Supplemental material

In this PDF file:

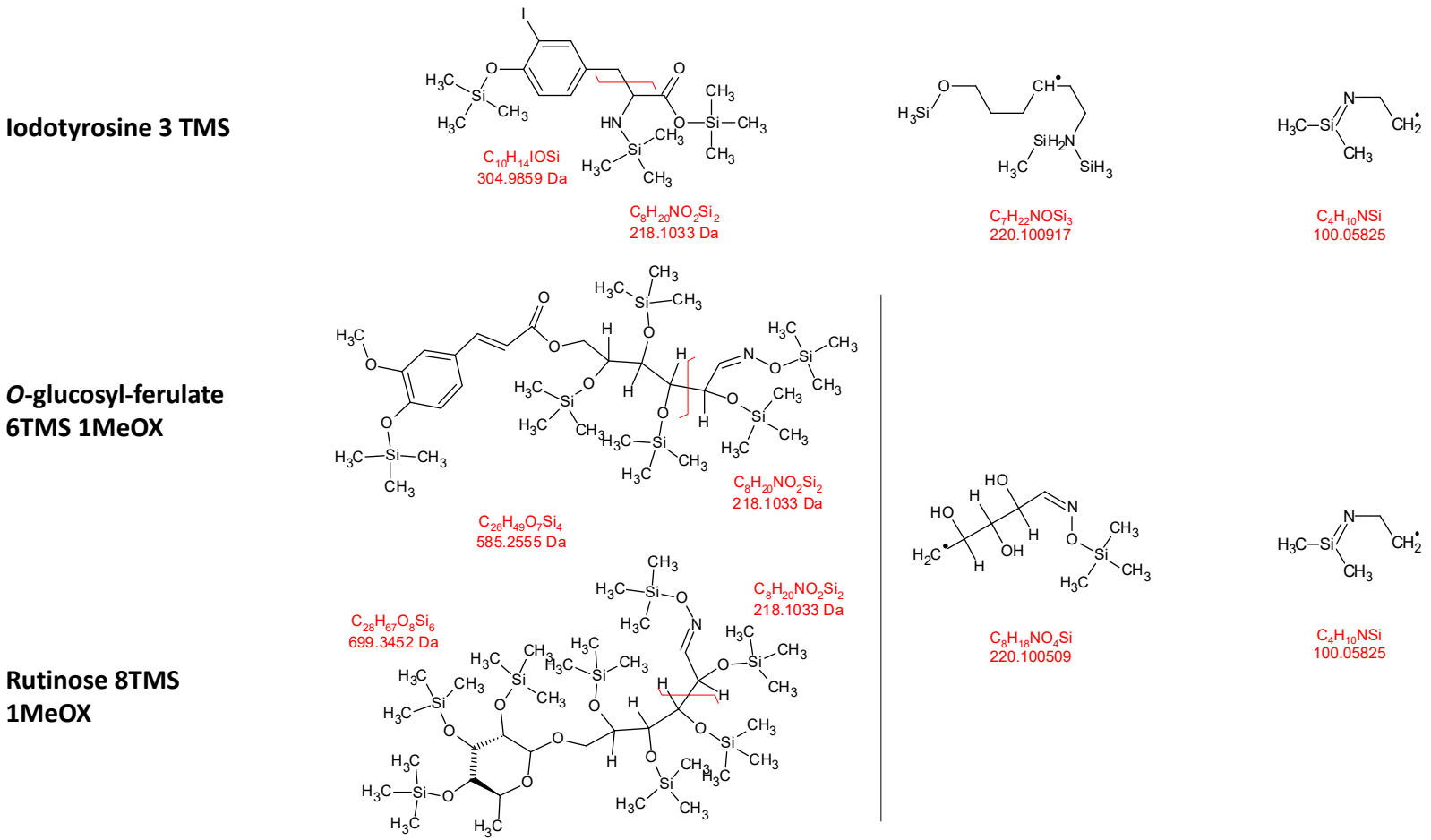
Supplementary Fig. S1. Fragmentation of iodotyrosine generating prevalent fragments similar to sugar derivatives.

Protocol P1. Exact mass GC-MS analyses (with emphasis on MS and data extraction).

Other files available on the *Plant Cell and Environment* website:

Supplementary File 1. Metabolite database in source (NIST) format.

Supplementary File 2. Excel file of metabolite list.



Supplementary Fig. S1. Fragmentation of iodotyrosine generating prevalent fragments (m/z 218.1033, 220.10091, 100.0582 Da) and two examples of alternative compounds that can generate fragments with rather similar retention time and ions with the same exact mass: glycosyl-ferulate and rutinose (rhamnosyl-1→6-α-glucose). In both cases, it involves an aldose (like glucose) and thus a methoxime group which is essential to generate the exact masses of interest.

Protocol P1. GC-MS analyses

REAGENTS

- Methoxyamine hydrochloride for GC derivatization, Lichropur (Supelco: 89803-1G) 20 $mg.mL^{-1}$ pyridine
- Pyridine anhydrous, 99.8% (Sigma-aldrich: 270970-4*25ml)
- N-methyl-N-trimethylsilyltrifluoroacetamide, Lichropur (MSTFA, Supelco: M7891-10*1ml)
- Hexane hypergrade for LC-MS, Lichrosolv (Supelco: 1037012500)
- Methanol suitable for HPLC, ≥99.9% (Sigma-aldrich: 34860-2.5L-R)
- Isopropanol suitable for HPLC, ≥99.9% (Sigma-aldrich: 34863)
- Mix alkanes, 14 alkanes from C_9 to C_{36} (Supelco: Connecticut n-Hydrocarbon Mix, 46827-U), solution in vial: 3 $mg.mL^{-1}$ hexane. (Decane C9, Docosane C10, Dodecane C12, Dotriacontane C14, Eicosane C16, Hexacosane C18, Hexadecane C20, Hexatriacontane C22, Nonane C24, Octacosane C26, Octadecane C28, Tetracosane C30, Tetradecane C32, Tetratriacontane C34, Triacontane C36).
- Internal Mass Standard: 2-Isopropyl-4-methylthiazole ≥97%, (Sigma-aldrich: W355518)
- Internal Standard: Adonitol ≥99% (Sigma-aldrich: A5502), 55µM in extraction solution methanol/water (90%/10%)
- H_2O milliQ
- Helium gas 5.0 B50 P200 (Messer: 102532501)
- Nitrogen gas 4.5 B50 P200 (Messer: 100512501)

CHEMICALS/STANDARDS

From Sigma Aldrich (Merck): Mass Spectrometry Metabolite Library (MSMLS, IROA Technologies), 2-Oxoadipic acid (75447), 3- Iodo-L-Tyrosine (I8250), 5-Aminolevulinic acid (08339), α-Tocopherol (T3251), β-Gentiobiose (G3000), Betain (W422312), Chlorogenic acid (C3878), Choline (C1879), Citraconic acid (C82604), D-2-Aminobutyric acid (116122), D-3 phosphoglyceric acid (P8877), D-Erythronic acid (75025), D-Erythronic acid γ-lactone (374385), D-Galacturonic acid (48280), D-Glucuronic acid (G5269), Diaminopimelate (33240), D-Melibiose (92413), D-Pinitol (441252) , D-Quinic acid (138622), D-Threitol (377619), D-Xylonic acid γ-lactone (89339), Dulcitol (D0256), Galactinol (79544), Hydrocinnamic acid (135232), L-α-Glycerophosphocholine (G5291), Itaconic acid (I29204), L-carnitine (C0283), L-Cystathionine (C7505), Levoglucosan (06724), L-Fucose (F2252), L-Pyroglutamic acid (83160), Nicotinic acid (N4126), Norleucine (N6877), O-Acetyl-Serine (CDS020792), Phosphocholin (P0378), Sigmasterol (S2424), Sinapinic Acid (D7927), Mystiric acid (M3128), Tryptamine (193747), Urea (U5378), Xanthosine (CDS020790).

EQUIPMENT

Q Exactive GC Mass Spectrometer (ThermoFisher Scientific) composed of:

- Trace 1300 Series GC (ThermoFisher Scientific)

Chromatographic separations are performed on a TG-5 SILMS column (30m x 0.25mm x 250µm; Thermo Scientific). To evaluate the GC column efficiency, a Grob Test (Sigma-aldrich: 86501) is regularly executed. Several blanks with hexane (Splitless mode) are performed before and after a batch analysis. A column blank is realized by injecting air at the very end of each analysis to check column integrity and possible MS contamination.

- Exactive GC – Orbitrap MS (ThermoFisher Scientific):

Mass performances are controlled every 24 hours with an EI calibration gas which contains perfluorotributylamine (PFTBA, FC43 calibration compound ThermoFisher Scientific) to perform the following tests:

- Positive Ion Evaluation (Spectral Mass Accuracy), Evacuate Vacuum Inlet and Leak check (<6%)
- Positive Mass calibration
- Mass tuning: AutoTune on mass 414 with an electron energy of 70eV to optimize: the Repeller voltage, the 3 different Lenses voltages, the electron lens voltage and the emission current.

Furthermore, Fore Vacuum and Ultra-High Vacuum pressures are controlled every day and must be kept under $6e^{-2}$ and $9e^{-10}$mbar respectively.

- Triplus RSH Autosampler (ThermoFisher Scientific):

The Triplus RSH autosampler allows to inject directly derivatized metabolites or volatile metabolites at the injection temperature (230°C) like the MS standard (IMT). Furthermore, the Triplus RSH can perform all the derivatization reactions for all samples in a batch analysis by using the following tools.

LS1	Tool Liquid Sampler LS1 with syringe 10µL (SYT365D0291)
LS3	Tool Liquid Sampler LS3 with syringe 100µL (SYT365H2141)
Agitator	Incubation Oven (capacity: 6 vials), Speed: 250 to 750rpm, Temperature 30 to 200°C
Trayplate	Trayplate with 3 trays (3 X 54-position Tray)
Vortexer	Vortexer module 1 vial (2mL) maximum speed 2000 rpm
Standard Wash	Standard Wash Station (5 vials 10mL: Hexane, Methanol, Isopropanol, 1 empty vial and a waste vial)

CONSUMABLES FOR SAMPLE PREPARATION

- 2mL vials, Crimp Top Vials with Clear Round Base (ThermoFisher Scientific: Chromacol 2-CRV) and inserts 0.4mL Clear Glass Conical Pulled Point (ThermoFisher Scientific: 6PME04C1)
- 10mL vials and caps for standard wash station (ThermoFisher Scientific: kit WSKT-25, caps: WSKTCL-100)
- Magnetic caps (ThermoFisher Scientific: 11-MC-ST101)
- Electronic Handheld Crimper for 20mm Flip-off (ThermoFisher Scientific: 60180-ECR20FO)

SOFTWARE TO MONITOR THE Q EXACTIVE GC MASS SPECTROMETER AND TO PERFORM DATA ANALYSIS

- ThermoFisher Scientific XCalibur
- Triplus RSH Sampling Workflow editor (ThermoFisher Scientific)
- TraceFinder 4.1 General Quan with Deconvolution plugin (ThermoFisher Scientific)

PROCEDURE

- **Batch design**

Blanks are realized at beginning and at the end of the batch analysis by injecting hexane. At least three standard quality control (QC) samples are integrated in the batch: at the beginning, in the middle and at the end of the batch. The QC standard is representative of the sample type under analysis, here we used pooled QC in which small aliquots of each biological sample to be studied are pooled. Furthermore, a derivation blank (Sample without any biological material but that received the same treatment from extraction to injection) is also added to the batch to control any contamination in the extraction solution or in the derivatization reagents.

- **Chemical derivatization for GC-MS analysis** (time required: about 130 min)

Samples are derivatized using the Triplus RSH Autosampler with a custom method allowing anticipation of derivatization reactions for each sample to inject fresh derivatised sample every 30min. Prior every use, the syringe is rinsed with hexane, then by the solution to be dispensed into the vial and finally rinsed by methanol.

1) Derivatization starts by adding 20µL methoxyamine (20mg.mL^{-1} pyridine) by the LS3 tool (100µL syringe). Vial is then transported to the vortexer tool (10s at 1200rpm) prior to be placed in the agitator tool at 37°C, continuous agitation (300 rpm) for 90min.

2) 30µL of N-methyl-N(trimethylsilyl)trifluoroacetamide (MSTFA) are added to reaction mixture by the LS3 tool before to be transported to the vortexer (10s at 1200rpm) and placed in the agitator at 37°C, continuous agitation (300rpm) for 30min.

3) 5µL of alkane mix (14 alkanes from C9 to C36, 3 µg.µL^{-1}, Connecticut n-Hydrocarbon Mix, Supelco) were added in the reaction mixture by the LS1 tool (syringe 10µL) to compute the retention index in every sample.

4) Finally, 1µL of Internal Mass Standard (2-Isopropyl-4-methylthiazole) is added by the LS1 tool . The final volume of the reaction for every sample vial is 56µL.

5) Injection of 1µL in splitless mode in GC Injector at 230°C (injector temperature).

At the end of a batch analysis, after injecting 2 hexane samples, both syringes are rinsed 3 times with isopropanol and then methanol. Finally, 1 µL air is injected to do a column blank.

- **Separation by gas chromatography**

Chromatographic separations are performed on a TG-5 SILMS with helium as gas carrier with a constant flow of 1 mL min^{-1}.

6) Oven temperature gradient begins at 70°C with a hold time of 4min

7) Temperature rises then to 325°C at 15°C.min^{-1}

8) Finally, temperature is hold at 325°C for 4 min

At the end of the run, the oven temperature is allowed to cool down to 70°C before the next injection.

- **Mass spectrometry analysis**

MS analyses were operated in positive polarity in full MS scan mode with the following source settings:

- Mass scan range 50-750 m/z
- Resolution 60,000
- AGC target $1E^6$
- MS transfer line 300°C
- Filament solvent delay 4.12 min
- Ionisation by electron (IE) impact (70 eV) was performed at 250°C ion source temperature.

- **Design of the "IsoSeed" Mass Exact Compound Database for High through-put Targeted Screening**

The MSMLS standards and some additional metabolites were analysed according the same protocol presented previously. A deconvolution plugin was used combined with the unknown screening TraceFinder 4.1 software. Unknown compound mass-spectrum can be then exported to NIST to confirm their identification. The deconvolution plugin can also automatically deconvolve coeluted chromatographic peaks into multiple components by aligning spectral peaks according to their slightly different retention times. This software can also identify compounds at low levels whose mass traces would otherwise be part of the noise without showing a well-defined chromatographic peak.

Automated steps taken by the deconvolution plug-in software for every peak detected:

1) Extract all ions and generate XICs (Extracted-Ion Chromatograms)
2) Peak detection and pick up the most intense peak for any RT and m/z
3) Ion overlay window % analysis to determine which of the XIC peaks are binned together as one component
4) Library search

Once a metabolite is clearly identified a "Target mass" and at least two "confirming mass fragments" are chosen and exported to our custom Compound database. Alternatively, the Deconvolution plug-in can export the binned mass traces to NIST format to create an exact-mass user-defined libraries in NIST with the exact mass of all detected fragments. Our compound database in the NIST format is available to download. Once unzipped, the folder must be placed in the following location on the computer: C:\NIST11\MSSEARCH.

- **Target Screening**

Peaks identification is carried out with Trace Finder® (Thermo Scientific) 4.1 software by using a Target Screening Master Method to quickly compare raw datas to the custom exact mass library (IsoSeedFragment). Metabolites were identified automatically by their retention time, major characteristic fragment (m/z ion) and two confirming fragments, with a maximum tolerance of 0.00007 Da.

- **Processing**
  - Peaks: S/N Ration Threshold 5.0 and Mass Tolerance 5.00ppm
  - Retention time window: 5sec
  - Confirmation by Fragment Ions (One fragment minimum), Intensity Threshold 200 and mass tolerance 5.00 ppm
  - Compound database: IsoSeedFragment

- **Peak detection**
    - Detection Algorithm Avalon
    - Detection method Highest peak
    - Smoothing 5
    - Start Threshold: 100
    - End Threshold: 100
    - Area Threshold: 200
    - P-P Resolution: 1
    - Bunch Factor: 1
    - Negative Peaks: Off
    - Tension: 1

Alternatively, an unknown screening method can be carried out and peaks can be identified with the NIST mainlib library and by our Compound database in the NIST format. The unknown screening is rather longer comparatively to the screening method.

Finally, all targeted peaks areas are reported to the internal standard peak area (adonitol) in each sample for relative quantification.