

High throughput acquisition and processing of tandem mass spectra

Steffen Neumann*

Andrea Thum†

January 31, 2014

Contents

1	Introduction	1
2	Acquisition	1
3	Processing	6
4	Session information	9

1 Introduction

?

2 Acquisition

```
> ##  
> ## finish the faahKO data  
> ##  
>  
> faahko <- group(faahko)
```

```
262 325 387 450 512 575
```

```
> faahko <- fillPeaks(faahko)
```

```
/vol/local/lib/R/site-library/faahKO/cdf/KO/ko15.CDF  
/vol/local/lib/R/site-library/faahKO/cdf/KO/ko16.CDF  
/vol/local/lib/R/site-library/faahKO/cdf/KO/ko18.CDF  
/vol/local/lib/R/site-library/faahKO/cdf/KO/ko19.CDF  
/vol/local/lib/R/site-library/faahKO/cdf/KO/ko21.CDF  
/vol/local/lib/R/site-library/faahKO/cdf/KO/ko22.CDF
```

*sneumann@ipb-halle.de

†Andrea Thum <andrea.thum@informatik.uni-halle.de>

```

/vol/local/lib/R/site-library/faahKO/cdf/WT/WT15.CDF
/vol/local/lib/R/site-library/faahKO/cdf/WT/WT16.CDF
/vol/local/lib/R/site-library/faahKO/cdf/WT/WT18.CDF
/vol/local/lib/R/site-library/faahKO/cdf/WT/WT19.CDF
/vol/local/lib/R/site-library/faahKO/cdf/WT/WT21.CDF
/vol/local/lib/R/site-library/faahKO/cdf/WT/WT22.CDF

> ##
> ## Create the annotation for putative [M+H]+
> ##
> xa_pos <- xsAnnotate(faahko)
> xa_pos <- groupFWHM(xa_pos)

Start grouping after retention time.
Created 133 pseudospectra.

> xa_pos <- findIsotopes(xa_pos)

Generating peak matrix!
Run isotope peak annotation
% finished: 10 20 30 40 50 60 70 80 90 100
Found isotopes: 58

> xa_pos <- groupCorr(xa_pos)

Start grouping after correlation.
Generating EIC's ..

Calculating peak correlations in 133 Groups...
% finished: 10 20 30 40 50 60 70 80 90 100

Calculating graph cross linking in 133 Groups...
% finished: 10 20 30 40 50 60 70 80 90 100
New number of ps-groups: 321
xsAnnotate has now 321 groups, instead of 133

> xa_pos <- findAdducts(xa_pos,
+                       ppm=100, mzabs=0.1,
+                       polarity="positive")

Generating peak matrix for peak annotation!

Calculating possible adducts in 321 Groups...
% finished: 10 20 30 40 50 60 70 80 90 100

> p <- getPeaklist(xa_pos)
> ## Find Interesting peaks
> dr <- diffreport(faahko, sortpval=FALSE)
> ## Interesting groups have to be
> ## 1) potential [M+H]+
> ## 2) differential
> ## 3) with a minimum intensity

```

```

>
> targetgroups <- which ( grepl("[M+H]", p[, "adduct"], fixed=TRUE)
+                          & dr[, "fold"] > 1
+                          & dr[, "wt15"] > 1000 )
> if (length(targetgroups) == 0) {
+   message("Sorry, nothing of interest left :-(")
+ }
> priorities <- order(dr[targetgroups, "fold"])
> ##
> reporttab <- groups(faahko)[targetgroups[priorities],]
> ## Fix column names
> ## Need to fix code later ;- )
>
> colnames(reporttab) <- sub("^rt$", "rtmed", colnames(reporttab))
> colnames(reporttab) <- sub("^mz$", "mzmed", colnames(reporttab))
> ##
> ## The exclusion peakID list is for iterative
> ## method generation
> ##
>
> templateFile <- system.file("20minKalibpos_Startermethod_MSMS.m/microTOFQAcquisition.met
+                             package = "MetShot")
> collisionEnergy <- c(20)
> methodname <- paste(tempdir(), "/MSMS-faahKO-20eV", sep="")
> picklist <- xcms2method(reporttab, methodPrefix=methodname,
+                          widthFactor=1.5, minWidth=4,
+                          template=templateFile,
+                          MSMSManual_ListCollisionEnergy=collisionEnergy,
+                          MSmode="positive")

[1] 545
[1] 2 3 5 6 7 8 9
      mzmed mzmin   mzmax   rtmed   rtmin   rtmax npeaks KO WT
9 315.0000   315 315.0000 2520.939 2498.248 2554.589     8 4 4
9 313.0012   313 313.0293 2787.765 2780.332 2796.762    12 6 6
[1] 148
[1] 1 2
[1] mzmed mzmin mzmax rtmed rtmin rtmax npeaks KO WT
<0 rows> (or 0-length row.names)
[[1]]
      mzmed   mzmin   mzmax   rtmed   rtmin   rtmax npeaks KO WT
2 219.0848 219.0488 219.1000 2524.852 2515.853 2532.286     9 4 5
5 343.0000 343.0000 343.0324 2686.042 2677.044 2693.478    12 6 6
1 279.0180 279.0000 279.0602 2791.676 2783.852 2797.939    12 6 6
8 506.1921 506.1000 506.2000 3393.401 3364.450 3453.652    13 6 6
6 330.1491 330.1054 330.2000 3494.342 3487.299 3506.078     6 2 4
3 479.2268 479.2000 479.2856 3626.580 3616.799 3637.925    12 6 6
4 255.1596 255.1118 255.1954 3677.440 3653.575 3698.179    12 6 6

[[2]]

```

	mzmed	mzmin	mzmax	rtmed	rtmin	rtmax	npeaks	KO	WT
7	315.0000	315	315.0000	2520.939	2498.248	2554.589	8	4	4
9	313.0012	313	313.0293	2787.765	2780.332	2796.762	12	6	6

>

>

The picklist

> picklist

[[1]]

	mzmed	mzmin	mzmax	rtmed	rtmin	rtmax	npeaks	KO	WT
2	219.0848	219.0488	219.1000	2524.852	2515.853	2604.665	9	4	5
5	343.0000	343.0000	343.0324	2686.042	2604.665	2738.665	12	6	6
1	279.0180	279.0000	279.0602	2791.676	2738.665	3081.194	12	6	6
8	506.1921	506.1000	506.2000	3393.401	3081.194	3470.476	13	6	6
6	330.1491	330.1054	330.2000	3494.342	3470.476	3561.438	6	2	4
3	479.2268	479.2000	479.2856	3626.580	3561.438	3645.750	12	6	6
4	255.1596	255.1118	255.1954	3677.440	3645.750	3698.179	12	6	6

[[2]]

	mzmed	mzmin	mzmax	rtmed	rtmin	rtmax	npeaks	KO	WT
7	315.0000	315	315.0000	2520.939	2498.248	2667.460	8	4	4
9	313.0012	313	313.0293	2787.765	2667.460	2796.762	12	6	6

It is also possible to create manymany MSMS methods:

```
> collisionEnergies <- c(20)
> pickLists <- list()
> reporttabTargetRun <- rep(NA, nrow(reporttab))
> ## write bruker method
> for (j in seq(length = length(collisionEnergies))) {
+   methodname <- paste(tempdir(), "/MSMS-faahKO-20eV",
+                         collisionEnergies[j], ".m", sep="")
+
+   pickLists <- xcms2method(reporttab, method=methodname,
+                             widthFactor=1.5, minWidth=4,
+                             template=templateFile,
+                             MSMSManual_ListCollisionEnergy=collisionEnergies[j],
+                             MSmode="positive")
+
+   message(paste("Created ", methodname,
+                 "with", length(pickLists), "Runs"))
+
+   for (i in 1: length(pickLists)) {
+     filename <- paste(methodname,"_",i, ".csv", sep="")
+     write.csv(pickLists[[i]], file=filename)
+   }
+ }
```

```

[1] 545
[1] 2 3 5 6 7 8 9
      mzmed mzmin  mzmax  rtmed  rtmin  rtmax npeaks KO WT
7 315.0000   315 315.0000 2520.939 2498.248 2554.589     8 4 4
9 313.0012   313 313.0293 2787.765 2780.332 2796.762    12 6 6
[1] 148
[1] 1 2
[1] mzmed mzmin mzmax rtmed rtmin rtmax npeaks KO    WT
<0 rows> (or 0-length row.names)
[[1]]
      mzmed  mzmin  mzmax  rtmed  rtmin  rtmax npeaks KO WT
2 219.0848 219.0488 219.1000 2524.852 2515.853 2532.286     9 4 5
5 343.0000 343.0000 343.0324 2686.042 2677.044 2693.478    12 6 6
1 279.0180 279.0000 279.0602 2791.676 2783.852 2797.939    12 6 6
8 506.1921 506.1000 506.2000 3393.401 3364.450 3453.652    13 6 6
6 330.1491 330.1054 330.2000 3494.342 3487.299 3506.078     6 2 4
3 479.2268 479.2000 479.2856 3626.580 3616.799 3637.925    12 6 6
4 255.1596 255.1118 255.1954 3677.440 3653.575 3698.179    12 6 6

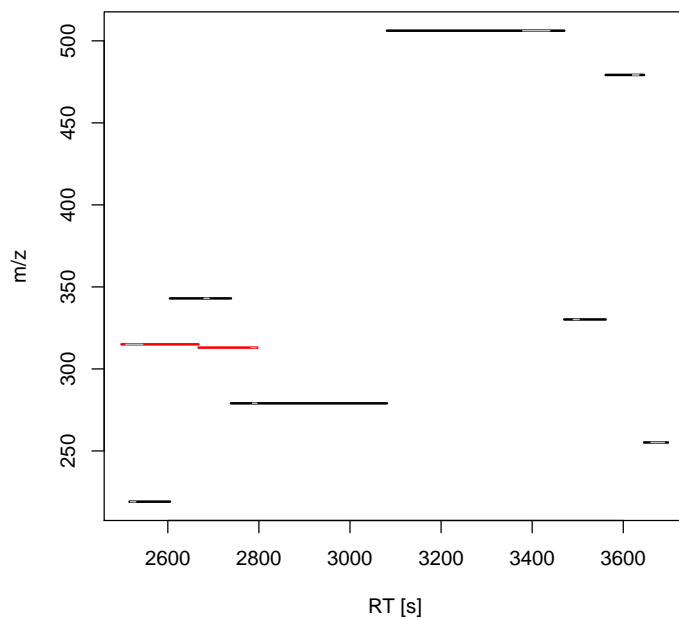
[[2]]
      mzmed mzmin  mzmax  rtmed  rtmin  rtmax npeaks KO WT
7 315.0000   315 315.0000 2520.939 2498.248 2554.589     8 4 4
9 313.0012   313 313.0293 2787.765 2780.332 2796.762    12 6 6

> for (j in 1:length(pickLists)) {
+   reporttabTargetRun[match(rownames(pickLists[[j]]), rownames(reporttab))] <- j
+ }
> reporttab<-cbind(reporttab, reporttabTargetRun)
>

Plot peaks and target windows superimposed.

> ## plot overview
>
> for (i in seq(1, length(pickLists))) {
+   plotMS2windows(reporttab, pickLists[[i]],
+                   peaks=ifelse(i==1, TRUE, FALSE),
+                   col=i)
+ }
>

```



3 Processing

Now a file was measured, we need to extract the MS2 spectra. Instead of “just” taking a spectrum, we do peak picking and CAMERA grouping.

```
> #####
> ##
> ## Parameters common to MS1 and MS2 picking
> ##
>
> peakwidth=c(6,15)
> ppm=25
> snthresh=2
> prefilter=c(2,20)
> #####
> ##
> ## Read MS1 peaks
> ##
>
> ms1peaks <- faahko
> #####
> ##
> ## Read MS2 peaks
> ## Obtain MS2 peaks (ab)using normal peak picker
> ##
```

```

>
> snthresh=2
> ## Create a pseudo-MS1 xraw
> filename <- system.file("microtof/MM48_MSMSpos_MH1_15ev_1-A,2_01_9310.mzXML",
+                           package = "MetShot")
> xraw <- msn2xcms(xcmsRaw(filename, includeMSn=TRUE))
> ms2peaks <- findPeaks(xraw, method="centWave",
+                       snthresh=snthresh, prefilter=prefilter,
+                       peakwidth=peakwidth, ppm=ppm, verbose.columns=T)

Detecting mass traces at 25 ppm ...
% finished: 0 10 20 30 40 50 60 70 80 90 100
817 m/z ROI's.

Detecting chromatographic peaks ...
% finished: 0 10 20 30 40 50 60 70 80 90 100
726 Peaks.

> ## for each peak
> ms2precursor <- cbind(rt=xraw@msnRt, mz=xraw@msnPrecursorMz)
> ##
> ## CAMERA annotation of spectra
> ##
>
> xs <- new("xcmsSet")
> peaks(xs) <- ms2peaks
> filepaths(xs) <- c(filename)
> sampnames(xs) <- c("MM48_MSMSpos")
> xs@peaks <- cbind(xs@peaks, sample=1)
> xa <- xsAnnotate(xs)
> xa <- groupFWHM(xa)

Start grouping after retention time.
Created 82 pseudospectra.

> ## xa <- groupCorrhack(xa,xraw,cor_eic_th=0.75)
>

```

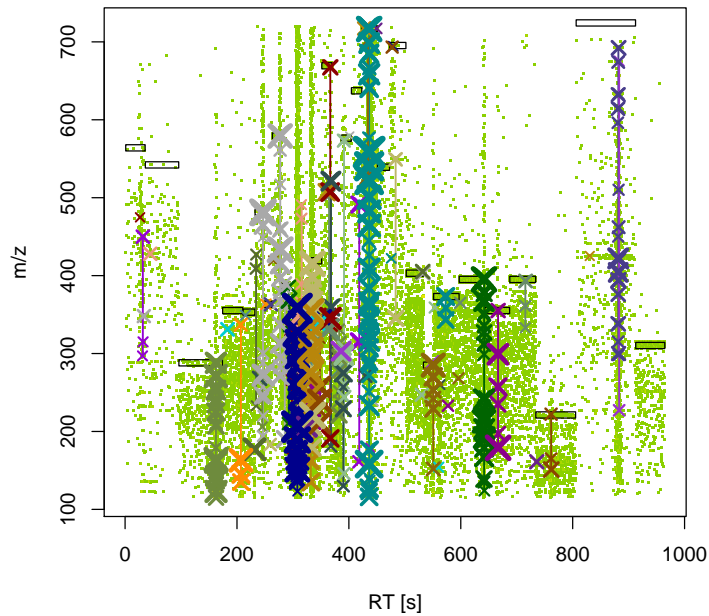
You can plot the overlaid MS2 raw data and the picked and grouped peaks:

```

> plotGroups(xraw, ms2precursor, xa)
>

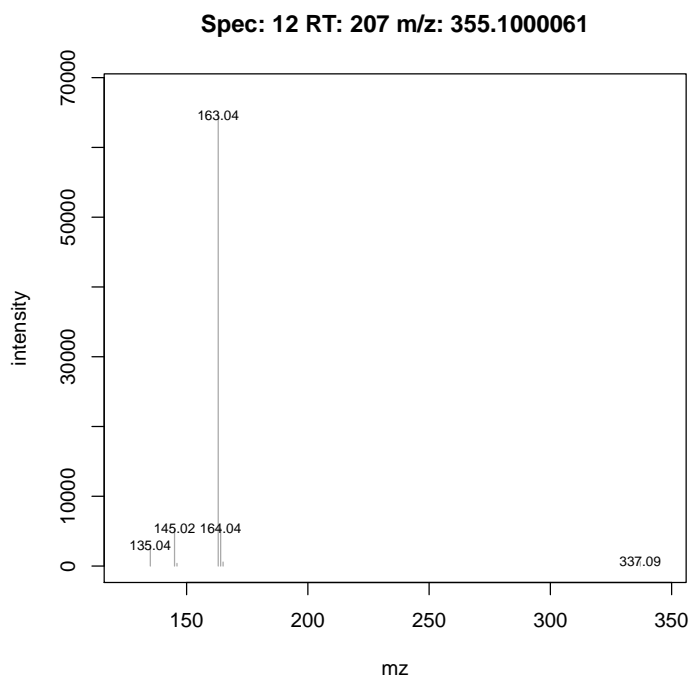
```

MM48_MSMSpos_MH1_15ev_1-A,2_01_9310.mzXML



You can plot the individual retrieved MS2 Spectrum:

```
> pspec <- 12
> rt <- median(getpspectra(xa, pspec)[,"rt"])
> mz <- ms2precursor[which(ms2precursor[,"rt"] == rt), "mz"]
> sp <- getpspectra(xa, grp=pspec)[,c("mz", "maxo")]
> plotPsSpectrum(xa, pspec=pspec,
+               maxlabel=5, log=FALSE, sleep=0,
+               title=paste("Spec:", pspec, "RT:", as.integer(rt), "m/z:", mz))
>
```

4 Session information

- R version 3.0.2 (2013-09-25), x86_64-pc-linux-gnu
- Locale: LC_CTYPE=en_US.UTF-8, LC_NUMERIC=C, LC_TIME=en_US.UTF-8, LC_COLLATE=C, LC_MONETARY=en_US.UTF-8, LC_MESSAGES=en_US.UTF-8, LC_PAPER=en_US.UTF-8, LC_NAME=C, LC_ADDRESS=C, LC_TELEPHONE=C, LC_MEASUREMENT=en_US.UTF-8, LC_IDENTIFICATION=C
- Base packages: base, datasets, grDevices, graphics, methods, parallel, stats, utils
- Other packages: Biobase 2.20.0, BiocGenerics 0.6.0, CAMERA 1.16.0, MetShot 0.1.23, Rcpp 0.10.6.1, XML 3.96-1.1, class 7.3-9, faahKO 1.2.13, igraph 0.6.5-1, multtest 2.16.0, mzR 1.7.3, xcms 1.39.1
- Loaded via a namespace (and not attached): Hmisc 3.10-1.1, MASS 7.3-26, RBGL 1.36.2, cluster 1.14.4, codetools 0.2-8, graph 1.38.2, grid 3.0.2, lattice 0.20-24, splines 3.0.2, stats4 3.0.2, survival 2.37-4, tools 3.0.2

References