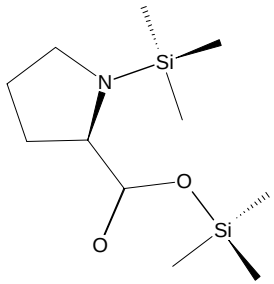
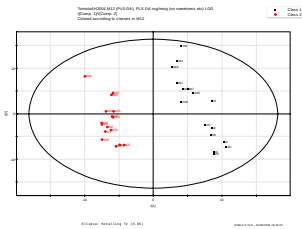
Metabolomic analysis of plant and plasma samples at UPSC

Homogenization » Extraction » Derivatization » Analysis » Data evaluation »

An overview of GC/MS based metabolomics strategies at UPSC



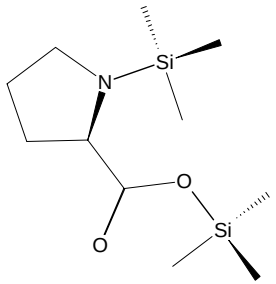


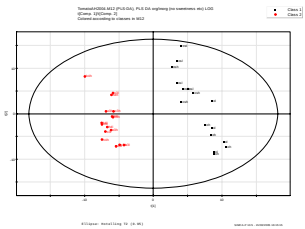
Authors of presentation
Krister Lundgren (krister.lundgren@genfys.slu.se)
Jonas Gullberg
Thomas Moritz (thomas.moritz@genfys.slu.se)

UPSC

Metabolomic analysis plant samples

Homogenization » **Extraction »** **Derivatization »** **Analysis »** **Data evaluation »**





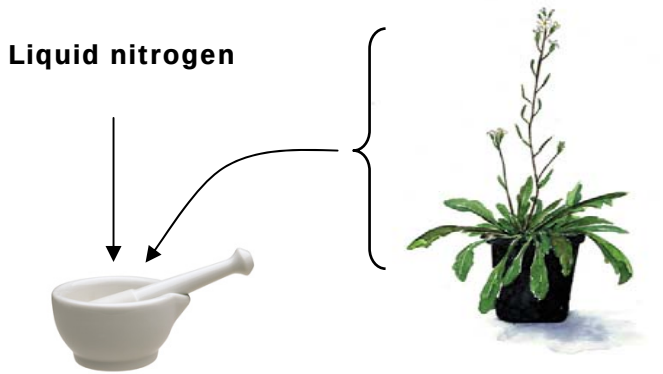
UPSC

Metabolomic analysis plant samples

Homogenization » Extraction » Derivatization » Analysis » Data evaluation »

To stop and “freeze” the metabolome the plant is covered with liquid nitrogen.

The plant is grinded in liquid nitrogen to homogenize.



UPSC

Metabolomic analysis plant samples

The metabolic composition in plants is a wide spectra of compounds regarding their physical properties and different functional groups.

To cover this range different solvents have to be used. For polar metabolites methanol, ethanol and water is often used, and for more lipophilic compounds chloroform is most commonly applied.

Therefore a mix of water, methanol and chloroform is used for the extraction. This mix is one homogenus phase.

The extraction is performed using an vibration mill.

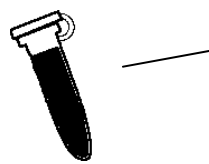
Add Extraction Mix:

-Water (20%v/v)
-Methanol (60% v/v)
-Chloroform(20% v/v)
-Internal standards (7,5 ng/µl)

plant ”powder”

~10 mg

Powder



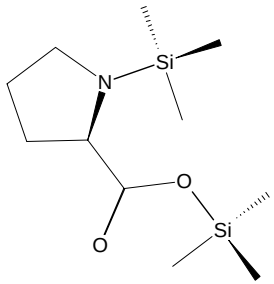
Vibration mill

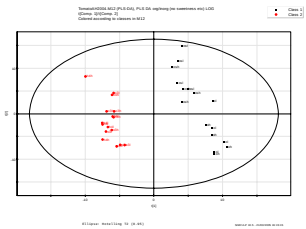
After the extraction, the supernatant is used in the derivatisation step.

UPSC

Metabolomic analysis plasma

Extraction » **Derivatization »** **Analysis »** **Data evaluation »**





Metabolomic analysis plasma

Extraction » Derivatization » Analysis » Data evaluation »

The metabolic composition in plasma is a wide spectra of compounds regarding their physical properties and different functional groups.

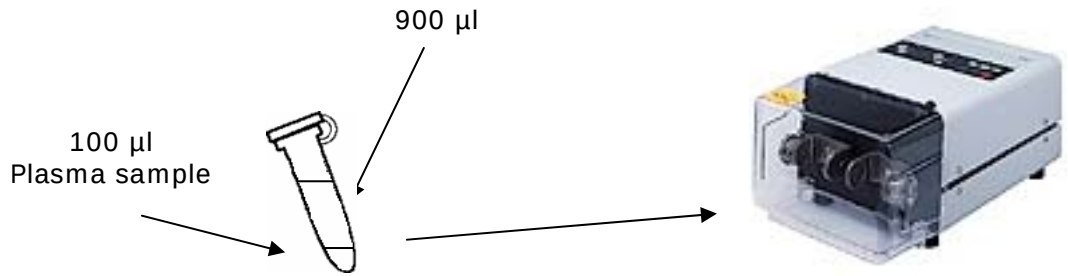
To cover this range different solvents have to be used. For polar metabolites, water is used, and for more lipophilic compounds methanol is used.

A mix containing of water and methanol and 11 internal standards is added for the extraction step.

The extraction is performed using an vibration mill.

Extraction Mix:

-Water (10%v/v)
-Methanol (90% v/v)
-Internal standards (7 ng/µl)



Vibration mill

After the extraction, the supernatant is used in the derivatisation step.

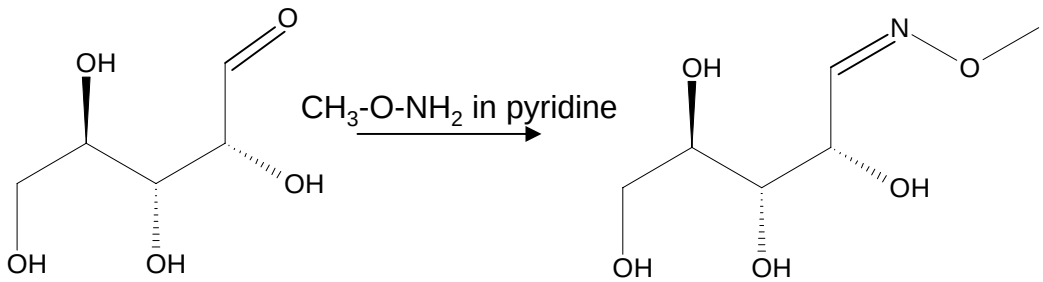
UPSC

Metabolomic analysis

Extraction » **Derivatization »** Analysis » Data evaluation »

Since only volatile and thermo-stable compounds can be analysed by GC/MS derivatisation is performed if GC/MS is to be used to analyse metabolites that are insufficiently volatile in their native state. The standard procedure in plant metabolomics is to first derivatise them using methoxyamine (CH_3-O-NH_2) in pyridine to stabilize carbonyl moieties in the metabolites, thereby suppressing keto-enol tautomerism and the formation of multiple acetal- or ketal-structures. Methoxyamination helps to reduce the numbers of derivatives of reducing sugars, and generates only two forms of the –N=C< derivative, (syn and anti-forms).

Methoxyamination



Metabolomic analysis

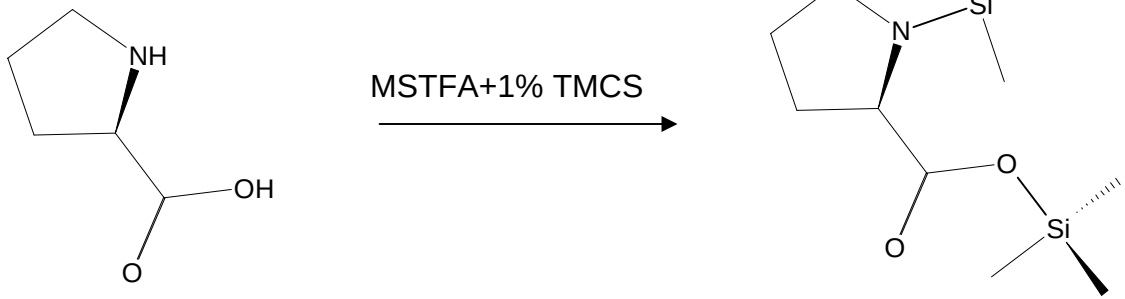
Extraction » **Derivatization »** Analysis » Data evaluation »

After methoxyamination functional groups, such as –OH, -COOH, –SH or –NH groups, are converted into TMS-ethers, TMS-esters, TMS-sulfides or TMS-amines, respectively, using a trimethylsilyl (TMS) reagent, usually N-methyl-N-trimethylsilyl-trifluoroacetamide (MSTFA)

Examples

Alcohols
Amino acids
Amines
Acids
Fatty acids
Sterols
Sugars

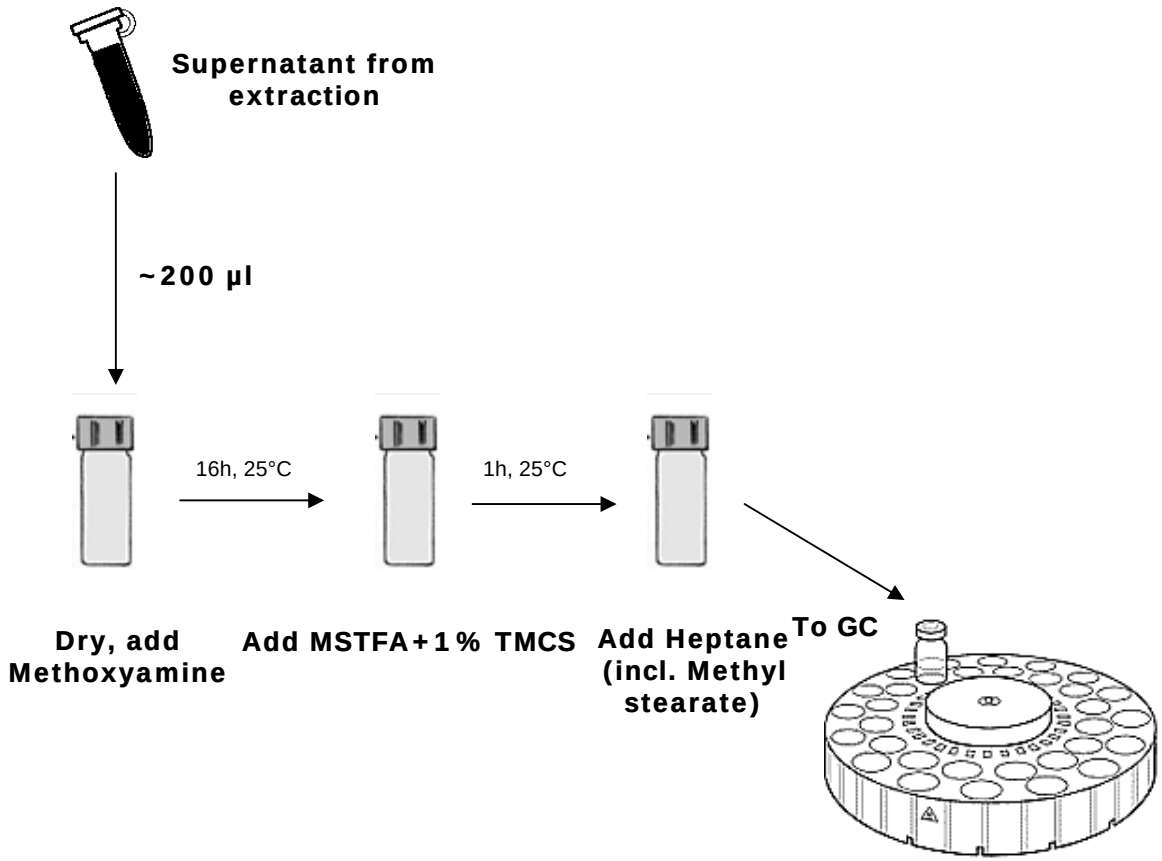
Silylation



Metabolomic analysis

Extraction » **Derivatization »** Analysis » Data evaluation »

The procedure of derivatization



UPSC

Metabolomic analysis

Extraction » Derivatization » **Analysis »** Data evaluation »

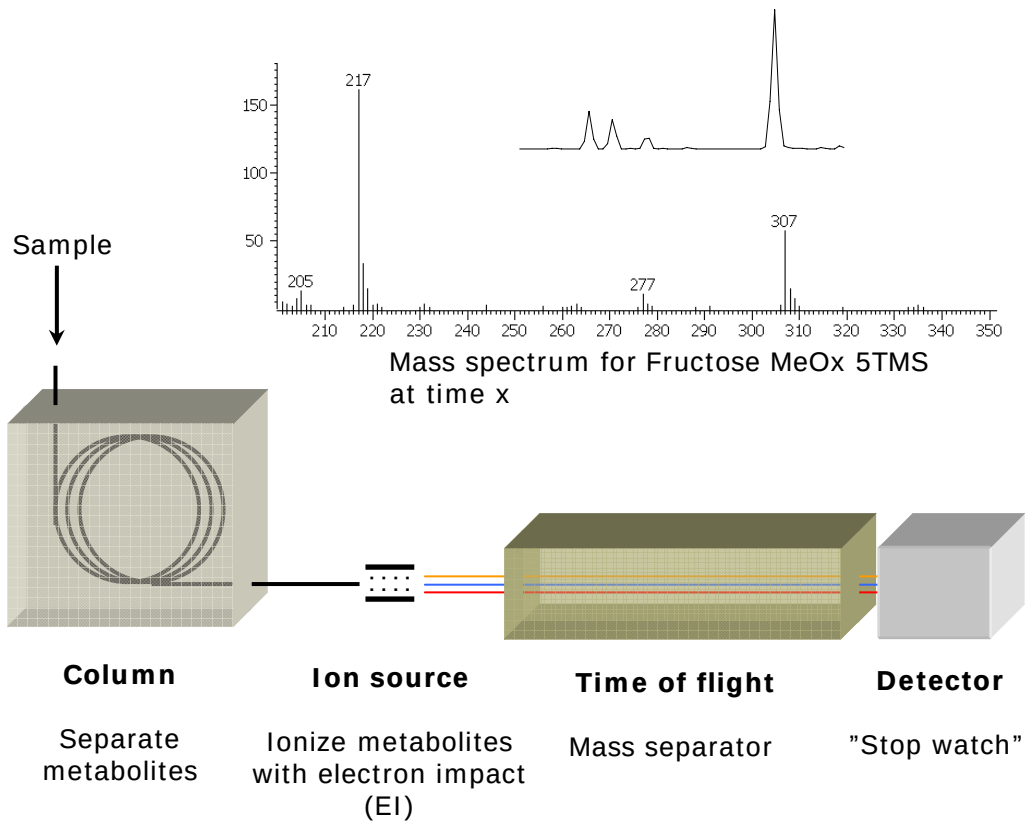
The derivatives of the metabolites are then separated on a GC column and analysed using a time of flight (TOF) mass spectrometer.

All the analyses are performed using LECO's GC-TOF MS Pegasus III.

GC program:

Splitless mode
70°C for 1 minute
70°C to 320°C, 40°C/min
320°C for 2 minutes
GC-column: 10m x 0.18mm ID; 0.18µm DB5-MS

Spectra accumulation (MS): 30 spectra/sec.
Total analyse time including cooling and stabilization: ~15 min



UPSC

Metabolomic analysis

Extraction » Derivatization » Analysis » **Data evaluation »**

The data evaluation is performed in four steps:

1) Calculation of internal standards (LECO ChromaTof™ or In-house script)

2) Hierarchical Multivariate Curve Resolution, H-MCR (In-House script, Matlab)

3) Multivariate evaluation (SIMCA)

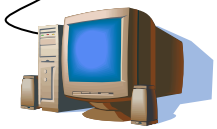
4) Identification of significant compounds, using mass spectra library databases. (NIST, public available and In-House)

GCTOF-computer

SMP-export

NetCDF-export

Ext. Harddrive



Process computer

First the data files are exported in two file-formats (*.smp and NetCDF) from the instrument computer to an external hard drive.

UPSC

Metabolomic analysis

The data evaluation is performed in four steps:

1) Calculation of internal standards

2) Hierarchical Multivariate Curve Resolution (H-MCR)

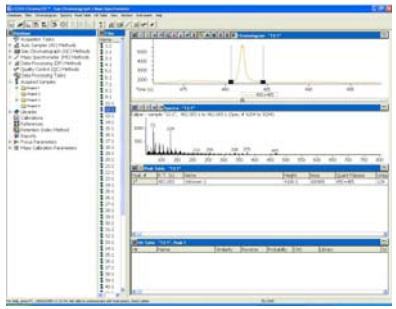
3) Evaluation in SIMCA

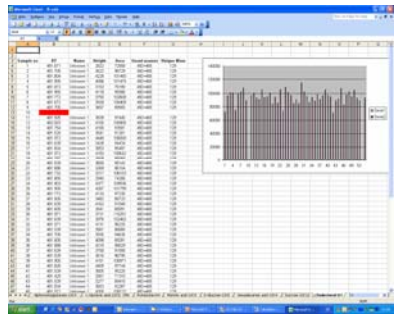
4) Identification of significant compounds.

Calculation of internal standards (LECO ChromaTof™ or in-house script)

Peak areas for each internal standard is calculated for all samples. The relative standard deviation is calculated for each internal standard to check the variation in the data.

11 internal standards are used with different chromatographic and structural properties.





First, this is a control to see that the analysis was ok (RSD% is low enough). Some “outlier” samples are discovered in this step and are removed before the H-MCR-calculation.
Secondly, we use the internal standard data for normalisation in the multivariate analysis (performed after the H-MCR).

UPSC

Metabolomic analysis

The data evaluation is performed in four steps:

1) Calculation of internal standards

2) Hierarchical Multivariate Curve Resolution (H-MCR)

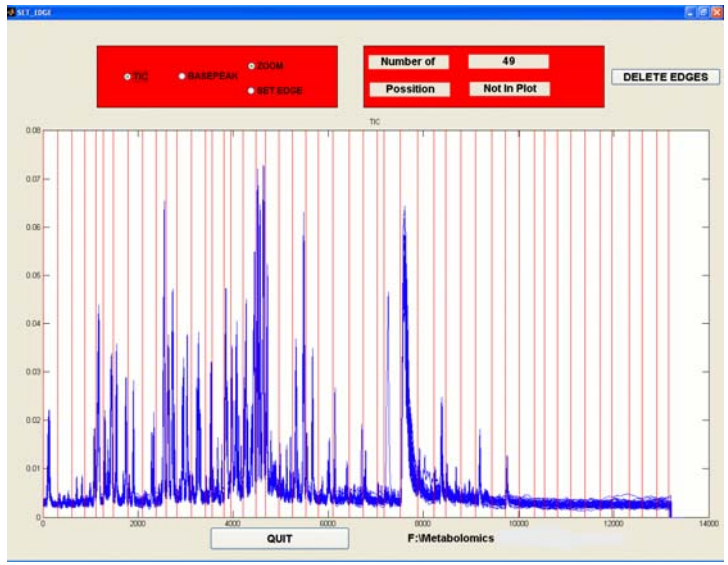
3) Evaluation in SIMCA

4) Identification of significant compounds.

Hierarchical Multivariate Curve Resolution (H-MCR) developed by Pär Jonsson

The data are imported to the in-house matlab script (*.cdf) and the data files are converted to *.mat format.

After smoothing and aligning, all chromatograms are placed upon each other, and divided into time windows by manually or automatically setting edges, where the edge is set at a local minimum.



UPSC

Metabolomic analysis

The data evaluation is performed in four steps:

1) Calculation of internal standards

2) Hierarchical Multivariate Curve Resolution (H-MCR)

3) Evaluation in SIMCA

4) Identification of significant compounds.

Hierarchical Multivariate Curve Resolution (H-MCR)

After the “edge setting” is done, two methods can be used:

Method 1: AR-Compression (Jonsson et al. 2004)	Rapid, and very good for initial classification using PCA or PLS-DA
Method 2: H-MCR (Jonsson et al. 2005, 2006)	Slow, but the outcome is deconvoluted data including peak areas and mass spectra in NIST format (database format).

UPSC

Metabolomic analysis

The data evaluation is performed in four steps:

1) Calculation of internal standards

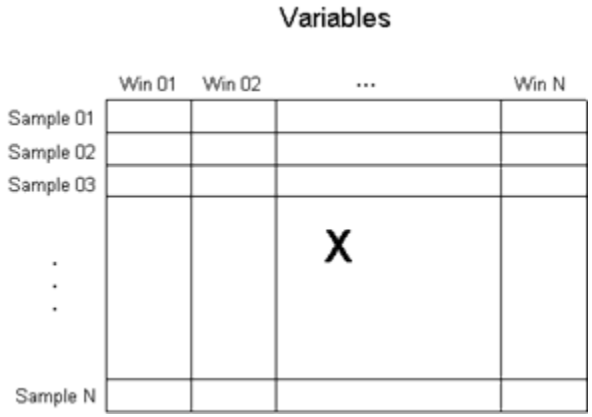
2) Hierarchical Multivariate Curve Resolution (H-MCR)

3) Evaluation in SIMCA

4) Identification of significant compounds.

Hierarchical Multivariate Curve Resolution (H-MCR)*

Both method 1 and 2 produces a data file, containing a matrix that includes variables (time windows) and sample names and their corresponding peak areas.



In method 1, each window give information about the variation for the sum of all mass channels, where Win01_T01 represents the most variation, Win01_T02 represents the second most variation etc.

In method 2, each window represents a specific compound, where Win01_C01 is one detected compound, Win01_C02 is another compound and so on.

*Only method 2 is H-MCR, i.e. deconvolution of peaks.

UPSC

Metabolomic analysis

Extraction » Derivatization » Analysis » **Data evaluation »**

The data evaluation is performed in four steps:

1) Calculation of internal standards

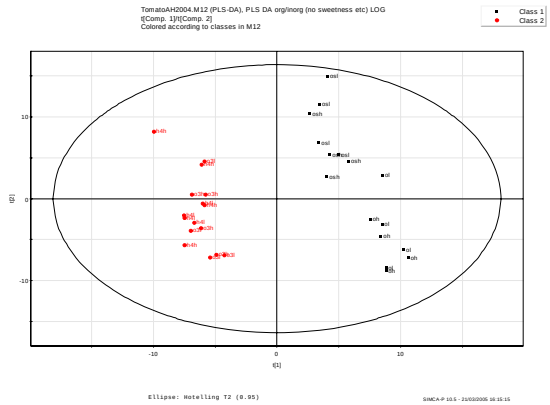
2) Hierarchical Multivariate Curve Resolution (H-MCR)

3) Evaluation in SIMCA

4) Identification of significant compounds.

Evaluation in SIMCA

The data matrices can be evaluated in SIMCA (Umetrics AB, Umeå, Sweden).
First of all classification can be made with PCA and PLS-DA/OPLS-DA (method 1 & 2).



Metabolomic analysis

The data evaluation is performed in four steps:

1) Calculation of internal standards

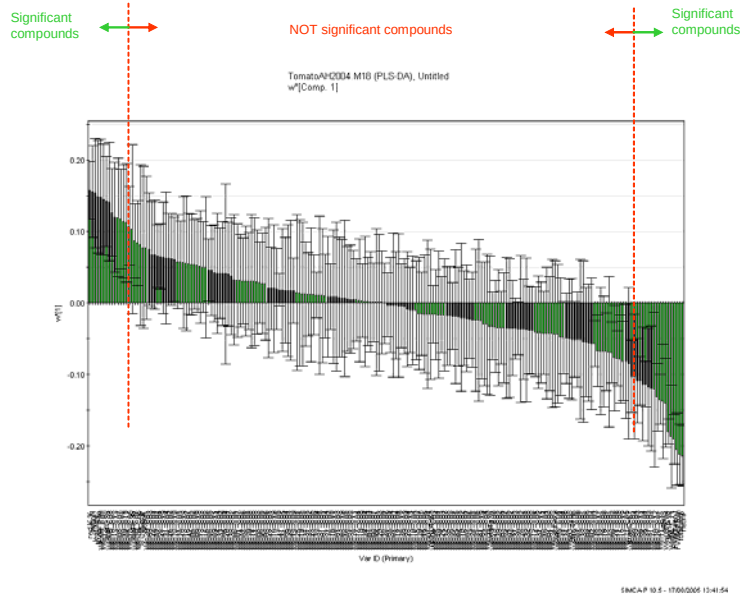
2) Hierarchical Multivariate Curve Resolution (MCR)

3) Evaluation in SIMCA

4) Identification of significant compounds.

Evaluation in SIMCA

Secondly, significant metabolites (WinXX_C0X) are determined using the loadings (method 2).



For each component (putative metabolite) a confidence level bar is calculated. If the confidence bar overlap the zero line, there is no significance. The other compounds are significant and contributes to the separation between the classes.

A list of the significant compounds is easily created in SIMCA. This list is imported into the H-MCR-script for the identification process.

UPSC

Metabolomic analysis

The data evaluation is performed in four steps:

1) Calculation of internal standards

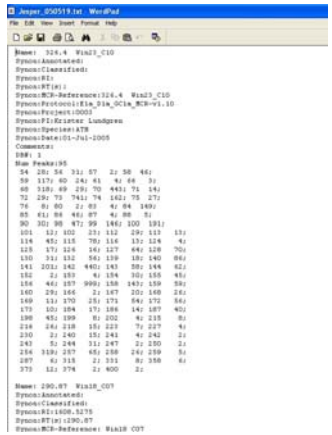
2) Hierarchical Multivariate Curve Resolution (MCR)

3) Evaluation in SIMCA

4) Identification of significant compounds.

Identification of significant compounds

The list of significant compounds is imported into the H-MCR-script and a text file containing mass spectra, retention times, retention index and other information is produced.



Metabolomic analysis

The data evaluation is performed in four steps:

1) Calculation of internal standards

2) Hierarchical Multivariate Curve Resolution (MCR)

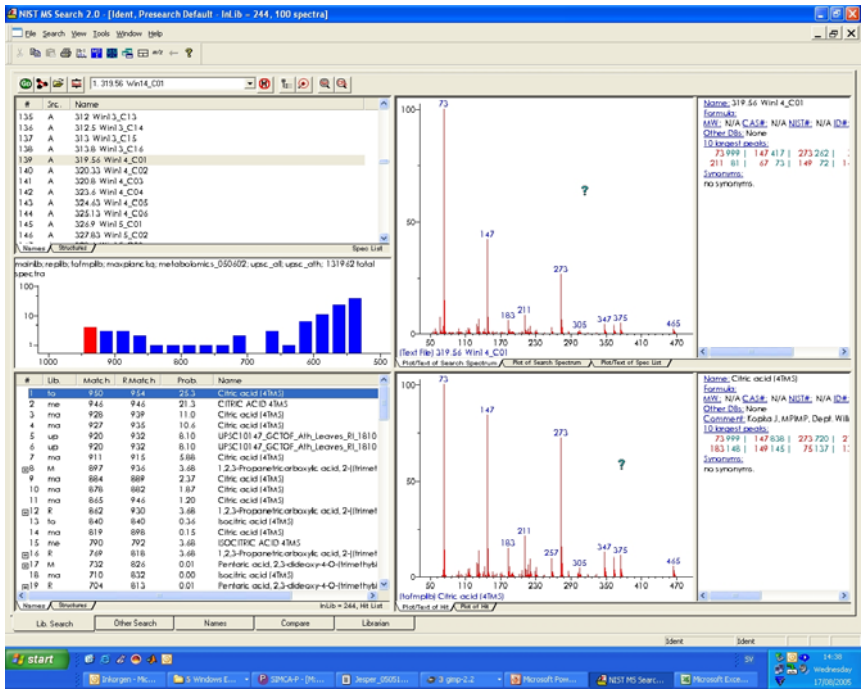
3) Evaluation in SIMCA

4) Identification of significant compounds.

Identification of significant compounds

The text-file containing mass spectral information of several significant metabolites is imported into the NIST database software for identification.

Several databases, including: In-house databases from UPSC and Max Planck Institute (MPI) library in Golm (http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html) and a commercial database from NIST are used.



UPSC

References Metabolomic analysis

Extraction and derivatization

Gullberg J, Jonsson P, Nordström A, Sjöström M, Moritz T (2004) Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of Arabidopsis thaliana samples in metabolomic studies with gas chromatography/mass spectrometry. Anal Biochem 331: 283-295.

A, J.; Trygg, J.; Gullberg.; Johansson, A.I.; Jonsson, P.; Antti, A.; Marklund, S.L.; Moritz, T. Extraction and GC/MS analysis of the human blood plasma metabolome. Anal. Chem. 2005, 77, 8086-8094.

Data processing H-MCR

Jonsson P, Gullberg J, Nordström A, Kusano M, Kowalczyk M, Sjöström M, Moritz T (2004) A strategy for identifying differences in large series of metabolomic samples analyzed by GC/MS. Anal Chem 76: 1738-1745

Jonsson P, Johansson AI, Gullberg J, Trygg J, A J, Grung B, Marklund S, Sjöström M, Antti H, Moritz T (2005) High-throughput data analysis for detecting and identifying differences between samples in GC/MS-based metabolomic analyses. Anal Chem 77: 5635-5642

Jonsson P, Johansson ES, Wuolikainen A, Lindberg J, Schuppe-Koistinen I, Kusano M, Sjöström M, Trygg J, Moritz T, Antti H (2006) Predictive metabolite profiling applying hierarchical multivariate curve resolution to GC-MS data - A potential tool for multi-parametric diagnosis. J Prot Res 5: 1407-1414

Mass spectra library identification

Schauer N, Steinhauser D, Strelkov S, Schomburg D, Allison G, Moritz T, Lundgren K, Roessner-Tunali U, Forbes MG, Willmitzer L, Fernie AR, Kopka J (2005) GC-MS libraries for the rapid identification of metabolites in complex biological samples. Febs Letters 579: 1332-1337

Multivariate analysis

Trygg J, Gullberg J, Johansson AI, Jonsson P, Moritz T (2006) Chemometrics in metabolomics-An introduction. In Plant metabolomics (Ed Saito K, Dixon RA, Willmitzer L) Spinger-Verlag.

Trygg J, Holmes E, Lundstedt T (2007) Chemometrics in metabonomics. J Proteome Res 6: 469-479

UPSC