Processing Tandem-MS and MS^n data with with xcms

S. Neumann, J. Kutzera

April 30, 2010

Introduction

This document describes how to use *xcms* for processing of Tandem-MS and MS^n data from e.g. triple-quad, QTOF, ion trap or orbitrap mass spectrometers.

```
> library(xcms)
> library(msdata)
```

1 Raw Data File Preparation

The *xcms* package reads full-scan LC/MS data and associated Tandem-MS and MS^n spectra from mzData and mzXML files (CDF does not support MS^n).

For the purposes of demonstration, we will use a some sample files included in the *msdata* package. The (truncated) raw data files are contained in the `data` directory. To access the mzData files, we first locate the *data* directory in the *msdata* package.

```
> mzdatapath <- system.file("iontrap", package = "msdata")
> list.files(mzdatapath, recursive = TRUE)
```

```
[1] "extracted.mzData"
```

2 Accessing, combining and Visualising MS^n spectra

The MS^n spectra are read into xcms just like plain LC-MS files. Inclusion of the additional scans has to be explicitely enabled. The xcmsRaw summary shows the additional content:

```
> library(xcms)
> mzdatafiles <- list.files(mzdatapath, pattern = "extracted.mzData",
+     recursive = TRUE, full.names = TRUE)
> xraw <- xcmsRaw(mzdatafiles[1], includeMSn = TRUE)
> xraw
```

```
An "xcmsRaw" object with 100 mass spectra

Time range: 1190-1309.2 seconds (19.8-21.8 minutes)
Mass range: 500-849.9333 m/z
Intensity range: 100008-828487000

MSn data on  22  mass(es)
        with  44  MSn spectra
Profile method: intlin
Profile step: 1 m/z (351 grid points from 500 to 850 m/z)

Memory usage: 2.71 MB
```

There is also a new peak picker, which actually returns all parent ions as annotated in the mzData / mzXML files. A warning will be issued if this is not present, either because of a poor converter or if xraw is read from a CDF file:

```
> peaks <- findPeaks(xraw, method = "MS1")
```

By giving larger windows, subsets of spectra can be returned, and these can be combined to obtain so called compound spectra containing e.g. MS^2 to MS^5.

3 xcmsFragments

A xcmsFragments-Object contains a Table with informations about all Peaks from MS1 to MS^n of one Experiment, including mz-value, retentiontime, MSlevel, and parentpeak. The data can be used for visualization or (later) computing alignments between different Experiments. The object is created as empty and can be filled with collect(). This method requires a peaktable or a xcmsSet for the MS1peaks and the original xcmsRaw for the msn-Data.

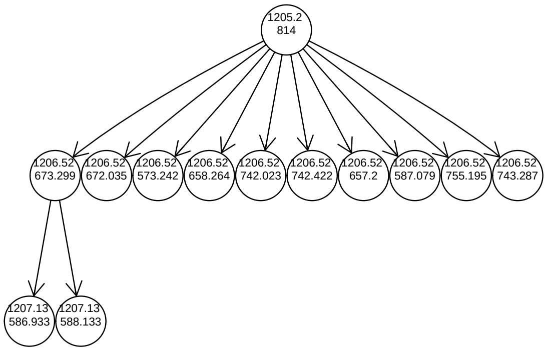
```
> xs <- xcmsSet(mzdatafiles, method = "MS1")
```

```
extracted:
```

```
> xfrag <- xcmsFragments(xs)
```

```
394 Peaks picked, 0 MSn-Specs ignored.
```

```
> xfrag
```



```
|#6_M814T1205.2F814
|     #75_M814T1206.52F673.2994
|          #85_M814T1207.13F586.9333
|          #86_M814T1207.13F588.1333
|     #76_M814T1206.52F672.0354
|     #77_M814T1206.52F573.2417
|     #78_M814T1206.52F658.2642
|     #79_M814T1206.52F742.0226
|     #80_M814T1206.52F742.4218
|     #81_M814T1206.52F657.1997
|     #82_M814T1206.52F587.0795
|     #83_M814T1206.52F755.1951
|     #84_M814T1206.52F743.2867
```

Figure 1: The MS^n fragmentation tree for the MS1 peak 6.

```
An "xcmsFragments" object with  394  peaks in 77 Spectra
From Level 1 to 3 Number of Samples:  1 .

Sample 1 :
    37 Peaks in Level 1
    351 Peaks in Level 2
    6 Peaks in Level 3

Memory usage: 0.0258 MB
```

plotTree prints out a simplyfied tree of the peaks in a xcmsFragments. The user can define borders for the retentionTime and for the MassRange, he can also print the tree for a single MS1-Peak. PlotTree uses the package RgraphViz and plots the tree text-based if requested or if the package cannot be found.

```
> plotTree(xfrag, xcmsFragmentPeakID = 6)
```