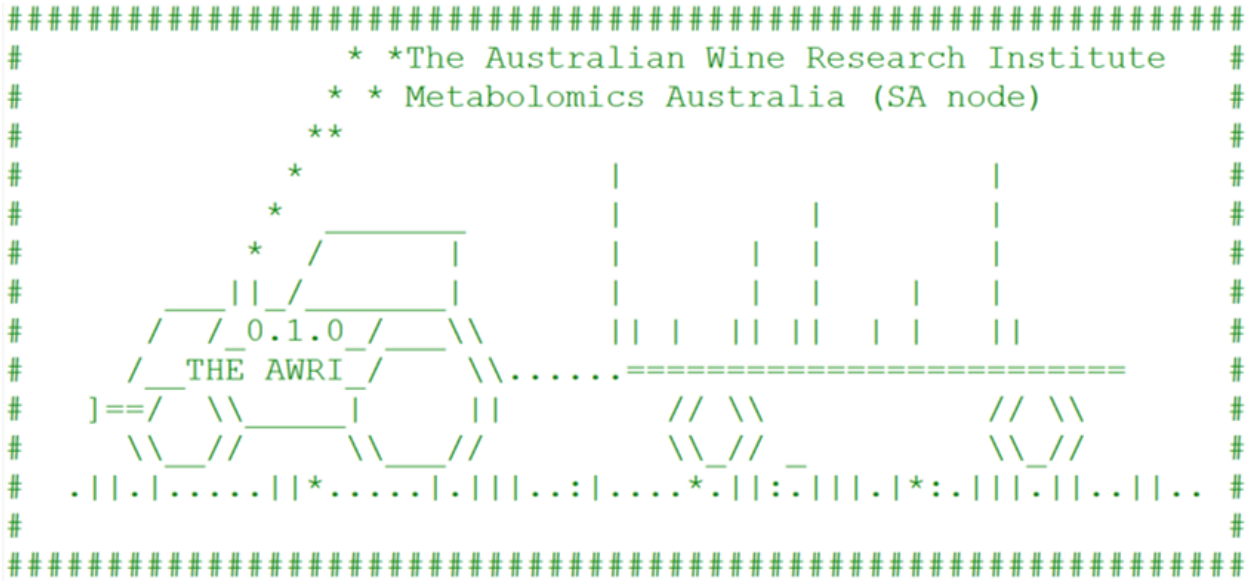
MStractor Workflow

Luca Nicolotti, The Australian Wine Research Institute, Metabolomics Australia

10/06/2020



Introduction

MStractor combines the functions for molecular feature extraction present in XCMS and CAMERA with user friendly dedicated GUIs for LC and MS parameters input, graphical QC outputs and descriptive statistics calculation.

A schematic representation of the workflow is displayed in **Figure 1**, reported below. The workflow consists of 10 steps with the possibility of running the workflow using either the most recent xcms functions specific for *XCMSnExp* object (available from xcms verison 1.51.5) (**branch a**) or the set of functions for *xcmsSet* objects (**branch b**). However, the 2 branches perform the same data processing steps and produce similar outputs. The reason of developing 2 parallel frameworks mainly depends on the availability of computational power on the user's end. Specifically, the authors observed that step **6a** tends to be quite time-consuming in particular when dealing with complex datasets. Therefore, in case of limited computing resources and large datasets, the authors suggest using branch b of the workflow. Step 8 which is common to both workflow branches, where the data set is processed using CAMERA. Since the latest release available for CAMERA does not support XCMSnEXP objects, the data set object is required to be converted to the xcmsSet format prior to this step.


1
Acquisition Parameters
Set Project, Chromatpography and Mass Spectrometry parameters
Adjust settings
2
Explore Reference File
Evaluate feature extraction parameters on a reference file via graphical outputs
3
Create Data set
Create a XCMSnEXP object and perform peak picking and grouping on all the data files present in the data set
Branch A
XCMSnEXP workflow
High Computing Power
Branch B
xcmsSet workflow
Low Computing Power
4a
Print raw EICs and TICs
Print Extracted ion chromatograms of all the features prior to retention time alignment
4b
Print raw EICs and TICs
Print Extracted Ion chromatograms of all the features prior to retention time alignment
5a
Retention time alignment
Align molecular features, perform regrouping and filling of missing data
5b
Retention time alignment
Align molecular features, perform regrouping and filling of missing data
6a
Print aligned EICs and TICs
Print Extracted ion chromatograms of all the features after retention time alignment
6b
Print aligned EICs and TICs
Print Extracted ion chromatograms of all the features after retention time alignment
7
Create summary data matrix
Create data matrix collating feature definition data and instrument response data
8
CAMERA
Group ions that are highly correlated in time and intensity domain
9
Export results
Create graphical outputs for peak curation*, export Base Peak Matrix
10
Descriptive Statistics
Perform median normalisation calculate descriptive statistics for each sample class
* Manual Curation Needed

Figure 1 = 'Schematic representation of MStractor Workflow'

Each step of the framework displayed in Figure 1 involves using 1 or more functions. MStractor functions can be classified into 3 different groups:

1.MStractor specific functions: developed by the authors to provide the user with a friendly and seamless experience. The functions belonging to this class mainly involve the use of GUIs allowing parameters input and storage to be used later in the workflow. This group also include functions to normalise data and calculate descriptive statistics.

2.Wrapper functions: includes wrapper functions of xcms and CAMERA. The function backbone is the xcms (or CAMERA) function and additional code is used to automate routine operations and produce graphical outputs within the selected working directory.

3.xcms and CAMERA functions: these functions are basic CAMERA and xcms function that are used along the workflow. However, the user is not required to update the function arguments because all the necessary parameters are input in the previous steps of the workflow

Data Preparation and File Naming

Preparing the data and creating the correct working directory and folder structure (outside the R environment) is required prior to executing the workflow. The raw data files need to be converted to an appropriate format which can be easily achieved using various tools freely available, such as Proteowizard (ProteWizard, http://proteowizard.sourceforge.net/) and Massmatrix (MassMatrix, http://www.massmatrix.net/). MStractor supported file formats include CDF, mzDATA, mzML and mzXML. After conversion, files should be grouped in folders according to their class of belonging. Every class-folder needs to be stored within the directory 'MSfiles'. The script will use the 'MSfiles' subfolder downstream to automatically determine the sample classes required for processing. If the folder 'MSfiles' is not present, the user will encounter errors. For each class-folder a minimum of 2 replicates is required. An example dataset (see Figure2) is provided with the package.

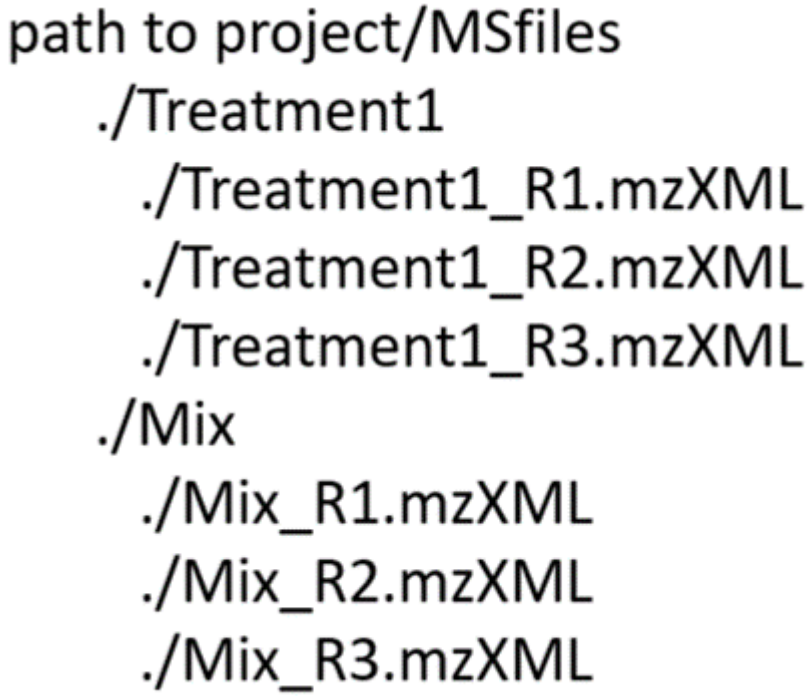


Figure 2 = 'Project subfolder structure'

File naming is a key aspect for correct visualisation of graphical output downstream. A specific function of the workflow assigns different colours and symbols to the different sample classes. In this respect, the filename has to correspond to the folder name followed by an underscore and a number inidicating the biological replicate, as displayed in **Figure 2**.

Installation

To install the package from GitHub, make sure the package 'remotes' is installed and run the following

```
library(remotes)

Sys.setenv(R_REMOTES_NO_ERRORS_FROM_WARNINGS="true")

remotes::install_github("MetabolomicsSA/MStractor")

# Alternatively, Download the tar.gz package from
# https://github.com/MetabolomicsSA/MStractor/releases
# and run the following

library(remotes)

Sys.setenv(R_REMOTES_NO_ERRORS_FROM_WARNINGS="true")

setRepositories(ind=1:2)

remotes::install_local("C:/pathtoPackage/MStractor_0.1.0.tar.gz",
 '               dependencies=NA)
```

```
library(MStractor)
```

1. *Acquisition Parameters*

The first step of the workflow consists in running 6 functions that allow the input of chromatographic, mass spectrometry and peak picking parameters as well as loading the data and automatically defining the visualization settings for graphical outputs.

The standard input values provided in each function are based on the settings of the acquisition instrument used by the authors and, therefore, suitable for processing the example data set available within the package. These criteria should be changed according to the user's LCMS system configuration, since they can dramatically influence the data processing outcome.

1.1 *Define Working Directory and Reference Files*

```
Project()
#If using the dataset provided within the package, set the working directory only and skip the rest
```

use instead:

```
path<-system.file("extdata",package = "MStractor")
files <- dir(path, pattern = ".mzXML", full.names = TRUE)
```

The function doesn't require arguments. Its execution opens a GUI (Figure 3) allowing the selection of the working directory and 2 reference fileschosen from one of the folders within the MSfiles directory (e.g.: 2 'Mix' replicates in the example dataset). The function also creates a QC folder where graphical outputs to evaluate the progressing of the workflow are stored .

Figure 3: GUI for project folder selection

1.2 *Input of Chromatographic and Mass Spectrometry Parameters*

ChromParam() allows the user to input the chromatographic parameters related to the data set to be processed and that will be stored and used in the later stages of the framework. The values to be entered are the retention time range of the data set (rt start and rt end), the maximum retention time drift observed and the minimum and maximum fullwidth at half maximun (FWHM min and max).

MassSpecParam() is mainly used to input mass spectrometry related parameters. The criteria to be entered are the acquisition mode (negative as default); the mass range to be considered (mz start and mz end); the

number of expected charges (default value set at 3); the file type (default set to '.mzXML'); the maximum number of chromatographic peaks expected for a single EIC and the sensitivity.

Both the functions return the input parameteres.

For more details about more details about the parameters consult the xcms manual: https://www.bioconductor.org/packages/release/bioc/manuals/xcms/man/xcms.pdf)

```
ChromParam()
```

```
## [[1]]
## [1] 1
##
## [[2]]
## [1] "max"
##
## [[3]]
## [1] 10
##
## [[4]]
## [1] 90
##
## [[5]]
## [1] 32
```

```
MassSpecParam()
```

```
## [[1]]
## [1] "negative"
##
## [[2]]
## [1] 100
##
## [[3]]
## [1] 1650
##
## [[4]]
## [1] 0.01
##
## [[5]]
## [1] 3
##
## [[6]]
## [1] 30
##
## [[7]]
## [1] 0.7
##
## [[8]]
## [1] ".mzXML"
```

```
# leave default values if using the package dataset
```

Values are entered using GUIs as displayed in Figure 4.

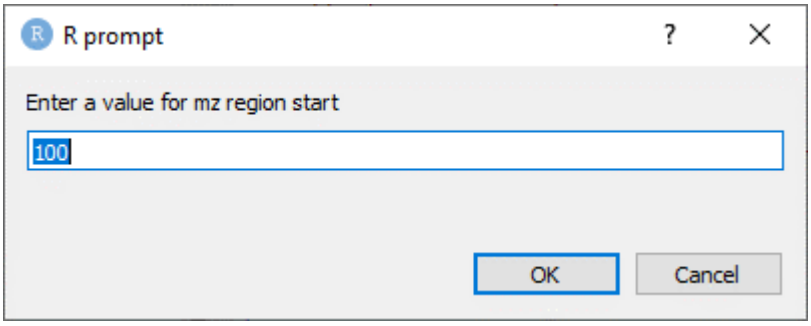


Figure 4: example of GUI input parameters

1.3 *Load the Dataset*

```
LoadData()
#Skip this step if using the dataset provided within the package
```

The function returns a complete list of the loaded files. It also starts recording the time necessary for data processing and creates the object 'ClassType' which associates every data file with the class of belonging.

1.4 *Define Class Identifiers*

```
ClassType<-c('Mix','Treatment1')
## don't run if using the the provided dataset
DefineClassAttributes(ClassType)
```

```
## [[1]]
## [1] "Mix"        "Mix"        "Mix"        "Treatment1" "Treatment1"
## [6] "Treatment1"
##
## [[2]]
## [1] "#FF0000" "#FF0000" "#FF0000" "#00FFFF" "#00FFFF" "#00FFFF"
##
## [[3]]
## [1] 1 1 1 2 2 2
```

The argument of the function is the object 'ClassType'. The function defines and returns specific symbols and colours for each sample class, which are then used to produce graphical outputs.

1.5 *Define Peak Picking Parameters*

```
#leave default values if using the package dataset
PeakPickingParam()
```

```
## [[1]]
## [1] 10
##
## [[2]]
## [1] 20
```

```
##
## [[3]]
## [1] 3.030303
##
## [[4]]
## [1] 50
##
## [[5]]
## [1] 26.51515
##
## [[6]]
## [1] 0.002
##
## [[7]]
## [1] FALSE
##
## [[8]]
## [1] 1
```

The function uses GUIs to define xcms peak picking parameters using the centwave algorithm. These include: minimum and maximum peakwidth, minimum and maximum m/z error (in ppm), values for the integration threshold (set at 5000) and signal to noise threshold, the integration method to be used (1 or 2 as defined by the xcms documentation) and whether the pick picking should fit the gaussian curve. The default value for minimum and maximum m/z error and signal to noise (S/N) threshold is 'none' and correspond, respectively, to 3.03 and 50 ppm and 1000 for S/N value. Detailed information about the mentioned parameters can be found in the xcms documentation https://www.bioconductor.org/packages/release/bioc/manuals/xcms/man/xcms.pdf

2. *Parameter Evaluation*

This step allows testing the suitability of the input parameters on the reference files selected in the first step.

```
exploreRefs()
```

Note: In case the package dataset is used, the following needs to be run

```
dir.create("./QC")
SampleGroup<-c("Mix", "Mix")
symbol<-c(1,1)
ClassCol<-c("#FF0000FF", "#FF0000FF")
path<-system.file("extdata",package = "MStractor")
files <- dir(path, pattern = ".mzXML", full.names = TRUE)
test<-files[1:2]
pd <- data.frame(sample_name = sub(basename(test), pattern = filetype,
    replacement ="", fixed = TRUE),sample_group = SampleGroup,
    stringsAsFactors = FALSE)
ref_data <- readMSData(test, pdata = new("NAnnotatedDataFrame",
    pd), mode = "onDisk")
expRefData(ref_data)
```

First, a GUI allows to define the settings of xcms functions findChromPeaks and groupChromPeaks which perform peak picking and grouping of molecular features. Secondly, data files are read using the xcms function readMSData and, lastly, peak picking and grouping is carried out. The output of the function is the XCMSnEXP object 'x_refs'.

```
refTic(x_refs)

## null device
##           1
```

The refTic function returns a plot of the overlaid TICs for the reference files

Figure 5: non-aligned overlaid TICs

```
get100(x_refs)
```

The get100 function randomly picks 100 features detected across the retention time domain and plots them into a matrix. Using this output, it is easy to check whether all the features are correctly integrated.

Figure 6: EICs of 100 features randomly picked across the retention time

3. *Create XCMSnExp Dataset*

An XCMSnEXP dataset containing all the raw files to be processed is automatically created based on the folder structure defined in
the data preparation step. The function perform the same steps described for exploreRefs(), with the only difference that the processing is applied to the whole dataset.

```
peakPickGroup() # don't run if using the example dataset
```

In case the example dataset fullDataSet is used, the following needs to be run

```
ClassType<-c('Mix', 'Treatment1')
SampleGroup<-c("Mix",        "Mix",        "Mix",        "Treatment1", "Treatment1", "Treatment1")
symbol<-c(1 ,1, 1, 2,2, 2)
ClassCol<- c("#FF0000FF", "#FF0000FF", "#FF0000FF", "#00FFFFFF", "#00FFFFFF", "#00FFFFFF")
path<-system.file("extdata",package = "MStractor")
files <- dir(path, pattern = ".mzXML", full.names = TRUE)
pd <- data.frame(sample_name = sub(basename(files), pattern = filetype,
    replacement = "", fixed = TRUE), sample_group = SampleGroup,
    stringsAsFactors = FALSE)
raw_data <- readMSData(files =files, pdata = new("NAnnotatedDataFrame",
    pd), mode = "onDisk")
ppgExData(raw_data)
```

It is important that the input parameters used in this step match the ones defined for the reference files. After performing peak picking on each datafile,peaks are matched across all samples and grouped using the xcms functions findChromPeaks and groupChromPeaks. The results are stored in the XCMSnEXP object 'xdata'

Workflow Branching

The workflow forks at step 4 in branch a and branch b, as displayed in **Figure 1**. From this point on, the user can choose to either proceed with the XCMSnEXP object or perform a conversion into a xcmsSet object. Workflow branch A requires higher computing powere than branch b. For demonstration purpose, please run each branch separately

This step is dedicated to plot outputs useful for evaluating the data processing progress and visualizing macroscopic differences among chromatographic profiles. It includes 2 functions for each branch of the workflow.

Branch A

4a *Print raw EICs and TICs*

```
OverlaidTICs(xdata, 'raw') #raw indicates non-retention time aligned signals
printEICs(xdata, 'raw')
```

For both branches plots of overlaid TIC (**Figure7**) chromatograms and for individual EICs are stored in the QC directory.

Figure 7

5a *Retention time alignment, grouping and peak filling*

Retention time correction parameters are entered via GUI. The retention time alignment method currently supported is 'loess' (see xcms manual), while Obiwarp will be implemented in the future releases. A plot of the retention time deviation is also generated and stored in the QC folder. After correction, features are regrouped and the filling of missing signals is performed.

```
RTalign(xdata,'loess')

xdata <- groupChromPeaks(xdata, param = PeakDensityParam(sampleGroups =
        xdata$sample_group, minFraction = 0.3, bw = 20))

xfilled <- fillChromPeaks(xdata, param =(FillChromPeaksParam(ppm = 50,
        expandMz = 0.5)))
```

6a *Print aligned EICs and TICs*

Step 6 replicates the series of functions already described for stage 4, with the only difference that retention time aligned TICs and EICs are generated

```
OverlaidTICs(xdata, 'aligned')
printEICs(xfilled, 'filled')
#'aligned' and 'filled' indicate retention time aligned signals
```

7a *Create Summary Datamatrix*

Step 7 is peculiar for **branch a** and collates features information generated by the xcms featureDefinition() function with their corresponding response (either the integrated peak area or maximum intensity). The arguments of the function are the filled XCMSnEXP object and the type of instrumental response desired (either 'into' or 'maxo'). After running step seven, the workflow merges again and the user has to proceed with step 8.

```
CreateDM(xfilled,'maxo')
```

Following this, the xsetConvert function (already described in step 4) reverts the XCMSnEXP object into the xcmsSet one. This is necessary for using CAMERA in the subsequent steps.

```
xsetConvert(xfilled)

## Note: you might want to set/adjust the 'sampclass' of the returned xcmSet object before proceeding w

sampnames(xset)<-spn
```

Branch B

The user needs to revert the object to a xcmsSet because it is required to execute the following code.

```
xsetConvert(xdata)
sampnames(xset)<-spn
```

4b *Print raw EICs and TICs*

```
getTICs(xcmsSet= xset, pngName= "./QC/TICs_raw.png", rt= "raw")
#raw indicates non-retention time aligned signals
printEICsXset(xset,'raw')
```

5b *Retention time alignment, grouping and peak filling*

Retention time correction parameters are entered via GUI. The retention time alignment method currently supported is 'loess' (see xcms manual), while Obiwarp will be implemented in the future releases. A plot of the retention time deviation is also generated and stored in the QC folder. After correction, features are regrouped and the filling of missing signals is performed.

```
RTalign_xset(xset,'loess')
xsAlign <- group(xsAlign, method= "nearest", mzVsRTbalance= 10, mzCheck=
        mzErrAbs,rtCheck= rtDelta, kNN=10)
xsFilled <- fillPeaks(xsAlign, method="chrom", expand.mz=0.5)
```

6b *Print aligned EICs and TICs*

Step 6 replicates the series of functions already described for stage 4, with the only difference that retention time aligned TICs and EICs are generated

```
getTICs(xcmsSet= xsAlign, pngName= "./QC/TICs_Aligned.png", rt= "corrected")
printEICsXset(xsFilled,'corrected')
# 'corrected' indicates retention time aligned signals
```

Workflow Merging

8. *Spectra Reconstruction*

From step 8, branches a and b merge into a unique workflow. Spectra reconstruction is carried out using the function autoCAMERA()that includes the 3 CAMERA functions groupFWHM(), findIsotopes() and groupCorr(). A GUI allows the input of the function parameters. Consult CAMERA manual for more information: https://www.bioconductor.org/packages/release/bioc/manuals/CAMERA/man/ CAMERA.pdf. Using autoCamera, the peak table embedded within the xcmsSet object is extracted and features arranged into pseudospectra groups according to their retention times. Then, based on the maximum expected charge test, isotopic patterns are located using and the features belonging to coeluting compounds are resolved using a correlation matrix.

```
xset
```

```
## An "xcmsSet" object with 6 samples
##
## Time range: 11.7-8505 seconds (0.2-141.8 minutes)
## Mass range: 96.9573-1565.3684 m/z
## Peaks: 15284 (about 2547 per sample)
## Peak Groups: 2242
## Sample classes: Mix, Treatment1
##
## Feature detection:
##  o Peak picking performed on MS1.
##  o Scan range limited to  1 - 4400
## Profile settings: method = bin
##                   step = 0.1
##
## Memory usage: 6.81 MB

autoCamera(xset)

## Start grouping after retention time.
## Created 649 pseudospectra.
## Generating peak matrix!
## Run isotope peak annotation
##  % finished: 10  20  30  40  50  60  70  80  90  100
## Found isotopes: 310
## Start grouping after correlation.
## Generating EIC's ..
## Warning: Found NA peaks in selected sample.
##
## Calculating peak correlations in 649 Groups...
##  % finished: 10  20  30  40  50  60  70  80  90  100
##
## Calculating peak correlations across samples.
##  % finished: 10  20  30  40  50  60  70  80  90  100
##
## Calculating isotope assignments in 649 Groups...
##  % finished: 10  20  30  40  50  60  70  80  90  100
## Calculating graph cross linking in 649 Groups...
##  % finished: 10  20  30  40  50  60  70  80  90  100
## New number of ps-groups:  1054
## xsAnnotate has now 1054 groups, instead of 649
```

The function generates the PksAn datamatrix, containing the results of the spectra reconstruction.

9. *Export Matrix*

Step 9 provide a series of 4 functions that generate the final results of the workflow. The second argument ofthe function "collectBP_EICs" is either "filled" for workflow branch a or "corrected" for workflow branch b

```
FilterDM(PksAn, xset)

#the second argument of the function below is either 'filled' (branch a) or 'corrected' (branch b)
CollectBP_EICs(BasePks,'filled')
```

```
## Time difference of 58.15745 mins


#after this step manual curation is necessary
BasePks_Curated(BasePks)
MedianNormalize(BasePksCur, xset)
```

By using FilterDM() pseudospectra containing less than a fixed number of ions (2 by default) are removed from the data matrix. Then, only the most intense ion for each pseudospectrum is retained (base peak response). This function was implemented to obtain a simplified data matrix avoiding redundant information and allowing for quick and efficient downstream data treatment.

The output is the BasePks object, a matrix summarizing the results of data filtering.

CollectBP_EICs() prints the extracted ion chromatograms related to the base peak matrix in a .png format. The function also returns the time required for processing the dataset. The pngs files are duplicated into 2 directories, named **'EICs_BasePeaks'** and **'EICs_BasePeaks_Curated'**. The former is used as data back-up, while the latter allows the user to perform a final curation on the EICs, by removing those that do not contain useful information. This represents the only step of the entire workflow where a manual input of the user is required. An example is displayed in **Figure 8**.

Figure 8: curation example

BasePks_Curated() generates an updated data matrix (BasePksCur) with the discarded features removed. The final matrix is saved in a '.tsv' file called 'Pks,BPs_Curated' which is saved within the working directory.

Lastly, by using MedianNormalize() the curated matrix is normalized on the median value. This is done to minimize possible minimal run-to-run variation in the instrument performances. The output is saved in the .tsv file 'Normalized Matrix'.

10. *Descriptive Statistics*

An sdditional step of the workflow allows calculating descriptive statistics (average value, standard deviation and % CV) on the analytical replicates of each sample class. The output is stored in separate .tsv file for each class.

```
StatsByClass(ClassType, xset)
```

Session information

```
sessionInfo()
```

```
## R version 4.0.1 (2020-06-06)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 18363)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=C                          LC_CTYPE=English_Australia.1252
## [3] LC_MONETARY=English_Australia.1252 LC_NUMERIC=C
## [5] LC_TIME=English_Australia.1252
##
## attached base packages:
## [1] stats4    parallel  stats     graphics  grDevices utils     datasets
```

```
## [8] methods   base
##
## other attached packages:
##  [1] MStractor_0.1.0     xcms_3.10.1         MSnbase_2.14.2
##  [4] ProtGenerics_1.20.0 S4Vectors_0.26.1    mzR_2.22.0
##  [7] Rcpp_1.0.4.6        BiocParallel_1.22.0 Biobase_2.48.0
## [10] BiocGenerics_0.34.0
##
## loaded via a namespace (and not attached):
##   [1] snow_0.4-3                  backports_1.1.7
##   [3] Hmisc_4.4-0                 sm_2.2-5.6
##   [5] plyr_1.8.6                  igraph_1.2.5
##   [7] CAMERA_1.44.0               splines_4.0.1
##   [9] fda_5.1.4                   GenomeInfoDb_1.24.0
##  [11] ggplot2_3.3.1               digest_0.6.25
##  [13] foreach_1.5.0               htmltools_0.4.0
##  [15] fansi_0.4.1                 magrittr_1.5
##  [17] checkmate_2.0.0             cluster_2.1.0
##  [19] doParallel_1.0.15           limma_3.44.1
##  [21] matrixStats_0.56.0          stabledist_0.7-1
##  [23] prettyunits_1.1.1           jpeg_0.1-8.1
##  [25] colorspace_1.4-1            xfun_0.14
##  [27] dplyr_1.0.0                 tcltk_4.0.1
##  [29] callr_3.4.3                 crayon_1.3.4
##  [31] RCurl_1.98-1.2              graph_1.66.0
##  [33] roxygen2_7.1.0              impute_1.62.0
##  [35] survival_3.1-12             iterators_1.0.12
##  [37] glue_1.4.1                  gtable_0.3.0
##  [39] zlibbioc_1.34.0             XVector_0.28.0
##  [41] DelayedArray_0.14.0         pkgbuild_1.0.8
##  [43] DEoptimR_1.0-8              abind_1.4-5
##  [45] scales_1.1.1                vsn_3.56.0
##  [47] htmlTable_1.13.3            clue_0.3-57
##  [49] foreign_0.8-80              svDialogs_1.0.0
##  [51] preprocessCore_1.50.0       Formula_1.2-3
##  [53] htmlwidgets_1.5.1           timeSeries_3062.100
##  [55] RColorBrewer_1.1-2          acepack_1.4.1
##  [57] ellipsis_0.3.1              spatial_7.3-12
##  [59] pkgconfig_2.0.3             XML_3.99-0.3
##  [61] nnet_7.3-14                 tidyselect_1.1.0
##  [63] rlang_0.4.6                 munsell_0.5.0
##  [65] tools_4.0.1                 cli_2.0.2
##  [67] generics_0.0.2              evaluate_0.14
##  [69] stringr_1.4.0               yaml_2.2.1
##  [71] mzID_1.26.0                 processx_3.4.2
##  [73] knitr_1.28                  robustbase_0.93-6
##  [75] purrr_0.3.4                 RANN_2.6.1
##  [77] ncdf4_1.17                  RBGL_1.64.0
##  [79] xml2_1.3.2                  compiler_4.0.1
##  [81] rstudioapi_0.11             png_0.1-7
##  [83] svGUI_1.0.0                 testthat_2.3.2
##  [85] affyio_1.58.0               MassSpecWavelet_1.54.0
##  [87] tibble_3.0.1                stringi_1.4.6
##  [89] statip_0.2.3                ps_1.3.3
```

```
##  [91] desc_1.2.0                 modeest_2.4.0
##  [93] lattice_0.20-41            fBasics_3042.89.1
##  [95] Matrix_1.2-18              vctrs_0.3.1
##  [97] pillar_1.4.4               lifecycle_0.2.0
##  [99] BiocManager_1.30.10        MALDIquant_1.19.3
## [101] data.table_1.12.8          bitops_1.0-6
## [103] GenomicRanges_1.40.0       R6_2.4.1
## [105] stable_1.1.4               latticeExtra_0.6-29
## [107] pcaMethods_1.80.0          affy_1.66.0
## [109] gridExtra_2.3              IRanges_2.22.2
## [111] codetools_0.2-16           MASS_7.3-51.6
## [113] gtools_3.8.2               assertthat_0.2.1
## [115] pkgload_1.1.0              SummarizedExperiment_1.18.1
## [117] rprojroot_1.3-2            withr_2.2.0
## [119] GenomeInfoDbData_1.2.3     berryFunctions_1.19.1
## [121] grid_4.0.1                 rpart_4.1-15
## [123] timeDate_3043.102          tidyr_1.1.0
## [125] rmarkdown_2.2              rmutil_1.1.5
## [127] base64enc_0.1-3
```

References

1)Smith, C.A. and Want, E.J. and O'Maille, G. and Abagyan,R. and Siuzdak, G.: XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching and identification, Analytical Chemistry, 78:779-787 (2006)

2)Ralf Tautenhahn, Christoph Boettcher, Steffen Neumann: Highly sensitive feature detection for high resolution LC/MS BMC Bioinformatics, 9:504 (2008)

3)H. Paul Benton, Elizabeth J. Want and Timothy M. D. Ebbels Correction of mass calibration gaps in liquid chromatography-mass spectrometry metabolomics data Bioinformatics, 26:2488 (2010)

4)Kuhl, C., Tautenhahn, R., Boettcher, C., Larson, T. R. and Neumann, S. CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. Analytical Chemistry, 84:283-289 (2012)