<td>11971.525671295</td>
<td>24170.4565255728</td>
<td>5043.8787274869</td>
</tr>
<tr>
<td>M102T379</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M102T59</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M103T1003</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M101T1012</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M103T152</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M103T45</td>
<td>26872.0684617598</td>
<td>51858.8307837991</td>
<td>1335.1123434995</td>
<td>15915.78182274</td>
<td>13919.992158649</td>
</tr>
<tr>
<td>M103T50</td>
<td>26002.5859959567</td>
<td>100654.744002913</td>
<td>6153.3602236106</td>
<td>82775.7546772999</td>
<td>17210.019273697</td>
</tr>
<tr>
<td>M103T63</td>
<td>26143.9616699194</td>
<td>66858.1859951143</td>
<td>45073.1929194364</td>
<td>65013.382999986</td>
<td>15636.7248046881</td>
</tr>
<tr>
<td>M105T50</td>
<td>37864.7519066144</td>
<td>172016.779677334</td>
<td>104572.584541783</td>
<td>186717.19434361</td>
<td>0</td>
</tr>
<tr>
<td>M105T57</td>
<td>22972.800197443</td>
<td>175038.512313167</td>
<td>125815.0563963625</td>
<td>232961.533116152</td>
<td>0</td>
</tr>
<tr>
<td>M107T348</td>
<td>39111.6763561207</td>
<td>111455.640695432</td>
<td>63522.713826157</td>
<td>94976.4961542975</td>
<td>13711.7442593179</td>
</tr>
<tr>
<td>M107T379</td>
<td>0</td>
<td>66740.2961367195</td>
<td>109739.70423633</td>
<td>113615.95868816</td>
<td>2009.0318237782</td>
</tr>
<tr>
<td>M108T336</td>
<td>2875.5310622199</td>
<td>65289.224125976</td>
<td>59553.329442304</td>
<td>162505.87408746</td>
<td>13579.44</td>
</tr>
<tr>
<td>M108T379</td>
<td>507.94377920085</td>
<td>52885.4503151381</td>
<td>6633.9879785059</td>
<td>140119.471455469</td>
<td>519.706044359327</td>
</tr>
<tr>
<td>M109T294</td>
<td>860.607251283829</td>
<td>84109.0329020055</td>
<td>30997.4207160858</td>
<td>137467.156183397</td>
<td>506.70818482066</td>
</tr>
<tr>
<td>M109T51</td>
<td>3163.7089178629</td>
<td>14569.0031063036</td>
<td>23613.85464709</td>
<td>264370.764749398</td>
<td>830.56049530485</td>
</tr>
<tr>
<td>M110T294</td>
<td>1324.5742670715</td>
<td>15356.52852816</td>
<td>55063.9665250007</td>
<td>86087.9042967472</td>
<td>0</td>
</tr>
<tr>
<td>M110T313</td>
<td>0</td>
<td>11985.72034715</td>
<td>72124.8941731575</td>
<td>71361.001601532</td>
<td>2461.35172321908</td>
</tr>
<tr>
<td>M110T55</td>
<td>2572.37712822047</td>
<td>13158.478684824</td>
<td>2853.7133660608</td>
<td>16036.85592733</td>
<td>0</td>
</tr>
<tr>
<td>M111T273</td>
<td>16799.0249707129</td>
<td>130599.8238562573</td>
<td>42009.1631329325</td>
<td>78842.833811576</td>
<td>3967.62968038254</td>
</tr>
<tr>
<td>M111T338</td>
<td>1946.9046446183</td>
<td>47811.3018666406</td>
<td>7387.0006640949</td>
<td>25405.690116671</td>
<td>466.6908431769</td>
</tr>
<tr>
<td>M111T51</td>
<td>3395.04200209943</td>
<td>5670.3821937001</td>
<td>9834.6297351838</td>
<td>25928.8676690728</td>
<td>7642.3468529675</td>
</tr>
<tr>
<td>M111T58</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

---

The image shows a screenshot of a web interface for managing datasets, with various tools and options for data analysis, preprocessing, and quality control. The dataset appears to be related to metabolomics, with a focus on numerical data and statistical analysis. The interface includes features for uploading and exporting files, as well as tools for converting formats, normalizing data, and performing statistical analyses. The history section on the right shows recent actions, including the annotation of a variable metadata file and data matrix.
Dataset

Renaming and annotation
Change the Datatype of the Dataset

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.
Dataset

Graphics
Dataset

Re-run a job
Cleanup

DATASET
Delete a dataset
The dataset isn’t really deleted
It’s in the Trash

CAMERA.annotate (version 2.0.0)

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

Convert retention time (seconds) into minutes:

Convert the columns rmed, rmin and rmax 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
[mintrac]

General ppm error:
5
[ppm]

This dataset has been deleted
Undelete it
Permanently remove it from disk
"Empty Trash" : to free up disk space
Workflow

• A workflow is a sequence of tool operations and parameters

• Can match the experiment protocol

• A workflow is built to be replayed (more or less strict)
Workflow

- Our workflow
Our workflow with Galaxy
From a history

Workflow

version 2.1


Changelog

Tutorials

Past events
From a history
From a history

Workflow

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

Upload File
This tool cannot be used in workflows

xcms.xcmsSet
 Include "xcms.xcmsSet" in workflow

xlcm.group
 Include "xcms.group" in workflow

xcms.retcor
 Include "xcms.retcor" in workflow

History items created

1: sacuri.zip
Treat as input dataset

2: xset.RData

3: sampleMetadata.tsv

4: xset.TICs_raw.pdf

5: xset.BPCs_raw.pdf

6: xset.log.txt

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

13: xset.group.retcor.group.RData

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

15: xset.group.retcor.group.annote.RData

16: xset.group.retcor.group.annote.dataMatrix.tsv

17: xset.group.retcor.group.annote.variableMetadata.tsv

18: xset.group.retcor.group.annote.variableMetadata.tsv

19: xset.group.retcor.group.annote.annotate_negative.RData

Sacuri
19 shown
289.7 MB
# Workflow

## Workflow manager

### Your workflows

<table>
<thead>
<tr>
<th>Name</th>
<th># of Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>LS-MS</td>
<td>7</td>
</tr>
<tr>
<td>Copy of 'gigaXml' shared by '<a href="mailto:ethevenot@sb-roscoff.fr">ethevenot@sb-roscoff.fr</a>'</td>
<td>13</td>
</tr>
<tr>
<td>Workflow LC/MS</td>
<td>6</td>
</tr>
<tr>
<td>Community</td>
<td>10</td>
</tr>
<tr>
<td>Full_workflow</td>
<td>19</td>
</tr>
<tr>
<td>Workflow XCMS</td>
<td>8</td>
</tr>
</tbody>
</table>

### Workflows shared with you by others

<table>
<thead>
<tr>
<th>Name</th>
<th>Owner</th>
<th># of Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>demo_workflow_06_annotation</td>
<td>mlandisb-roscoff.fr</td>
<td>6</td>
</tr>
<tr>
<td>cohort</td>
<td>ethevenotsb-roscoff.fr</td>
<td>15</td>
</tr>
<tr>
<td>gigaRaw-convert</td>
<td>ethevenotsb-roscoff.fr</td>
<td>1</td>
</tr>
</tbody>
</table>

### Other options

- Configure your workflow menu
### Your workflows

<table>
<thead>
<tr>
<th>Name</th>
<th># of Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>8</td>
</tr>
</tbody>
</table>

### Workflows shared with you by others

<table>
<thead>
<tr>
<th>Name</th>
<th>Owner</th>
<th># of Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>demo_workflow_06_annotation</td>
<td><a href="mailto:mlandi@sb-roscoff.fr">mlandi@sb-roscoff.fr</a></td>
<td>6</td>
</tr>
<tr>
<td>cohort</td>
<td><a href="mailto:ethevenot@sb-roscoff.fr">ethevenot@sb-roscoff.fr</a></td>
<td>15</td>
</tr>
<tr>
<td>gigaRaw-convert</td>
<td><a href="mailto:ethevenot@sb-roscoff.fr">ethevenot@sb-roscoff.fr</a></td>
<td>1</td>
</tr>
</tbody>
</table>

### Other options

- Configure your workflow menu
Edit a workflow: drag and drop
Workflow

Edit a workflow: drag and drop
Edit a workflow: delete a noodle
Edit a workflow: add a tool

**Workflow Canvas | Workflow XCMS**

- **xcms.retcor** Retention Time Correction using retcor function from xcms R package
- **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.

**Workflow**
Add the tool
Select random lines from a file
Edit a workflow: add a noodle
Edit a workflow: set or release a parameter
Edit a workflow: set or release a parameter

1 - Select Select random lines from a file
2 - Set Randomly select as 100
Run a workflow

Running workflow "LS-MS"

Step 1: xsms xsmsSet (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 Lf

Full width at half maximum of matched filtration gaussian model peak
4 Lf

Advanced options
show

Maximum number of peaks per extracted ion chromatogram
50

Signal to noise ratio cutoff
3 Lf

Number of steps to merge prior to filtration
2 Lf

Step 2: xsms_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 Lf

This history is empty. You can load your own data or set data from an external source
Run a workflow: HOP!

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.recor.RData
9: xset.group.recor.TICs_corrected.pdf
10: xset.group.recor.log.txt
11: xset.group.recor.group.RData
12: xset.group.recor.group.Rplots.pdf
13: xset.group.recor.group.log.txt
14: xset.group.recor.group.recor.RData
15: xset.group.recor.group.recor.TICs_corrected.pdf
16: xset.group.recor.group.recor.log.txt
17: xset.group.recor.group.recor.group.RData
18: xset.group.recor.group.recor.group.Rplots.pdf
19: xset.group.recor.group.recor.group.log.txt
20: xset.group.recor.group.recor.group.fillPeaks.RData
21: xset.group.recor.group.recor.group.fillPeaks.log.txt
22: xset.group.recor.group.recor.group.fillPeaks.annotateDiffreport.variableMetadata.tsv
23: xset.group.recor.group.recor.group.fillPeaks.annotateDiffreport.dataMatrix.tsv
24: xset.group.recor.group.recor.group.fillPeaks.annotateDiffreport.zip
25: xset.group.recor.group.recor.group.fillPeaks.annotateDiffreport.log.txt
Workflow

- 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
Datasets

Saved Histories

<table>
<thead>
<tr>
<th>Name</th>
<th>Datasets</th>
<th>Tags</th>
<th>Sharing</th>
<th>Size on Disk</th>
<th>Created</th>
<th>Last Updated</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preprocessing</td>
<td>8</td>
<td>1</td>
<td></td>
<td>45.6 MB</td>
<td>~18 hours ago</td>
<td>~less than ago</td>
<td>current history</td>
</tr>
<tr>
<td>Switch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>View</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share or Publish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delete</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delete Permanently</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After_Preprocessing</td>
<td>3</td>
<td></td>
<td></td>
<td>1.4 MB</td>
<td>~37 minutes ago</td>
<td>~7 minutes ago</td>
<td></td>
</tr>
<tr>
<td>Unnamed history</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.
**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**
  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**
  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@workflow4metabolomics.org)*

*Restricted community*

*All the Galaxy server users*
Share

- Get shared histories

**Individual**

**Public**

Galaxy / METABO

Histories shared with you by others

<table>
<thead>
<tr>
<th>Name</th>
<th>Datasets</th>
<th>Created</th>
<th>Last Updated</th>
<th>Shared by</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmonsoor</td>
<td>6</td>
<td>Apr 28, 2014</td>
<td>~2 days ago</td>
<td><a href="mailto:mmonsoor@sb-roscoff.fr">mmonsoor@sb-roscoff.fr</a></td>
</tr>
</tbody>
</table>

For 0 selected histories: Copy Unshare

Galaxy / METABO

Published Histories

<table>
<thead>
<tr>
<th>Name</th>
<th>Annotation</th>
<th>Owner</th>
<th>Community Tags</th>
<th>Last Updated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preprocessing</td>
<td>mlandi</td>
<td></td>
<td></td>
<td>~14 seconds ago</td>
</tr>
<tr>
<td>TP1 xcms sacuri</td>
<td>mmonsoor</td>
<td></td>
<td></td>
<td>~1 day ago</td>
</tr>
<tr>
<td>TP1 xcms sacuri</td>
<td>jfmartin</td>
<td></td>
<td></td>
<td>Apr 28, 2014</td>
</tr>
</tbody>
</table>
Share

- Get shared workflows

Individual

Workflow

Workflows shared with you by others

Name

Owner

mmensoor@sb-roscoff.fr

# of Steps

7

Public

Published Workflows

Name

Annotation

Owner

Search name, annotation, owner, and tags

Advanced Search

complete_workflow_RFME

Published Workflows

Data Libraries

Published Workflows

Published Visualizations

Published Pages

~17 hours ago
• Import shared

Histories

Workflows
Share

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 2-4