MS-Xelerator™

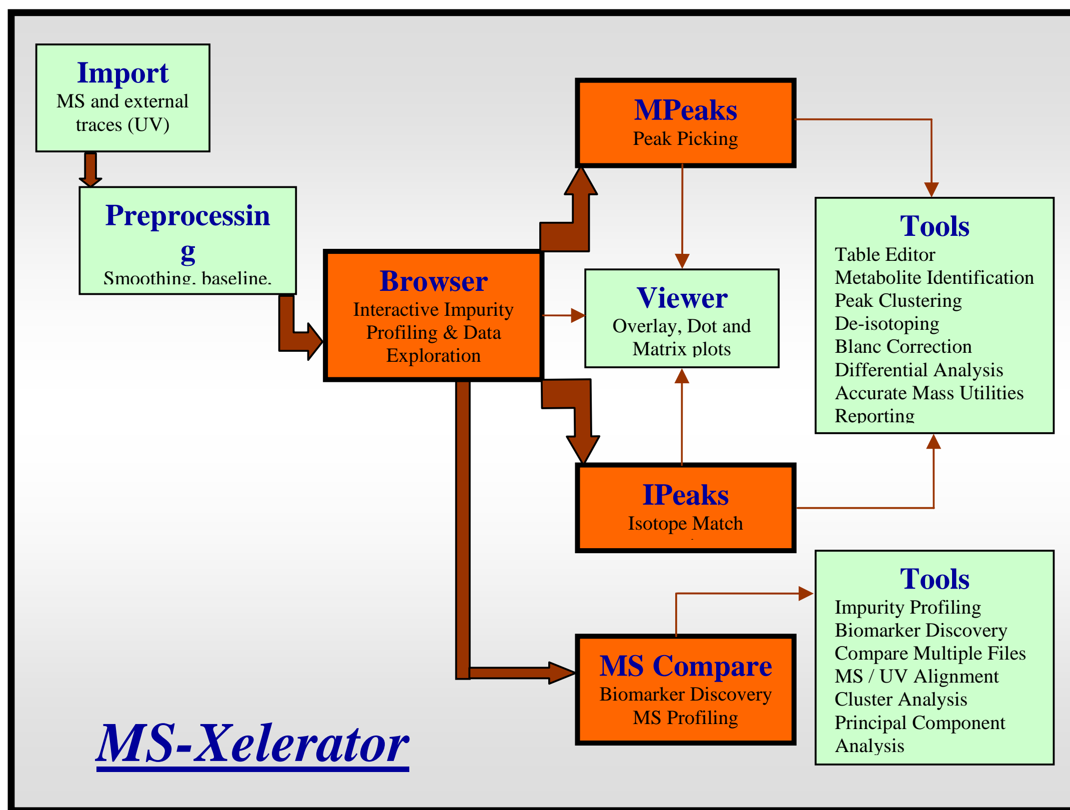
Accelerating Data Analysis in LC-MS Profiling Studies

Due to automation of sample preparation and acquisition, the amount of data from LC-MS experiments, coming from Impurity Profiling, Degradation Profiling, and Metabolite Profiling, has increased extremely. The processing of data can be divided in two parts. The first step consists of finding all significant components in the data. The second step involves the use of additional information to pick out relevant components from the list of significant ones. Especially the part of finding significant components in a data set is a time consuming task, which is strongly dependent of the problem and the knowledge of the analyst involved. Many significant peaks however are not relevant (^{13}C isotope peaks, adduct peaks, matrix components, etc., etc.). These non-relevant peaks will have to be removed using dedicated algorithms.

MS-Xelerator is an extremely fast and comprehensive software product for peak finding and data processing of LC-MS, GC-MS and CE-MS data sets in the area of:

- **Impurity Profiling**
- **Degradation Profiling**
- **Metabolite Profiling**
- **Differential Analysis**
- **Metabonomics**
- **Quantitative Proteomics &**
- **Biomarker Discovery**

The program's unique algorithms are unparalleled in speed, sensitivity and ease of use. MS-Xelerator offers a large number of filtering and visualization possibilities directly interacting with your data.



MS-Xelerator Overview

Fast and powerful algorithms for peak picking:
Most LC-MS data sets can be analyzed with MS-Xelerator in less than 10 seconds, after which the user will have a number of unique filter operations to enhance the search for relevant peaks and components.

Unique Filter Operations:
A large number of filter operations can be used to further reduce the data, e.g.: Peak list editing and sorting, de-isotoping of peaks, clustering of fragment ions into components, blanc/reference correction, accurate mass filters, Na & K adduct identification, isotope matching (Cl, Br, ^{13}C, etc), identification of Metabolite and Degradation products, and construction of user specified peak lists. Charge State Calculation and determination of monoisotopic peaks.

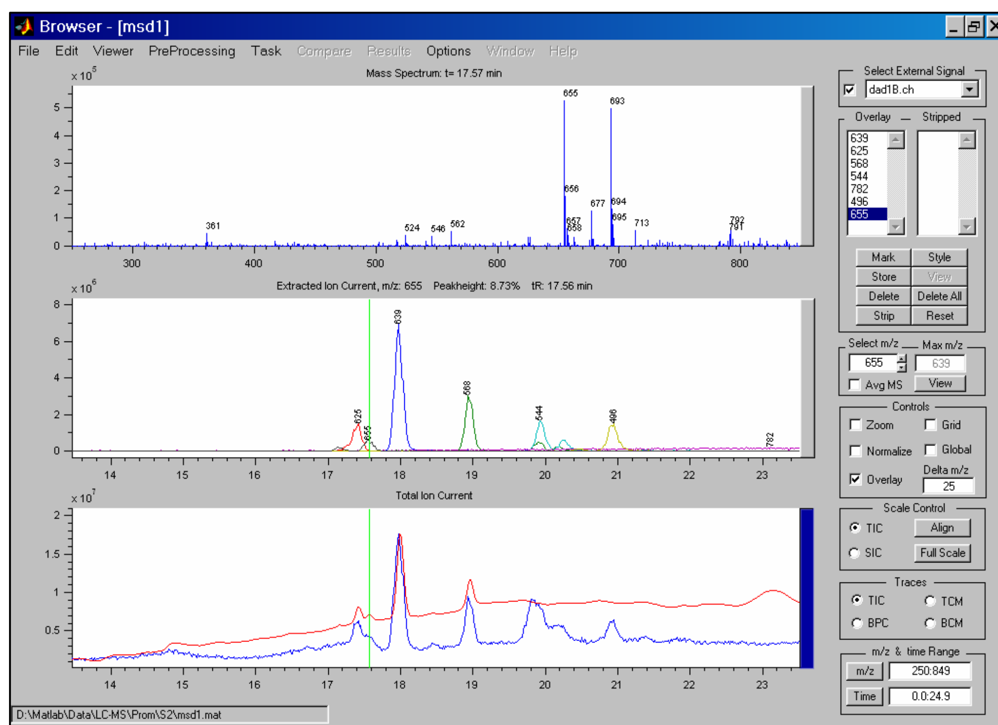
Data Import Formats

MS-Xelerator can import the following data formats: NetCDF, ASCII, Thermo Xcalibur, Waters Masslynx, Agilent MSD*, and mzXML. *Galactic conversion utility required.

The Browser

The Browser Module consists of a very fast interactive graphical environment to explore mass chromatograms and mass spectra based on either the Total Ion Current or the Base Peak Chromatogram. Different mass chromatograms can be overlaid, normalized, rescaled and at the same time compared to external signals, e.g. UV-traces.
The module is especially useful for the exploration of samples having not too much peaks, e.g. Impurity Profiling and Peak Purity Analysis. Two features greatly enhance the search for small peaks that cannot easily be recognized from the Total Ion Current: ***Peak Stripping*** and ***Local Screening Exploration*.**

Peak Stripping subtracts any mass chromatogram from the Total Ion Current. By subtracting the larger peaks, the smaller ones become visible. *Peak Stripping* can be done automatically but also manually. *Local Screening Exploration* is a technique where series of Total Ion Currents are constructed over small m/z ranges. Small components that were hidden in your data now become visible. A Local Screening Exploration can be done in less than a minute and can be guided by external traces, e.g. UV-signals. Detected peaks can be stored to a list, and exported to the Viewer for further visual exploration.



Browser Module showing TIC in overlay with UV-signal (bottom), extracted mass chromatograms in overlay (middle), and the mass spectrum of the selected scan (top)

Data Preprocessing

MS-Xelerator has a number of preprocessing utilities for data enhancement. If your data are influenced by noise, spikes and baselines, preprocessing can greatly improve the quality of your data making peak detection much easier, in particular when using the Browser.

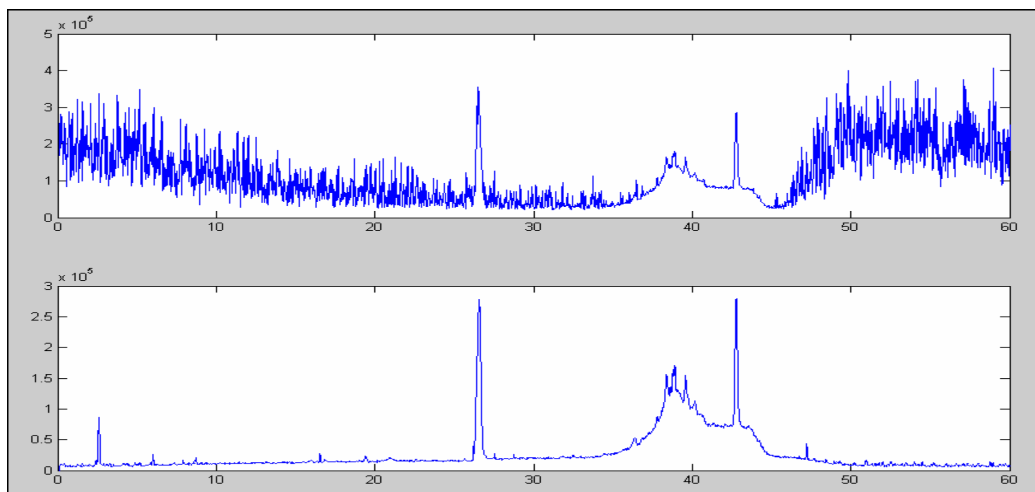
- ***Smoothing:*** the program offers two algorithms for smoothing: Savitsky-Golay smoothing and Gaussian smoothing.

- ***Baseline correction:*** noisy ions with high baseline levels can lead to very noisy Total Ion Currents. Subsequently, peaks will be difficult to detect based on the TIC. The program offers three fast algorithms for baseline correction: the Peak Filter Method, a Minimum Envelope Technique and a Running Median Filter baseline correction routine. The MPeaks algorithm for peak finding is not influenced by horizontal baselines.

- ***Synchronization:*** differences in elution time between detectors - e.g. MS versus UV – can make a mutual comparison difficult. MS-Xelerator offers an easy to use *Synchronization Module* for data alignment.

- ***Mass Defect Filter:*** by removal of ions with m/z values outside certain decimal ranges much background material and non-relevant ions are removed. Applying a Mass Defect Filter can simplify the analysis to a great extend, especially in the case of Metabolite Profiling studies.

- ***De-Spiking:*** spikes in the data can completely dominate the Total Ion Current, making peak detection extremely difficult. MS-Xelerator has a very fast de-spiking algorithm eliminating all spikes in the data set. Total Ion Currents become simpler and data exploration will be much easier. The MPeaks algorithm has a build-in spike-detection utility.

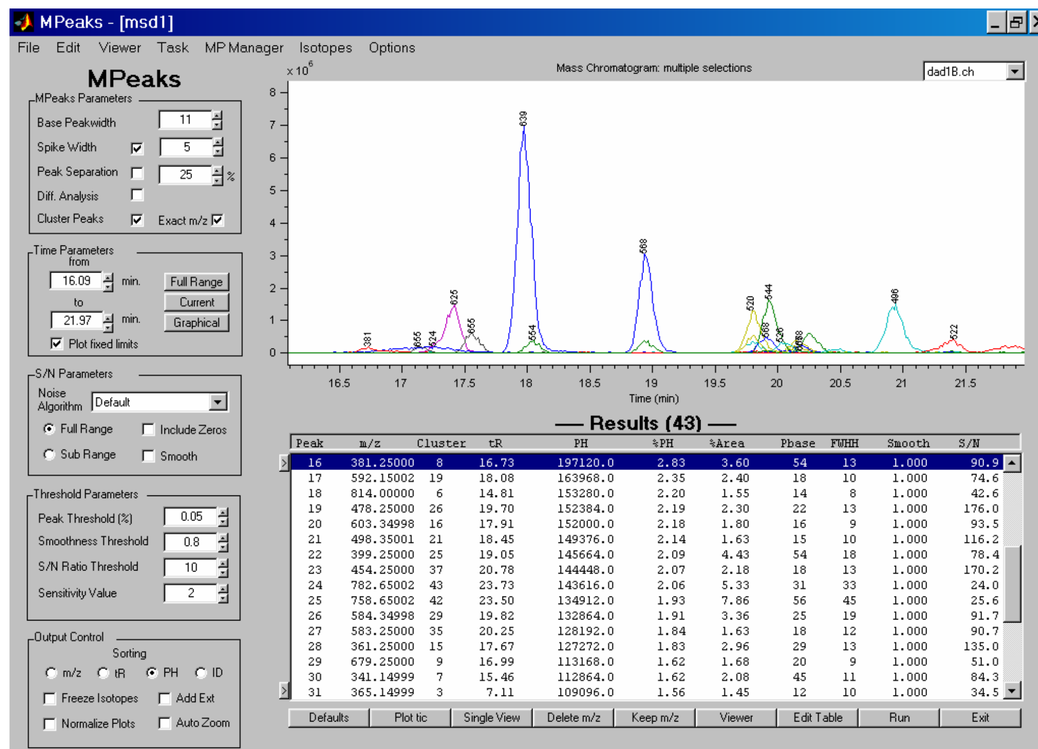


Spike Filtering Example: Comparison of raw TIC (top) with spike-filtered TIC (bottom)

MPeaks

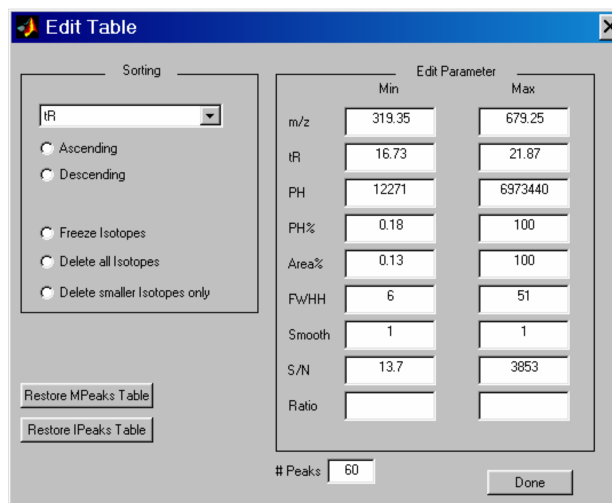
The MPeaks algorithm is the engine of the program. Basically the algorithm finds all significant peaks in your data. To convert significant peaks into relevant peaks a number of unique filter and presentation modes are available. The algorithm has been optimized for speed and sensitivity and runs in both nominal and accurate mass mode. Experience has shown that **almost any data set can be processed in less than 10 seconds**. Although a number of easy to understand parameters can be changed to tune the algorithm, basically only two of these are necessary to get started. The parameters are optimized for chromatographic and MS users. Even users inexperienced to the field of LC-MS will have no problems running the algorithm.

MPeaks outputs a compact table with a number of characteristics describing each peak (m/z value, retention time, peak height, % peak height, % area, etc.). Tables can be stored, printed, loaded and edited. The MPeaks module has its own specialized interactive plots, but results can also be exported to the Viewer.



MPeaks Module showing peak height sorted results from a default run

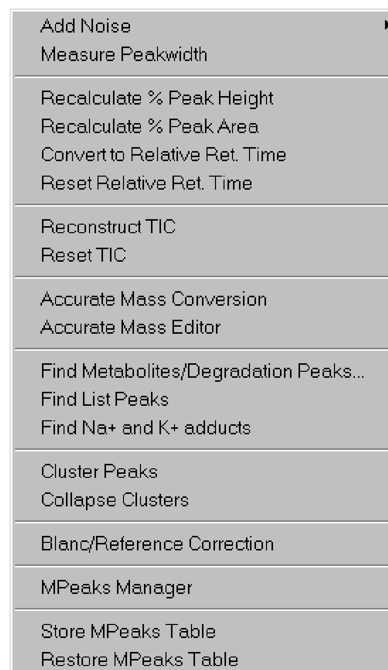
A large number of filter operations and utilities are available that help the user to search for relevant results and to give guidance to the interpretation of the results.

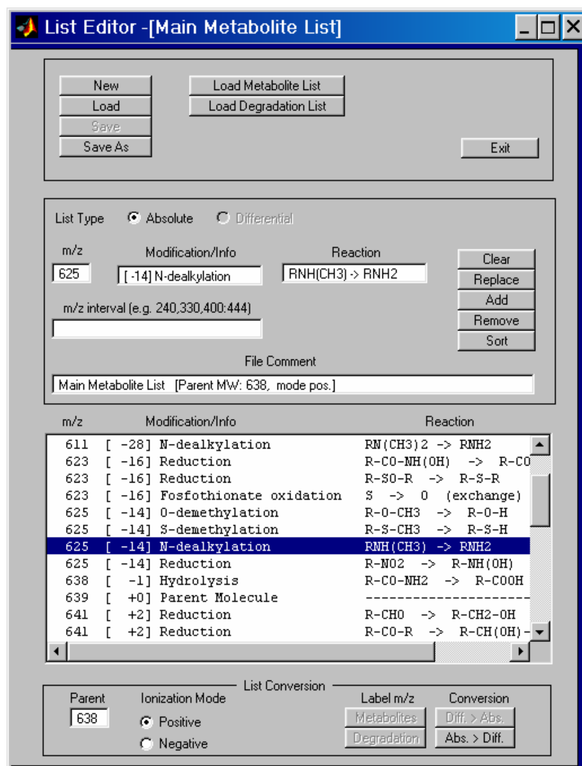


- **MPeaks Editor:** results can be sorted on any of the calculated peak characteristics. Peaks can be deleted based on easily constructed filters, e.g. find all peaks eluting between 10 and 20 minutes having peak heights larger than 0.2%, relative to the largest peak found.
- **^{13}C Isotope peak finder:** automatically delete all ^{13}C Isotope peaks from the result table.
- **TIC Reconstruction:** reconstruct the TIC based on selected mass chromatograms.
- **Calculate Charge States and Monoisotopic Mass Peaks**
- **Differential Analysis using auto-alignment:** very sensitive comparison of 2 samples.
- **Blanc/Reference Correction:** correct main table for peaks that are also present in your blanc or reference sample (label and/or delete blanc peaks).

- **Na & K adduct peak finder:** automatically find all Na and K adducts in the table (label and/or delete adducts).
- **Relative Retention Time:** convert retention times to relative retention times based on a selected peak.

- **Cluster Peaks, NEW!!** many of the peaks found are actually fragments or ^{13}C isotopes. The cluster algorithm finds all peaks that belong to each other. These clusters can be labeled or you will have the possibility to retain the largest peak within each cluster (Convert Peaks to Components).
- **Isotope Matching:** find peak pairs that comply with predefined isotope patterns.
- **Accurate Mass Utilities:** Perform accurate peak search based on exact mass values and subsequently create special filters designed to operate on accurate mass values.
- **Metabolite and Degradation labeling:** the table is searched and labeled for metabolites and degradation products based on differential prediction lists.
- **List Peak Identification:** search the table based on your own created peak lists.
- **Batch Processing:** automate the above steps using the Post Process Manager on many samples simultaneously.





Identification Lists: prediction lists can be edited and stored based on user requirements. Entering the molecular weight of the parent molecule converts the list to a prediction list that can be used for automatic labeling of the peaks found by MPeaks.

The List Editor has predefined Metabolites and Degradation lists. User defined lists can be created within the list editor or can be imported from any text file.
Lists can be created in nominal or accurate mass mode. MPeaks can search and identify peaks from both nominal and accurate mass lists. **If your mass spectrometer operates in accurate mass mode, you can use accurate peak identification!!**

Tracking new components in your sample (Differential Analysis) can be performed by comparison of a user defined peak list with peaks found by MPeaks.

List Editor showing Metabolite ID List

— Results (43) —

Peak	m/z	cluster	tR	PH	%PH	%Area	Metabolite ID MW(638)
1	639.25000	17	17.97	6973440.0	100.00	100.00	[+0] Parent Molecule
2	625.34998	12	17.42	1476096.0	21.17	23.55	[-14] O-demethylation
3	655.25000	13	17.56	608768.0	8.73	9.27	[+16] Hydroxylation
4	655.25000	10	17.13	213632.0	3.06	2.76	[+16] Hydroxylation
5	607.34998	34	20.18	84800.0	1.22	1.04	[-32] Decarboxylation
6	568.25000	23	18.93	3066368.0	43.97	43.12	
7	544.34998	31	19.93	1630208.0	23.38	20.66	
8	496.35001	39	20.94	1505280.0	21.59	26.01	
9	520.34998	28	19.79	1294336.0	18.56	19.20	
10	568.34998	30	19.91	486976.0	6.98	6.80	
11	522.34998	40	21.41	426688.0	6.12	8.08	
12	554.25000	18	18.05	396800.0	5.69	5.39	
13	526.25000	32	20.02	334272.0	4.79	4.75	
14	533.25000	2	7.03	237504.0	3.41	2.57	
15	558.25000	33	20.18	223616.0	3.21	3.11	

MPeaks Accurate Mass Table with nominal labeling of possible metabolite peaks

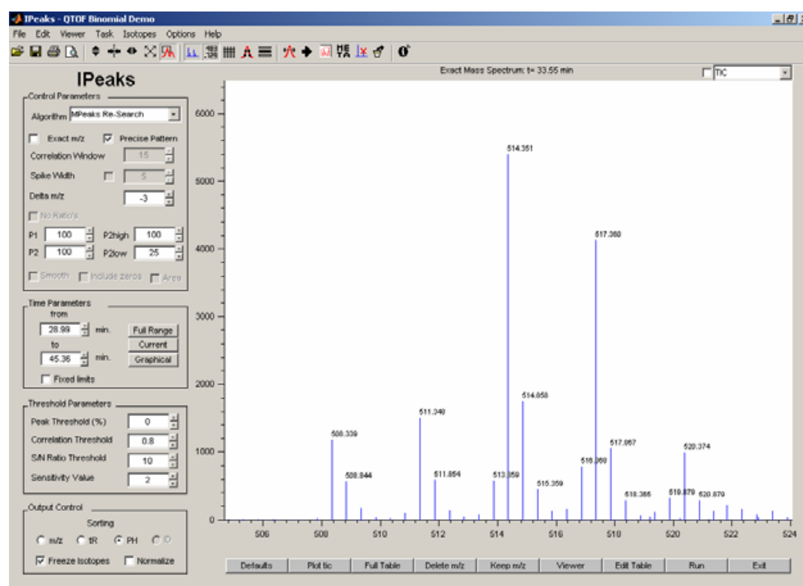
IPeaks: Isotope Pattern Recognition & Quantitative Proteomics

IPeaks is dedicated to finding peaks/components in your data having specific isotopic patterns. These patterns can be "natural" as will be the case in drugs containing e.g. Cl, Cl_2, Br or they may have been introduced by using special stable isotope labeling techniques like: SILAC, ICAT, ECAT, $^{14}N/^{15}N$, $^{16}O/^{18}O$, etc. The latter ones are often used in the field of (Expression) Proteomics.

SILAC: Stable Isotope Labeling for amino Acids in Cell Culture
ICAT: Isotope Coded Affinity Tag
ECAT: Element Coded Affinity Tag
SITE: Stable Isotope Tagging of Epitopes

Peaks having specific isotope patterns or peaks from labelled compounds are found by IPeaks in a fraction of the time compared to a manual screening of the data set. The search will be performed on the raw data. No identification (of peptides) using database searching will be necessary. The IPeaks module contains the following features:

- **Four different Isotope Matching algorithms, some require co-elution others not**
- **Choose from predefined isotope patterns: Cl, Cl_2, Br, ^{13}C, Binomial Distributions, $^{16}O/^{18}O$ labeling, SILAC $^{13}C/^{15}N$, $^{14}N/^{15}N$ labeling**
- **Create user specified Isotope Patterns**
- **Adjust ratio's between isotope peaks; ratio's based on peak height or peak area**
- **Special Isotope filtering algorithms and Plot handling.**



IPeaks Result Screen: detection of Binomial Isotope Patterns having delta m/z=3 spacing

The Viewer

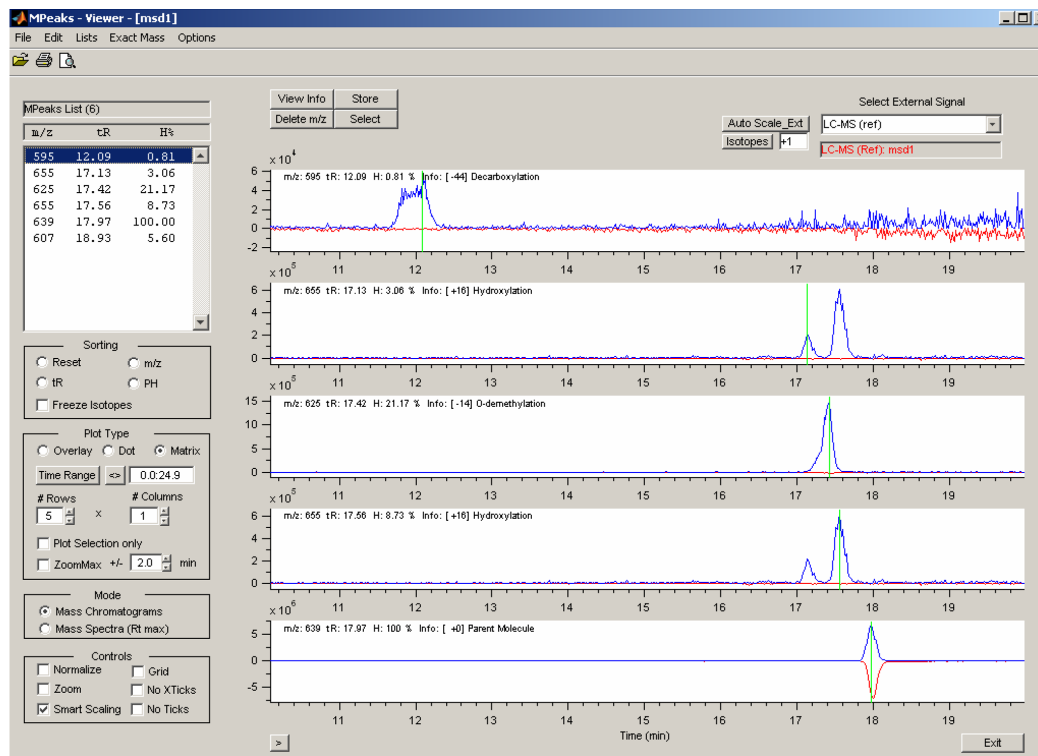
Results from MPeaks, IPeaks or the Browser can be exported to the Viewer, a dedicated module for visualization of LC-MS data. Mass Chromatograms or Mass Spectra can be viewed in three different plot modes: Overlay Plots, Matrix/Line Plots and Dot Plots. Series of Mass Chromatograms can easily be sorted on m/z value, retention time and peak height. Many of the plots are interactive and offer possibilities for smart scaling and overlays. Any view of your data set can be combined with other data, e.g. UV-traces, TIC/BPC traces, Isotope-traces or blanc and reference samples.

Within the Viewer lists of specific m/z values can be created, edited, commented and stored for later use (compare multiple samples). For identification and exploration of Metabolites and Degradation products the Viewer offers so-called predefined Prediction Lists. Based on these lists and the molecular weight of your parent molecule the data file is screened to find and identify possible metabolites and degradation products.

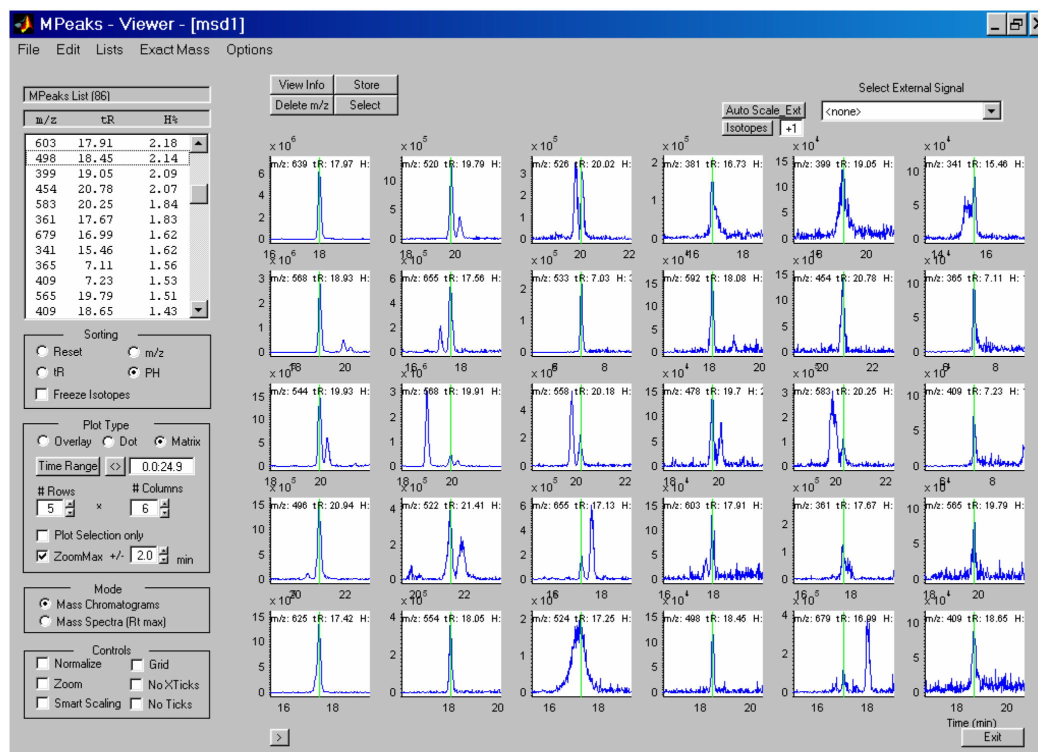
- **Overlay Plots:** plotting of single or multiple mass chromatograms or mass spectra. Create overlays with external signals, normalization, smart zooming, etc.

- **Matrix/Line Plots:** plotting of mass chromatograms or mass spectra in matrix mode, up to 100 plots. Plots can be overlaid with external signals, mass chromatograms from different samples or with isotope peaks. One column line plots (e.g. 5x1) are especially easy to use in combination with retention time sorting. All plots will line up with identical time scales.

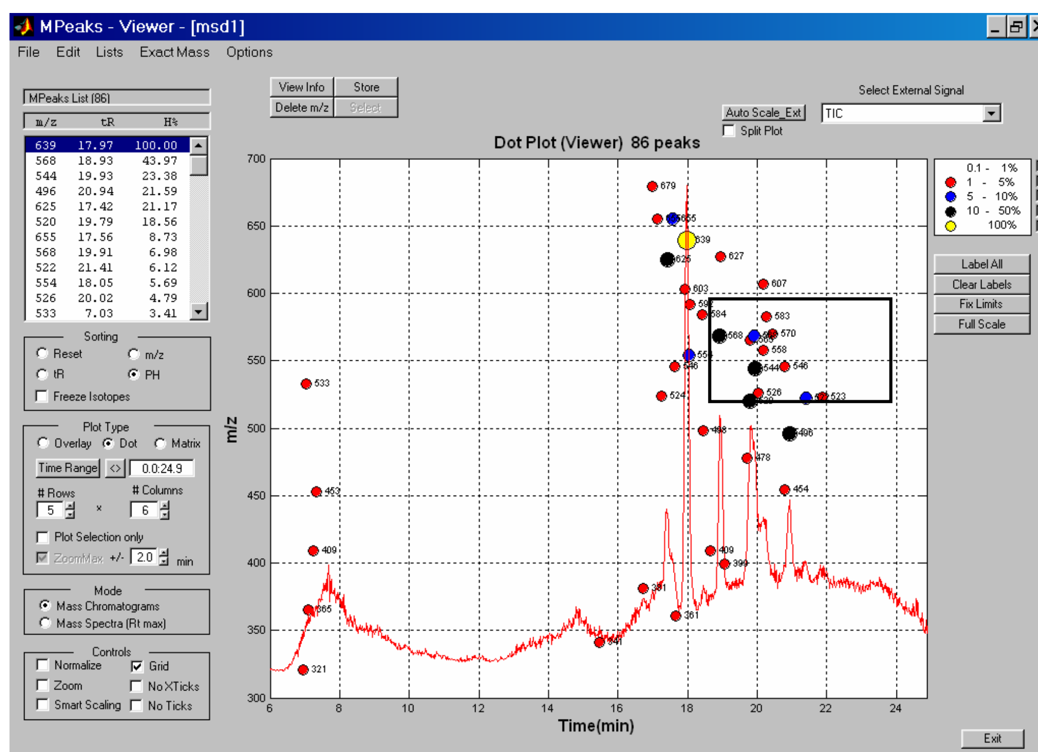
- **Dot Plot:** a dot plot is actually a 3-dimensional representation of your data set. Plotted are the peaks along the time and m/z direction. Peaks are represented by color-coded dots of which the size and color correspond to the peak height level. Peak height plot levels can be turned on and off. The dot plot can be overlaid with e.g. the TIC or the UV-signal and the peaks can be interactively labeled and viewed in any of the other plot modes on the fly.



MPeaks Viewer – Line Plot (5x1) in which the peaks are sorted by retention time and plotted with identical time axes, possible metabolites are identified and labeled in each subplot. Results are from a Differential Analysis between two samples. Components from second sample are plotted in red (major component excluded from analysis)



MPeaks Viewer – Matrix Plot (5x6), peaks sorted by peak height, smart scaling applied



MPeaks Viewer – Dot Plot in overlay with Total Ion Current (0.1 – 1.0 % level excluded). Each dot represents a peak found by MPeaks; dots are labeled by m/z value. Individual peaks or a group of peaks can be marked and viewed in overlay or matrix mode.

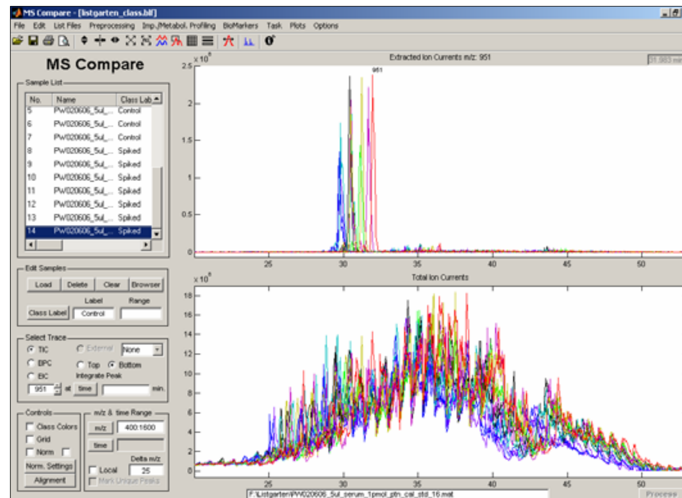
MS Compare: Biomarker Discovery - Metabonomics

Within MS Compare you can **compare series of LC/MS samples** coming from impurity profiling, metabolite profiling or biomarker discovery studies. The comparison can be done using mass chromatograms or mass spectra. Groups of samples (classes) can be colored to more easily detect differences between the groups. After a manual screening, important peaks can be stored to so-called profile tables. Based on these tables a number of different views can be constructed to either see similarities or differences between selected samples (Heat maps, 3D views, Local Screening Plots, Profile Plots etc). The tables are also used to perform more advanced chemometric techniques like **Cluster Analysis and Principal Component Analysis (PCA)**.

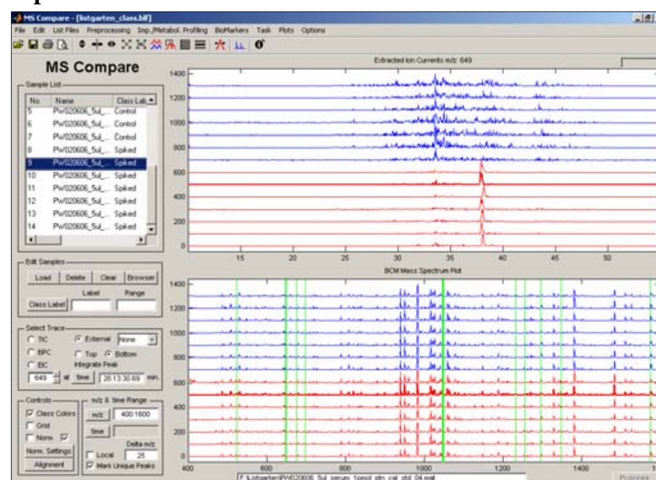
A number of preprocessing methods are available to enhance the quality of your data. Additionally, four alignment algorithms can be used to correct for retention time shifts between the samples: Offset Shifting, Cross Correlation, Correlation Optimized Warping and Peak Based Warping.

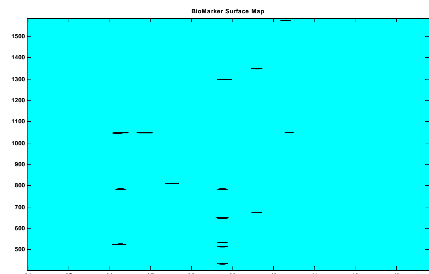
The MS Compare module is linked to all other modules in MS-Xelerator. Samples can be linked to MPeaks to do an in-depth full analysis. The MS Compare module also contains a **Biomarker Discovery** algorithm, optimized to perform automated two-class comparisons. Whenever you have two sample classes, e.g. control vs. disease or normal vs. treated, the Biomaker algorithm will find unique peaks discriminating the two classes. These peaks can be visualized in so-called Biomarker Surfaces Maps or exported to tables.

Comparison of multiple samples before alignment correction. Bottom: overlay of TIC's. Top: overlay of extracted ion currents showing severe alignment problem.

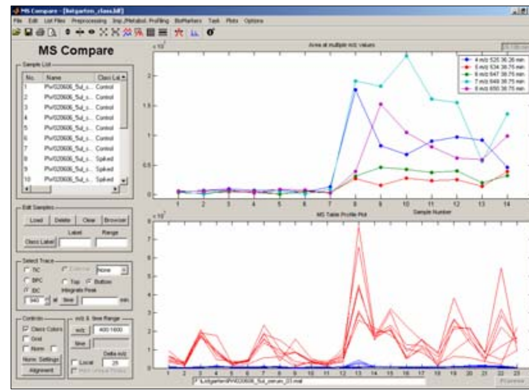


Direct comparison of Mass Spectra from two classes of samples. Unique features will be automatically identified. The unique features can be displayed in a 2D Biomarker Surface Map and exported to tables.

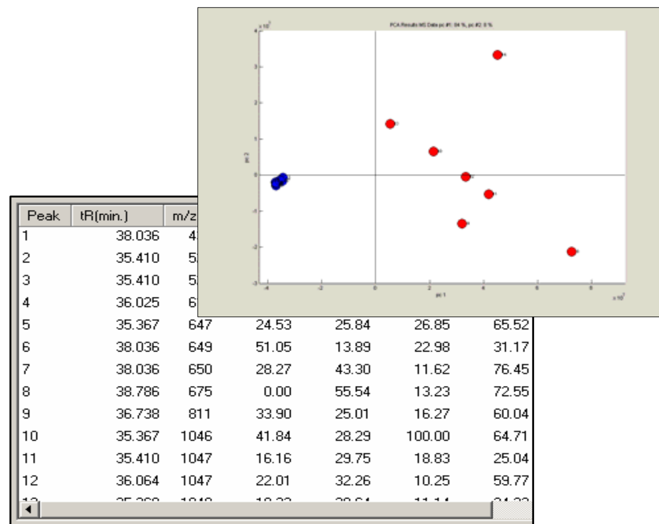




Results from Biomarker Discovery can be explored using profile plots (top – level of selected peaks versus sample number) or line plots (plot peak area for each selected peak for all detected peaks.



Convert results to tables and apply Principal Component Analysis to get a 2-dimensional overview of samples (detect clusters, outliers etc)



Peak	tR(min.)	m/z				
1	38.036	4				
2	35.410	5				
3	35.410	5				
4	36.025	6				
5	35.367	647	24.53	25.84	26.85	65.52
6	38.036	649	51.05	13.89	22.98	31.17
7	38.036	650	28.27	43.30	11.62	76.45
8	38.786	675	0.00	55.54	13.23	72.55
9	36.738	811	33.90	25.01	16.27	60.04
10	35.367	1046	41.84	28.29	100.00	64.71
11	35.410	1047	16.16	29.75	18.83	25.04
12	36.064	1047	22.01	32.26	10.25	59.77

MS-Xelerator™

MsMetrix

Accelerating Data Analysis in LC/MS