Leaf metabolomic data of eight sunflower lines and their sixteen hybrids under water deficit

Thierry Berton1,2,a, Stéphane Bernillon1,2,a, Olivier Fernandez1,b, Harold Durufle3,c, Amélie Flandin1,2, Cédric Cassan1,2, Daniel Jacob1,2, Nicolas B. Langlade3, Yves Gibon1,2 and Annick Moing1,2,*

1 INRAE, Univ. Bordeaux, Biologie du fruit et pathologie, UMR 1332, Centre INRAE de Nouvelle Aquitaine–Bordeaux, 33140 Villenave d’Ornon, France
2 Bordeaux Metabolome, MetaboHUB**, PHENOME, IBVM, Centre INRAE de Nouvelle Aquitaine–Bordeaux, 33140 Villenave-d’Ornon, France
3 LIPME, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France

Received 20 May 2021 – Accepted 13 July 2021

Abstract – This article describes how metabolomic data were produced on sunflower plants subjected to water deficit. Twenty-four sunflower (Helianthus annuus L.) genotypes were selected to represent genetic diversity within cultivated sunflower and included both inbred lines and their hybrids. Drought stress was applied at the vegetative stage to plants cultivated in pots using the high-throughput phenotyping facility Heliaphen. Here, we provide untargeted and targeted metabolomic data of sunflower leaves. These compositional data differentiate both plant water status and different genotype groups. They constitute a valuable resource for the community to study the adaptation of crops to drought and the metabolic bases of heterosis.

Keywords: Helianthus / abiotic stress / drought stress / LC-MS / metabolomic profiling


Mots clés : Helianthus / stress abiotique / stress hydrique / LC-MS / profils métabolomiques

Highlights

Leaf metabolomic data were produced on sunflower plants of inbred lines and their hybrids subjected to water deficit at the vegetative stage.

They differentiate both plant water status and different genotype groups.

They constitute a valuable resource to be combined with other omics data and study the adaptation to drought and the bases of heterosis.

---

*Correspondence: annick.moing@inrae.fr
Co-first authors.
* Present address: SFR Condorcet CNRS 3417, Université de Reims Champagne-Ardenne, Unité résistance induite et bioprotection des plantes, EA4707, Reims, France.
** https://doi.org/10.15454/1.5572412770331912E12.
1 Specifications table

<table>
<thead>
<tr>
<th>Subject area</th>
<th>Biology</th>
</tr>
</thead>
<tbody>
<tr>
<td>More specific subject area</td>
<td>Metabolomic data</td>
</tr>
<tr>
<td>Type of data</td>
<td>LC-MS: LC-MS acquisition files, R command text file for spectra processing, LC-MS/MS acquisition files, Word file for LC-MS annotation table, tab file for calculated data table</td>
</tr>
<tr>
<td>How data was acquired</td>
<td>The Heliaphen robot and targeted robotized analyses of major compounds or LC-MS analyses of polar extracts</td>
</tr>
<tr>
<td>Data format</td>
<td>Targeted-analyses processed data: txt</td>
</tr>
<tr>
<td>Experimental factors</td>
<td>24 genotypes of Helianthus annuus in two environmental conditions (irrigated or not) with three replicates</td>
</tr>
<tr>
<td>Experimental features</td>
<td>Absolute contents of major compounds of sunflower leaf</td>
</tr>
<tr>
<td>Data source location</td>
<td>The outdoor Heliaphen phenotyping platform at INRAE station, Auzeville-Tolosane, France (43°31'41.8&quot;N, 1°29'58.6&quot;E)</td>
</tr>
<tr>
<td>Data accessibility</td>
<td>The LC-MS data are publicly available in Data INRAE repository (<a href="https://data.inrae.fr/dataverse/sunflower">https://data.inrae.fr/dataverse/sunflower</a>, <a href="https://doi.org/10.15454/1.5572412770331912E12">https://doi.org/10.15454/1.5572412770331912E12</a>) under license etalab-2.0</td>
</tr>
<tr>
<td>Related research article</td>
<td>(Blanchet et al., 2018; Gody et al., 2020; Balliau et al., 2021)</td>
</tr>
</tbody>
</table>

2 Value of the data

Drought stress is a crucial issue for crop adaptation to climate change and sunflower is particularly impacted as it is mostly cultivated in marginal lands (Debaeke et al., 2017). In the present experiment, plants were subjected to two treatments (Well-Watered or Water-Deficit) during the vegetative stage. This experiment was performed in the outdoor high-throughput, semi-automated phenotyping facility Heliaphen (https://www6.inrae.fr/phenotoul_eng/WHO-ware/PhenoToul/HeliaPhen).

Heterosis is an outstanding phenomenon involved in natural selection and used in crop breeding to adapt plants to environmental constraints. Twenty-four genotypes of cultivated sunflower consisting in four maintainer lines, four restorer lines and their 16 corresponding hybrids are included in this experiment which allows studying heterosis effect on metabolism.

This dataset provides metabolomic data of sunflower leaves of lines and hybrids under control and water deficit conditions. These data consist in unique untargeted and targeted metabolomic profiles of sunflower responses to drought based on a large genetic variability.

3 Data

Climate change is affecting plant biodiversity, and crop choice and yields. A better knowledge of plant adaptation mechanisms to this recent phenomenon is, therefore, of major interest for crop science, agriculture and for feed and food security (Porter et al., 2019). Helianthus annuus L., the domesticated sunflower, is the fourth most important oilseed crop in the world (USDA, 2019). It seems promising for agriculture adaptation to global change because it can maintain stable yields across a range of environmental conditions, especially during stress induced by water limitation (Debaeke et al., 2017). It can be considered as an archetypical systems biology model with large drought stress response which involves many molecular pathways (Moschen et al., 2017) and subsequent metabolic and physiological processes.

In this data article, we are sharing the metabolomic data of 24 sunflower genotypes grown in two environmental conditions in an outdoor phenotyping facility. This dataset is part of a larger project that integrates other omics data (Blanchet et al., 2018; Gody et al., 2020; Balliau et al., 2021).

The LC-MS data and metadata associated with this article were deposited in the Data INRAE repository. The targeted analyses data were deposited in the Data INRAE repository.

4 Experimental design, plant material and growth conditions

The experiment was performed from May to July 2013 on the outdoor Heliaphen phenotyping facility at the Institut National de Recherche pour l’Agriculture, l’Alimentation et l’Environnement (INRAE) station, Auzeville, France (43°31’41.8”N, 1°29’58.6”E) as previously described (Blanchet et al., 2018; Gosseau et al., 2019). Briefly, germinated plantlets were transplanted into individual pots filled with 15-l potting soil and covered with a 3-mm-thick polystyrene sheet to prevent soil water evaporation. Plants were fertilized with Peters Professional fertilizer (17-07-27, 500 mL, 0.6 g/L) and an oligo-element mixture solution (Hortition, 0.46 g/L) at 17 days after germination (DAG), and treated with Polyaxe (5 mg/L applied on foliage) against thrips at 21 DAG.
In total, 144 plants, corresponding to 24 genotypes, four maintainer (SF009, SF092, SF109 and SF193) and four restorer (SF279, SF317, SF326 and SF342) lines and their corresponding hybrids obtained by crossing, were grown in two conditions: well-watered (WW) and water-deficit (WD) with three biological replicates (Blanchet et al., 2018; Gody et al., 2020). Before the beginning of the water deficit application at 35 DAG, pots were saturated with water and excessive water was drained. Pots were weighed to obtain the full soil water retention mass. At 38 DAG, irrigation was stopped (approximately 20-leaf stage) for WD plants as described previously (Gosseau et al., 2019). Plants were weighed by the Heliaphen robot to estimate transpiration (Gosseau et al., 2019). WW plants were re-watered at each weighing to reach soil water full retention capacity. Pairs of WD and WW plants were harvested when the fraction of transpirable soil water of the stressed plant reached 0.1 (occurring between 42 and 47 DAG). Two out of the three SF342 line plants died under the WW condition. The corresponding plant samples could not be harvested and data could not be obtained.

At harvest, leaves for metabolome analyses were cut without their petiole and immediately frozen in liquid nitrogen from 11 a.m. to 1 p.m. The harvested leaf was the leaf above the leaf that had reached its maximum size the most recently, as from 11 a.m. to 1 p.m. The suspended pellet, starch was determined and expressed in mg bovine serum albumine equivalents per g DW. After neutralisation of the suspended pellet, starch was determined and expressed in glucose equivalents per g DW (Hendriks et al., 2003). Absorbencies were read at 340 or 595 nm using an MP96 microplate reader (SAFAS, Monaco). For fluorescence, 405 nm excitation and 485 nm emission were used with a Xenius multifunction microplate reader (SAFAS, Monaco). All chemicals and substrates for targeted analyses were purchased from Sigma-Aldrich Ltd. (Gillingham, United Kingdom). All enzymes were purchased from Roche Applied Science (Meylan, France).

5.3 LC-MS based metabolomic profiling

LC-MS-based metabolomic profiling of extracts was performed using the same extracts as for targeted analyses. The sample injection order was randomized. The QC sample was injected every 12 samples to correct for mass spectrometer signal drift. The extracts were analysed using LC-MS (Ultimate 3000 − LTQ-Orbitrap Elite, ThermoScientific, Bremen, Germany), using a C18 chromatographic column (C18-Gemini 2.0 × 150 mm, 3 μm, 110 Å, Phenomenex, Torrance, CA, USA), a 18 min acetonitrile gradient in acidified water (solvent A: ultrapure water + 0.1% formic acid, solvent B: LC-MS grade acetonitrile) with a 300 μL min⁻¹ flow rate and the following elution gradient: 0-0.5 min, 3% B; 0.5-1 min, 3-10% B; 1-9 min, 10-50% B; 9-13 min, 50-100% B; 13-14 min, 100% B; 14-14.5 min 100-3% B; 14.5-18 min, 3% B. The column temperature was 30°C. The injection volume was 5 μL. The LC-MS instrument was equipped with an HESI source operated in the positive-ion mode. Source parameters were the following: source voltage, 3.2 kV; sheath gas, 45 arbitrary units (a.u.); auxiliary gas, 15 a.u.; sweep gas, 0 a.u.; capillary temperature, 350°C; heater temperature, 350°C. Full Scan MS spectra were acquired at 240k resolution power with a 50-1000 mass range. Data dependent MS/MS spectra were acquired at 60k resolution mass. The selected ions were fragmented in CID mode at a 35% normalized collision energy. The MS data were processed using R (R Core Team, 2018) with XCMS (Smith et al., 2006) and MetNormalizer (Shen et al., 2016) packages. Briefly, the corresponding MS-based variables were named using their nominal masses in Da and retention time in s (MxxxTyyy). Variables detected in blank extracts were filtered out. Variables with m/z values varying by more than 0.005 Da or with retention time varying by more than 20 s between different samples were also filtered out. Variables with intensity coefficients of variation in QC greater than 20% were also removed. This resulted in a data matrix of 4843 variables. Intensity drift was corrected using support vector regression. Finally, intensities were normalized according to the sample powder mass used for extraction. Annotation of intense ions...
Table 1. Annotation of LC-MS signatures of sunflower leaf ethanolic extracts with LC-MS and LC-MS/MS data in positive ionization mode.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Metabolite Rt a (min)</th>
<th>Metabolite [M + H]+ a m/z</th>
<th>LC-MS/MS fragments m/z a,b</th>
<th>Putative name</th>
<th>Metabolite ID</th>
<th>Class</th>
<th>Neutral molecular formula</th>
<th>Calculated molecular formula [M + H]+ m/z</th>
<th>MSI level c</th>
</tr>
</thead>
<tbody>
<tr>
<td>M284T195</td>
<td>3.26</td>
<td>284.0997</td>
<td>152.0568</td>
<td>Guanosine</td>
<td>CHEBI:16750</td>
<td>Nucleosides</td>
<td>C10H13N5O5</td>
<td>284.09894</td>
<td>2</td>
</tr>
<tr>
<td>M166T209</td>
<td>3.30</td>
<td>166.0864</td>
<td>120.0806</td>
<td>Phenylalanine</td>
<td>CHEBI:28044</td>
<td>Amino acids</td>
<td>C9H11NO2</td>
<td>166.0862</td>
<td>2</td>
</tr>
<tr>
<td>M328T210</td>
<td>3.50</td>
<td>328.1730</td>
<td>220.1195; 202.1091; 136.0618</td>
<td>Pantothenic acid-hexose</td>
<td>-</td>
<td>Amino compounds</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>M356T315</td>
<td>4.68</td>
<td>356.1024</td>
<td>179.0340</td>
<td>5-O-Caffeoylquinic acid</td>
<td>CHEBI:16384</td>
<td>Cinnamic acids</td>
<td>C16H18O9</td>
<td>355.10236</td>
<td>2</td>
</tr>
<tr>
<td>M356T341</td>
<td>5.68</td>
<td>356.1023</td>
<td>ND</td>
<td>3-O-Caffeoylquinic acid</td>
<td>CHEBI:16112</td>
<td>Cinnamic acids</td>
<td>C16H18O9</td>
<td>355.10236</td>
<td>2</td>
</tr>
<tr>
<td>M611T367</td>
<td>6.10</td>
<td>611.1630</td>
<td>465.1028; 303.0502</td>
<td>Rutin</td>
<td>CHEBI:28527</td>
<td>Flavonoids</td>
<td>C21H20O12</td>
<td>465.10275</td>
<td>3</td>
</tr>
<tr>
<td>M370T374</td>
<td>6.21</td>
<td>370.1179</td>
<td>207.0995; 177.0550; 145.0286</td>
<td>Quercetin hexoside</td>
<td>-</td>
<td>Flavonoids</td>
<td>C15H14O7</td>
<td>369.11800</td>
<td>2</td>
</tr>
<tr>
<td>M642T389</td>
<td>6.47</td>
<td>642.1736</td>
<td>495.1133; 333.0603</td>
<td>Pentahydroxy, methoxyflavone hexoside-deoxyhexoside</td>
<td>-</td>
<td>Flavonoids</td>
<td>C28H32O17</td>
<td>641.17122</td>
<td>3</td>
</tr>
<tr>
<td>M518T468</td>
<td>7.74</td>
<td>518.1334</td>
<td>ND</td>
<td>3,4-Dicaffeoylquinic acid</td>
<td>CID: 5281780</td>
<td>Cinnamic acids</td>
<td>C25H24O12</td>
<td>517.13405</td>
<td>3</td>
</tr>
<tr>
<td>M549T502</td>
<td>8.36</td>
<td>549.1256</td>
<td>ND</td>
<td>Trihydroxy, methoxyflavone malonylhexoside</td>
<td>-</td>
<td>Flavonoids</td>
<td>C26H23O14</td>
<td>549.1288</td>
<td>3</td>
</tr>
<tr>
<td>M379T524</td>
<td>8.71</td>
<td>379.1747</td>
<td>ND</td>
<td>4-O-Caffeoylquinic acid</td>
<td>CHEBI:18107</td>
<td>Nucleosides</td>
<td>C10H12N4O6</td>
<td>285.08296</td>
<td>3</td>
</tr>
<tr>
<td>M274T582</td>
<td>9.68</td>
<td>274.0759</td>
<td>255.0655; 163.0392; 137.0234</td>
<td>Butein</td>
<td>CHEBI:3237</td>
<td>Flavonoids</td>
<td>C13H12O5</td>
<td>273.07574</td>
<td>2</td>
</tr>
</tbody>
</table>

a Rt and m/z measured with the raw data files. Only the [M + H]+ ion was considered as it was the most abundant ion in each acquired MS spectrum.

b MS/MS fragmentation data from data dependent scan in CID mode at a 35% normalized collision energy value; ND: spectrum was not acquired.

c Level of metabolite identification (Sumner et al., 2007).

d This variable corresponds to M1 of the [M + H]+ ion.
(Fernandez et al., 2019; Stelzner et al., 2019) was performed using RT, accurate m/z and fragment ions from an MS/MS acquisition of an aliquot of the QC sample. This resulted in the annotation of 18 compounds belonging to eight compound families (Tab. 1). All chemicals for LC-MS analyses were purchased from Sigma Aldrich (Saint-Quentin Fallavier, France) and Extrasynthèse (Genay, France).

Finally, due to plant death or the lack of leaf material for several plants, 121 and 125 samples out of the 144 initial ones were analysed by the targeted (6 variables) and LC-MS based metabolomic (4843 variables) approaches, respectively. To get an overview of each data set, a principal component analysis (PCA) was performed using BioStatFlow web tool (Jacob et al., 2020) on data mean-centred and scaled to unit variance. The two treatments tended to separate along PC2 explaining about 28% of total variability for the targeted analyses (Fig. 1A) and about 9% for the LC-MS based data (Fig. 1B). The lines and hybrids tended to separate along PC1 explaining about 10% of total variability for the LC-MS based data (Fig. 1B). These metabolome data can be combined with other omic and phenotypic data of the same samples (Blanchet et al., 2018; Gody et al., 2020; Balliau et al., 2021) to get deeper insights into drought effects and heterosis.

**Supplementary material**

**DATA-TargetedAnalyses-SunflowerLeaf.txt:** This file contains targeted measurements of major compounds for each genotype and their three biological replicates (in columns) for WW and WD conditions.

**DATA-LCMS-SunflowerLeaf.txt:** This file contains the intensities of LC-MS-Orbitrap metabolite signatures for each genotype and their three biological replicates (in columns) for WW and WD conditions.

The Supplementary Material is available at [http://www.ocljournal.org/10.1051/ocl/2021029/olm](http://www.ocljournal.org/10.1051/ocl/2021029/olm).

**Acknowledgments.** We thank the Heliaphen team (especially Nicolas Blanchet) for plant culture. These data were produced with the funding of the French National Research Agency (SUNRISE ANR-11-BTBR-0005, Metabohub ANR-11-INBS-0010, PHENOME ANR-11-INBS-0012). This work was part of the “Laboratoire d’Excellence (LABEX)” TULIP (ANR-10-LABX-41).

**Conflicts of interest.** The authors declare that they have no conflicts of interest in relation to this article.

**References**


