Best-Matched Internal Standard Normalization in Liquid Chromatography-Mass Spectrometry Metabolomics Applied to....

Article *in* Analytical Chemistry · December 2017

DOI: 10.1021/acs.analchem.7b04400

CITATION
1

READS
47

4 authors, including:

Katherine R. Heal
University of Washington Seattle
60 PUBLICATIONS **165** CITATIONS
SEE PROFILE

Anitra E Ingalls
University of Washington Seattle
118 PUBLICATIONS **2,435** CITATIONS
SEE PROFILE

analytical chemistry

Article

Cite This: Anal. Chem. XXXX, XXX, XXX−XXX

pubs.acs.org/ac

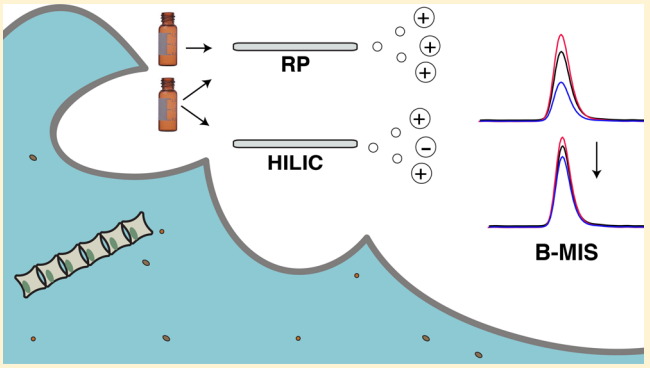
Best-Matched Internal Standard Normalization in Liquid Chromatography−Mass Spectrometry Metabolomics Applied to Environmental Samples

Angela K. Boysen,[†,‡] Katherine R. Heal,[†,‡] Laura T. Carlson,[†] and Anitra E. Ingalls*[,†]

[†]School of Oceanography, University of Washington, Seattle, Washington, United States

Supporting Information

ABSTRACT: The goal of metabolomics is to measure the entire range of small organic molecules in biological samples. In liquid chromatography−mass spectrometry-based metabolomics, formidable analytical challenges remain in removing the nonbiological factors that affect chromatographic peak areas. These factors include sample matrix-induced ion suppression, chromatographic quality, and analytical drift. The combination of these factors is referred to as obscuring variation. Some metabolomics samples can exhibit intense obscuring variation due to matrix-induced ion suppression, rendering large amounts of data unreliable and difficult to interpret. Existing normalization techniques have limited applicability to these sample types. Here we present a data normalization method to minimize the effects of obscuring variation. We normalize peak areas using a batch-specific normalization process, which matches measured metabolites with isotope-labeled internal standards that behave similarly during the analysis. This method, called best-matched internal standard (B-MIS) normalization, can be applied to targeted or untargeted metabolomics data sets and yields relative concentrations. We evaluate and demonstrate the utility of B-MIS normalization using marine environmental samples and laboratory grown cultures of phytoplankton. In untargeted analyses, B-MIS normalization allowed for inclusion of mass features in downstream analyses that would have been considered unreliable without normalization due to obscuring variation. B-MIS normalization for targeted or untargeted metabolomics is freely available at https://github.com/IngallsLabUW/B-MIS-normalization.



Liquid chromatography−mass spectrometry (LC-MS)-based metabolomics has emerged as an important scientific tool over the past decade, with the potential to identify and quantify thousands of compounds resulting from cellular activity.[1−3] Quantification relies on an analyte's concentration being proportional to its peak area. However, independent of analyte concentration, peak areas can be affected by sample-matrix-induced ion suppression, injection volume, chromatographic quality, or analytical drift over time; the overall effects of these nonbiological factors are referred to here as obscuring variation since they may mask biological differences of interest.[4] Performing metabolomics analyses on environmental samples is especially challenging because complex matrices such as natural seawater or culture media introduce variability in ionization efficiency, but there are no universally accepted normalization techniques for minimizing obscuring variation in environmental samples. Here we present a normalization method that addresses this outstanding challenge in LC-MS-based metabolomics.

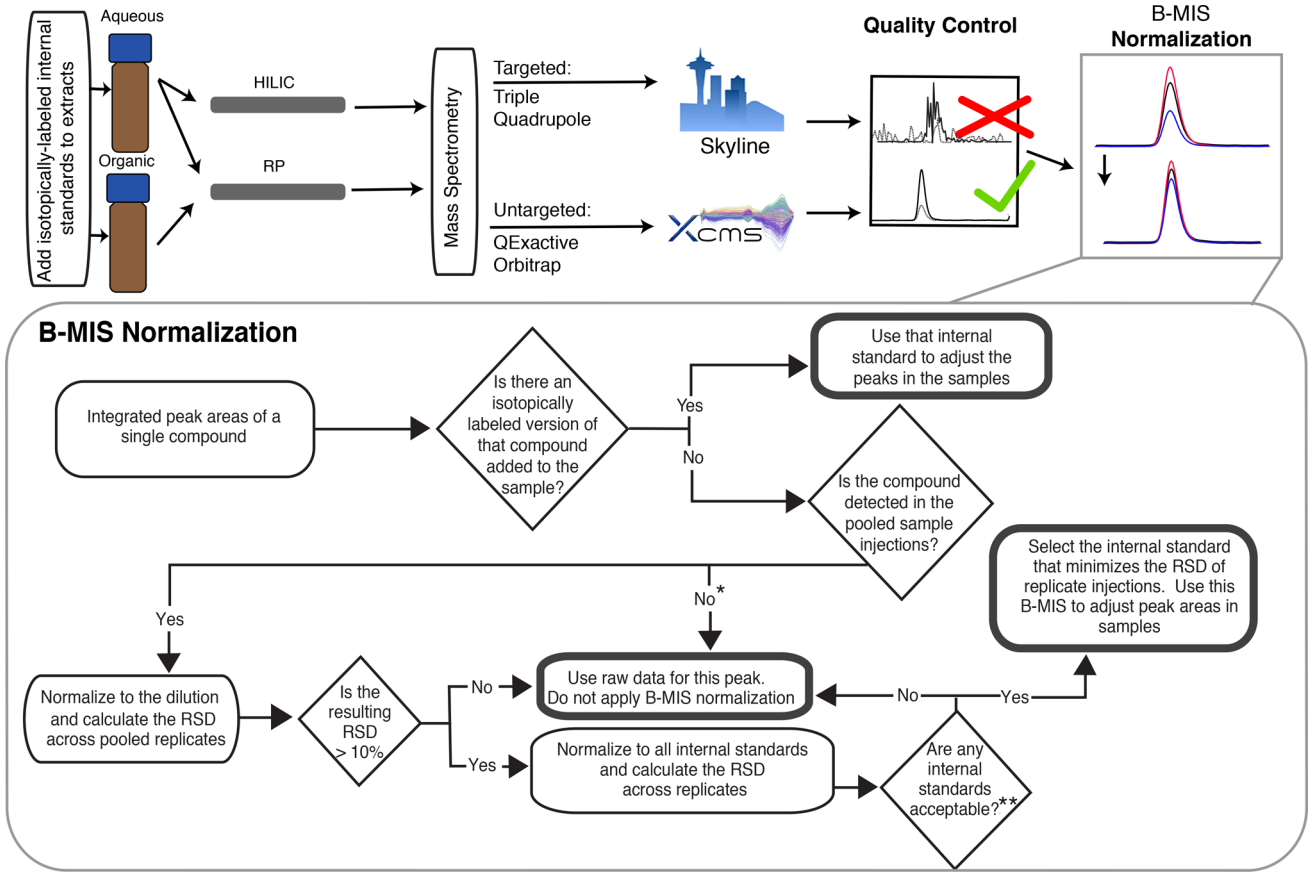
In biomedical metabolomics, there are several approaches to minimize obscuring variation that rely on pre- or post-acquisition normalization. However, application of these approaches to environmental metabolomics is limited, and little attention has been given to removing obscuring variation in environmental samples.[3,5−12] Preacquisition normalization approaches ensure that samples are at similar concentrations before analysis to minimize the effect of matrix-induced ion suppression. Postacquisition normalization techniques fall into two general categories, those that use a quality control sample injected several times over the course of a run to correct for analytical drift, and those that employ isotope-labeled internal standards.

Analytical drift over time can be normalized postacquisition by applying a quality control-based locally weighted scatterplot smoothing (LOESS)[13,14] or vector regression;[15] these approaches attempt to minimize one aspect of obscuring variation (analytical drift) and assume that the sample-matrix induced ion suppression is consistent between samples. This approach may be appropriate in biomedical studies when prenormalization can be carefully applied (for instance, by adjusting sample concentration or injection volume so that there is a consistent amount of creatinine in each sample, as

Received: October 25, 2017
Accepted: December 14, 2017
Published: December 14, 2017

DOI: 10.1021/acs.analchem.7b04400
Anal. Chem. XXXX, XXX, XXX−XXX

Scheme 1. Overview of Data Acquisition and Processing[a]



[a]After extraction, aqueous fractions were run on both RP and HILIC, while organic fractions were run only on RP. These fractions were run on both the TQS and QE. Skyline[27] was used for integrations for targeted analyses. XCMS[29−31] was used for peak-picking and integrations for untargeted analyses. Data were subjected to quality control filtering as described in the text. For each compound detected, the B-MIS was chosen via the steps shown. RSD = relative standard deviation. *Only applicable in targeted analyses. **This cut-off is user defined; in our analyses, an internal standard was considered acceptable if it decreased the RSD of the pooled injections by 40%.

done routinely in urine metabolomics[16,17]). For complex matrices, prenormalization is often impossible due to the variability in matrix among samples or lack of information about sample size. For instance, if cell density is not known until after sample collection, extracting a consistent amount of biomass is not possible. Furthermore, these approaches require large sample sets and tens of injections of a pooled sample as a quality control, which is not always possible in metabolomics studies with limited sample size and few samples to compare.

Other postacquisition strategies use isotope-labeled internal standards to minimize obscuring variation. Some studies normalize each metabolite peak area to a single isotope-labeled internal standard.[18] This approach is very sensitive to the internal standard used for normalization and assumes that all compounds experience the same obscuring variation as the internal standard; these caveats make this approach limited in utility. Other studies use multiple internal standards, applying each to analytes that elute within a prescribed retention time window.[19,20] Although this may be an improvement over a single internal standard, these approaches do not consider the fact that obscuring variation is not fully described by retention time. Our group has previously used a hybrid approach that matches isotope-labeled internal standards to analytes for which commercially available isotope-labeled internal standards are unavailable,[21,22] but this has not been applied to full-scale metabolomics analyses.

In laboratory studies, growing an organism on a ^{13}C-labeled substrate can yield a valuable ^{13}C-labeled stock used to calculate accurate concentrations of both targeted and untargeted metabolites and aid in metabolite identification,[9,23] but this approach cannot be applied to studies of mixed communities in the environment. More sophisticated approaches attempt to model the variability of each peak by a combination of internal standards.[4,24] These methods require tens of injections of a pooled sample to obtain sufficient data for model training and make use of logarithmic transformations, which may limit quantitative comparison and hinder intuitive interpretation. To our knowledge, these methods have not been applied to targeted metabolomic analyses.

Given the paucity of normalization approaches that are appropriate for minimizing obscuring variation in samples with a high degree of complexity in both analytes and matrices, we present a method called best-matched internal standard (B-MIS) normalization. This method has the following advantages over existing techniques: (1) B-MIS normalization can be applied to targeted and untargeted metabolomics studies with sample sets of any size or nature. (2) B-MIS normalization can be customized to accommodate any number and variety of

DOI: 10.1021/acs.analchem.7b04400
Anal. Chem. XXXX, XXX, XXX−XXX

internal standards of the user's choice. (3) B-MIS normalization does not assume a uniform matrix effect among samples and minimizes obscuring variation induced by sample-specific matrix-induced ionization. (4) B-MIS normalization is independent of and can be paired with other sample normalization pre- or postacquisition. (5) B-MIS normalization is simple to apply, intuitive to interpret, and open source.

To demonstrate both the need for this method and its functionality, we used two sample sets: marine environmental samples and laboratory cultures of marine phytoplankton. We analyzed these samples on two mass spectrometers: a Waters Xevo TQ-S triple quadrupole, for targeted analyses, and a Thermo QExactive HF, for untargeted analyses, using both reversed phase (RP) and hydrophilic interaction liquid chromatography (HILIC). A comparative analysis of these platforms revealed the strengths of each and highlighted the need for careful normalization and application of B-MIS normalization regardless of platform.

EXPERIMENTAL SECTION

The workflow for sample analysis and data processing is shown in Scheme 1. Additional details on sample preparation, chromatography, mass spectrometry, and quality control parameters are described in the Supporting Information.

Sample Preparation. We evaluated two sets of complex samples in triplicate: cultures of one species of diatom grown under four experimental conditions (salinity of 32 g/kg at both −1 and 4 °C, salinity of 41 g/kg at both −1 and 4 °C), and a natural marine microbial community at four depths in the North Pacific subtropical gyre (24° 33.284′ N, 156° 19.790′ W; samples collected at 15, 45, 75, and 125 m, all with salinities of 35 g/kg). Polar and nonpolar metabolites were extracted using a modified Bligh−Dyer extraction[25,26] using 1:1 methanol/water (aqueous phase) and dichloromethane (organic phase). Methodological blanks were extracted and analyzed along with each sample set. Isotope-labeled internal standards were added either before or after the extraction to all samples, blanks, and pooled samples (Table S1). To evaluate the effect of obscuring variation due to different matrix strengths and analytical drift, pooled samples were run at both full and half concentration (diluted with water) at least three times throughout a sample set.

Liquid Chromatography. As in previous metabolomic studies,[2] we paired RP with HILIC (with separate injections) to maximize the number of compounds we were able to detect since many compounds that eluted early on the RP column were better retained and provided higher quality peaks on the HILIC column (Figures S1 and S2). All chromatographic separations were carried out on a Waters Acquity I-Class UPLC (Waters Corporation, Milford, MA). For targeted analysis, we monitored 105 analytes with RP and 126 analytes with HILIC, according to which column demonstrated better retention; 11 analytes were monitored on both columns to compare the performance of RP and HILIC chromatography in complex matrices (Table S2).

Mass Spectrometry. For targeted metabolomics, we used a Waters Xevo TQ-S triple quadrupole (TQS) with electrospray ionization (ESI) in selected reaction monitoring mode (SRM) with polarity switching. SRM conditions for each compound (collision energy, cone voltage, precursor, and product ions, Table S2) were optimized by infusion of each metabolite standard. For most metabolites, two SRM transitions were selected based on maximum peak areas.

We used a Thermo QExactive HF (QE) with ESI for untargeted analyses. For HILIC, a full scan method with polarity switching was used. For RP, positive ionization mode was chosen over positive/negative switching because few of the compounds in our targeted method ionized in the negative mode and using a single mode enabled higher resolution.

Data Processing. For targeted analysis, we integrated peaks using Skyline for small molecules.[27] After integration, we passed the data through an in-house quality control (QC) filter to ensure proper metabolite identification as described in the Supporting Information.

For untargeted metabolomics data, we converted Thermo.RAW files to .mzxml using MSConvert[28] and processed each data set using XCMS[29−31] (using parameters obtained via an Isotopologue Parameter Optimization[32]), yielding retention-time corrected mass features. We processed both data sets in separate XCMS runs according to injection and polarity (RP-aqueous, RP-organic, HILIC-aqueous (positive polarity), and HILIC-aqueous (negative polarity)). Data were quality controlled, as described in the Supporting Information, prior to normalization.

B-MIS Normalization. The best normalization for each analyte was determined using repeat injections of a quality control sample to search for an isotope-labeled internal standard whose obscuring variation matched the observed obscuring variation of the analyte, as shown in Scheme 1. For our sample sets, multiple injections of the pooled sample at full and half strength were used in order to capture the obscuring variation due to variable matrix strength. After correcting for dilution, if the relative standard deviation (RSD) of an analyte's peak area in these repeat injections was less than 10%, the raw data were used. Otherwise, each analyte-internal standard pair was evaluated for normalization. An internal standard was deemed an acceptable match for an analyte if the analyte's RSD improved by 40% over the raw RSD, as discussed in the following section. If multiple internal standards were acceptable for normalization of a single analyte, the one that minimized the RSD was selected. When a B-MIS was chosen, the peak area was divided by the ratio of the peak area of the B-MIS to the average peak area of the B-MIS across the whole sample set, resulting in an adjusted peak area for the analyte. This adjusted peak area served as a relative quantification across samples in the sample set. If the goal of an analysis is absolute quantification and standards are available, analytes can be quantified via standard addition in a subset of samples and then applied across the sample set using the relative quantification of the B-MIS adjusted peaks.

For internal standards with only one or two labeled atoms in the compound, we examined the methodological blank to ensure the signal was comparable to that in the samples, indicating that there was not significant contribution from naturally occurring isotopologues. For future users, we suggest choosing isotope-labeled internal standards with at least a +2 amu label to avoid any naturally occurring isotopologue contamination. In the sample sets analyzed here, the isotope-labeled internal standards were significantly more abundant that naturally occurring isotopologues, so contamination was negligible.

DOI: 10.1021/acs.analchem.7b04400
Anal. Chem. XXXX, XXX, XXX−XXX

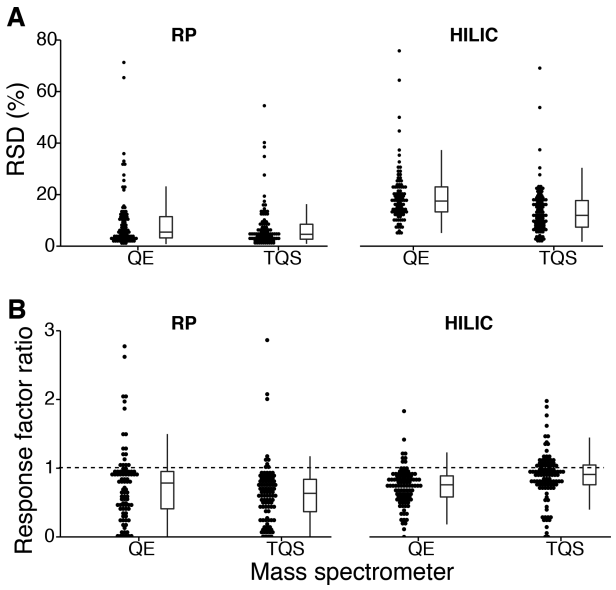


Figure 1. Box-and-whisker plots of (A) relative standard deviation (RSD) of peak areas determined by repeated injections of standards spiked into our environmental sample matrix and (B) response factor ratio on four LC−MS configurations, with individual dots representing single compounds (expanded horizontally for visualization of data density, n = 101 analytes in RP, 110 in HILIC). Response factor ratio is the ratio of the intensity of a standard injected in environmental matrix (less the ambient matrix signal, if applicable) to the intensity of the standard injected in water. Response factor ratio < 1 indicates ion suppression, response factor ratio > 1 indicates ion enhancement.

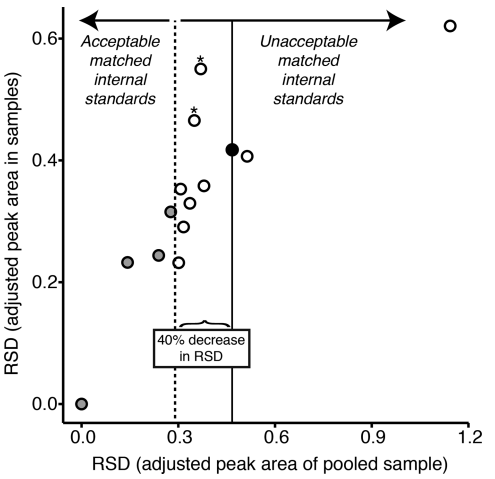


Figure 2. RSD of D_3-alanine in replicate injections of the pooled sample (x-axis) and across all samples in the environmental sample set (y-axis) after normalization to all possible internal standards. The black point represents the RSD of raw peak areas. The point at the origin is D_3-alanine normalized to itself. Filled points show the normalization results with an acceptable matched internal standard according to our cutoff of 40% RSD improvement (left of dotted line); open points show normalization results using unacceptable internal standards (right of dotted line, Scheme 1). Results from other internal standards are in Figure S3.

RESULTS AND DISCUSSION

Evaluation of Data Quality and Instrument Performance. As in other studies,[33−35] we found the HILIC column produced more variable chromatography than the RP column (higher median RSD of peak area, Figure 1A, see the example of choline in Figure S1, full results in Table S3). Despite lower chromatographic reproducibility, compounds analyzed on the HILIC column generally exhibited less ion suppression than those analyzed on the RP column. RP chromatography displayed a bimodal distribution of response factor ratios (Figure 1B), demonstrating significant ion suppression for many analytes (response factor ratio < 1). Ion suppression was especially dramatic for polar compounds that elute early in RP but were retained on the HILIC column (Figures S1 and S2). The range of ion suppression and RSD for a given LC-MS configuration indicated that careful selection of a normalization technique was necessary to account for obscuring variation. This is particularly important when comparing samples in complex matrices such as the environmental matrix highlighted in Figure 1. Furthermore, the wide range in response factor ratios and RSD of peak areas for different analytes in the same matrix demonstrated that normalizing all compounds to a single internal standard[18,20] would likely introduce more obscuring variation than it removed.

Evaluating B-MIS Normalization. Most environmental metabolomics analyses do not attempt to remove variability introduced during analysis.[5,6,8,36] B-MIS normalization was designed to minimize obscuring variation in order to allow robust comparisons within and between sample sets. We evaluate B-MIS normalization by (a) demonstrating that lowering the RSD of multiple injections of a pooled sample improves data precision and arriving at a cutoff value at which to apply normalization, (b) comparing the results from B-MIS normalization with the "gold standard" of correcting with an isotopologue for each analyte of interest, and (c) assessing the selection of internal standards used for normalization.

To demonstrate the fundamental principle and evaluate the effectiveness of B-MIS normalization, we applied the technique to the internal standards themselves. In other words, we evaluated how well the internal standards' variabilities were matched by each of the other internal standards. Figure 2 is an example of this application, where each point is the resulting RSD of D_3-alanine in replicate injections of the pooled sample (x-axis) and across all samples in the environmental sample set (y-axis) after normalization to each possible internal standard; since we add the same amount of the internal standard to each sample, both of these values should be near zero. Figure 2 highlights that internal standards that reduced RSD of the pooled injections generally resulted in a reduced RSD across samples (toward zero in both the x- and y-axes). However, in this example case, there were two instances where a reduction in RSD of pooled sample injections increased the RSD of the samples (marked on Figure 2 with an asterisk); this introduction of variability can be avoided by applying normalization only if the pooled RSD was improved beyond a cutoff value.

Our goal was to maximize the number of compounds that can be improved by B-MIS normalization while minimizing the likelihood of introducing error. After assessing all internal standards, we determined that requiring a 40% improvement of pooled RSD (that is, $(RSD_{final} - RSD_{raw})/RSD_{raw}$) avoided the majority of situations where using an internal standard would result in an increase of sample variability (Figures 2 and S3 and Table S6). Therefore, we chose this cutoff as the criteria to apply B-MIS normalization (see Scheme 1). Additionally, we did not normalize any compounds or mass features that had a raw pooled RSD of below 10% because the chances of

DOI: 10.1021/acs.analchem.7b04400
Anal. Chem. XXXX, XXX, XXX−XXX

significant improvement are low. Before applying B-MIS normalization, we suggest users perform a similar sensitivity analysis to determine an appropriate cutoff using a representative sample set and internal standard suite. Tools for this analysis are available with B-MIS.

The gold standard for absolute quantitation is an isotopologue of each analyte, but isotope-labeled internal standards are costly, consume duty time on the MS (in targeted analyses) and are not commercially available for many compounds. In our targeted analyses, analytes with corresponding isotope-labeled internal standards were normalized to those internal standards (see Scheme 1). If this was not prescribed, the algorithm occasionally selected a different internal standard, though the lowered RSD was often similar to that of the prescribed B-MIS. For example, in the environmental data set, the amino acid valine could have selected the D_7-proline internal standard based on the RSD reduction (raw peak area RSD was 30.8%, which was minimized to 3.5% with normalization to D_7-proline). However, the "gold standard" internal standard, D_8-valine, reduced the RSD to a similar 4.8%, indicating proline and valine experienced nearly identical obscuring variation in this matrix. For the other analytes that have an isotope-labeled internal standard, the majority included their corresponding internal standard as an acceptable B-MIS (>60% in both data sets, Table S5). Most cases where the corresponding internal standard was not considered acceptable were due to the stringency of the 40% improvement cutoff. Nearly all of the analytes with an isotope-labeled internal standard were less variable after normalization to their corresponding internal standard, with one exception in each data set (pyridoxal in the environmental data set, cysteic acid in the culture data set; these exceptions were both improved by normalization in the other data sets).

The success of B-MIS normalization relies heavily on having an appropriate suite of internal standards. However, if a mass feature does not match well with any internal standard, B-MIS normalization defaults to using the raw data, which a user can choose to include in subsequent analysis. Increasing the number and chemical variety of internal standards results in fewer features whose obscuring variation are not improved by B-MIS normalization (Figure S4). In our two data sets, the B-MIS for the targeted analytes were typically close in retention time (in RP, Figure S5) or similar in chemistry (in HILIC). For instance, amino acids generally matched with an isotope-labeled amino acid B-MIS while sulfonates generally matched with a sulfonate B-MIS (Table S5). Figure S4 shows that the internal standards used in HILIC chromatography covered a wide enough range of chemical diversity to effectively reduce obscuring variation in the majority of analytes detected under these conditions.

The specific choice of internal standards used in B-MIS normalization is not prescribed and can vary from user to user; we suggest a selection of internal standards that spans retention time, m/z, and functional groups of interest. Before analyzing a full sample set, users should determine the proportion of mass features which select each internal standard as "acceptable" and "best" matches by running a pooled or representative sample at different dilutions. This will reveal which, if any, internal standards are matching with a large number of mass features and may indicate that more internal standards with similar chemical diversity could be beneficial (see Table S4). For example, in our untargeted diatom culture data $^{13}C_3$–vitamin B_1 (thiamine) was selected as acceptable for nearly 10% of mass features detected in the aqueous fraction analyzed with RP chromatography and it was selected as the B-MIS for all of these features. It is likely that normalization of these compounds could improve if we included more internal standards similar to $^{13}C_3$-vitamin B_1. Some internal standards were infrequently selected as appropriate (such as ^{15}N-isoleucine), while others were appropriate for many but only the best for a few (such as D_7-proline), indicating sufficient coverage to effectively normalize compounds that experienced similar obscuring variation.

Application of B-MIS Normalization. B-MIS normalization only adjusts areas of metabolite peaks with high obscuring variation if normalization improves the technical reproducibility above a specified cutoff (in our case, a 40% improvement over the raw technical reproducibility); therefore, a subset of analytes are not normalized. Figure 3 shows an example of two analytes (taurine and tyrosine) detected in the pooled diatom-culture sample whose RSD were improved by B-MIS normalization. For the sample sets analyzed here, 54% (diatom samples) and 79% (environmental samples) of analytes in the targeted data selected an internal standard for normalization (Table S5 and Figure S5).

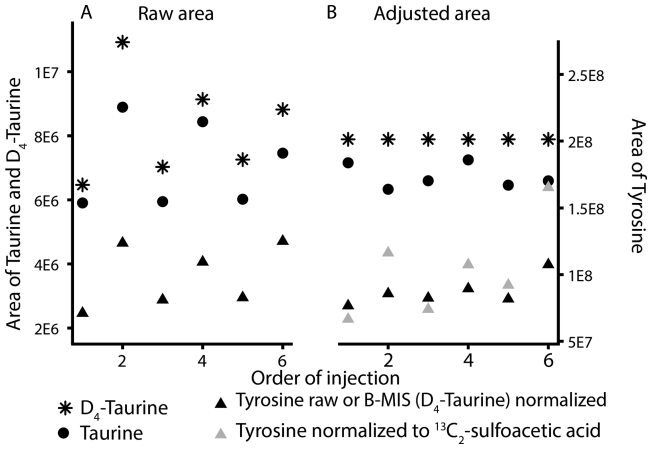


Figure 3. Example of B-MIS normalization for taurine and tyrosine in the diatom sample set. (A) Areas of D_4-taurine, taurine, and tyrosine in pooled samples after adjusting for dilution, in order of injection. (B) After normalization, areas are adjusted, with the B-MIS normalization reducing the RSD of taurine from 19% to 5%, while the RSD tyrosine decreases from 27% to 12% (black triangles). Normalizing to an unacceptable internal standard ($^{13}C_2$-sulfoacetic acid, gray triangles) did not reduce the RSD of tyrosine as effectively.

For each compound, the RSD of the pooled injections depends on the sample matrix, though the sample sets showed a similar range of RSD after B-MIS normalization (Figure S6). Of the 25 targeted analytes that were measured in both sample sets and found acceptable internal standards for normalization, 20 shared acceptable internal standards, while five analytes did not (Table S5 and Figure S6). The compounds which did not improve with B-MIS normalization were not always consistent from environmental to culture sample sets. For example, there were 18 compounds that did not pick a B-MIS in the targeted environmental sample set. In the culture sample set, six of these compounds selected a B-MIS, four did not select a B-MIS, and the remaining eight were not detected (Table S5). These differences between sample sets demonstrated that, while fundamental chemistry has some predictive capacity to

DOI: 10.1021/acs.analchem.7b04400
Anal. Chem. XXXX, XXX, XXX−XXX

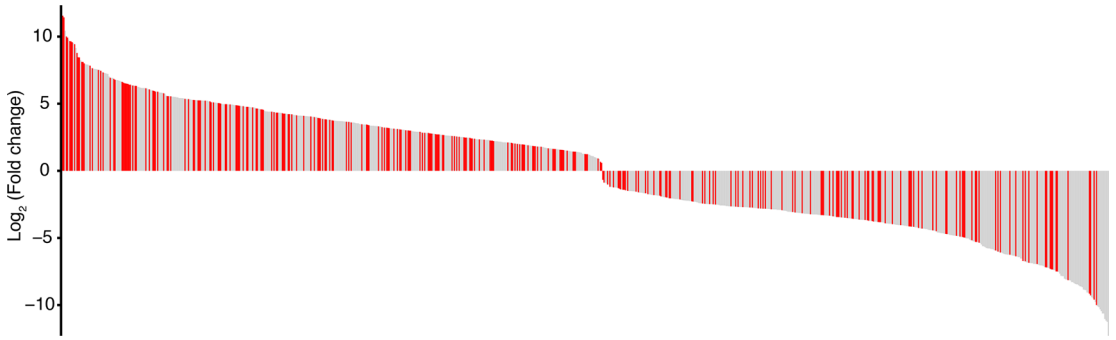


Figure 4. Log_2-fold change between surface (15 m) and deep (125 m) environmental samples for the 500 most abundant mass features in each sample type that were significantly different between depths (false discovery rate[38] adjusted *p*-values < 0.05, Student's *t*-test). A positive fold change indicates enrichment in the 125 m samples. Many mass features (292 of the 883 shown here) may not have been considered reliable peaks (RSD > 20%) due to high obscuring variation which was minimized through B-MIS normalization; these mass features are highlighted in red.

describe obscuring variation, the complexity of the sample matrix makes definitive prediction unrealistic.

In untargeted metabolomics, it is common to remove any mass features that do not exhibit reliable data.[37] One way this is achieved is by calculating the RSD of each mass feature in a multiply injected pooled sample and removing features that have a RSD greater than 20%.[13,37] In the case of our untargeted data, B-MIS normalization increased the number of mass features that passed this RSD filter from 3836 to 5028 (in the diatom samples) and from 5568 to 9372 (in the environmental samples), preserving many high-quality peaks that would otherwise have been discarded (Figure S7). In our environmental untargeted analysis, 54% and 53% of the mass features detected on the HILIC and RP columns, respectively, matched with an acceptable internal standard and showed decreased variability with a B-MIS (Figure S4). In this data set, a third of the most abundant and significantly different mass features would have been excluded if this RSD filtering had been performed prior to B-MIS normalization (Figure 4). A total of 7 of the 10 mass features most enriched in the sample collected at 125 m compared to the sample collected at 15 m would have been excluded without B-MIS normalization before filtering (Figure 4). This comparison highlights that meaningful biological interpretation of the differences between samples hinges on proper normalization. We therefore recommend applying B-MIS normalization before filtering out mass features based on RSD.

CONCLUSIONS

Published environmental metabolomics analyses have not attempted to minimize the obscuring variation inherent to LC-MS-based data acquisition. In both targeted and untargeted environmental samples, we demonstrate the need for careful normalization due to matrix-induced ion suppression and analytical drift. This manuscript introduces, evaluates, and applies a batch-specific, simple, and customizable normalization process: Best-Matched Internal Standard (B-MIS) normalization. By applying B-MIS normalization, we are able to retain many more LC-MS mass features during untargeted RSD filtering. For future users, we have made the tools to customize, apply, and evaluate B-MIS normalization freely available at https://github.com/IngallsLabUW/B-MIS-normalization.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.7b04400.

- Supplementary Methods; Table S1; Figures S1−S7; and captions for Tables S2−S6 PDF.
- Tables S2−S6 (XLSX)

AUTHOR INFORMATION

Corresponding Author

*E-mail: aingalls@uw.edu. Phone: (206) 221-6748.

ORCID

Katherine R. Heal: 0000-0002-4504-1039

Author Contributions

‡These authors contributed equally to this work, order determined by eyeglass prescription strength.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge L. Hmelo, J. Collins, R. Lundeen, R. Lionheart, A. Weid, and N. Kellogg for assistance with lab work, data analysis, or manuscript writing; J. Young and H. Dawson for providing the diatom cultures; B. Durham for providing several standards; and the crew and scientific party of the R/V Kilo Moana during HOE-Legacy 2A. Thanks to four anonymous reviewers whose comments greatly improved this manuscript. This work was supported by grants from the Simons Foundation (LS Award ID: 385428, A.E.I.; SCOPE Award ID 329108, A.E.I.), NSF OCE-1228770 and OCE-1205232 to A.E.I., NSF GRFP to K.R.H., NSF IGERT Program on Ocean Change to A.K.B.

REFERENCES

(1) Gika, H. G.; Theodoridis, G. A.; Plumb, R. S.; Wilson, I. D. *J. Pharm. Biomed. Anal.* **2014**, *87*, 12−25.
(2) Hounoum, B. M.; Blasco, H.; Emond, P.; Mavel, S. *TrAC, Trends Anal. Chem.* **2016**, *75*, 118−128.
(3) Bundy, J. G.; Davey, M. P.; Viant, M. R. *Metabolomics* **2009**, *5*, 3−21.
(4) Sysi-Aho, M.; Katajamaa, M.; Yetukuri, L.; Orešič, M. *BMC Bioinf.* **2007**, *8*, 93.
(5) Kido Soule, M. C.; Longnecker, K.; Johnson, W. M.; Kujawinski, E. B. *Mar. Chem.* **2015**, *177*, 374−387.

DOI: 10.1021/acs.analchem.7b04400
Anal. Chem. XXXX, XXX, XXX−XXX

(6) Llewellyn, C. A.; Sommer, U.; Dupont, C. L.; Allen, A. E.; Viant, M. R. *Prog. Oceanogr.* **2015**, *137*, 421−433.
(7) Johnson, W. M.; Kido Soule, M. C.; Kujawinski, E. B. *ISME J.* **2016**, *10*, 2304−2316.
(8) Barofsky, A.; Vidoudez, C.; Pohnert, G. *Limnol. Oceanogr.: Methods* **2009**, *7*, 382−390.
(9) Swenson, T. L.; Jenkins, S.; Bowen, B. P.; Northen, T. R. *Soil Biol. Biochem.* **2015**, *80*, 189−198.
(10) Baran, R.; Bowen, B. P.; Bouskill, N. J.; Brodie, E. L.; Yannone, S. M.; Northen, T. R. *Anal. Chem.* **2010**, *82*, 9034−9042.
(11) Paul, C.; Mausz, M. A.; Pohnert, G. *Metabolomics* **2013**, *9*, 349−359.
(12) Johnson, W. M.; Kido Soule, M. C.; Kujawinski, E. B. *Limnol. Oceanogr.: Methods* **2017**, *15*, 417−428.
(13) Dunn, W. B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-McIntyre, S.; Anderson, N.; Brown, M.; Knowles, J. D.; Halsall, A.; Haselden, J. N.; Nicholls, A. W.; Wilson, I. D.; Keil, D. B.; Goodacre, R.; Consortium, T. H. S. M. *Nat. Protoc.* **2011**, *6*, 1060−1083.
(14) Calderón-Santiago, M.; López-Bascón, M. A.; Peralbo-Molina, Á.; Priego-Capote, F. *Talanta* **2017**, *174*, 29−37.
(15) Shen, X.; Gong, X.; Cai, Y.; Guo, Y.; Tu, J.; Li, H.; Zhang, T.; Wang, J.; Xue, F.; Zhu, Z.-J. *Metabolomics* **2016**, *12*, 1−12.
(16) Gagnebin, Y.; Tonoli, D.; Lescuyer, P.; Ponte, B.; de Seigneux, S.; Martin, P.-Y.; Schappler, J.; Boccard, J.; Rudaz, S. *Anal. Chim. Acta* **2017**, *955*, 27−35.
(17) Wu, Y.; Li, L. *J. Chromatogr. A* **2016**, *1430*, 80−95.
(18) Bromke, M. A.; Sabir, J. S.; Alfassi, F. A.; Hajarah, N. H.; Kabli, S. A.; Al-Malki, A. L.; Ashworth, M. P.; Méret, M.; Jansen, R. K.; Willmitzer, L. *PLoS One* **2015**, *10*, e0138965.
(19) Fei, F.; Bowdish, D. M.; McCarry, B. E. *Anal. Bioanal. Chem.* **2014**, *406*, 3723−3733.
(20) Bijlsma, S.; Bobeldijk, I.; Verheij, E. R.; Ramaker, R.; Kochhar, S.; Macdonald, I. A.; Van Ommen, B.; Smilde, A. K. *Anal. Chem.* **2006**, *78*, 567−574.
(21) Heal, K. R.; Carlson, L. T.; Devol, A. H.; Armbrust, E.; Moffett, J. W.; Stahl, D. A.; Ingalls, A. E. *Rapid Commun. Mass Spectrom.* **2014**, *28*, 2398−2404.
(22) Heal, K. R.; Qin, W.; Ribalet, F.; Bertagnolli, A. D.; Coyote-Maestas, W.; Hmelo, L. R.; Moffett, J. W.; Devol, A. H.; Armbrust, E. V.; Stahl, D. A.; Ingalls, A. E. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 364−369.
(23) Bueschl, C.; Krska, R.; Kluger, B.; Schuhmacher, R. *Anal. Bioanal. Chem.* **2013**, *405*, 27−33.
(24) Katajamaa, M.; Orešič, M. *BMC Bioinf.* **2005**, *6*, 179.
(25) Bligh, E. G.; Dyer, W. J. *Can. J. Biochem. Physiol.* **1959**, *37*, 911−917.
(26) Canelas, A.; ten Pierick, A.; Ras, C.; Seifar, R.; van Dam, J.; van Gulik, W.; Heijnen, J. *Anal. Chem.* **2009**, *81*, 7379−7389.
(27) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. *Bioinformatics* **2010**, *26*, 966−968.
(28) Chambers, M. C.; et al. *Nat. Biotechnol.* **2012**, *30*, 918−920.
(29) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Anal. Chem.* **2006**, *78*, 779−787.
(30) Tautenhahn, R.; Böttcher, C.; Neumann, S. *BMC Bioinf.* **2008**, *9*, 504.
(31) Benton, H. P.; Want, E. J.; Ebbels, T. M. *Bioinformatics* **2010**, *26*, 2488−2489.
(32) Libiseller, G.; Dvorzak, M.; Kleb, U.; Gander, E.; Eisenberg, T.; Madeo, F.; Neumann, S.; Trausinger, G.; Sinner, F.; Pieber, T.; Magnes, C. *BMC Bioinf.* **2015**, *16*, 118.
(33) Contrepois, K.; Jiang, L.; Snyder, M. *Mol. Cell. Proteomics* **2015**, *14*, 1684−1695.
(34) Cubbon, S.; Antonio, C.; Wilson, J.; Thomas-Oates, J. *Mass Spectrom. Rev.* **2010**, *29*, 671−684.
(35) Spagou, K.; Tsoukali, H.; Raikos, N.; Gika, H.; Wilson, I. D.; Theodoridis, G. *J. Sep. Sci.* **2010**, *33*, 716−727.
(36) Poulson-Ellestad, K. L.; Jones, C. M.; Roy, J.; Viant, M. R.; Fernández, F. M.; Kubanek, J.; Nunn, B. L. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 9009−9014.
(37) Vinaixa, M.; Samino, S.; Saez, I.; Duran, J.; Guinovart, J. J.; Yanes, O. *Metabolites* **2012**, *2*, 775−795.
(38) Benjamini, Y.; Hochberg, Y. *J. R. Stat. Soc. Ser. B* **1995**, 289−300.

DOI: 10.1021/acs.analchem.7b04400
Anal. Chem. XXXX, XXX, XXX−XXX