Journal of Chromatography A, 1389 (2015) 112–119

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Isotopologue ratio normalization for non-targeted metabolomics☆

Daniel Weindl, André Wegner, Christian Jäger, Karsten Hiller*

Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 7, Avenue des Hauts-Fourneaux, L-4362 Esch-Belval, Luxembourg

ARTICLE INFO

Article history:
Received 3 December 2014
Received in revised form 9 February 2015
Accepted 10 February 2015
Available online 17 February 2015

Keywords:
IDMS
NTFD
Non-targeted metabolomics
Normalization
Quantification

ABSTRACT

Robust quantification of analytes is a prerequisite for meaningful metabolomics experiments. In non-targeted metabolomics it is still hard to compare measurements across multiple batches or instruments. For targeted analyses isotope dilution mass spectrometry is used to provide a robust normalization reference.

Here, we present an approach that allows for the automated semi-quantification of metabolites relative to a fully stable isotope-labeled metabolite extract. Unlike many previous approaches, we include both identified and unidentified compounds in the data analysis. The internal standards are detected in an automated manner using the non-targeted tracer fate detection algorithm. The ratios of the light and heavy form of these compounds serve as a robust measure to compare metabolite levels across different mass spectrometric platforms. As opposed to other methods which require high resolution mass spectrometers, our methodology works with low resolution mass spectrometers as commonly used in gas chromatography electron impact mass spectrometry (GC–EI-MS)-based metabolomics.

We demonstrate the validity of our method by analyzing compound levels in different samples and show that it outperforms conventional normalization approaches in terms of intra- and inter-instrument reproducibility. We show that a labeled yeast metabolite extract can also serve as a reference for mammalian metabolite extracts where complete stable isotope labeling is hard to achieve.

1. Introduction

Metabolomics, the attempt to measure the levels of all metabolites of a given system under the given conditions, has become increasingly important in biomedical research [1,2]. Metabolomics data can be the basis for biomarker discoveries [3], biotechnological applications, or metabolic flux analysis [4–7].

However, analytical variance poses problems to the comparison of measurements from different runs or instruments, especially in non-targeted metabolomics. Common data treatments like total ion current normalization cannot be used for cross-platform comparisons and only account for certain types of errors like fluctuations in overall sensitivity. Often these techniques are limited to a set of very similar metabolite profiles. Normalization on pool samples can be performed, but this does not take into account the potentially different metabolite profiles with different matrix effects.

Analytical variance is best addressed by adding stable isotope-enriched internal standards to the sample. The addition of stable isotope-enriched compounds to a sample before mass spectrometric analysis is referred to as isotope dilution mass spectrometry (IDMS). IDMS is commonly used for targeted quantitative metabolomics. In non-targeted metabolomics many compounds remain unidentified and can, thus, not be included in any standard mixture. However, this shortcoming can be circumvented by using fully labeled metabolite extracts of a similar sample as reference. For example, metabolite extracts of fully ^{13}C-enriched yeast, bacteria, plant, algae, and filamentous fungi have been used successfully as complex standard mixtures for large scale metabolite quantification or determination of sum formulas [8–13]. So far, they have not been used for automated non-targeted metabolomics.

For liquid chromatography electrospray ionization high resolution mass spectrometry (LC–ESI–HRMS) data, there are methods for non-targeted IDMS available for both semi-quantification and identification of analytes. Bueschl et al. [13] applied complete isotopic enrichment, whereas the isotopic ratio outlier analysis (IROA) [14] uses partial stable isotopic enrichment. Pairs of labeled and unlabeled compounds are automatically detected from the typical isotopic peak patterns. However, these methods are not applicable for low resolution mass spectrometers and hard ionization techniques like electron ionization (EI) which produce a large

☆ Presented at the 30th International Symposium on Chromatography (ISC 2014), Salzburg, Austria, 14–18 September 2014.

* Corresponding author. Tel.: +352 4666446136.
E-mail address: karsten.hiller@uni.lu (K. Hiller).

http://dx.doi.org/10.1016/j.chroma.2015.02.025

number of fragment ions. Without accurate mass measurements, mass spectral peak patterns arising from fragmentation often cannot clearly be distinguished from isotopic peak patterns. Therefore, other means are necessary for the automated and non-targeted detection of stable isotope-labeled compounds in such data.

Here, we present an approach for GC–EI-MS metabolomics that allows for the robust normalization or semi-quantification of both identified and unidentified metabolites relative to a spiked-in stable isotope-labeled metabolite extract. We used a similar approach as Wu et al. [9] who applied fully ^{13}C-labeled yeast metabolite extract as internal standard. However, their analysis has been very targeted and did not make use of the information on unidentified analytes. We overcome this limitation by employing the non-targeted tracer fate detection (NTFD) algorithm [15] to detect all isotopically enriched compounds within a reference mixture in an automated manner. The intensity ratios of native compounds and the corresponding references are then used to normalize analyte levels in the sample of interest. Additionally, the number of carbon and nitrogen atoms of the unidentified compounds can be obtained. Using this experimental setup, absolute quantification of identified compounds is possible as shown by others [9]. We demonstrate the validity of our methodology by comparing intra- and inter-instrument variation to conventional methods.

2. Materials and methods

2.1. Materials

Chemicals were purchased from Sigma–Aldrich, unless indicated differently. All solvents used were of grade *Chromasolv* or better.

2.2. Culture conditions

To produce the fully labeled reference mixture, *Saccharomyces cerevisiae* strain S90 mating type α was grown on YPD agar at 30 °C for 48 h. A single colony was transferred to 5 mL of liquid YPD medium for an overnight culture, and then to YNB medium containing [$^{15}N_2$]ammonium sulfate and D-[U-^{13}C]glucose (Cambridge Isotope Laboratories, 99% isotopic purity) as sole nitrogen and carbon source again over night. Cultures were incubated on a rotary shaker (Infors Multitron) at 30 °C and 200 rpm. Following another 5 mL YNB labeling culture over night, culture volume was increased to 100 mL. Cultures were inoculated at OD_{600} = 0.1, cell growth was monitored using a cell density meter (Biowave CO8000) and metabolites were extracted in mid-exponential growth phase.

S. cerevisiae strain YJM789 was grown on YPD agar at 30 °C for 48 h. After an over night culture in 5 mL liquid YPD medium, a 10 mL YPD culture was prepared and extracted in mid-exponential growth phase.

A549 cells (ATCC CCL-185) were grown in multi-well plates in DMEM medium (Invitrogen) supplemented with 10% (v/v) FBS and 1% (v/v) penicillin/streptomycin in an incubator (Sanyo) at 21% O_2, 5% CO_2 at 37 °C.

2.3. Metabolite extraction and standard addition

The yeast culture was centrifuged at 3900 × *g* for 3 min at −10 °C, the pellet resuspended in 2 mL extraction fluid (50%, v/v, methanol in water, −20 °C) and transferred to a reaction tube, prefilled with 600 mg acid-washed glass beads (ø150–212 μm, Sigma–Aldrich). 10 mL of the YPD and 25 mL of the YNB culture were harvested at $OD_{600} \approx 2$. Cell lysis was performed using a Precellys24 (Bertin) homogenizer, equipped with a Cryolys cooling option held at 0 °C, and the following program: 2 × 30 s at 6800 rpm with 30 s pause in-between. After adding 500 μL chloroform, thorough mixing, and centrifugation at 14,000 × *g* for 5 min at 4 °C, the upper aqueous phase was used for analysis of polar metabolites. The labeled polar metabolite extract was diluted 1:10 in methanol:water (1:1, v:v) and stored at −80 °C until use. The interphase forming during the extraction was hydrolysed in 1.5 mL of 6N hydrochloric acid at 99 °C over night. The supernatant was evaporated and the residue was extracted with 1.5 mL methanol:water (1:1, v:v) and diluted 1:10 with methanol:water (1:1, v:v).

To generate the library of labeled compounds 30 μL of unlabeled metabolite extract and 4 μL of the unlabeled hydrolysate were measured separately, and in mixture with 30 μL and 8 μL of $^{13}C^{15}N$-labeled polar extract and interphase.

As internal standards for the yeast YJM789 samples 6 μL of $^{13}C^{15}N$-labeled yeast S90 polar extract and 10 μL interphase hydrolysate were spiked into 100 μL of the polar extract of interest.

A549 cell extract was prepared from 4×10^5 cells. Cells were washed with 1 mL 0.9% (w/v) NaCl and quenched with 400 μL methanol (−20 °C). After adding 400 μL water (4 °C), the cells were scraped off with a cell scraper and the cell suspension was transferred into an Eppendorf tube containing 400 μL chloroform at −20 °C. Tubes were shaken for 20 min at 1400 rpm and 4 °C and centrifuged for 5 min at 16,100 × *g* at 4 °C. A detailed protocol is available in [16]. To 300 μL of the aqueous phase, 6 μL of uniformly $^{13}C^{15}N$-labeled S90 polar extract and 10 μL of interphase hydrolysate were added.

2.4. Sample preparation & GC–MS measurement

The metabolite extracts were transferred to glass vials with micro inserts and dried in a CentriVap vacuum evaporator (Labconco) at −4 °C. Automated sample derivatization was performed by using a multi-purpose sampler (GERSTEL). Dried samples were dissolved in 15 μL pyridine, containing 20 mg/mL methoxyamine hydrochloride and incubated at 40 °C for 60 min under shaking. In a second step, 15 μL *N*-methyl-*N*-trimethylsilyl-trifluoroacetamide (MSTFA) were added to the samples and they were further incubated at 40 °C for 30 min under continuous shaking.

GC–MS analysis was performed on an Agilent 7890A GC coupled to an Agilent 5975C inert XL Mass Selective Detector (Agilent Technologies). A sample volume of 1 μL was injected into a split/splitless inlet, operating in splitless mode at 270 °C. The gas chromatograph was equipped with a 30 m DB-35MS capillary column with a 5 m DuraGuard capillary in front of the analytical column (Agilent J&W GC Column).

Helium was used as carrier gas with a constant flow rate of 1.0 ml/min. The GC oven temperature was held at 80 °C for 6 min and increased to 300 °C at 6 °C/min. After 10 min, the temperature was increased at a rate of 10 °C/min to 325 °C and held for 4 min. The total run time was 59.167 min.

The transfer line temperature was set to 280 °C. The MS was operating under electron ionization at 70 eV. The MS source was held at 230 °C and the quadrupole at 150 °C. Full scan mass spectra were acquired from *m/z* 70 to *m/z* 800.

For inter-instrument comparison the samples were also measured on an Agilent 7890B gas chromatograph coupled to an Agilent 5977A mass spectrometer using the same column type and temperature program.

2.5. Chromatogram preprocessing

Deconvolution of mass spectra, peak picking, integration, and retention index calibration were performed using the MetaboliteDetector software [17]. Compounds were identified using an in-house mass spectra library. The following deconvolution settings were applied: Peak threshold: 5; Minimum peak height: 5; Bins per scan: 10; Deconvolution width: 5 scans; No baseline

adjustment; Minimum 20 peaks per spectrum; No minimum required base peak intensity. Retention index calibration was based on an C_{10}–C_{40} even *n*-alkane mixture.

2.6. Generation of compound library for quantification

A library of all detected compounds present in the labeled yeast extract was generated using an adapted implementation of the NTFD algorithm [15,8] which implements the following filters and generates compound libraries for MetaboliteDetector. For each compound, the isotopically enriched fragments were determined. Therefore, the yeast S90 extracts have been measured in triplicate. The *m/z* of the M+0 peak and the highest isotopic peak M+*N* were considered as potential quantification ions for the unlabeled and labeled form of the corresponding compound. The following NTFD settings were applied: Minimal number of labeled fragments: 1; Minimum (maximum) amount of label: 0.1 (0.9); M1 correction: 0; Maximum fragment deviation: 0.1. Signals at $m/z \leq 147$ were excluded. As a filter for proper isotope clusters, the unlabeled spectrum was required to have an M+1 peak with an intensity of $0.01 \cdot M_0 < M_1 < M_0$. Fragments with an M−1 peak present with $M_{-1} > 0.2 \cdot M_0$ indicating overlapping fragment ion clusters were excluded.

Of the labeled fragments detected, only those which had their labeled and unlabeled peaks separated by three mass units (M+*N* with $N \geq 3$) and had an M+*N* intensity in the unlabeled spectrum of $M_N < 0.05 \cdot M_0$ were considered for further analysis. The mass spectra recorded for the mixture of light and heavy compounds, as well as their corresponding retention indices and quantification ions were collected for quantification of the analyte and reference compound in the sample of interest. We used the spectrum of the light and heavy mixture instead of those of the pure light or heavy form, because it ensures the best spectrum match with the same analyte in the sample of interest in which the labeled and unlabeled form are ideally present in equal amounts.

2.7. IDMS normalization

For the IDMS normalization of analyte levels we calculated the ratio of the summed heavy and light ion intensities. The peak areas were obtained from the MetaboliteDetector batch quantification in targeted-mode using the compound library generated in the previous step and the following settings: ΔRI: 5; Scoring method: RI + Spec; Req. score: 0.7; Compound reproducibility: 1; Required *S*/*N*: 5; Minimum number of ions: 15; No extended SIC scan.

2.8. Validation

We compared our isotopologue ratios to M+0 intensities normalized to total ion current. For the latter, all intensity values were divided by the summed intensity of all peaks in all mass spectra. This was performed within MetaboliteDetector. The normalized intensities of all light quantification ions that were chosen for the isotopologue ratios were summed up. For single internal standard normalization all intensities were divided by the summed intensities of the M_N peaks of ([U-^{13}C, U-^{15}N]ornithine) 4TMS (used for YJM789, *m/z* 192, 250, 264, 336, 355, 427) or ([U-^{13}C]malic acid) 3TMS (used for A549, *m/z* 236, 249, 339, 354).

To determine the injection-to-injection variability, the same derivatized sample was injected three times in a row. For all metabolites present in the reference library, we calculated the relative standard deviation of the isotopologue ratios as well as those of the TIC- and single internal standard- normalized intensities.

For the inter-instrument comparison, a derivatized sample was injected into two different GC–MS models using the same column type and temperature program. The intensities of instrument A were plotted over those of instrument B to show the correlation (Fig. 4B). Normalization was performed for visualization of the quantification results from the three approaches in a single plot. For this purpose, every data point was divided by the range of values of the respective normalization method.

3. Theory

3.1. Method overview

Our non-targeted IDMS normalization approach is based on a complex stable isotope labeled metabolite mixture as internal standard and involves the following steps (Fig. 1):

- Generation of a stable isotope labeled reference mixture.
- Determination of all stable isotope-enriched compounds within the reference mixture in a non-targeted manner.
- Selection of suitable quantification ions for those compounds.
- Spiking the reference mixture into a sample of interest prior to GC–EI-MS measurement.
- Quantification of the native compound relative to the corresponding labeled internal standard.

3.2. Generation of reference mixtures

As a reference mixture, we used a metabolite extract from a fully isotopically enriched yeast culture, because it provides a reference for a large number of known and unknown compounds. For that purpose, we cultivated yeast in a batch culture on defined minimal medium containing ^{13}C and ^{15}N substrates. As opposed to earlier studies [19,9,8], we performed simultaneous ^{15}N- and ^{13}C-labeling in an attempt to further separate high and low mass variant of our analytes. If isotopic peak clusters of the high and low mass variant of a fragment are overlapping, this fragment cannot be used for quantification. This matters for subsequent GC–MS analysis where polar analytes are often alkylsilylated to increase their volatility. The relatively high natural abundance of silicon isotopes and the large number of alkylcarbons introduced into the molecule increase the abundance of isotopic peaks. Simultaneous labeling of both ^{15}N as well as ^{13}C reduces the number of cases where isotopic peaks of the derivatized labeled and unlabeled metabolites overlap and, therefore, cannot be used for quantification. Apart from this reduced number of quantification fragments, the presented method can also be used with ^{13}C-labeling alone.

This isotopically enriched yeast culture was homogenized and metabolites were extracted using a methanol, water, chloroform mixture. During the extraction process three phases form: A chloroform phase containing non-polar metabolites, an aqueous phase containing polar metabolites, and an interphase containing precipitated proteins and nucleic acids. We were only interested in metabolites of the polar phase and used this phase as the reference mixture. Additionally, we performed an acid hydrolysis of the interphases formed during the extraction (see Section 2 for details) and supplemented the previous polar extract with this mixture to increase the concentration of free amino acids and nucleobases.

3.3. Detection of labeled compounds

We detected all labeled compounds within the spike-in extract in a non-targeted and automated manner using the NTFD algorithm (Fig. 1a–e) [15]. For this purpose, an unlabeled yeast extract as well as a mixture of labeled and unlabeled extract were measured. NTFD matches and subtracts the spectra of each analyte found in both samples and detects isotopic enrichment as peaks

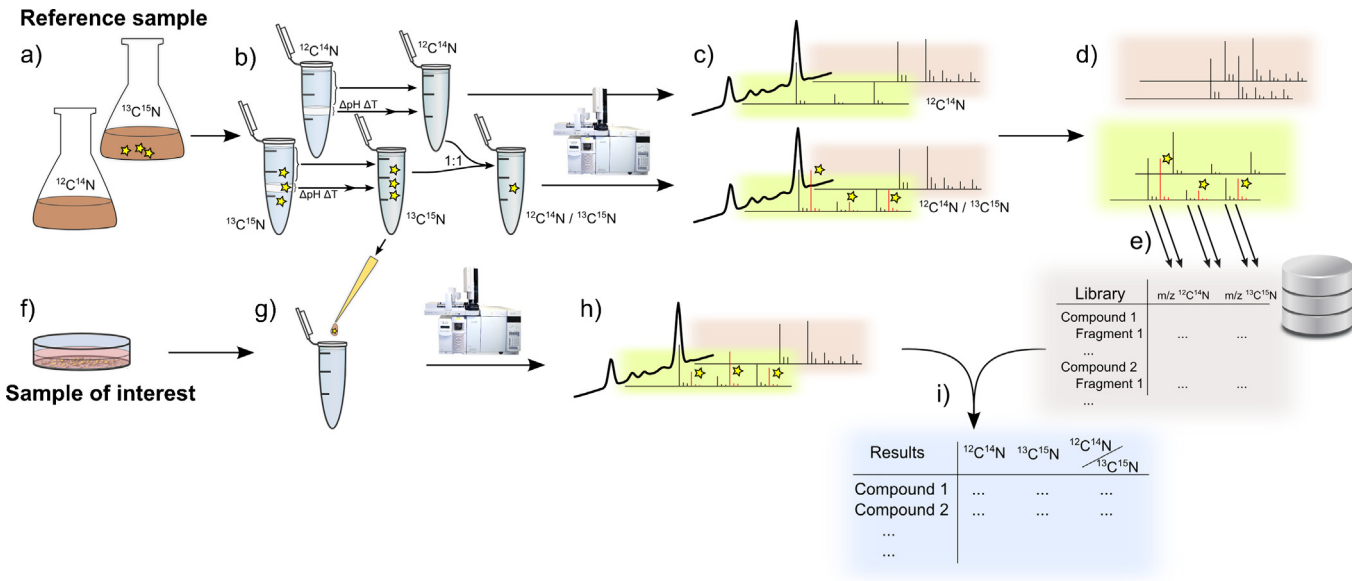

Fig. 1. Experimental setup. (a) A reference organism is grown simultaneously in defined medium and in a medium where all carbon and nitrogen sources are substituted by their fully stable isotope-labeled analogues. (b) Metabolites are extracted using water:methanol:chloroform. The protein- and nucleic acid- containing interphase is hydrolyzed, pooled with the polar metabolites and used as reference extract. (c, d) NTFD is used to detect all stable isotope-labeled compounds and fragments, as well as the *m/z* ratios of their light and heavy isotopologues. Therefore, the unlabeled extract and a mixture of labeled and unlabeled extract are measured with GC–EI-MS. (e) The spectra of all these labeled compounds and selected quantification ions are collected in a reference library to be used to match and quantify compounds within a sample of interest. (f, g) An aliquot of the labeled reference mixture is added to a sample of interest. (h, i) The previously determined ions are used for quantification of the detected compounds. The ratios of the intensities of light and heavy forms for each analyte provide a robust measure for the comparison of metabolite amounts across experiments.

in the resulting difference spectrum [15]. The output is a list of all stable isotope-enriched compounds and the *m/z* ranges as well as the mass isotopomer distributions for all enriched fragments

3.4. Selection of quantification ions

Once we determined the *m/z* ranges of the labeled fragments via NTFD, we selected potential quantification ions for the labeled and unlabeled compounds (Fig. 1e). The first peak (M+0, in formulas M_0) of the isotope cluster arises from the unlabeled isotopologue and, thus, represents the native compound from the sample of interest. As corresponding reference quantification ion, we selected the most abundant peak (M+*N*, in formulas M_N) of the remaining isotope cluster, which arises from the maximal ^{13}C- and ^{15}N-incorporation (Fig. 2A-c, d).

In low-mass fragments or in spectra of small metabolites, there is often an overlap of the natural isotope clusters of the unlabeled and $^{13}C^{15}N$-labeled compound (Fig. 2A-a). Such overlapping fragments impair the IDMS-based quantification and, hence, were excluded if the M+*N* abundance in the spectrum of the unlabeled

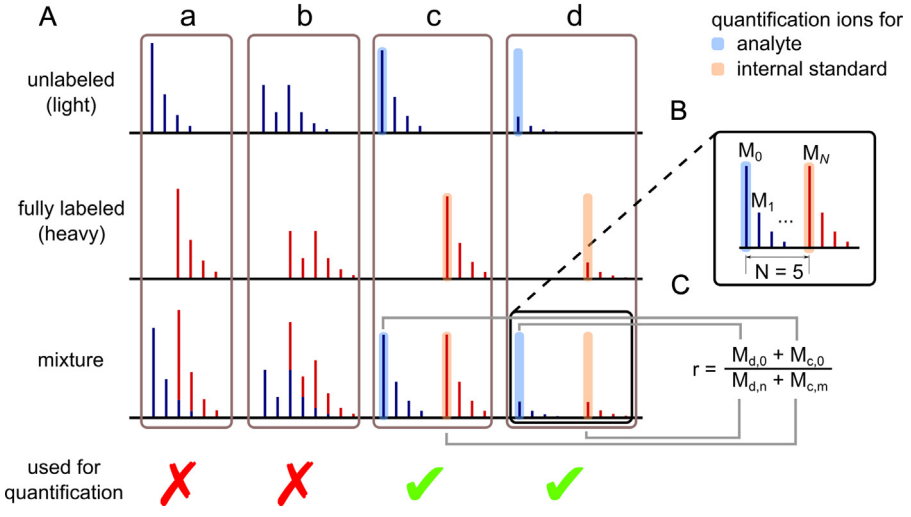

Fig. 2. (A) Schematic mass spectra of an analyte in the unlabeled sample, in the fully isotopically labeled reference mixture, and in an equimolar mixture thereof. In the mass spectrum of the heavy isotopologue, the labeled fragments are shifted towards higher mass. Due to the natural isotope abundance in the derivatization reagents, isotopic peaks are also present in the mass spectrum of the fully labeled metabolite. (a) The small *m/z* difference between light (M_0) and heavy (M_N) isotopologues causes an overlap of the natural isotope clusters of the forms. Such fragments were not used for quantification. (b) Multiple fragments of the unlabeled compound are overlapping, rendering these ions unsuitable for quantification. (c, d) Natural isotope clusters of light and heavy forms are well-separated. These fragments can be used for quantification. (B) The *m/z* difference of light and heavy form provides the number of isotopes contained in the given fragment. (C) Analyte levels are quantified relatively to the corresponding internal standard. The information from all suitable fragments are combined for a more robust result. $M_{i,j}$ denotes the intensity of the *j*th isotopic peak of fragment *i*; *m* and *n* represent *N* of the M_N peak in the respective fragments.

compound was over 5% of M+0. A slight overlap was accepted to not exclude spectra solely because of possible small impurities. However, if this M+*N* signal is not analytical noise but an isotopic peak, this needs to be considered for large M_0/M_N ratios in the sample of interest. In case, the abundance of the labeled compound is very low as compared to the unlabeled compound, even low natural isotopologue contribution will heavily influence the intensity ratio. As additional filter, quantification ions in overlapping fragments within the spectrum of the unlabeled compound (Fig. 2A-b) were excluded. All remaining ion pairs of each spectrum were used for quantification. The number of quantification ion pairs retained per compound ranged from one to seven, with an average of two.

Furthermore, the selected quantification ions hold information on the elemental composition. The mass difference *N* of the detected light (M+0) and heavy (M+*N*) ions provides the combined number of carbon and nitrogen atoms in each of these fragments. This information is very valuable for unidentified compounds. However, due to the hard electron ionization this number does not necessarily correspond to the number of carbon and nitrogen atoms of the underivatized compound. Nevertheless, it is helpful as a lower bound and in practice the number of atoms is often correct, because in the heavier fragment ions the atoms lost during fragmentation are often derived from the derivatization reagent which does not contribute to the mass shift. For fragments of which the elemental composition was known [19], the number of carbon and nitrogen atoms matched the mass difference of the light and heavy ions (Supplemental Table S1). The low mass spectrometric resolution did not have any negative impact on the determination of the number of carbon and nitrogen atoms.

3.5. Quantification

All isotopically enriched spectra and quantification ions selected in the previous step were collected in a reference library. This library was subsequently used in a targeted analysis to quantify the M+0 and M+*N* abundance of those compounds within the sample of interest (Figs. 1i and 2C).

From these data, we calculated the ratio of the summed intensities of all M+0 and M+*N* peaks:

$$r = \frac{\sum M_0}{\sum M_N}$$

The use of multiple fragments makes the quantification more robust. The relative contribution of the different fragments increases with their signal intensity.

4. Results and discussion

4.1. Analysis of the reference mixture

The yeast strain used, grows in a medium containing one single carbon and nitrogen source, so that we obtained nearly completely ^{13}C- and ^{15}N-labeled metabolite extracts from yeast grown on [U-^{13}C]-D-glucose and [$^{15}N_2$]ammonium sulfate. The remaining $^{12}C^{14}N$ fraction arises from impurities of the tracers (Supplemental Figure S2). To determine compounds in our reference mixture that qualify as internal standards, we performed GC–MS measurements of the $^{13}C^{15}N$-labeled yeast extract to create a reference library as described above. We generated two datasets on two different GC–MS instruments to later be able to assess inter-instrument variation. Below, numbers are presented as "result instrument A" ("result instrument B"). In the reference yeast extract, 179 (163) deconvoluted mass spectra were detected in each of three replicate measurements (Fig. 3A). Of these mass spectra 109 (134) were isotopically enriched and had at least one pair of

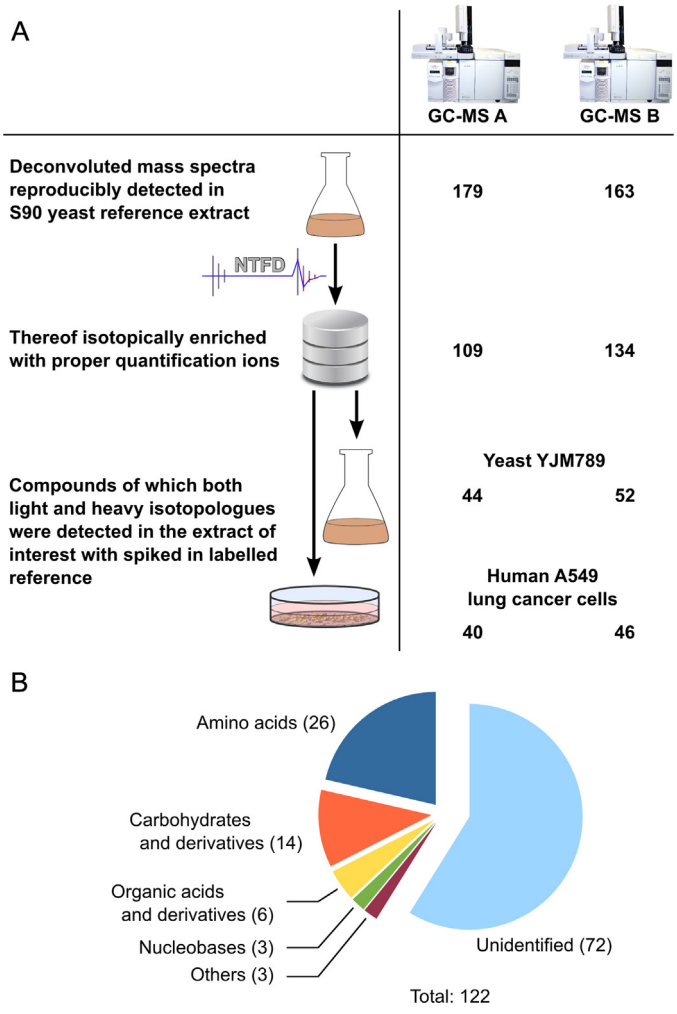

Fig. 3. (A) Number of potential internal standards in reference mixture and number of matching compounds in different sample types. Comparison of measurements from two separate GC–MS instruments. (B) Classes of metabolite derivatives which were detected in the reference mixture and for which suitable quantification ions were found (compound numbers are means of measurements on two GC–MS instruments).

quantification ions meeting the requirements described above. Among the excluded compounds, there were known contaminants and isotopically enriched compounds with overlapping fragments. The generated yeast extract, thus, provides internal standards for more than 100 compounds and its preparation is relatively low priced and easy. About one third of the potential internal standards have been identified using an in-house mass spectra reference library as methyloxime- and trimethylsilyl-derivatives of about 40 different metabolites (Supplemental Table S1). Among the identified compounds, amino acid derivatives were the most prominent class (Fig. 3B).

For long-term use of the same reference mixture, the stability of its constituents needs to be considered. We did not assess long-term stability of the yeast extract, but for other matrices metabolite levels have been shown to be stable for storage periods of at least 6 months and multiple freeze thaw cycles [20]. Therefore, we trust that the majority of compounds is stable also over extended storage periods if aliquots are frozen at −80 °C, and repeated freeze–thaw-cycles and exposure to higher temperatures are avoided.

The complete stable isotope enrichment of a reference organism could even be circumvented by a different labeling strategy: Any unlabeled reference mixture could be derivatized with a stable isotope-labeled silylating reagent suitable for GC–MS sample

preparation before spiking it into the conventionally derivatized sample of interest. Such a differential derivatization and mixing of two samples has been presented before by Huang and Regnier [21]. It has the advantage, that always the sample type of interest can be used as a reference and that no fully isotopically enrichable organism as reference is required. A disadvantage, however, is that also any contamination introduced during the sample workup or exogenous compounds, for example stemming from growth medium, will also be derivatized and cannot be distinguished from endogenous metabolites. Furthermore, currently, there are only deuterated but no ^{13}C-labeled silylating reagents commercially available. The pronounced isotope effect of deuterium on chromatographic retention increases chromatographic sample complexity, which is not the case with ^{13}C- and ^{15}N-labeling (Supplemental Figure S3). Moreover, highly deuterated derivatization agents can lead to very dissimilar mass spectra and render spectrum matching more complex, thereby hampering any automated non-targeted analyses.

4.2. Applicability to different sample types

As a test case, we used a fully labeled yeast extract to quantify compounds within a metabolite extract of a different yeast strain. For 44 (52) compounds out of the 109 (134) compounds in the reference library, we detected both the light and heavy form in the sample of interest (Fig. 3A, Supplemental Table S1). Many of the analytes for which there was a labeled analogue in the reference mixture were not found in the sample of interest. This is mostly due to two facts: Firstly, the sample of interest was only the polar metabolite extract alone without the hydrolysis products added to the reference mixture. Secondly, for the generation of the quantification library, the injected amount of reference mixture was higher than that of the sample of interest. We chose to rather overload the measurements of the reference mixture to obtain a more comprehensive library, because analytes which were not detected at that stage, could not have been included in subsequent analyses. Considering the price and availability of stable isotope labeled analogues which would have to be acquired otherwise, the number of reference compounds is respectable.

As the overlap of the reference mixture with the yeast sample was reasonably good, we were interested in the applicability to mammalian cell extracts. Mammalian cells cannot easily be completely labeled with stable isotopes because of their complex nutrient requirements. To assess the applicability of a yeast extract for the quantification of metabolites in mammalian samples, we spiked the fully labeled reference mixture into a polar metabolite extract derived from human A549 lung cancer cells. This way we were able to normalize 40 (46) compounds present in the A549 extract. Thereof, 32 compounds were identified (Supplemental Table S1). The number and identity of compounds common to reference mixture and sample of interest were similar for yeast and human extracts. Although generally the two organisms differ significantly, the difference in detected analytes was relatively small, because GC–MS mostly covers primary metabolites which are highly conserved across species. This makes our approach applicable for a wide variety of samples.

This normalization approach is restricted to compounds that are present in both sample of interest and reference mixture, for compounds not present in the reference mixture only the raw signal intensities are available. Therefore, an adequate reference metabolome has to be chosen. Ideally, the same type as the sample of interest but isotopically enriched would be used. However, this is impossible for body fluids, tissue samples, or cells with complex substrate requirements. To optimize the reference mixture and to increase the overlap with the sample of interest, different extracts from different organisms or different growth conditions can be combined. For individual missing but important compounds, a stable isotope labeled analogue can be added to the reference mixture. It is preferable to choose a reference as similar as possible to the sample of interest to maximize the overlap of both composition and metabolite concentrations. If metabolite concentrations differ strongly, it is less likely that both of them will lie within the linear range of the detector. We showed, however, that yeast metabolite preparations can also be used, to a certain degree, as internal standard for different samples like mammalian cells. In the analyzed samples, the levels of most analytes and respective internal standards differed by less than factor 5 (Supplemental Figure S1).

4.3. Better intra-instrument reproducibility

Our quantification approach aims at making non-targeted metabolomics analyses more robust and comparable across measurements performed at different times or on different instruments. Thus, we chose intra- and inter-instrument variation as performance measure to compare our IDMS approach to conventional methods.

As reference methods we have chosen normalization to a single internal standard and normalization to summed signal or total ion current (TIC). TIC normalization divides the intensity of every compound in a sample by the summed signal of all compounds in this sample. This normalization approach is commonly used in metabolomics analyses [22]. For single internal standard normalization all signal intensities are divided by that of the isotope labeled internal standard. For the single internal standard we chose one compound from the reference mixture: ([U-^{13}C,U-^{15}N]ornithine) 4TMS for the yeast and ([U-^{13}C]malic acid) 3TMS for the A549 sample. These two were selected because they showed a good peak shape and abundance in the respective samples, and there were no other derivatives of these metabolites detected.

To assay intra-instrument variation, we determined the relative standard deviations (RSDs) across three injections of the same sample (Fig. 4A). For some compounds RSD was rather high, because of their very low abundance. Overall, the IDMS normalization was more robust than the two conventional approaches. With the yeast samples, mean injection-to-injection variability across analytes (RSD) decreased significantly by 8.56 percentage points as compared to TIC normalization (paired t-test, $p = 0.00402$, $n = 52$). Single internal standard normalization performed better than TIC normalization, but RSDs from the IDMS normalization were still significantly lower by 5.55 percentage points ($p = 0.0048$). For human cell samples there was a similar trend. RSDs after IDMS normalization were 8.82 percentage points lower than after TIC normalization ($p = 0.0219$, $n = 40$) and 1.43 percentage points lower than after normalization to single internal standard (not significant). These results validate our automated choice of quantification ions and the subsequent quantification term. The improvement for both, yeast and human samples demonstrates that the yeast extract can also be successfully applied to different sample types.

The IDMS approach would improve reproducibility even more on platforms without automated sample preparation where different derivatization times occur. Ratios of different derivatives of a single analyte change over time and impact quantification results. However, the individual internal standards are subject to the same conditions and correct for such biases.

4.4. Better inter-instrument reproducibility

A strong motivation for this quantification approach was the need for a comparable measure of metabolite levels across different instruments. Measurements on different instruments are subject to differently aged inlets, chromatography columns, ion sources or detectors which can heavily influence the results and

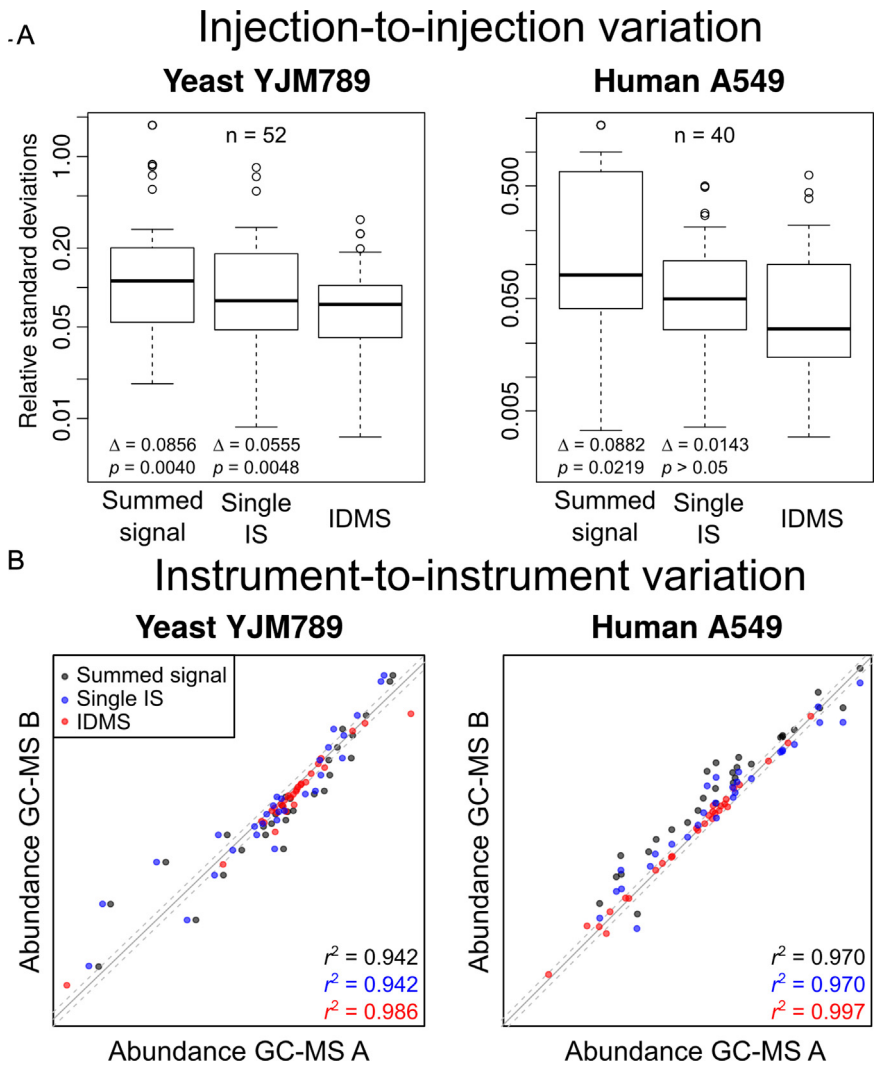

Fig. 4. Comparison of the described IDMS approach with summed sample signal normalization and single internal standard (IS) for two different sample types. (A) Isotopologue ratios with individual internal standards (IDMS) show lower relative standard deviations across analytes than normalization to summed signal or single internal standard. The same sample was injected three times. Axes are logarithmic. Δ: Difference of the means of the respective method to IDMS, *p*: paired *t*-test *p*-value. (B) Metabolite levels determined on instrument A plotted over those determined on instrument B. The diagonal corresponds to identical quantification results. Inter-instrument-variation is indicated by the distance of the points from the diagonal and is mostly lower for the isotopologue ratios. Relative sample signal and isotopologue ratios were divided by their range to match the different scales. Axes are logarithmic. Summed signal normalization can introduce systematic errors, as visible by the global "upward shift" of the points in the A549 plot. Normalizing to a single IS reduces this shift. r^2: Pearson product-moment correlation coefficient of the natural logarithm of the normalized intensities. Dotted lines represent 10% deviation from the diagonal.

hamper comparability with other measurements. To assess inter-instrument-variations we analyzed two measurements of the same sample performed on two different GC–MS instruments (Fig. 4B). The measurements were performed on two GC–MS instruments with the same configuration so that TIC normalization would not be excluded *per se*. Quantification results of the yeast samples using the IDMS-normalization showed lower random differences between the two instruments as compared to TIC- and single internal standard normalization (Fig. 4B). In the analysis of the A549 samples, TIC-normalization seemed to introduce a systematic error: In the measurement on instrument B, most compounds showed higher levels than on instrument A. This was partially corrected for with the single internal standard normalization. Isotopologue ratio- normalized metabolite levels showed the best correlation across the two instruments.

Summed signal normalization is sensitive to detector saturation and background signals from contaminations as for example introduced from derivatization reagents. These factors alter the summed signal intensity in a sample-independent manner and may introduce an error in the quantification results. Single internal standard addition is less sensitive to overall background signal and to individual high abundant compounds leading to detector saturation. However, a single internal standard cannot sufficiently represent all analyte-specific analytical discrimination. In contrast, the isotopologue ratios are very robust, because the analyte of interest and the internal standard are subject to the very same analytical conditions and the ratios are not influenced by discrimination of certain compound classes due to for example deteriorating inlet or column performance.

5. Conclusion

The method presented here allows for the robust quantification of both identified as well as unidentified metabolites relative to a spiked-in complex mixture of stable isotope-labeled compounds. Such a mixture can easily be produced in-house from any prototrophic organism [9,8] or commercial ^{13}C-enriched extracts can be used [10]. All potential internal standards are detected

using the NTFD algorithm [15] and suitable quantification ions are assigned automatically. Hence, it is easy to use, also in cases of very complex reference mixtures. Later on, absolute quantification of known compounds can be achieved by including compound mixtures of known concentration in the measurements as shown in previous studies [9].

To our knowledge, the approach presented in this manuscript is the first description of an automatable workflow for non-targeted IDMS for GC–EI-MS analysis or low resolution mass spectrometry in general. Previous approaches employing labeled metabolite extracts were always targeted and did not consider unidentified compounds. Operating in a non-targeted manner would otherwise require significant user effort to manually examine all deconvoluted mass spectra for isotopic enrichment and to compare labeled and unlabeled mass spectra to pick proper quantification ions.

Often, initially unidentified compounds turn out to be of interest in subsequent experiments. With the technical advancement, growing compound libraries, and better matching algorithms [23] they can be identified later on. When such compounds happen to be identified at a later date, also their absolute quantification is possible retroactively with our method as all measurements already contain the respective internal standard.

The isotopic peak ratios are not only more robust than conventional normalization methods in a run-to-run comparison, but also allow for comparison of analyte levels across measurements performed at different times and on different instruments or even laboratories. Low analytical variance is crucial for meaningful non-targeted metabolomics and with better inter-instrument comparability larger metabolomics studies can be realized better.

We implemented the described procedure for the automated generation of the reference library with appropriate quantification ions as a new feature in the NTFD application [8] which is freely available for download from http://ntfd.mit.edu/. The generated compound library can be used with the freely available MetaboliteDetector software [17] to integrate the signal intensities of these quantification ions in the samples of interest. Automated calculation of the isotopologue ratios will be included in subsequent releases of MetaboliteDetector.

Supplemental information

Supplemental Figures S1–S3 and a list of detected compounds and quantification ions (Supplemental Table S1) are available online.

Acknowledgements

The authors thank Lars Steinmetz for providing the yeast strains used in this project. This project was supported by the Fonds National de la Recherche (FNR) Luxembourg (ATTRACT A10/03).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.chroma.2015.02.025.

References

[1] R. Dallmann, A.U. Viola, L. Tarokh, C. Cajochen, S.A. Brown, The human circadian metabolome, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 2625–2629, http://dx.doi.org/10.1073/pnas.1114410109.
[2] R. Wei, Metabolomics and its practical value in pharmaceutical industry, Curr. Drug Metab. 12 (2011) 345–358.
[3] J.R. Mayers, et al., Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development, Nat. Med. 20 (2014) 1193–1198, http://dx.doi.org/10.1038/nm.3686.
[4] A. Michelucci, T. Cordes, J. Ghelfi, A. Pailot, N. Reiling, O. Goldmann, T. Binz, A. Wegner, A. Tallam, A. Rausell, M. Buttini, C.L. Linster, E. Medina, R. Balling, K. Hiller, Immune-responsive gene 1 protein links metabolism to immunity by catalyzing itaconic acid production, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 7820–7825, http://dx.doi.org/10.1073/pnas.1218599110.
[5] A.K. Krug, et al., Transcriptional and metabolic adaptation of human neurons to the mitochondrial toxicant MPP+, Cell Death Dis. 5 (2014) e1222, http://dx.doi.org/10.1038/cddis.2014.166.
[6] M.L. Reaves, J.D. Rabinowitz, Metabolomics in systems microbiology, Curr. Opin. Biotechnol. 22 (2011) 17–25, http://dx.doi.org/10.1016/j.copbio.2010.10.001.
[7] T.M. Wasylenko, G. Stephanopoulos, Metabolomic and ^{13}C-metabolic flux analysis of a xylose-consuming *Saccharomyces cerevisiae* strain expressing xylose isomerase, Biotechnol. Bioeng. (2014), http://dx.doi.org/10.1002/bit.25447.
[8] M.R. Mashego, L. Wu, J.C. Van Dam, C. Ras, J.L. Vinke, W.A. Van Winden, W.M. Van Gulik, J.J. Heijnen, MIRACLE: mass isotopomer ratio analysis of U-^{13}C-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites, Biotechnol. Bioeng. 85 (2004) 620–628, http://dx.doi.org/10.1002/bit.10907.
[9] L. Wu, M.R. Mashego, J.C. van Dam, A.M. Proell, J.L. Vinke, C. Ras, W.A. van Winden, W.M. van Gulik, J.J. Heijnen, Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly ^{13}C-labeled cell extracts as internal standards, Anal. Biochem. 336 (2005) 164–171, http://dx.doi.org/10.1016/j.ab.2004.09.001.
[10] O. Vielhauer, M. Zakhartsev, T. Horn, R. Takors, M. Reuss, Simplified absolute metabolite quantification by gas chromatography-isotope dilution mass spectrometry on the basis of commercially available source material, J. Chromatogr. B 879 (2011) 3859–3870, http://dx.doi.org/10.1016/j.jchromb.2011.10.036.
[11] P. Kiefer, J.-C. Portais, J.A. Vorholt, Quantitative metabolome analysis using liquid chromatography-high-resolution mass spectrometry, Anal. Biochem. 382 (2008) 94–100, http://dx.doi.org/10.1016/j.ab.2008.07.010.
[12] P. Giavalisco, K. Köhl, J. Hummel, B. Seiwert, L. Willmitzer, ^{13}C isotope-labeled metabolomes allowing for improved compound annotation and relative quantification in liquid chromatography-mass spectrometry-based metabolomic research, Anal. Chem. 81 (2009) 6546–6551, http://dx.doi.org/10.1021/ac900979e.
[13] C. Bueschl, B. Kluger, M. Lemmens, G. Adam, G. Wiesenberger, V. Maschietto, A. Marocco, J. Strauss, S. Bödi, G.G. Thallinger, R. Krska, R. Schuhmacher, A novel stable isotope labelling assisted workflow for improved untargeted LC-HRMS based metabolomics research, Metabolomics 10 (2014) 754–769, http://dx.doi.org/10.1007/s11306-013-0611-0.
[14] F.A. de Jong, C. Beecher, Addressing the current bottlenecks of metabolomics: Isotopic Ratio Outlier Analysis, an isotopic-labeling technique for accurate biochemical profiling, Bioanalysis 4 (2012) 2303–2314, http://dx.doi.org/10.4155/bio.12.202.
[15] K. Hiller, C.M. Metallo, J.K. Kelleher, G. Stephanopoulos, Nontargeted elucidation of metabolic pathways using stable-isotope tracers and mass spectrometry, Anal. Chem. 82 (2010) 6621–6628, http://dx.doi.org/10.1021/ac1011574.
[16] S.C. Sapcariu, T. Kanashova, D. Weindl, J. Ghelfi, G. Dittmar, K. Hiller, Simultaneous extraction of proteins and metabolites from cells in culture, MethodsX 1 (2014) 74–80, http://dx.doi.org/10.1016/j.mex.2014.07.002.
[17] K. Hiller, J. Hangebrauk, C. Jäger, J. Spura, K. Schreiber, D. Schomburg, MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis, Anal. Chem. 81 (2009) 3429–3439, http://dx.doi.org/10.1021/ac802689c.
[8] K. Hiller, A. Wegner, D. Weindl, T. Cordes, C.M. Metallo, J.K. Kelleher, G. Stephanopoulos, NTFD–a stand-alone application for the non-targeted detection of stable isotope-labeled compounds in GC/MS data, Bioinformatics (2013), http://dx.doi.org/10.1093/bioinformatics/btt119.
[19] A. Wegner, D. Weindl, C. Jäger, S.C. Sapcariu, X. Dong, G. Stephanopoulos, K. Hiller, Fragment formula calculator (FFC): determination of chemical formulas for fragment ions in mass spectrometric data, Anal. Chem. 86 (2014) 2221–2228, http://dx.doi.org/10.1021/ac403879d.
[20] H.G. Gika, G.A. Theodoridis, I.D. Wilson, Liquid chromatography and ultra-performance liquid chromatography-mass spectrometry fingerprinting of human urine: sample stability under different handling and storage conditions for metabonomics studies, J. Chromatogr. A 1189 (2008) 314–322, http://dx.doi.org/10.1016/j.chroma.2007.10.066.
[21] X. Huang, F.E. Regnier, Differential metabolomics using stable isotope labeling and two-dimensional gas chromatography with time-of-flight mass spectrometry, Anal. Chem. 80 (730) (2008) 107–114, http://dx.doi.org/10.1021/ac071263f.
[22] A. Hutschenreuther, A. Kiontke, G. Birkenmeier, C. Birkemeyer, Comparison of extraction conditions and normalization approaches for cellular metabolomics of adherent growing cells with GC-MS, Anal. Methods 4 (2012) 1953–1963, http://dx.doi.org/10.1039/C2AY25046B.
[23] A. Wegner, S.C. Sapcariu, D. Weindl, K. Hiller, Isotope cluster based compound matching in Gas Chromatography/Mass Spectrometry for non-targeted metabolomics, Anal. Chem. (2013), http://dx.doi.org/10.1021/ac303774z.