MS data processing
Report creation and Annotations
Yann GUITTON
30/05/2017
The R-package **CAMERA** is a Collection of Algorithms for MEtabolite pRofile Annotation.

Its primary purpose is the annotation and evaluation of LC-MS data. It includes algorithms for annotation of isotope peaks, **adducts and fragments** in peak lists.

Additional methods cluster mass signals that originate from a single metabolite, based on rules for mass differences and peak shape comparison.
xCMS diffreport & CAMERA

CAMERA.annotate = CAMERA::annotateDiffreport

In details:

1- xcms::diffreport : Generates features list, EICs, BoxPlot and statistics
xCMS diffreport & CAMERA

CAMERA.annotate = CAMERA::annotateDiffreport

In details:

1- xcms::diffreport : Generates features list, EICs, BoxPlot and statistics
CAMERA.annotate = CAMERA::annotateDiffreport

In LC-MS ESI Features are usually not alone

Number of features is not equal to number of detected molecules
CAMERA.annotate = CAMERA::annotateDiffreport

In details:
2- CAMERA::xsAnnotate: read xcms object
3- CAMERA::groupFWHM: search co-eluting features RT based
4- CAMERA::findIsotopes: search for isotopic realtion between features (C\textsubscript{12}/C\textsubscript{13})
5- CAMERA::groupCorr : try to improve co-elution separation
6- CAMERA::findAdduct: search for known adducts and fragments [M+Na]^+, [M+H-H\textsubscript{2}O]^+, ....

Non annotated, but low intensity
CAMERA adduct annotation: defining rules

[\text{Green= Annotation OK}]
[\text{Red= No annotation or wrong or adduct not in CAMERA}]

\[ [\text{M+H-NH}_3^+] \]

\[ [\text{M+H-NH}_3\text{-H}_2\text{O}]^+ ? \]
xcms diffreport & CAMERA

CAMERA.annotate = CAMERA::annotateDiffreport

In details:
2- CAMERA::xsAnnotate: read xcms object
3- CAMERA::groupFWHM: search co-eluting features RT based
4- CAMERA::findIsotopes: search for isotopic relation between features \((C_{12}/C_{13})\)
5- CAMERA::groupCorr: try to improve co-elution separation
6- CAMERA::findAdduct: search for known adducts and fragments \([M+Na]^+\), \([M+H-H_2O]^+\), ...

Some times …. Co-elution are not fully resolved, have a look to your data
Additionnal information added to the diffreport by CAMERA
xcms diffreport & CAMERA

CAMERA.annotate = CAMERA::annotateDiffreport

In details:

1- xcms::diffreport : Generates features list, EICs, BoxPlot and statistics

2- CAMERA::xsAnnotate: read xcms object

3- CAMERA::groupFWHM: search co-eluting features RT based

4- CAMERA::findIsotopes: search for isotopic relation between features (C_{12}/C_{13})

5- CAMERA::groupCorr : try to improve co-elution separation

6- CAMERA::findAdduct: search for known adducts and fragments [M+Na]^+, [M+H-H_2O]^+, ....

Many steps = quite a lot of parameters
**Format Conversion**

**Preprocessing**

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

**RData file**

- No rdata.xcms.fillpeaks or rdata dataset available.
- Output file from another function xcms (fillPeaks)

**Group co-eluted peaks based on RT [groupFWHM]**

- Multiplier of the standard deviation
  - 6
  - [sigma]

- Percentage of FWHM width
  - 0.6
  - [perfwhm]

**Annotation general options**

- General ppm error
  - 5
  - [ppm]

- General absolute error in m/z
  - 0.015
  - [mzabs]
CAMERA, annotate function. Returns annotation results (isotope peaks, adducts and fragments) and a diffreport if more than one condition.

**Annotate Isotopes [findIsotopes]**

- **Max. ion charge**
  - 3
  - [maxcharge]

- **Max. number of expected isotopes**
  - 4
  - [maxiso]

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

**Mode**

- All functions
  - [quick] If TRUE, use only groupFWHM and findIsotopes functions. Else if FALSE, use also groupCorr.

**Verifying grouping co-eluted peaks [groupCorr]**

- **groupCorr**: correlation threshold (0..1)
  - 0.75
  - [cor_eic_th]

- **groupCorr**: Method selection for grouping peaks after correlation analysis into pseudospecies
  - hcs
  - [graphMethod]

- **groupCorr**: significant correlation threshold
  - 0.05
  - [nyal]
CAMERA.annotate function. Returns annotation results (isotope peaks, adducts and fragments) and a diffreport if more than one condition. (Galaxy Version 2.2.0)

- **groupCorr**: Use correlation inside samples for peak grouping
- **groupCorr**: Use isotopic relationship for peak grouping
- **groupCorr**: Use correlation across samples for peak grouping

**Annotate Adducts [findAdducts]**

Which polarity mode was used for measuring of the ms sample

- **negative**
- **polarity**

How much peaks will be calculated in every thread using the parallel mode

- **100**
- **[max_peaks]**

Use a personal ruleset file

- **TRUE**
- **[rules]**

User defined ruleset

- **No csv dataset available.**
CAMERA annotate

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

Statistics and results export: [diffreport]

Number of condition

One condition

Export options

Convert retention time (seconds) into minutes

Yes  No

Convert the columns rtmed, rtmin and rtmax into minutes

Number of decimal places for mass values reported in ions’ identifiers.

4

A minimum of 4 decimal places is recommended. Useful to avoid duplicates within identifiers

Number of decimal places for retention time values reported in ions’ identifiers.

0

Useful to avoid duplicates within identifiers

General used intensity value

into

[intval] See the help section below

Resubmit your raw dataset or your zip file

Execute
Output files

- **xset.annotate.variableMetadata.tsv**
  
  For each metabolite (row): the value of the intensity in each sample, fold, anova, mzmed, mzmin, mzmax, rtmed, rtmin, rtmax, npeaks, isotopes, adduct and pcgroup

- **xset.annotate.dataMatrix.tsv**
  
  A tabular file which represents for each metabolite (row), the value of the intensity in each sample (column).

- **xset.annotate.zip**
  
  It contains filebase_eic, filebase_box and filebase.tsv for one condition vs another (Anova analysis).

- **xset.annotate.Rdata rdata.camera.quick or rdata.camera.positive or rdata.camera.negative**
  
  Rdata file, that be used outside Galaxy in R.
Output files

• xset.annotate.variableMetadata.tsv

For each metabolite (row):  the value of the intensity in each sample, fold, anova, mzmed, mzmin, mzmax, rtmed, rtmin, rtmax, npeaks, isotopes, adduct and pcgroup

<table>
<thead>
<tr>
<th>isotopes</th>
<th>adduct</th>
<th>pcgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>[M+H]+</td>
<td>102.032</td>
<td>1126</td>
</tr>
<tr>
<td>[M+CO2]+</td>
<td>146.021</td>
<td>110</td>
</tr>
<tr>
<td>[M+H]-</td>
<td></td>
<td>715</td>
</tr>
<tr>
<td>[M+CO2]-</td>
<td></td>
<td>1280</td>
</tr>
<tr>
<td>[M+H]-</td>
<td></td>
<td>867</td>
</tr>
<tr>
<td>[M+CO2]-</td>
<td></td>
<td>1057</td>
</tr>
<tr>
<td>[M+H]-</td>
<td></td>
<td>931</td>
</tr>
<tr>
<td>[M+CO2]-</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>[M+H]-</td>
<td>70.0546</td>
<td>237</td>
</tr>
<tr>
<td>[M+CO2]-</td>
<td></td>
<td>791</td>
</tr>
<tr>
<td>[M+H]-</td>
<td></td>
<td>716</td>
</tr>
</tbody>
</table>
And next...database search!