

Article

Enhanced metabolome coverage and evaluation of matrix effects by the use of experimental-condition-matched ¹³C-labeled biological samples in isotope-assisted LC-HRMS metabolomics

Asja Čeranić¹, Christoph Bueschl¹, Maria Doppler¹, Alexandra Parich¹, Kangkang Xu¹, Marc Lemmens², Hermann Buerstmayr², Rainer Schuhmacher^{1*}

- ¹ Institute of Bioanalytics and Agro-Metabolomics, Department of Agrobiotechnology, IFA-Tulln, University of Natural Resources and Life Sciences Vienna (BOKU), Konrad-Lorenz-Strasse 20, Tulln an der Donau, 3430, Upper Austria, Austria; asja.ceranic@boku.ac.at (A.C.); christoph.bueschl@boku.ac.at (C.B.); maria.doppler@boku.ac.at (M.D.); alexandra.parich@boku.ac.at (A.P.); kangkang.xu@boku.ac.at (K.X.)
- ² Institute of Biotechnology in Plant Production, Department of Agrobiotechnology, IFA-Tulln, University of Natural Resources and Life Sciences Vienna (BOKU), Konrad-Lorenz-Strasse 20, Tulln an der Donau, 3430, Upper Austria, Austria; marc.lemmens@boku.ac.at (M.L.); hermann.buerstmayr@boku.ac.at (H.B.)
- * Correspondence: rainer.schuhmacher@boku.ac.at (R.S.); Tel.: +43 1 47654-97307 (R.S.)

A.C. and C.B. contributed equally to the work

Abstract: Stable isotope-assisted approaches can improve untargeted LC-HRMS metabolomics studies. Here, we demonstrate at the example of chemically stressed wheat that metabolome-wide internal standardization by globally ¹³C-labeled metabolite extract (GLMe-IS) of experimental-condition-matched biological samples can help to improve coverage of treatment-relevant metabolites and aid in the post-acquisition assessment of putative matrix effects in samples obtained upon different treatments. For this, native extracts of toxin- and mock-treated (control) wheat ears were standardized by the addition of uniformly-¹³C-labeled wheat ear extracts that were cultivated under similar experimental conditions (toxin-treatment and control) and measured with LC-HRMS. The results show that 996 wheat-derived metabolites were detected with the non-condition-matched ¹³C-labeled metabolite extract, while another 68 were only covered by the experimental-condition-matched GLMe-IS. Additional testing is performed with the assumption that GLMe-IS enables compensation for matrix effects. Although on average no severe matrix differences between both experimental conditions were found, individual metabolites may be affected as is demonstrated by wrong decisions with respect to the classification of significantly altered metabolites. When GLMe-IS was applied to compensate for matrix effects, 272 metabolites showed significantly altered levels between treated and control samples, 42 of which would not have been classified as such without GLMe-IS.

Keywords: Untargeted metabolomics, internal standard, Deoxynivalenol, abiotic stress of wheat, matrix effects

1. Introduction

Alongside with NMR and GC-MS, the most dominant technique in metabolomics studies is LC-HRMS. It facilitates high throughput measurements with high sensitivity and selectivity and can also be tailored for many different applications as well as chemical compound classes. Typically, metabolomics is employed to study metabolic alterations in biological extracts obtained from different experimental conditions (i.e. treatments and controls). While targeted approaches are focused on the evaluation of known metabolites from a predefined list, untargeted workflows aim at evaluation of the global detectable metabolome and deal with hundreds to thousands of metabolites that are mainly unknown. Therefore, statistically based data evaluation is essential in untargeted

metabolomics in order to maintain high throughput also in all post-measurement evaluation steps. A typical data set of an untargeted metabolomics study that consists of control and treatment samples may be statistically evaluated by a comparison of their metabolic profiles (i.e. type and abundances of the detected compounds) that actually reflects the metabolic response to the treatment(s). For a reliable comparison and meaningful statistical analysis the detected compounds should truly be metabolites of the studied biological system [1]; the detector response has to be approximately linear [2] and ideally the extent of matrix effects between treatment and control should be similar to avoid a large portion of technical bias and to minimize the variance [3].

With respect to the origin of the detected signals and metabolites, the measured chromatograms do not contain only compounds of biological origin but also many contaminants such as extraction solvents and plasticizers. Such contaminant signals may be reliably filtered out using stable isotope-assisted (SIA) approaches. In these, highly enriched global stable isotopically labeled biological extracts can be used for metabolome-wide internal standardization of native biological extracts under investigation. Together, the native and labeled metabolite features that originate from the biological sample will show highly characteristic artificial isotopolog patterns in the obtained mass spectra, which can be efficiently detected as biology derived metabolites with automated software (e.g. MS-DIAL, geoRge, HiTIME, MetExtract II or X¹³CMS [4-8]). As many wheat metabolites like nitrogen containing compounds, sugars, organic acids or phenolics can be newly induced upon stress [9]), SIA-LC-HRMS approaches may benefit from closely “experimental-condition-matched” globally labeled metabolite extracts as internal standards (GLMe-IS) rather than an unspecific reference metabolome in order to effectively detect newly stress-induced metabolites. Such extracts are obtained from ¹³C-labeled biological samples or organisms that are generated under similar experimental conditions as the non-labeled of the actual biological experiment, ideally with the only difference being the isotopic enrichment of the applied labeled nutrients. Besides plants, also microbes like bacteria and fungi as well as mammalian cell cultures and tissues were successfully labeled [10, 11].

Another factor that can potentially limit the reliability, accuracy and interpretation of LC-HRMS data in comparative metabolomics studies are matrix effects. They occur in the ionization source of the mass spectrometer, in particular with electrospray ionization, when co-eluting metabolites alter the ionization efficiency of the analyte(s). As a consequence, either an increase or decrease in the ionization efficiency may occur that results in falsely higher or lower detected metabolite abundances compared to single compounds in pure solvents. While the exact mechanism of matrix effects remains unclear, it is generally assumed to occur as a result of competition between co-eluting metabolites for charge transfer and access to the surface of analyte containing eluent droplets in the ionization source of the mass spectrometer [12]. In untargeted metabolomics applications, matrix effects represent a challenge as each co-eluting metabolite is either considered as the analyte while the others can be seen as the matrix and vice versa. Therefore, all low molecular weight metabolites in the analyzed extract are of interest and some of those that cause significant matrix effects cannot simply be removed from the system by clean up or other types of sample preparation.

Consequently, each metabolite may be affected by or contribute to alterations in ionization efficiency, the extent of which can change for the very same metabolite depending on the qualitative and quantitative composition of the co-eluting metabolites. This may represent an additional challenge for comparative studies that deal with biological samples of different conditions and thus metabolic profiles among the compared samples. Consequently, this may result in inaccurate assignment of metabolite abundance and eventually wrong judgment about differences in the metabolic composition of samples/organisms under investigation.

Unfortunately, it is currently impossible to quantify the matrix effects exactly [13, 14], but at least for targeted approaches, there are different ways to estimate the extent of or account for ion suppression/enhancement such as the comparison between solvent- and matrix calibration or injection of a matrix sample and post column infusion of standard compounds into the column effluent as has been reviewed elsewhere [15, 16]. For the reason explained above, it is however very difficult or even impossible to assess the existence or extent of matrix effects in case of untargeted

approaches. The best method to take account for matrix effects is the addition of labeled metabolic constituents in the form of an internal standard to the native biological extract [3]. Potential differences in matrix effects between the biological samples under investigation can be corrected by using abundance ratios of native and labeled forms (normalized abundances) as a relative quantitative measure for each metabolite in the comparative study. Based on the assumption that native and the corresponding labeled metabolites always co-elute and experience ion suppression/enhancement to the same extent, the abundance ratio of the two is not affected as long as every native sample contains the same amount of labeled internal standard. In untargeted metabolomics studies that deal with hundreds to thousands of metabolites at the same time, it can be challenging to find a representative labeled internal standard. If only one or several labeled metabolites are used to internally standardize the entire investigated biological extract [17] most of the matrix effects could not be compensated for, since they can vary within seconds during analysis and also strongly depend on the physicochemical properties of the (unknown) sample constituents [14]. Consequently, globally labeled ^{13}C (or ^{15}N or ^{34}S)-labeled biological samples are best suited as internal standards since here, the ratio of native and labeled forms of every metabolite that is shared between labeled and native samples is considered for compensation of ion suppression/enhancement [11].

The present study aims to assess the performance of globally labeled biological samples for metabolome-wide and experimental-condition-matched internal standardization. At the example of the comparative study of wheat, abiotic stress, (i.e. toxin treatment) was used to show the benefits of experimental-condition-matched ^{13}C -GLMe-IS for the global detection of biology derived metabolites as well as for the fast assessment of matrix effects as a measure of the accuracy in comparative quantification and judgement of significance.

2. Results and discussion

The experimental-condition-matched GLMe-IS has potential to improve untargeted metabolomics approaches by increasing the detection coverage of truly biology derived metabolites and by compensating putatively different matrix effects between samples obtained upon different treatments.

To test this, metabolic profiles of toxin-stressed and mock (H_2O)-treated control wheat ears in native and ^{13}C -labeled form were analyzed with a well-established SIA untargeted metabolomics workflow [1, 8].

The experimental setup used to generate the analyzed data set is shown in **Figure 1**. Wheat (*Triticum durum*) has either been grown in a native CO_2 atmosphere (98.9% ^{12}C) in the glasshouse or in a ^{13}C -enriched atmosphere (99% ^{13}C) in a customized cultivation chamber [18] to generate native and uniformly ^{13}C -labeled (L) wheat, respectively. Both were either treated with water (Control and L-Control) or toxin (Toxin and L-Toxin). Extracts of labeled wheat ears, L-Toxin and L-Control, were mixed 1+1 (v/v) to obtain the experimental-condition-matched GLMe-IS that was subsequently spiked to each of the native sample extracts. By this, two sample types were obtained, which are termed Toxin//GLMe-IS and Control//GLMe-IS (see **Figure 1a**).

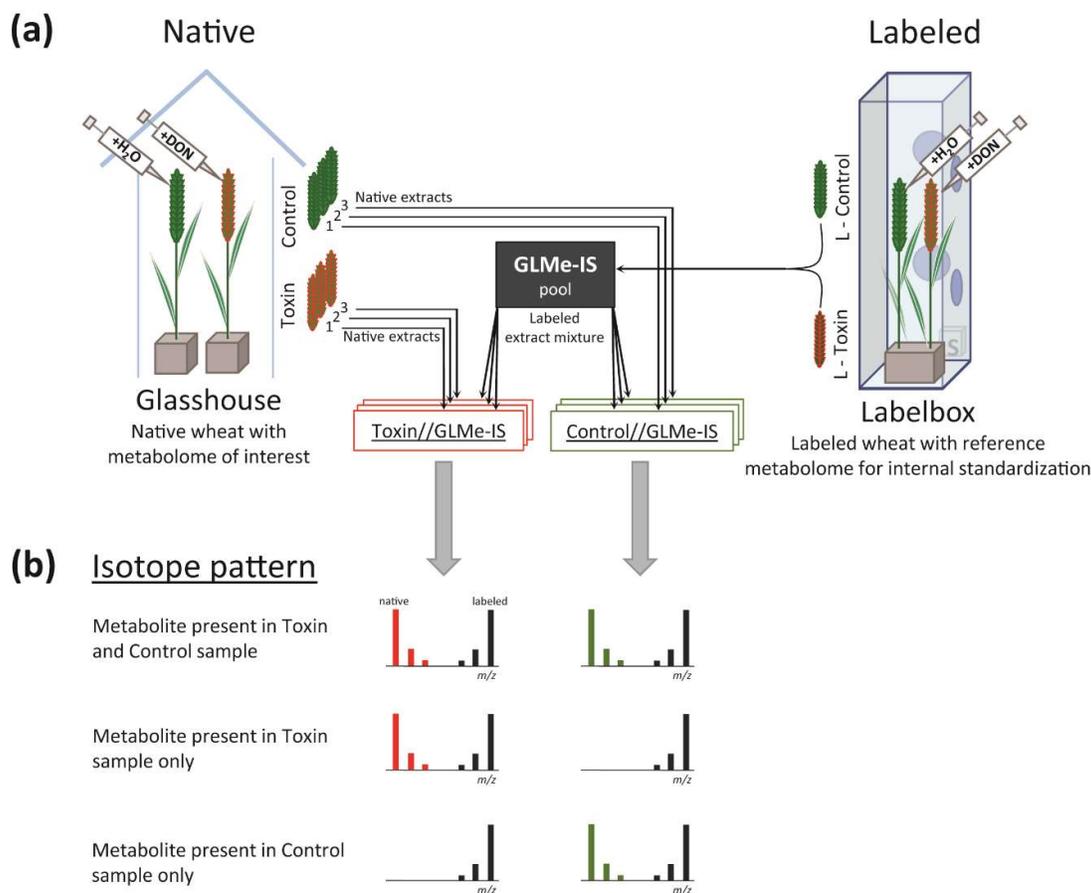


Figure 1. (a) Experimental setup to generate native and ^{13}C -labeled wheat ear sample extracts. Deoxynivalenol (DON) was used for toxin-treatment while water served as control. (b) Isotope pattern in the mass spectrum for three exemplary metabolites. The labeled form is present for all metabolites as it derives from the GLMe-IS pool while the presence of native forms may depend on the sample type (i.e. Toxin and Control).

2.1. Comparison of the metabolic composition between native Control and Toxin samples

In order to compare the qualitative metabolic composition of control and toxin-stressed wheat ear samples, a list of metabolic features (characterized by m/z , RT, and abundances) was generated using MetExtract II [8]. According to the SIA approach, each metabolic feature in the obtained list is present in a native form as well as in a labeled form (originating from GLMe-IS) in at least one sample type (see isotope pattern in **Figure 1b**). Metabolic features of the same metabolite were grouped into one feature group and the most abundant ion of the group was used to represent the respective metabolite.

2.1.1. Qualitative comparison of Control and Toxin treatment

In total, 1080 metabolites were detected, for which the distribution of retention times and m/z are shown in form of a feature plot in **Figure 2a**. 990 metabolites were shared between Control and Toxin treatment, while 62 metabolites were detected exclusively in the Toxin samples and 6 in the Control samples, respectively (**Figure 2b**). The remaining 22 metabolites were inconclusively detected (only in a fraction of the samples of the Toxin or Control groups) and are thus not counted towards either of the two groups.

A detailed annotation and identification of the detected metabolites was not the purpose of this study. However, to illustrate the biological relevance, the additionally detected treatment-specific metabolites shall briefly be summarized here. More than one-third of the 62 stress-specific

compounds (Toxin treatment additional metabolites) are secondary metabolites derived from the aromatic amino acids phenylalanine [19, 20], tryptophan and tyrosine [20, 21]. Amongst them, feruloyltryptamine is one of the detected and identified toxin-treatment specific metabolites. It belongs to hydroxycinnamic acid amides that are well-characterized defense-related compounds in wheat from the phenylpropanoid pathway, which are produced by the plant for cell-wall strengthening [22] to prevent the plant pathogenic *Fusarium* from spreading beyond the initially infected wheat spikelet ([22, 23], [24]).

These treatment-specific and biologically presumably important metabolites were only detected with the application of the experimental-condition-matched GLMe-IS. If any other labeled reference matrix would be used e.g. ^{13}C -labeled control samples or commercially available ^{13}C -labeled wheat ears which are not experimental-condition-matched, these treatment-specific metabolites could not have been detected with the SIA-LC-HRMS approach. Although treatment-specific metabolites might in principle also be detected without SIA application, they would still largely be hidden among numerous non-biology derived compounds as has been demonstrated earlier [1]. Furthermore, the presented approach supplements metabolomics studies as it facilitates annotation and structure elucidation of detected metabolites by taking advantage of the specific isotope pattern of native and labeled metabolite analogs from MS and MS/MS spectra (for more details see also [20]).

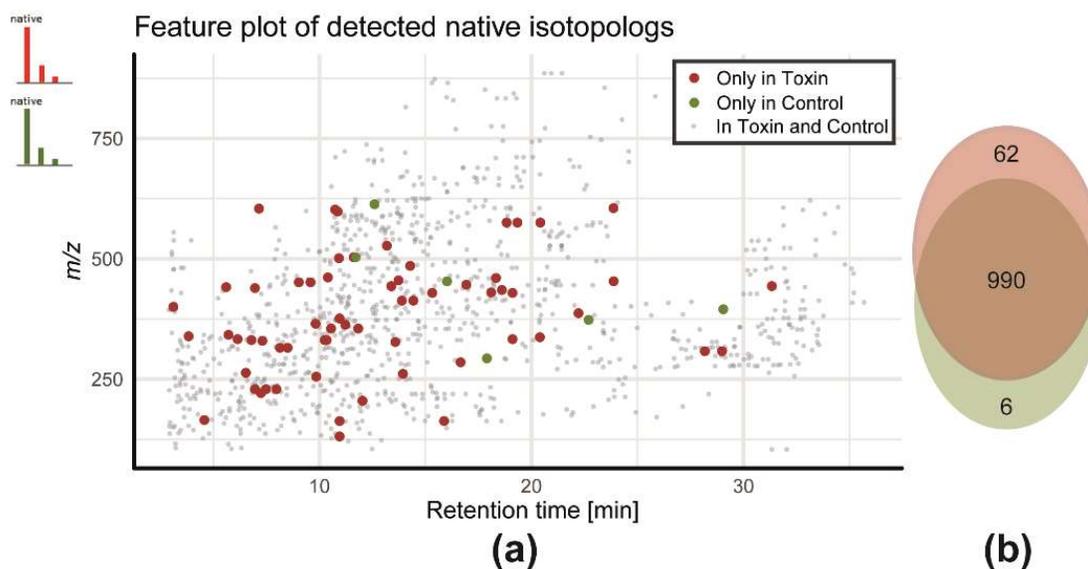


Figure 2. (a) Feature plot of all detected native metabolic features and (b) qualitative distribution of native metabolites in control- and toxin-treated wheat ear extracts, Toxin only (red), Control only (green) or common to both sample types (grey). Only those native features that have a labeled analog in GLMe-IS could be detected by MetExtract II.

2.1.2. Comparative quantification of Control- and Toxin samples

The differential quantitative analysis showed that the toxin treatment has a significant effect on the metabolite profiles of the wheat extracts. This includes both treatment- and control-specific compounds as well as relative abundances of shared metabolites as shown in **Figure 3**. Both sample types are separated well along the first principal component, which covers 61% of the total variance between the two experimental groups (**Figure 3a**). A univariate comparison revealed that 282 (26%) of all native metabolites (common for Control and Toxin) were significantly differing with a fold change of at least 2 or 0.5 and a p-value below 0.05 (**Figure 3b**). From these, 89 native metabolites were significantly more abundant in Control samples while the majority (n=193) showed higher levels in Toxin samples.

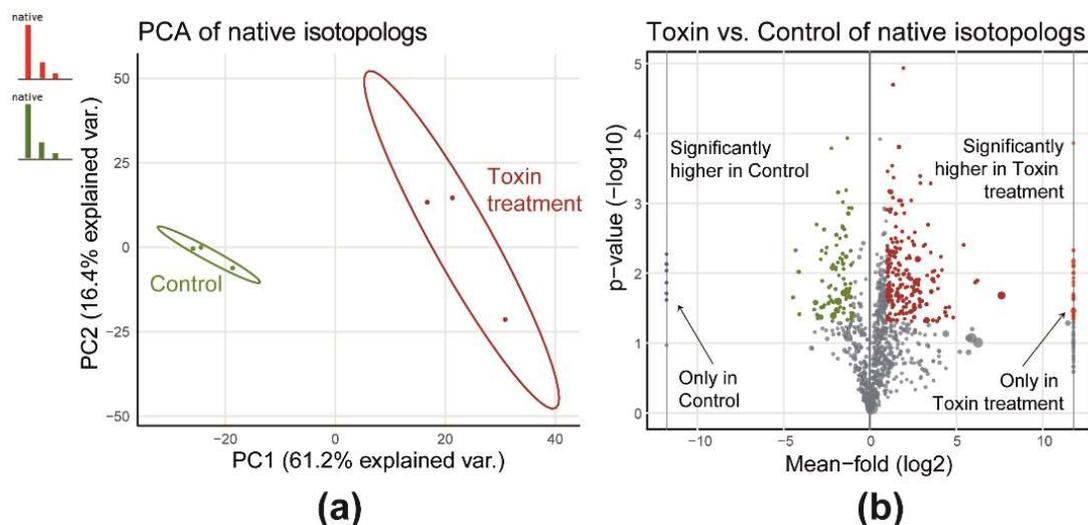


Figure 3. (a) PCA-scores plot considering all commonly detected ^{12}C native metabolic features for Control and Toxin, and (b) volcano plot for a Toxin and Control comparison of native metabolite features

For being able to fully rely on the measured abundances of the native metabolites and to draw a meaningful conclusion from a typical metabolomics study, ion suppression/enhancement (i.e. matrix effects) have to be evaluated. While ion suppression cannot be fully avoided in complex biological extracts, they should at least be similar for the samples derived upon application of different experimental conditions. However, the matrix effects, which derive from competition for charges in the ion source of the mass spectrometer vary with the physicochemical properties of co-eluting metabolites (i.e. proton affinity and surface activity [25]) and might therefore be stronger or weaker, depending on the metabolic composition of the control- and toxin-treated samples. Thus, in the presented study, in which the two sample types have different metabolic composition with respect to the native metabolites as well as different abundances of metabolites shared between the two experimental treatments (shown in **Figure 3b**) comparative quantification may be inaccurate if the matrix effects are not properly accounted for.

2.2. Investigation of putative matrix effects and applicability of the data in comparative metabolome studies

In a typical untargeted metabolomics experiment matrix effects remain hidden as all biochemical small molecules constitute the matrix and at the same time are the objects under investigation. However, matrix effects can be investigated by the use of globally ^{13}C -labeled material for internal standardization. To this end absolute and normalized (internal standardized) abundances of the native signals are compared in different ways.

2.2.1. Improvement of quantification by internal standardization

Native metabolite abundances and ratios of native to labeled (contained in the pooled ^{13}C extract (GLMe-IS)) were used to test whether comparative quantification of native ^{12}C metabolites can be improved by internal standardization.

To assess quantification on a single metabolite level, first the fold-change values (Toxin/Control) were calculated after normalization of ^{12}C peak areas (A_{Control} , A_{Toxin}) to ^{13}C peak areas ($A_{\text{GLMe-IS}}$) and compared to the fold-changes derived from absolute ^{12}C peak areas. For this, the ratio of the two fold-change values of the ^{12}C peak areas (with and without internal standardization) were plotted against $(A_{\text{Toxin}} / A_{\text{GLMe-IS}}) / (A_{\text{Control}} / A_{\text{GLMe-IS}})$ (**Figure 4a**). Values close to 100% on the y-axis indicate that the obtained fold-change values before and after internal standardization are similar and therefore the data suggest that the two experimental sample types Control and Toxin (i.e. the two matrices) produce similar matrix effects on average. The relatively low discrepancy between the fold-change

values with and without normalization can be explained by the high extent of similarity between the matrices tested here. Only 6% of the detected metabolites were exclusively found in one of the sample types. In other cases, where sample composition differs much more it has been shown that the ionization efficiency of many investigated metabolites were affected, especially when the same metabolites were compared in different biofluid matrices such as urine, saliva, blood plasma, serum etc. [3, 26]. In case of such largely differing matrices, normalization of abundances to those of the labeled internal standards enabled to compensate for differences in ionization efficiency of the investigated metabolites [26-28]. Compensation of ion suppression by GLMe-IS enables to reliably compare the metabolic composition of largely differing wheat tissues like root, stem, leaf, spikelet, rachis etc. [29, 30]. It shall be mentioned that this type of reliability assessment in comparative quantification can only be accomplished by the application of experimental-condition-matched GLMe-IS, as all detected biology derived metabolites (incl. treatment-specific) are present in both a native and ^{13}C -labeled form in the investigated samples.

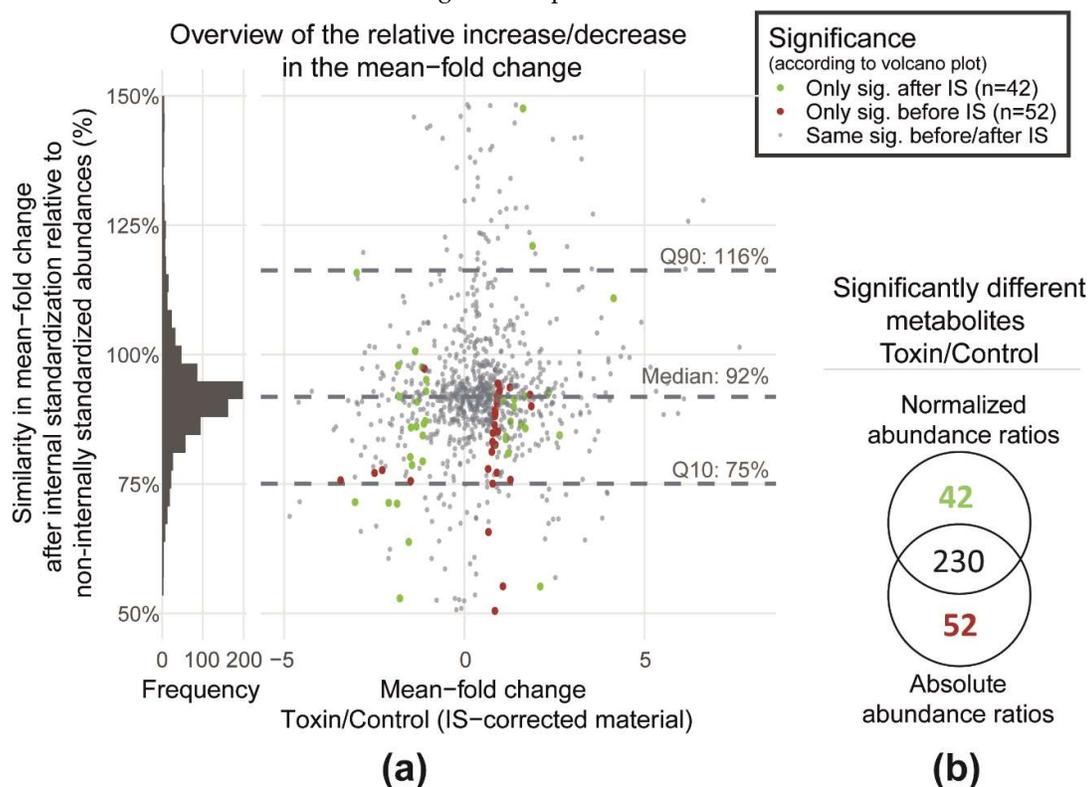


Figure 4. (a) Correlation of abundance ratios between Control and Toxin constituents with and without normalization to ^{13}C -metabolites of GLMe-IS (i.e. absolute and normalized abundances). (b) The Venn diagram shows the distribution of significantly differing metabolites of the Control and Toxin comparison between two groups of the comparative significance test that is obtained with the application of native absolute and normalized abundance ratios

2.2.2. Comparative statistical analysis of native metabolome abundances in Toxin and Control performed with absolute and normalized abundances

While on average, matrix effects were similar in both sample types (**Figure 4a**), another aspect that also has to be demonstrated for being able to rely on the comparative quantification based on ^{12}C feature abundances (as is the case for a typical state-of-the-art metabolomics study) is to critically evaluate the outcome of significance testing. Since the purpose of a typical untargeted metabolomics study is to find metabolites that are significantly altered by the treatment, it is crucial to verify that the comparison of toxin-treated and control samples leads to similar results when absolute metabolite abundances compared to abundance ratios obtained after normalization are used. In this context,

normalized abundances (i.e. matrix effect corrected) can be considered as a reference for the comparison. Any deviation in the classification of significance with and without internal standardization (i.e. absolute abundances versus normalized abundances) may indicate a bias due to matrix effects.

To test this, metabolites, whose abundance was significantly altered under stress were evaluated by the use of ^{12}C peak areas with and without internal standardization (**Figure 4b**). In total, 230 metabolites are reported to be significantly different between Toxin and Control samples by both approaches. 42 of those were only reported to be significantly different when internal standardized and thus normalized abundances were employed, while another 52 metabolites exclusively with absolute abundances. This clearly indicates that, if matrix effects are not corrected, around 19% of truly significantly differing metabolites (i.e. the one obtained with normalized abundances) would be falsely excluded from this group.

It can be seen from the feature map (Supplementary Materials, Figure S1) that these metabolites are distributed throughout the whole chromatogram. This indicates that metabolites of different polarities and m/z range may be affected. Further, **Figure 4a** shows that the majority of these metabolites are distributed around the mean fold change ratio threshold (2 or 0.5), while only a few showed a mean fold change ratio that is higher than 2 or lower than 0.5.

In conclusion, our results demonstrate that if abundances of detected native metabolites are not corrected for matrix effects, the reliability or correctness of the statistically assisted assessment of significantly different metabolites between two experimental conditions can be affected. In order to really prove that internal standardization of native peak areas by ^{13}C areas is not biased and thus significance testing is reliable, absolute matrix effects would have to be estimated by comparison with authentic standards in pure solvent followed by significance testing. Since this is not feasible in current untargeted applications, internal standardization with GLMe-IS is considered to be the best choice at present.

2.2.3. Evaluation of matrix effects on the abundances of labeled metabolite signals

GLMe-IS does also offer another way to test for matrix effects. For the following section, a statistical analysis was carried out that compared the peak abundances of labeled metabolites in the two sample types Toxin//GLMe-IS and Control//GLMe-IS. Both contain the same aliquot of the pooled GLMe-IS but differ in the type of native biological extracts. Therefore, the native metabolome can be regarded as the “matrix”, while the ^{13}C -labeled biochemical constituents can be seen as the metabolites of interest that are the same in both sample types with respect to identity and abundance. If the two native matrices (i.e. Control or Toxin) do not influence the abundance of the ^{13}C labeled metabolites differently, then no quantitative differences should be observed between the two conditions. If on the other hand quantification of ^{13}C -features results in differing abundances for the same metabolite, then differences in the metabolic composition of ^{12}C compounds (see sections 2.1.1. and 2.1.2.) must be the reason for this deviation.

To test this, the same statistical analysis as used in section 2.1.2 was carried out for the labeled metabolite signals (**Figure 5**). The scores plot of the PCA shows that the two sample types overlap and no clear separation of the samples is achieved (**Figure 5a**). Similarly, a univariate comparison of labeled metabolite abundances showed that the majority of metabolites were not significantly different ($p > 0.05$ and fold change < 0.5 or > 2) when paired with control or toxin-treated matrix (**Figure 5b**). Only 7 metabolites were significantly more abundant. Moreover, on average a minor relative increase of around 13% towards higher abundances for those in the toxin matrix was observed. Based on these results, it can be concluded, that there was low matrix effect when using unsupervised statistical methods. This is in contrast to an earlier study of Bueschl, Kluger et al. [1]. There, the investigated samples of a *Fusarium* wildtype- and knock-out strain strongly differed in the qualitative metabolic composition (2/3 of the metabolites were not shared between the wild type and the knock-out strain), which was reflected in a clear separation of samples in the PCA scores plot caused by large differences in the observed matrix effects [1].

It shall be noted however that the effect of different matrices on comparative quantification and subsequent statistical analysis has to be evaluated from case to case and might be severe, especially if the metabolic composition is largely altered by experimental perturbations or if different tissue types are compared in biological studies. Thus, the proper evaluation of ion suppression remains a big challenge in untargeted metabolomics studies. Without globally stable isotope enriched biological samples the metabolome-wide compensation of potential matrix effects is not possible.

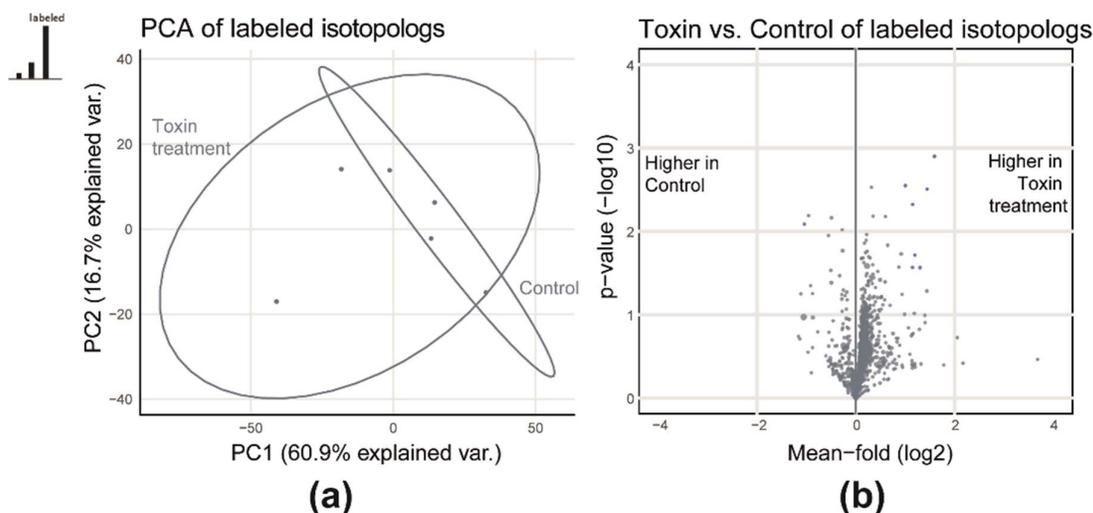


Figure 5. (a) PCA and (b) volcano plot to demonstrate that the native matrices have similar effects on absolute abundances of labeled metabolites in both sample types.

3. Materials and Methods

3.1. Chemicals

Methanol (MeOH, LC-MS CHROMASOLV®), acetonitrile (ACN, LC-MS CHROMASOLV®) and formic acid (FA, MS grade, ~ 98% purity) were purchased from Riedel-de Haën, Honeywell (Seelze, Germany). The ultra-pure water was obtained from an ELGA Purelab system Veolia Water (Ultra AN MK2, Vienna, Austria). $^{13}\text{CO}_2$ (99% purity) was purchased from Eurisotop (St-Aubin, France) and synthetic air was obtained from Messer (Gumpoldskirchen, Austria). Deoxynivalenol (DON) was purified from *Fusarium graminearum* cultivars and kindly provided by the Institute of Biotechnology in Plant Production (University of Natural Resources and Life Sciences, Vienna (BOKU), Department of Agrobiotechnology (IFA-Tulln), Tulln, Austria).

3.2. Plant material

The uniformly ^{13}C -labeled wheat (*Triticum durum*) was produced in the plant cultivation chamber termed labelbox (PhytolabelBox, ECH, Halle). The same wheat cultivar was grown in the glasshouse (native, unlabeled wheat) at BOKU, IFA-Tulln.

3.3. Plant cultivation

The uniformly ^{13}C -labeled wheat was grown in the labelbox (PhytolabelBox, ECH, Halle) in hydroponic cultures under controlled atmosphere, nutrient delivery, light intensity and temperature regime. To produce globally ^{13}C -labeled plants, the atmosphere in the labelbox contained $^{13}\text{CO}_2$ instead of $^{12}\text{CO}_2$. The complete protocol for the wheat cultivation is detailed in [18]. At flowering stage (day 87 after the placement into labelbox) the wheat ears were treated inside the labelbox with the fungal toxin deoxynivalenol (DON) or H_2O , respectively. These two materials are designated L-Toxin and L-Control. 144 h after the treatment the plants were harvested and shock frozen in liquid N_2 . For

long time storage, the ears were freeze-dried (Labconco FreeZone 6Plus; Labconco, Kansas City, MO, USA) until 5% rest moisture and stored at -80°C until further use.

The cultivation of native wheat started with the planting of vernalized wheat seedlings in soil filled pots. The native wheat was grown in a glasshouse under controlled temperature and light regime as described in [21]. Similarly to the labeled plants, the native wheat plants were treated with DON and H₂O at the onset of flowering. These two materials are designated with Toxin and Control. 144h after the treatment the plants were harvested, shock frozen in liquid N₂ and stored at -80°C until further use.

3.4. Sample preparation for LC-HRMS analysis

The fresh native (Toxin and Control) and dried labeled (L-Toxin and L-Control) wheat ears were milled separately under cooled condition for 30s at 40Hz to a fine powder with a ball mill (MM400, Retsch, Haan, Germany). The dried L-Toxin and L-Control wheat powders were rehydrated with H₂O (70 µl/30 mg dried wheat powder) in frozen condition immediately before extraction. Further, the native (100 mg) and rehydrated labeled (30 mg) wheat ear powders were extracted, each with 1 ml of a mixture of ACN/MeOH/H₂O (3/3/2) + 0.1% FA according to [29]. The L-Control and L-Toxin extracts were mixed 1/1 (v/v) in one pool (GLMe-IS). Three biological replicates of Control and Toxin extracts were spiked with the same volume aliquot of the GLMe-IS (1/1 (v/v), see **Figure 1a**). Additionally, H₂O + 0.1 % FA was added to achieve 1:1 organic solvent:water ratio in the diluted extracts for LC-HRMS analysis.

3.5. LC-HRMS measurement

The LC-HRMS measurements were performed on the QExactive HF-Orbitrap high-resolution mass spectrometer coupled to the Vanquish UHPLC system (Thermo Fisher Scientific). The following settings were used:

3.5.1. Liquid chromatography

The reversed-phase separation was performed on a C18 column (3.5 µm; 2.1 x 150 mm, Waters XBridge®) coupled to a C18 precolumn (4 x 3 mm, Phenomenex). The eluent system consisted of (A) H₂O + 0.1 % FA and (B) MeOH + 0.1 % FA that was passed through the column at a flow rate of 0.25 ml/min with a chromatographic gradient according to [1].

3.5.2. Mass spectrometry

Full scan MS measurements were performed in positive mode and resolving power setting 120,000 at *m/z* 200 and *m/z* range of 100 to 1000 were chosen. The HESI source was operated at 320°C and +3.5 kV was applied to the spray needle. The auxiliary and sheath gas flow rates were 5 and 55 units and the setpoint for the auxiliary gas heater was set to 350°C.

3.6. Data processing and statistical evaluation

The LC-HRMS raw data were converted into the mzXML file format via MSConvert of the ProteoWizard open-source software-package [31]. The files were further processed by the in-house developed MetExtract II software according to the procedure described in [8]. Briefly, the software searched for pairs of co-eluting chromatographic peaks, which represent both the native and the uniformly ¹³C-labeled metabolite form. All other LC-HRMS features were discarded. For each such found chromatographic peak pair, the *m/z* value for the monoisotopic form, retention time, charge number, mass ion species and neutral loss (if possible) and its total number of carbon atoms (derived from the difference between the monoisotopic and fully ¹³C-labeled isotopologs) were automatically calculated from the data. Moreover, the software matched ions from different samples into a comprehensive data matrix and integrated the peak areas of the detected monoisotopic and fully ¹³C-labeled isotopologs.

All statistical analyses were implemented in the R programming language for statistical computing [19; version 3.5.3]. A metabolite had to be present in all samples of a biological group (Toxin or Control) for being considered to be present in that biological group. Moreover, any metabolite that was detected in more than 0 samples but less than all samples of a group was excluded from the statistical analysis. Missing values were replaced with 0. For multivariate analysis, the data matrix was mean-centered and auto-scaled [32]. Ellipses in the PCA scores plot indicate the 95% confidence interval. For univariate analysis, a non-paired, two-sided t-test was calculated and a critical alpha value of 0.05 was chosen together with a minimum mean-fold-change of ≥ 2 or ≤ 0.5 respectively, as a significance threshold. Metabolites only detected in the Control or the Toxin group were set to a fold change of infinite (i.e. only present in the Toxin group) or 0 (i.e. only present in the Control group).

4. Conclusions

The reliable and efficient assessment of the performance of untargeted LC-HRMS based metabolomics approaches remains to be a major challenge. The present study demonstrates that stable isotope-assisted techniques can be successfully used to evaluate and improve metabolome coverage as well as the accuracy of comparative quantification of the detected metabolites. An experimental-condition-matched ^{13}C -labeled biological extract (GLMe-IS) was used to detect all truly biology derived metabolites as well as to assess and compensate putatively different ion suppression/enhancement effects in that may occur in samples representing different experimental conditions. Here, plant treatment with the mycotoxin deoxynivalenol was used to exemplify abiotic stress conditions and to illustrate the induction of numerous treatment-specific compounds. The presented approach can be easily extended to both other biological systems as well as further experimental treatments. Here, a total of 68 metabolites (making up 6% of the total detected metabolome) were uniquely found either in toxin-treated or H_2O -treated control samples. These metabolites were reliably detected as treatment-related metabolites with the aid of GLMe-IS. Further, such an application may also benefit the subsequent analytical steps aiming at sum formula elucidation and annotation of unknown metabolites. Amongst the metabolites which are shared between both experimental conditions, roughly a 26% found to have significantly different abundances. As those metabolites are assumed to be significantly affected by the toxin treatment, they would be of most interest in a typical biological study. Hence, reliable significance testing is crucial in untargeted studies, as based on this, candidate metabolites would be selected and annotated for biological interpretation and even be purified for full structure characterization as well as further biological studies. While under the conditions applied in the present study, matrix effects (i.e. ion suppression/enhancement) turned out to be only moderate, still a considerable proportion of the metabolites, which were assumed to differ significantly between the two conditions were wrongly classified. It is important to mention that the effect of different matrices on comparative quantification and subsequent statistical analysis is a potential error source of untargeted metabolomics studies and has to be evaluated from case to case. Wrong classification of metabolites as a result of significance testing might be severe, especially if the metabolic composition is largely altered by experimental perturbations or e.g. if different tissue types are compared in biological studies. While the accurate assessment of (absolute) matrix effects remains challenging, GLMe-IS enables the metabolome wide compensation of matrix effects. This can be used for a more accurate comparative quantification, especially in untargeted metabolomics applications.

Supplementary Materials: S1: Feature map of detected metabolites. Different color code represents metabolites that have shown different results for the significantly differing metabolites in the comparative quantification test (Toxin versus Control). In red are highlighted those which are significantly different (between Toxin and Control) only if absolute native abundances are used, and in green only those obtained after normalized (native to labeled) abundances are applied

Author Contributions: Conceptualization, A.C., C.B., R.S., and M.D.; data curation, C.B., formal analysis, C.B.; funding acquisition, R.S., M.L., and H.B.; investigation, A.C.; methodology, A.C., M.D., C.B., and R.S.; project administration, R.S., C.B., and M.D.; resources, A.C., M.D., A.P., and K.X.; software, C.B.; supervision: R.S.;

validation, A.C., C.B., and M.D.; visualization, C.B., A.C., and M.D.; writing – original draft, A.C. and C.B.; writing – review and editing, R.S., M.D., M.L., and A.P.. All authors have contributed, read and agreed to the manuscript.

Funding: This research was funded by the Austrian Science Fund (project SFB Fusarium F3715 and F3711) and the Provincial Government of Lower Austria (projects NoBiTUM, OMICS 4.0). The APC was funded by Austrian Science Fund.

Acknowledgments: The authors would like to thank Barbara Steiner from the Institute of Biotechnology in Plant Production (BOKU, IFA-Tulln, Austria) for her help with the wheat plant cultivation.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Bueschl, C., Kluger, B., Lemmens, M., Adam, G., Wiesenberger, G., Maschietto, V., Marocco, A., Strauss, J., Bodi, S., Thallinger, G.G., Krska, R., and Schuhmacher, R. A novel stable isotope labelling assisted workflow for improved untargeted LC-HRMS based metabolomics research. *Metabolomics* **2014**, *10*, 754-769. DOI: 10.1007/s11306-013-0611-0.
2. Ouerdane, L., Meija, J., Bakirdere, S., Yang, L., and Mester, Z. Nonlinear signal response in electrospray mass spectrometry: implications for quantitation of arsenobetaine using stable isotope labeling by liquid chromatography and electrospray Orbitrap mass spectrometry. *Anal Chem* **2012**, *84*, 3958-3964. DOI: 10.1021/ac203137n.
3. Chamberlain, C.A., Rubio, V.Y., and Garrett, T.J. Impact of matrix effects and ionization efficiency in non-quantitative untargeted metabolomics. *Metabolomics* **2019**, *15*, 135. DOI: 10.1007/s11306-019-1597-z.
4. Huang, X., Chen, Y.J., Cho, K., Nikolskiy, I., Crawford, P.A., and Patti, G.J. X13CMS: global tracking of isotopic labels in untargeted metabolomics. *Anal Chem* **2014**, *86*, 1632-1639. DOI: 10.1021/ac403384n.
5. Leeming, M.G., Isaac, A.P., Pope, B.J., Cranswick, N., Wright, C.E., Ziogas, J., O'Hair, R.A., and Donald, W.A. High-resolution twin-ion metabolite extraction (HiTIME) mass spectrometry: nontargeted detection of unknown drug metabolites by isotope labeling, liquid chromatography mass spectrometry, and automated high-performance computing. *Anal Chem* **2015**, *87*, 4104-4109. DOI: 10.1021/ac504767d.
6. Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., VanderGheynst, J., Fiehn, O., and Arita, M. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods* **2015**, *12*, 523-526. DOI: 10.1038/nmeth.3393.
7. Capellades, J., Navarro, M., Samino, S., Garcia-Ramirez, M., Hernandez, C., Simo, R., Vinaixa, M., and Yanes, O. geoRge: A Computational Tool To Detect the Presence of Stable Isotope Labeling in LC/MS-Based Untargeted Metabolomics. *Anal Chem* **2016**, *88*, 621-628. DOI: 10.1021/acs.analchem.5b03628.
8. Bueschl, C., Kluger, B., Neumann, N.K.N., Doppler, M., Maschietto, V., Thallinger, G.G., Meng-Reiterer, J., Krska, R., and Schuhmacher, R. MetExtract II: A Software Suite for Stable Isotope-Assisted Untargeted Metabolomics. *Analytical Chemistry* **2017**, *89*, 9518-9526. DOI: 10.1021/acs.analchem.7b02518.
9. Saia, S., Fragasso, M., De Vita, P., and Beleggia, R. Metabolomics Provides Valuable Insight for the Study of Durum Wheat: A Review. *J Agric Food Chem* **2019**, *67*, 3069-3085. DOI: 10.1021/acs.jafc.8b07097.

10. Allen, D.K., Bates, P.D., and Tjellstrom, H. Tracking the metabolic pulse of plant lipid production with isotopic labeling and flux analyses: Past, present and future. *Prog Lipid Res* **2015**, *58*, 97-120. DOI: 10.1016/j.plipres.2015.02.002.
11. Freund, D.M. and Hegeman, A.D. Recent advances in stable isotope-enabled mass spectrometry-based plant metabolomics. *Curr Opin Biotechnol* **2017**, *43*, 41-48. DOI: 10.1016/j.copbio.2016.08.002.
12. Taylor, P.J. Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography-electrospray-tandem mass spectrometry. *Clin Biochem* **2005**, *38*, 328-334. DOI: 10.1016/j.clinbiochem.2004.11.007.
13. Wang, S., Cyronak, M., and Yang, E. Does a stable isotopically labeled internal standard always correct analyte response?: A matrix effect study on a LC/MS/MS method for the determination of carvedilol enantiomers in human plasma. *Journal of pharmaceutical and biomedical analysis* **2007**, *43*, 701-707.
14. Zhou, W., Yang, S., and Wang, P.G. Matrix effects and application of matrix effect factor. *Bioanalysis* **2017**, *9*, 1839-1844. DOI: 10.4155/bio-2017-0214.
15. Liu, G.W., Ji, Q.C., and Arnold, M.E. Identifying, Evaluating, and Controlling Bioanalytical Risks Resulting from Nonuniform Matrix Ion Suppression/Enhancement and Nonlinear Liquid Chromatography-Mass Spectrometry Assay Response. *Analytical Chemistry* **2010**, *82*, 9671-9677. DOI: 10.1021/ac1013018.
16. Matuszewski, B.K., Constanzer, M.L., and Chavez-Eng, C.M. Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS. *Analytical Chemistry* **2003**, *75*, 3019-3030. DOI: 10.1021/ac020361s.
17. Denery, J.R., Nunes, A.A., and Dickerson, T.J. Characterization of differences between blood sample matrices in untargeted metabolomics. *Anal Chem* **2011**, *83*, 1040-1047. DOI: 10.1021/ac102806p.
18. Ceranic, A., Doppler, M., Buschl, C., Parich, A., Xu, K., Koutnik, A., Buerstmayr, H., Lemmens, M., and Schuhmacher, R. Preparation of uniformly labelled (¹³C)- and (¹⁵N)-plants using customised growth chambers. *Plant Methods* **2020**, *16*, 46. DOI: 10.1186/s13007-020-00590-9.
19. Doppler, M., Kluger, B., Bueschl, C., Steiner, B., Buerstmayr, H., Lemmens, M., Krska, R., Adam, G., and Schuhmacher, R. Stable Isotope-Assisted Plant Metabolomics: Investigation of Phenylalanine-Related Metabolic Response in Wheat Upon Treatment With the Fusarium Virulence Factor Deoxynivalenol. *Front Plant Sci* **2019**, *10*, 1137. DOI: 10.3389/fpls.2019.01137.
20. Doppler, M., Bueschl, C., Kluger, B., Koutnik, A., Lemmens, M., Buerstmayr, H., Rechthaler, J., Krska, R., Adam, G., and Schuhmacher, R. Stable Isotope-Assisted Plant Metabolomics: Combination of Global and Tracer-Based Labeling for Enhanced Untargeted Profiling and Compound Annotation. *Front Plant Sci* **2019**, *10*, 1366. DOI: 10.3389/fpls.2019.01366.
21. Warth, B., Parich, A., Bueschl, C., Schoefbeck, D., Neumann, N.K., Kluger, B., Schuster, K., Krska, R., Adam, G., Lemmens, M., and Schuhmacher, R. GC-MS based targeted metabolic profiling identifies changes in the wheat metabolome following deoxynivalenol treatment. *Metabolomics* **2015**, *11*, 722-738. DOI: 10.1007/s11306-014-0731-1.
22. Macoy, D.M., Kim, W.Y., Lee, S.Y., and Kim, M.G. Biotic stress related functions of hydroxycinnamic acid amide in plants. *Journal of Plant Biology* **2015**, *58*, 156-163. DOI: 10.1007/s12374-015-0104-y.
23. Kikot, G.E., Hours, R.A., and Alconada, T.M. Contribution of cell wall degrading enzymes to pathogenesis of *Fusarium graminearum*: a review. *Journal of Basic Microbiology* **2009**, *49*, 231-241. DOI: 10.1002/jobm.200800231.

24. Contreras, M.D.M., Yopez, A.J.M., and Martinez, R.R. Determination of deoxynivalenol (DON) in wheat, barley and corn and its relationship with the levels of total molds, *Fusarium* spp., infestation percentage, and water activity. *Archivos Latinoamericanos De Nutricion* **2000**, 50, 183-186.
25. Silvestro, L., Tarcomnicu, I., and Rizea, S., *Matrix effects in mass spectrometry combined with separation methods—comparison HPLC, GC and discussion on methods to control these effects*. 2013: InTech.
26. Mizuno, H., Ueda, K., Kobayashi, Y., Tsuyama, N., Todoroki, K., Min, J.Z., and Toyo'oka, T. The great importance of normalization of LC-MS data for highly-accurate non-targeted metabolomics. *Biomed Chromatogr* **2017**, 31. DOI: 10.1002/bmc.3864.
27. Chen, D., Han, W., Su, X., Li, L., and Li, L. Overcoming Sample Matrix Effect in Quantitative Blood Metabolomics Using Chemical Isotope Labeling Liquid Chromatography Mass Spectrometry. *Anal Chem* **2017**, 89, 9424-9431. DOI: 10.1021/acs.analchem.7b02240.
28. Han, W. and Li, L. Matrix effect on chemical isotope labeling and its implication in metabolomic sample preparation for quantitative metabolomics. *Metabolomics* **2015**, 11, 1733-1742. DOI: 10.1007/s11306-015-0826-3.
29. Doppler, M., Kluger, B., Bueschl, C., Schneider, C., Krska, R., Delcambre, S., Hiller, K., Lemmens, M., and Schuhmacher, R. Stable Isotope-Assisted Evaluation of Different Extraction Solvents for Untargeted Metabolomics of Plants. *International Journal of Molecular Sciences* **2016**, 17. DOI: 10.3390/ijms17071017.
30. Redestig, H., Kobayashi, M., Saito, K., and Kusano, M. Exploring matrix effects and quantification performance in metabolomics experiments using artificial biological gradients. *Anal Chem* **2011**, 83, 5645-5651. DOI: 10.1021/ac200786y.
31. Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **2008**, 24, 2534-2536. DOI: 10.1093/bioinformatics/btn323.
32. van den Berg, R.A., Hoefsloot, H.C., Westerhuis, J.A., Smilde, A.K., and van der Werf, M.J. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **2006**, 7, 142. DOI: 10.1186/1471-2164-7-142.