

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/221848400>

# High-Performance Liquid Chromatography–Mass Spectrometry Analysis of Plant Metabolites in Brassicaceae

**Article** in *Methods in molecular biology* (Clifton, N.J.) · January 2012

DOI: 10.1007/978-1-61779-594-7\_8 · Source: PubMed

CITATIONS

11

READS

530

3 authors:



**Ric De Vos**

Wageningen University & Research

324 PUBLICATIONS 17,306 CITATIONS

[SEE PROFILE](#)



**Bert Schipper**

Wageningen University & Research

17 PUBLICATIONS 769 CITATIONS

[SEE PROFILE](#)



**Robert David Hall**

Wageningen University & Research

222 PUBLICATIONS 11,380 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Anthocyanin production in yeasts [View project](#)



Tomato Research [View project](#)

# Chapter 8

## High-Performance Liquid Chromatography–Mass Spectrometry Analysis of Plant Metabolites in *Brassicaceae*

Ric C.H. De Vos, Bert Schipper, and Robert D. Hall

### Abstract

The *Brassicaceae* family comprises a variety of plant species that are of high economic importance as vegetables or industrial crops. This includes crops such as *Brassica rapa* (turnip, Bok Choi), *B. oleracea* (cabbages, broccoli, cauliflower, etc.), and *B. napus* (oil seed rape), and also includes the famous genetic model of plant research, *Arabidopsis thaliana* (thale cress). Brassicaceae plants contain a large variety of interesting secondary metabolites, including glucosinolates, hydroxycinnamic acids, and flavonoids. These metabolites are also of particular importance due to their proposed positive effects on human health. Next to these well-known groups of phytochemicals, many more metabolites are of course also present in crude extracts prepared from *Brassica* and *Arabidopsis* plant material.

High-pressure liquid chromatography coupled to mass spectrometry (HPLC-MS), especially if combined with a high mass resolution instrument such as a QTOF MS, is a powerful approach to separate, detect, and annotate metabolites present in crude aqueous-alcohol plant extracts. Using an essentially unbiased procedure that takes into account all metabolite mass signals from the raw data files, detailed information on the relative abundance of hundreds of both known and, as yet, unknown semipolar metabolites can be obtained. These comprehensive metabolomics data can then be used to, for instance, identify genetic markers regulating metabolic composition, determine effects of (a)biotic stress or specific growth conditions, or establish metabolite changes occurring upon food processing or storage.

This chapter describes in detail a procedure for preparing crude extracts and performing comprehensive HPLC-QTOF MS-based profiling of semi-polar metabolites in Brassicaceae plant material. Compounds present in the extract can be (partially or completely) annotated based on their accurate mass, their MS/MS fragments and on other specific chemical characteristics such as retention time and UV-absorbance spectrum.

**Key words:** LCMS, Brassica, Arabidopsis, Secondary metabolites, Phenylpropanoids, Flavonoids, Glucosinolates

## 1. Introduction

The *Brassicaceae* family represents a highly interesting and contrasting collection of plants, including a number of major vegetables (e.g., *Brassica oleracea* and *B. rapa*), a source of vegetable proteins and oils (*B. napus*) as well as the model species of plant science, *Arabidopsis thaliana*. Even within *B. oleracea* alone, the diversity of vegetables is remarkable—cabbage, cauliflower, broccoli, Brussels sprouts. This species therefore represents an extremely interesting group of crops to study in the context of their varied biochemical composition and restricted genetic diversity. There is specifically, great interest in a number of secondary metabolite groups and their contrasting profiles within the different sub species of *Brassica*, due to their proposed link with having potentially health-promoting properties (1–3). These secondary metabolite groups include common compounds such as various phenylpropanoids (hydroxycinnamates) and flavonoids, as well as *Brassicaceae*-specific metabolites such as glucosinolates. Within these phytochemicals, there is marked variation between species and cultivars (4–7). However, next to these well-known and frequently analyzed target groups of compounds, *Brassicaceae* plants of course contain many more metabolites. Comprehensive, unbiased metabolomics approaches will reveal much more information on a wider range of metabolites which may also play key roles in other important phenotypic characteristics of *Brassica* plants and their products, such as disease/pest resistance and taste. For example, for this reason, broccoli was chosen as the representative vegetable crop within the EU research project META-PHOR (8) where current technologies and bioinformatics tools are being modified and advanced to increase our ability to generate biologically relevant information relating biochemical profile and phenotype.

The highly rich biochemistry of key secondary metabolites within the *Brassicaceae* covers many compounds that can be efficiently extracted in semi-polar aqueous-alcohol solutions. Of these semi-polar compounds not involved in primary metabolism, quite a number have already been shown to have phenotypic/physiological importance. It is also mainly secondary metabolites through for example their resistance effects, antioxidant properties, and color and flavor characteristics, which are attracting much attention from health, food, and nutrition groups (1–3, 9–11). High Pressure Liquid Chromatography (HPLC), using a  $C_{18}$ -reversed phase column, is the most frequently used and favored technique for analyzing semi-polar compounds present in plant extracts (2, 4, 12, 13). Due to the presence of one or more phenolic moieties, (poly)phenolic compounds and their conjugated forms typically exhibit ultraviolet and/or visible (UV/Vis) light absorbance that can be used for their detection. Many phenolic backbone structures

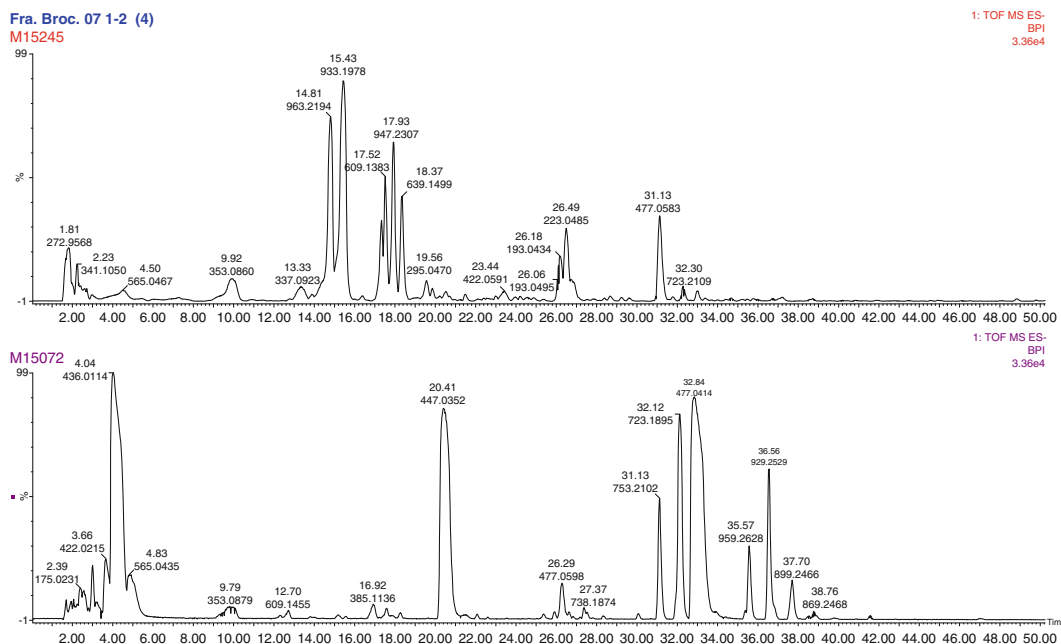


Fig. 1. Representative LC-QTOF MS chromatograms (ESI negative mode) of *B. rapa* Bok Choi leaves (*upper panel*) and *B. oleracea* Broccoli flower head (*lower panel*). The largest peaks represent glucosinolates and flavonoids, which are highly abundant in most *Brassica* species.

show specific absorbance spectra that can be recorded by photodiode array detectors (PDA or DAD) and can be used for (partial) metabolite identification. However, phenolic structures showing similar absorbance spectra, as well as many other secondary metabolites not displaying specific UV/Vis absorbance, need other detection and identification methods. Reversed-phase HPLC, hyphenated with mass spectrometry (HPLC-MS or, in short, LCMS) is nowadays a popular analytical technique for semi-polar compounds and a powerful platform for both targeted and untargeted profiling of plant secondary metabolites (see Fig. 1). High-resolution LCMS with exact mass determination can provide not only comprehensive mass-retention time fingerprints representing hundreds of known and, as yet, unknown metabolites, but in combination with MS-fragmentation abilities (see Fig. 2) and/or UV/Vis detection may also provide valuable structural information of the compounds detected (14–16).

LCMS with electrospray ionization (ESI) can be performed in either negative mode or positive mode, usually depending on the (classes of) compounds of primary interest. Due to their sulfate-glucose core structure, the *Brassicaceae*-specific glucosinolates ionize extremely well in negative electrospray ionization (ESI) mode and, therefore, can be very sensitively detected (17, 18). Other major secondary metabolites usually present in plant materials,

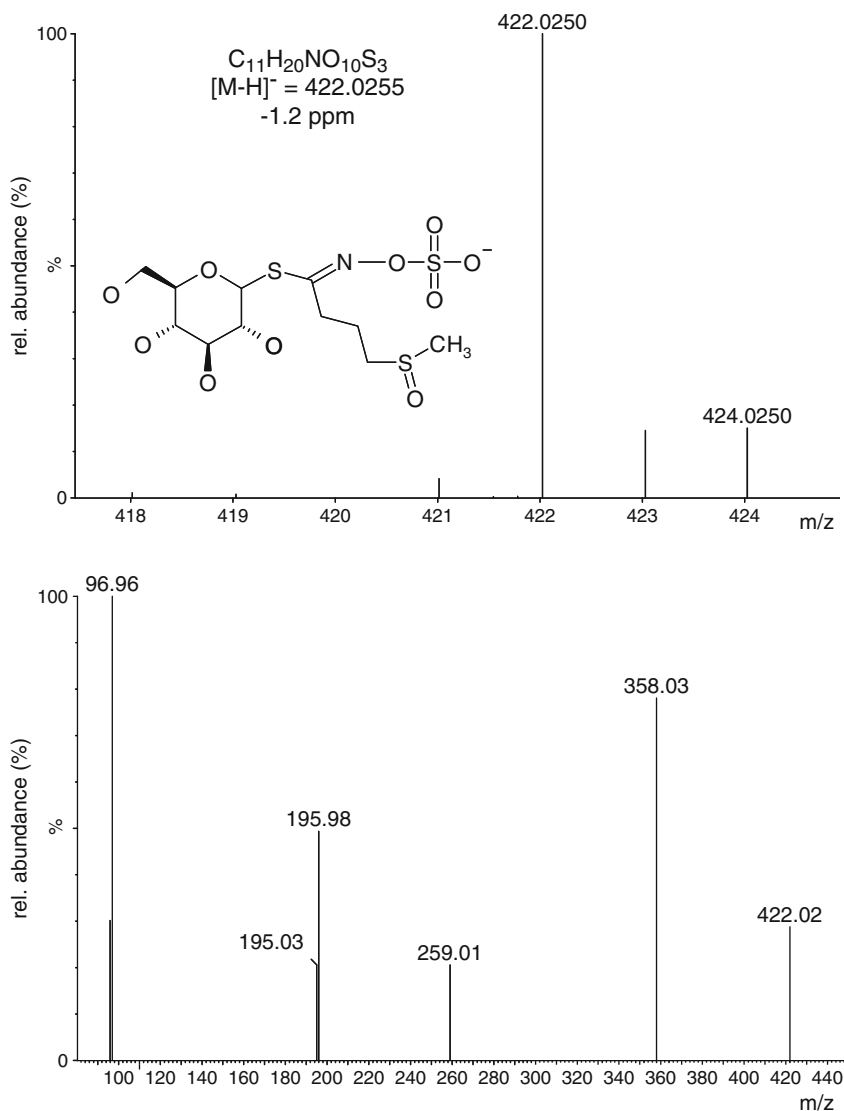


Fig. 2. Accurate mass detection of the molecular ion (*upper panel*) and collision-induced MS/MS fragments (*lower panel*) of glucoiberin (3-methylsulfinylpropylglucosinolate) present in Broccoli florets. Measured masses ( $[M-H]^-$ ) are indicated at the top of the mass peaks. The detected molecular ion deviated  $-1.2$  ppm ( $0.5$  mD) from the calculated mass of the elemental formula of glucoiberin. Glucosinolates show a characteristic  $HSO_4^-$  fragment of  $m/z$  96.96 upon MS/MS.

such as phenylpropanoids, flavonoids and a range of other glycosylated compounds, can also be readily detected in ESI-negative mode (2, 12, 13). On the other hand, polyamines and other compounds comprising chemical structures that easily form proton adducts, e.g., alkaloids and anthocyanins, can be better detected in ESI-positive mode. Thus, analysis of plant samples in both positive and negative ionization mode will provide the most comprehensive insight into their metabolic composition (16, 19, 20). Nevertheless, profiling in just a single ionization mode may already be sufficient

to obtain a global overview of the differences and similarities between samples and to relate metabolic variation to other traits or processes, such as genetic variation (4), gene function analyses (16, 21, 22), plant development (23) and food processing (24). Therefore, this chapter describes a user-friendly protocol to extract and analyze secondary metabolites in *Brassica* and *Arabidopsis* tissues, using accurate mass LCMS in negative mode. This general approach is however, more broadly applicable to most other plant species with the need for only minor modifications.

---

## 2. Materials

### 2.1. Plant Material Sampling

1. Plants, leaves, tissues, etc. of any *Brassicaceae* species.
2. Plastic bags or storage tubes resistant to liquid nitrogen, e.g., polypropylene 50 mL tubes with screw cap (Greiner), Eppendorf micro-test tubes, or 12 mL glass tubes with screw caps (Omnilabo).
3. Liquid nitrogen for sample quenching and grinding (see Note 1).
4. Protective insulating gloves for handling super-cooled objects.
5. Metal spatula or small spoon, precooled with liquid N<sub>2</sub>.

### 2.2. Chemicals (See Note 2)

1. Methanol absolute, HPLC supra-gradient grade.
2. Formic acid for analysis 98–100%.
3. Acetonitrile, HPLC supragradient grade.
4. Ultrapure water (e.g., MilliQ or its double-distilled equivalent).
5. Leucine enkaphaline, ≥95% pure, isolated by HPLC, or equivalent reference compound, to be used for online mass correction (so-called “lock mass”).
6. Phosphoric acid p.a. 85% in water solution (w/v) or equivalent compound mixture suitable for mass spectrometer calibration over the mass range of 100–1,500 Da.
7. Liquid nitrogen or nitrogen gas generator for supplying gas to the mass spectrometer ionization source.
8. Argon 5.0, at least 99.999% pure, for supplying gas to the mass spectrometer collision cell.

### 2.3. Reagents and Solvents

1. *Sample extraction solution*: 0.133% (v/v) formic acid (FA) in pure methanol. Prepare sufficient solution for extraction of the complete series of samples.
2. *HPLC mobile phase*: 0.1% FA (v/v) in ultrapure water (eluent A), and eluent B is 0.1% FA (v/v) in acetonitrile (eluent B). Since

the chromatographic behavior of some *Brassicaceae* metabolites, and especially of intact glucosinolates, is sensitive to even slight variations in the acidity of the mobile phase, prepare the mobile phases freshly by precisely adding 0.1% v/v FA to both water and acetonitrile. Prepare sufficient of both eluents to analyze the entire sample series at one time.

3. *MS calibration solution*: 1 mL of a 0.05% (v/v) phosphoric acid solution in 50% acetonitrile/ultrapure water. Load into the gas-tight glass syringe.
4. *Lock mass solution*: leucine encephaline in 50% (v/v) acetonitrile/ultrapure water to obtain a final concentration of 0.1 mg/mL. Prepare sufficient solution for analysis of the complete series of samples.

## 2.4. Equipment

1. Freezer at  $-80^{\circ}\text{C}$  for (long-term) storage of raw and ground plant materials or products.
2. Pipettes and tips suitable for handling organic solvents (Microman, Gilson).
3. Pestle and mortar or, preferably, a ball mill, e.g., Retsch Mixer Mill MM 301 (Retsch, Germany) for small *Arabidopsis* samples or a metal electric grinder, e.g., IKA A11 Basic Analytical mill (IKA, Germany), for larger samples.
4. Balance for accurate weighing of 100–500 mg frozen sample powder.
5. Ultrasonic bath.
6. Single-use sterile and nonpyrogenic latex-free syringes.
7. Single-use syringe filters free of polymers, such as Anotop 10 (diameter 10 mm, pore size 0.2  $\mu\text{m}$ ; Whatman) or Minisart RC4 (diameter 4 mm, pore size 0.45  $\mu\text{m}$ ; Sartorius). Filters for MS analyses should be resistant to the extraction solution used (i.e., 75% methanol + 0.1% FA) and free of polyethylene glycol or any other soluble polymer (see Note 3).
8. Crimp cap autosampler vials of 1–2 mL with aluminum crimp caps containing natural rubber/polytetrafluoroethylene septa.
9. Vacuum filtration unit for 96-wells format (see Note 4).
10. Protein filtration plates in 96-well format.
11. 96-well plates (Ritter style) with 700- $\mu\text{L}$  glass inserts (Waters) and a 96-square well PTFE-coated seal (Waters).
12. Analytical column Luna C18(2), 2.0 mm diameter, 150 mm length, 100 Å pore size, spherical particles of 3  $\mu\text{m}$  (Phenomenex).
13. Precolumns: Luna C18(2), 2.0 mm diameter, 4 mm length (Security Guard, Phenomenex).

14. PEEK in-line filter holder with PEEK frit 0.5- $\mu$ m pore size (UpChurch Scientific).
15. Alliance 2795 HT high performance liquid chromatography system, or comparable system, equipped with an internal degasser, sample cooler, and column heater (Waters).
16. Separate HPLC pump for continuously pumping the lock mass solution at 10  $\mu$ L/min.
17. Photodiode array detector (PDA) (Waters 2996).
18. High-resolution mass spectrometer: Quadrupole-time-of-flight (QTOF) Ultima V4.00.00 mass spectrometer equipped with an electrospray ionization (ESI) source and separate lock mass spray inlet (Waters) (see Note 5).
19. Syringe pump for injecting calibration solution.
20. Gas-tight glass syringe 0.1–1.0 mL.
21. MS data acquisition software: MassLynx 4.1 (Waters).
22. Mass signal extraction and alignment software such as MetAlign (25).
23. Optional: multivariate analyses software such as GeneMaths (26).

---

### 3. Methods

#### **3.1. Plant Growth and Sampling Conditions**

Samples to be prepared for metabolomics studies should be as specific and representative as possible for the plant, genotype, tissue, or cell type to be analyzed. For instance, if only specific cell types or tissues are known or suspected to be affected by a certain treatment or mutation, any possible effect of the treatment on the metabolome will be diluted out by other tissues. Thus, if the aim is to detect metabolic changes specifically occurring in root tips, start isolating the root tips from the nonresponding rest of the root system. In studies aiming to link metabolic variation to genetic variation, the epigenetic (biological) variation should be kept as low as possible by means of controlled plant growth and plant pooling. For instance, in the large-scale genetical metabolomics study in Arabidopsis RILs (4), seeds were sown on agar containing a nutrient solution, in Petri dishes with a density of a few hundred seeds per dish. Dishes were temperature-treated to promote uniform germination and were then all randomly placed in a single climate chamber in five blocks where each block contained one replicate dish of each line. After 6 days of controlled growth, the lids of the Petri dishes were removed to ensure that seedlings were free of condensed water on the day of harvest. On day 7, at 7 h into the light period, all seedlings were harvested within 2 h by submerging the complete Petridish briefly in liquid nitrogen and



scraping off the seedlings with a razor blade. Finally, per line material from two dishes was pooled to make one of the replicate samples and from the remaining three dishes to make the second. To obtain representative material from larger plants, such as leafy *Brassica* vegetables, a representative number of leaf disks from different leaves or at least three complete leaves should be pooled per plant. In the case of seeds, a large number of seeds (at least 50) should be taken as a representative sample of the genotype, developmental stage, or treatment.

Once harvested, metabolite changes must be kept to a minimum. Therefore, upon harvest, plants or tissues should be snap-frozen in liquid nitrogen, even in the field/greenhouse if at all possible. To obtain homogenous material from the plants, plant parts or products, the frozen material should be ground into a fine powder using liquid nitrogen. Take care that tissues remain fully frozen at all stages from harvest until metabolite extraction; otherwise, throw away the sample. Without knowing the effect of lyophilization on the metabolite profile, lyophilization of tissue is not recommended, unless for specific practical reasons.

### **3.2. Tissue Sampling**

1. Prelabeled bags or tubes with a freezer-proof marker pen or freezer-compatible labels. In the case of seeds or small seedlings (e.g., *Arabidopsis*) use 1.5- or 2.2-mL Eppendorf tubes; in the case of larger tissues use 50-mL Greiner tubes or plastic bags that are resistant to liquid nitrogen.
2. Harvest a representative amount of tissue (leaf, roots, flower head, etc.) in tubes or bags by rapid freezing in liquid nitrogen (see Note 6).
3. Homogenize the frozen tissue in liquid nitrogen into a fine powder using a pestle and mortar. For large series of samples, preferably use a ball mill for *Arabidopsis* or an analytical mill for larger tissue amounts. These should be precooled with liquid nitrogen. Homogenize for 20 s. Transfer the homogenized powder into precooled storage containers resistant to liquid nitrogen, using a precooled metal spatula or small spoon.
4. Weigh 100 mg frozen powder of *Arabidopsis* with an accuracy of better than 5% into a precooled Eppendorf tube, or 500 mg in the case of larger amounts of tissue into a 10-mL glass tube with screw cap (see Note 7). Smaller sample amounts can be used as well, but this is not advisable in view of the inherent higher weighing error using frozen material. Also weigh replicate samples of the same plant powder, to be included as quality control samples and technical replicates for extraction and analysis (see Note 8).

### **3.3. Metabolite Extraction**

1. Prepare extracts freshly at the beginning of a series of analyses, after ensuring the LCMS system has been prepared, tested and calibrated properly. Add ice-cold sample extraction solution

(99.867% methanol acidified with 0.133% FA) to the tube containing the weighed frozen powder, in a volume:fresh weight ratio of 3:1 in the case of leaf material. Close the lid and immediately vortex for 10 s. Assuming a tissue-water content of about 95%, this will result in a final concentration of about 75% methanol and 0.1% FA.

2. Extract the metabolites by 15-min sonication at maximum frequency (40 kHz) in a water bath at room temperature.
3. Centrifuge for 10 min at maximum speed (20,000×*g* for Eppendorf tubes; 3,000×*g* for glass tubes) at room temperature.
4. Filter supernatant through a 0.2-μm PTFE filter using a disposable syringe into a 1.8-mL glass vial and close vial with the crimp cap. All filters used should be free of aqueous-methanol soluble polymers, such as polyethylene glycol. In the case of large numbers of samples, if possible use suitable filtration plates in 96-well format and a vacuum filtration unit (see Note 9).
5. If necessary, extracts can be stored at +4 or, preferably at –20°C. After storage, always sonicate and/or filter each sample once more before analysis (see Note 10).

### **3.4. Conditioning of the HPLC-PDA System**

1. Prepare HPLC mobile phase solvents as described in item 2 of Subheading 2.3. Prime HPLC pump and tubing, and degas both solvents for at least for 10 min using the in-line degasser of the Alliance 2795 HT.
2. Install one PEEK in-line solvent filter between the injection system and the precolumn cartridge. Place two precolumns in tandem into the cartridge, place online in front of the analytical column, and place both cartridge and column in the column oven conditioned at 40°C.
3. Place the outlet from the column directly to a waste bottle and precondition the LC system and column system by increasing the percentage of eluent A stepwise (starting at 100% eluent B) until the initial gradient conditions are reached.
4. Program the HPLC method according to the gradient settings given below. In the standard setup, we use relatively long chromatographic runs of 60 min, including a linear gradient from 5 to 35% eluent B for 45 min, column washing at 75% eluent B for 5 min, return to 5% eluent B for 2 min, and reconditioning at 5% eluent B for 5 min, with a mobile phase flow rate of 0.19 mL/min into the analytical column (diameter of 2.0 mm). This flow rate corresponds to 1 mL/min on a 4.6-mm column, which is standard in most HPLC-UV/Vis applications.
5. After preconditioning the LC system, connect the PDA detector and subsequently the QTOF MS with the PEEK tubing. Program the PDA to acquire data every 1 s from 210 nm to 600 nm with a resolution of 4.8 nm. Wavelength range, scan

rate and resolution can be adjusted according to LC run time and the research aims. Always precondition the PDA-lamp, column oven temperature, and analytical column for at least 1 h before starting sample analyses.

6. Check the entire system for air bubbles and all connections for leakage by verifying the LC pressure stability.

### **3.5. Conditioning of the MS System**

Before each series of sample analyses, the mass spectrometer should be well-conditioned and calibrated to obtain good performance in terms of mass accuracy and resolution. In contrast to electron impact ionization, as used in most GC-(TOF) MS applications, detection sensitivity and mass spectra obtained by soft-ionization LCMS are highly dependent on the type of mass spectrometer, ionization source, and chromatographic system used. The procedure and settings described here are for a QTOF Ultima with ESI source and the TOF-tube in V-mode, in combination with the HPLC conditions described above. Depending upon samples and compounds of specific interest, settings and conditions may need specific adaptations.

1. Connect the outlet of the PDA, with an eluent flow rate of 0.19 mL/min, to the inlet of the mass spectrometer and set the capillary voltage to 2.75 kV, cone voltage to 35 V, source temperature to 120°C, and desolvation temperature to 250°C. Use a cone gas flow rate of 50 L/h and desolvation gas flow rate of 600 L/h. Precondition the MS for at least 2 h at these standard settings before sample analysis.
2. Disconnect the LC flow from the MS, and use the syringe pump to inject the MS calibration solution into the ESI source, at an initial flow rate of 10  $\mu$ L/min.
3. Acquire data from  $m/z$  80 to 1,500 at a scan rate of 0.9 s and an interscan delay of 0.1 s. A series of phosphoric acid cluster peaks should appear throughout the entire range of the mass spectrum. To obtain proper calibration and accurate mass calculations, none of the mass calibration peaks should exceed an intensity of 250 counts/s (in continuum mode) and the intensity of the clusters over the mass range should be as uniform as possible. Adjust pump flow, capillary voltage, cone voltage, desolvation gas flow, and/or collision energy until criteria are optimal.
4. Combine the spectra from 50 adjacent scans during acquisition mode at optimal settings in continuum mode, center the mass signals and check mass resolution of the machine at  $m/z$  488.8772 (negative ionization mode) or  $m/z$  490.8918 (positive ionization mode). Mass resolution is calculated by dividing the  $m/z$  value of the centered mass signal by the mass difference at half height of the Gaussian-shaped mass peak in

continuum mode, and should be better than 8,500 (with the QTOF Ultima in V-mode). Otherwise, retune the instrument and repeat the procedure.

5. Use the centered mass data for calibration of the instrument using a polynomial-5 fit. Mean residual mass deviation, according to the MassLynx calibration procedure, should be less than 1.0 ppm, otherwise adjust the calibration settings.
6. Reconnect the LC flow to the MS. Check the effluent from the complete LC-PDA system, including mobile phase, tubing, columns and PDA flow cell, by acquiring centroid data from  $m/z$  80 to 1,500 under the exact conditions of sample analysis. To prevent excessive ion suppression of sample compounds, individual mass signals at initial gradient conditions should preferably be less than 200 counts per scan (centroid data) in negative mode or less than 500 counts per scan in positive mode.
7. Prepare an MS method file to acquire mass data from  $m/z$  80 to 1,500, at a scan rate of 0.9 s and an interscan delay of 0.1 s and in centroid mode. The range of masses to be detected in sample extracts should fall within the range of calibration masses (see Note 11). During sample analyses, set the standard setting of collision energy to 10 eV in negative ion mode and to 5 eV in positive ion mode. If needed for optimal ionization of key compounds, the collision energy may be adjusted. The MS is programmed to switch from sample to lock spray every 10 s and to average two scans for lock mass correction ( $m/z$  556.2767 in positive mode and 554.2619 in negative mode). The lock mass solution is used for online calibration of the mass accuracy during sample analysis. Adjust the flow rate or concentration of the lock mass solution to obtain a stable intensity of about 600–800 counts per scan (in centroid mode) during LCMS runs (see Note 12).
8. The aqueous-methanol extracts are placed in trays inside the autosampler (20°C) during the analysis series, in a randomized order (see Note 8). Program the injection system to operate in sequential mode and to load the syringe with 5  $\mu$ L of sample, with 5  $\mu$ L of air both before and after the sample. The injection needle is washed with 100% methanol between injections.
9. Check for the presence of sufficient eluents, lock mass solution, nitrogen and argon gasses, and computer hard disk space for the entire sample series.
10. Start the sample series with at least four injections of one of the plant extracts, to stabilize the LC and MS systems. Check stability of retention times and mass accuracy of known compounds during these first runs. Deviations of observed known parent masses from their calculated masses should be less than 5 ppm (at signal intensities similar to that of the local lock mass), otherwise stop the series and recalibrate the MS.

11. Once all extracts have been run successfully, transfer data from the LCMS-data acquisition computer to a second computer on which both the acquisition software and data-processing software have been installed.

### 3.6. Data Processing

Depending on the aim of the research, the raw data may be processed in order to extract metabolite intensity signals in different ways. Relative metabolite intensities may be calculated from their corresponding chromatographic peaks and expressed either as maximal peak height or as area under curve, presuming a more or less Gaussian shape of the chromatographic peak.

In the case of interest in only specific classes of *Brassica* metabolites, e.g., glucosinolates, peak integration tools delivered with the data acquisition software may be used. We use the QuantLynx data processing package delivered with the MassLynx acquisition software (Waters) of the LC-QTOF MS. Since high mass resolution is used, the mass peak integration parameters can be set at a narrow mass window (e.g., 20 ppm) around the exact mass for each compound of interest, enabling specific detection and a high signal to noise ratio. For the untargeted approach, we routinely use the MetAlign software (25). Standard settings for processing of LC-QTOF Ultima MS data from *Brassica* samples, as collected according to the procedure described here, are given in Fig. 3 (see Notes 13 and 14). For further details on the MetAlign software, the reader is referred to the contribution of Dr. Lommen (see Chapter 15).

The MetAlign data output can be cleaned, if needed, for low abundant or misaligned signals and further processed according to the research aim. For instance, metabolite signals significantly differing between samples can be determined, or multivariate analyses techniques such as principal components analyses and hierarchical clustering can be applied to obtain a global view of overall metabolic differences and similarities between samples (4, 12, 22–24).

### 3.7. LC-MS/MS to Identify Selected Mass Signals

If needed or desired, compounds can be further identified using LC-QTOF MS/MS. For this purpose, masses of interest can be incorporated into a mass inclusion list (data-directed MS/MS) in the MassLynx software. Ten instead of 5  $\mu\text{L}$  of sample containing relatively high amounts of the compounds of interest are now injected, in order to obtain higher intensities of the parent ions and thus also their MS/MS fragments. The collision energy profile is programmed to increase sequentially from 5, 10, 20, to 30 eV (ESI positive mode) or 10, 15, 30, to 50 eV (ESI negative mode). If these settings are insufficient to obtain informative MS/MS information for the masses of interest, the collision energy profile can be adjusted. Also, if the intensities of essential MS/MS fragments are too low for exact mass calculation, the amount of compound

**a**

Accurate or Nominal?

SELECT DATA TYPE

☒ Accurate mass data

Mass Resolution is Defined as Mass Divided by the Mass Width at Half Height

Mass Resolution: 7500

WARNING: Wrong mass resolution leads to peak loss!

Amplitude range for accurate mass determination:

From 500 to 1200

☒ [TOF without DRE (saturation effects on mass)]

Echo suppression:

Interval around mass peak: 0.45

Percentage of amplitude of mass peak: 5

"Forrest" Suppression:

Interval around mass peak: 6

Percentage of amplitude of mass peak: 3

Interval offset from mass peak: 0.5

☐ Nominal mass data

Mass Bin Parameter for Conversion to Nominal: 0.65

A Value of 0.65 Gives a Mass Bin from -0.35 to 0.65

OK

**b**

Metalign = RIKILT program for full scan MS comparison; Free Permanent License Version 180309

PART A: PROGRAM CONFIGURATION, DATA SET SELECTION AND BASELINE CORRECTION

1A. Program configuration

1B. Mass Resolution/Bin

Accurate Mass Mode

Mass Res = 7500

SELECT INPUT DATA SETS

2A. Group 1: List of Data Sets

2B. Select

Clear

and / or

3A. Group 2: List of Data Sets

3B. Select

Clear

BASELINE AND NOISE ELIMINATION PARAMETERS

4. Retention Begin (Scan nr) 70

5. Retention End (Scan nr) 2500

6. Maximum Amplitude 25000

7. Peak Slope Factor (x Noise) 1.0

8A. Peak Threshold Factor (x Noise) 3

8B. Peak Threshold (Abs. Value) 25

9. Average Peak Width at Half Height (Scans) 20

☐ 10. Keep Peak Shape (no alignment)

11. Run Baseline Correction

PART B: SCALING AND ALIGNING DATA SETS

12. SCALING OPTIONS

☒ No Scaling

☐ Autoscaling on Total Signal

☐ Scale on Marker Peak

From 1st Data Set, Group 1:

Mass 100

Scan Nr. 200

13. INITIAL PEAK SEARCH CRITERIA

Scan Nr. Max. Shift

Begin of 1st Region 70 50

End of 1st Region 2500 50

Begin of 2nd Region 0 0

End of 2nd Region 0 0

14. TUNING ALIGNMENT OPTIONS AND CRITERIA

☐ No Pre-align Processing (Rough)

☒ Pre-align Processing (Iterative)

Calculation Criteria for Chromatography Shift Profiles

15. Maximum Shift per 100 Scans 35

Mass Peak Selection

1st Iteration Last Iteration

16. Min. Factor (x Noise) 5 3

17. Min. Nr. of Masses 3 2

SELECT MIN. NR. PER PEAK SET

18. Group 1: 0 max=189

or

19. Group 2: 0 max=0

20. Run Scaling and Alignment

21. Detailed Ascii Output  
Excel Compatible Output  
Differential Retention Display

PART C: PEAK SELECTION AND EXPORT TO MS SOFTWARE FORMAT FOR VISUALISATION

PEAK SELECTION CRITERIA

22. Significance Percentage 99.00000

23. Minimum Ratio between Means 2

24. Minimum S/N Ratio 3

☐ 25. Either in Gr. 1 or Gr. 2: >> 1 Masses/Compound

26. FILTER ON CONDITION

☒ Group 2 > Group 1

☐ Group 1 > Group 2

27. Run Peak Selection

28. Detailed Ascii Output  
Excel Compatible Output  
Masslynx Include List Output

TOTAL PROCESSING

29. Total Processing

30. Save and Exit

Fig. 3. Interface of MetAlign software with standard settings for peak extraction and alignment of Brassica LC-QTOF Ultima MS data. (a) Mass resolution and accurate mass calculation settings (see Note 13). (b) Baseline correction and peak alignment settings (see Note 14).

injected can be increased, for instance by drying the extract and dissolving again in a smaller volume of methanol, or by applying solid-phase extraction.

---

## 4. Notes

1. When working with liquid nitrogen, standard laboratory safety precautions (eye and hand/skin protection, wearing protective clothes, etc.) should be taken into account at all times.
2. Most organic solvents used in LCMS, such as methanol and acetonitrile, are toxic and highly flammable, while formic acid is volatile and corrosive. Therefore, all solutions should be handled in a fume hood with standard laboratory safety precautions (see also Note 1).
3. If you are unsure of the purity of the filters to be used, always wash two filters through with the blank solvent to be used for the sample extracts and check on the MS for contaminating peaks. If these are present, you should either wash all filters thoroughly with a suitable solvent before use or choose another supplier.
4. The use of a vacuum filtration unit and 96-wells format filter plates is optional, but highly recommended in cases of large numbers of samples. All filter plates should be prewashed at least three times with sample extraction solution to remove polymers, such as polyethylene, from the filters and housing. Check new batches or types of filter plates for recovery of compounds by comparing the metabolic profiles of a series of samples filtered using the filter plate with those filtered using manual syringe filters.
5. The LCMS method described here is specifically adapted for a QTOF Ultima MS (Waters). Other mass detector systems may need other specific procedures and settings for conditioning, calibration, and metabolite detection.
6. To prevent storage tubes or bags exploding, remove all liquid nitrogen by gently pouring off before closing and never screw tube caps firmly! Frozen tissue can be stored at  $-80^{\circ}\text{C}$  for more than 1 year.
7. The water content of the samples is an important issue, as it may determine the detection and abundance of metabolites present in the extracts. For instance, in experiments on water stress the restricted water supply will result in a higher dry weight content, including the metabolite concentration, which should be corrected for. In the case of samples such as those with highly variable (predetermined) water contents, freeze-



dried powders or dry seeds, pure water can be added to adjust each sample to always give a final solvent concentration of 75% methanol and 0.1% FA. If the water content is unknown or cannot easily be determined, freeze-dry the samples and prepare the extracts by adding 75% methanol + 0.1% FA. Similarly, if the concentration of metabolites is too high resulting in chromatographic saturation, the extracts can be diluted with 75% methanol + 0.1% FA. Alternatively, plant materials can be extracted with more volume, taking care that the final concentration is 75% methanol + 0.1% FA. For instance, we routinely extract 100 mg *Arabidopsis* seedlings with 9 volumes of 83.3% methanol + 0.11% FA.

8. It is strongly advised to include a series of extracts from the same plant material as technical replicates, e.g., in order to estimate technical variability, to correct for batch effects, to optimize settings for data processing software, etc. Therefore, prepare a large pooled sample of material from different plants to be analyzed and prepare at least 5 extracts as technical replicates. These technical replicates should be analyzed at least every ten samples, with one replicate at the start and one at the end of the sample entire series. When using multiple 96-well filtration plates, divide five technical replicates over each plate in order to correct for plate variability.
9. We use a TECAN Genesis Workstation 150 equipped with a four-channel pipetting robot and a TeVacS 96-wells filtration unit. Prewash the filtration plates (Captiva 0.45  $\mu\text{m}$ , Ansys Technologies) at least three times with 700  $\mu\text{L}$  75% methanol containing 0.1% FA. Dry the points of the filter tips by blotting onto filter paper. Place a 96-well plate with 700- $\mu\text{L}$  glass inserts (Waters) in the filtration unit under the prewashed filtration plate. Load each well with 700  $\mu\text{L}$  of extract and vacuum-filtrate until all filters are dry (2 times 20 s). Carefully remove air bubbles trapped at the bottom of the inserts and cover the plate with a 96-square well PTFE-coated seal.
10. The samples should be analyzed as soon as possible after extract preparation, to minimize loss of labile compounds. This, however, is not always practically feasible, for instance due to sudden malfunction of the LCMS system. In these cases, the extracts should be stored at  $-20^{\circ}\text{C}$  or  $+4^{\circ}\text{C}$ . Before analyzing, sonicate the vials or inserts for 15 min to dissolve possible precipitates before analysis. If needed, filter the samples again.
11. The present procedure describes the mass calibration of compounds within the mass range of  $m/z$  80–1,500, using a polynomial function. Extrapolation of this polynomial function towards higher or lower  $m/z$  values is not valid and results in an incorrect mass detection. So, if compounds outside this



mass range are of specific interest, ensure that the calibration of the machine covers the entire desired mass window.

12. Since the mass detected by the QTOF Ultima is dependent upon the intensity of the signal and most accurate at an intensity corresponding to that of the lock mass (11, 15), the lock mass signal should be as stable as possible during the analysis of the entire sample series. This intensity of the lock mass is used during data MetAlign-processing to calculate accurate masses within a user-defined intensity window (see *also* Note 13).
13. Within the MetAlign software (see Fig. 3b, button 1B), the resolution and amplitude (=intensity) range for accurate mass calculation have to be specified. The optimal settings for the accurate mass calculation are dependent upon the dynamic range of the MS with regard to accurate mass detection. For the QTOF Ultima MS, the most accurate range is between -50% and +50% of the recorded intensity of the lock mass. The more stable the lock mass is during analyses of the sample series, the more sample data points will fall within the selected amplitude range and thus the more reliable the accurate mass output will be. Variation in measured accurate mass across a chromatographic peak may also result in splitting of the metabolite signal into two or more accurate mass peaks. If this occurs, slightly lower the mass resolution value. For example, we routinely analyze Brassica samples by QTOF MS at a resolution of about 8,500, but use 7,500 as the setting for MetAlign.
14. We recommend selecting the sample that has been analyzed just in the middle of the entire LCMS series as the reference file for alignment by MetAlign, i.e., as the first sample in the entire sample list (button 2), to minimize the extent of retention profile correction between first and last samples analyzed. Always perform a test baseline correction and alignment on a few variable samples, to check whether the default settings are at least correct to extract and align mass peaks that are of specific interest (if any). Set parameters for peak extraction and noise (buttons 4–9) and run baseline correction (button 11). Manually check selected mass peaks at the beginning, middle and at the end of the baseline-corrected chromatograms and in the original raw data. If it is obvious that some mass signals from relatively broad chromatographic peaks are missing in the baseline-corrected data, set parameter 9 at a slightly higher value and rerun baseline correction. On the other hand, if closely eluting peaks of compounds with similar accurate mass have been extracted as single peaks, lower the value at button 9. Once peak extraction and baseline correction settings are satisfactory, run baseline correction for all samples. After baseline correction of the entire series, inspect retention shifts in the baseline-corrected data files of the reference sample and of

the first and last sample of the entire data set. Set maximum shift at initial peak searching criteria (button 13) according to default settings, or to a value at least a factor of 2 higher than visually observed retention shifts and higher than that set in parameter 9. After running the alignment (button 20), create the data output file (button 21). Check technical replicates for variation in mass signal intensities and misalignments, e.g., by making scatter plots and frequency distribution tables of signals detected in the replicate extracts. Adapt alignment settings if needed or filter out misaligned or other inappropriate signals from the dataset.

---

## Acknowledgements

This work was financed by the EU Framework VI program project META-PHOR (2006-FOODCT-036220) and additional financing from the Centre for Biosystems Genomics and The Netherlands Metabolomics Centre, both initiatives under the auspices of the Netherlands Genomics Initiative.

## References

1. Jahangir, M., Kim, H.K., Choi, Y.H., and Verpoorte, R. (2009) Health-affecting compounds in Brassicaceae. *Comprehensive Reviews in Food Science and Food Safety* **8**, 31–43.
2. Olsen, H., Aaby, K., and Borge, G.I.A. (2009) Characterization and quantification of flavonoids and hydroxycinnamic acids in curly kale (*Brassica oleracea* L. Convar. *acephala* Var. *sabellica*) by HPLC-DAD-ESI-MSn. *J. Agric. Food Chem.* **57**, 2816–2825.
3. Malíková, J., Swaczynová, J., Kolár, Z., and Strnad, M. (2008) Anticancer and antiproliferative activity of natural brassinosteroids. *Phytochemistry* **69**, 418–426.
4. Keurentjes, J.J.B., Fu, J.Y., De Vos, R.C.H., Lommen, A., Hall, R.D., Bino, R.J., Van der Plas, L.H., Jansen, R.C., Vreugdenhil, D., and Koornneef, M. (2006). The genetics of plant metabolism. *Nature Genetics* **38**, 842–849.
5. Bennett, R.N., Rosa, E.A.S., Mellon, F.A., and Kroon, P.A. (2006) Ontogenic profiling of glucosinolates, flavonoids, and other secondary metabolites in *Eruca sativa* (salad rocket), *Diplotaxis erucoides* (wall rocket), *Diplotaxis tenuifolia* (wild rocket), and *Bunias orientalis* (Turkish rocket). *J. Agric. Food Chem.* **54**, 4005–4015.
6. Jeffery, E.H., Brown, A.F., Kurilich, A.C., Keck, A. S., Matusheski, N., Klein, B.P., and Juvik, J. A. (2003). Variation in content of bioactive components in broccoli. *Journal of Food Composition and Analysis* **16**, 323–330.
7. Kurilich, A.C., Jeffery, E.H., Juvik, J.A., Wallig, M.A., and Klein, B.P. (2002) Antioxidant capacity of different broccoli (*Brassica oleracea*) genotypes using the oxygen radical absorbance capacity (ORAC) assay. *J. Agric. Food Chem.* **50**, 5053–5057.
8. <http://www.meta-phor.eu>.
9. Ferreres, F., Sousa, C., Pereira, D. M., Valentao, P., Taveira, M., Martins, A., Pereira, J. A., Seabra, R. M., and Andrade, P. B (2009) Screening of antioxidant phenolic compounds produced by in vitro shoots of *Brassica oleracea* L. var. *Costata* DC. *Combinatorial Chemistry & High Throughput Screening* **12**, 230–240.
10. Lopez-Berenguer, C., Carvajal, M., Moreno, D.A., and Garcia-Viguera, C. (2007) Effects of microwave cooking conditions on bioactive compounds present in broccoli inflorescences. *J. Agric. Food Chem.* **55**, 10001–10007.
11. Verkerk, R. and Dekker, M. (2004) Glucosinolates and myrosinase activity in red cabbage (*Brassica oleracea* L. var. *Capitata* f. *rubra* DC.) after various microwave treatments. *J. Agric. Food Chem.* **52**, 7318–7323.
12. De Vos, R.C.H., Moco, S., Lommen, A., Keurentjes, J.J.B., Bino, R.J. and Hall R.D.

- (2007) Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nature Protocols* **2**, 778–791.
13. Bottcher, C., von Roepenack-Lahaye, E., Schmidt, J., Schmotz, C., Neumann, S., Scheel, D. and Clemens, S. (2008) Metabolome analysis of biosynthetic mutants reveals a diversity of metabolic changes and allows identification of a large number of new compounds in Arabidopsis. *Plant Physiol.* **147**, 2107–2120.
  14. Matsuda, F., Yonekura-Sakakibara, K., Niida, R., Kuromori, T., Shinozaki, K. and Saito, K. (2009) MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. *Plant J.* **57**, 555–577.
  15. Moco, S., Bino, R. J., Vorst, O., Verhoeven, H. A., De Groot, J., Van Beek, T. A., Vervoort, J. and De Vos, R. C. H. (2006) A liquid chromatography-mass spectrometry-based metabolome database for tomato. *Plant Physiol.* **141**, 1205–1218.
  16. Von Roepenack-Lahaye, E., Degenkolb, T., Zerjeski, M., Franz, M., Roth, U., Wessjohann, L., Schmidt, J., Scheel, D. and Clemens, S. (2004) Profiling of Arabidopsis secondary metabolites by capillary liquid chromatography coupled to electrospray ionization quadrupole time-of-flight mass spectrometry. *Plant Physiol.* **134**, 548–559.
  17. Rochfort, S.J., Trenerry, V.C., Imsic, M., Panozzo, J. and Jones, R. (2008) Class targeted metabolomics: ESI ion trap screening methods for glucosinolates based on MSn fragmentation. *Phytochemistry* **69**, 1671–1679.
  18. Mellon, F.A., Bennett, R.N., Holst, B. and Williamson, G. (2002) Intact glucosinolate analysis in plant extracts by programmed cone voltage electrospray LC/MS: Performance and comparison with LC/MS/MS methods. *Anal. Biochem.* **306**, 83–91.
  19. Fait, A., Hanhineva, K., Beleggia, R., Dai, N., Rogachev, I., Nikiforova, V. J., Fernie, A. R. and Aharoni, A. (2008) Reconfiguration of the achene and receptacle metabolic networks during strawberry fruit development. *Plant Physiol.* **148**, 730–750.
  20. Hanhineva, K., Rogachev, I., Kokko, H., Mintz-Oron, S., Venger, I., Karenlampi, S., and Aharoni, A. (2008) Non-targeted analysis of spatial metabolite composition in strawberry (*Fragaria x ananassa*) flowers. *Phytochemistry* **69**, 2463–2481.
  21. Malitsky, S., Blum, E., Less, H., Venger, I., Elbaz, M., Morin, S., Eshed, Y., and Aharoni, A. (2008) The transcript and metabolite networks affected by the two clades of Arabidopsis glucosinolate biosynthesis regulators. *Plant Physiol.* **148**, 2021–2049.
  22. Bino R.J., De Vos, R.C.H., Lieberman, M., Hall, R.D., Bovy, A., Jonker, H. H., Tikunov, Y., Lommen, A., Moco, S. and Levin, I. (2005) The light-hyperresponsive *high pigment-2<sup>ab</sup>* mutation of tomato: alterations in the fruit metabolome. *New Phytol.* **166**, 427–438.
  23. Moco, S., Capanoglu, E., Tikunov, Y., Bino, R. J., Boyacioglu, D., Hall, R. D., Vervoort, J. and De Vos, R. C. H. (2007) Tissue specialization at the metabolite level is perceived during the development of tomato fruit. *J. Exp. Bot.* **58**, 4131–4146.
  24. Capanoglu, E., Beekwilder, J., Boyacioglu, D., Hall R.D. and De Vos R. C. H. (2008) Changes in antioxidant and metabolite profiles during production of tomato paste. *J. Agric. Food Chem.* **56**, 964–973.
  25. <http://www.metalign.wur.nl>.
  26. <http://www.applied-maths.com/genemaths/genemaths.htm>.