

Supporting information for:

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, WA*

[‡]*These authors contributed equally to this work, order determined by eyeglass prescription
strength*

E-mail: aingalls@uw.edu

Phone: (206) 221-6748

Table of Contents

| | |
|-----|---|
| S3 | Supplementary methods |
| S10 | Table S1: Isotope-labeled internal standards used in this study |
| S11 | Captions for Table S2: MS and LC conditions for each analyte in our targeted method; Table S3: Full results from sensitivity analysis; Table S4: Results of B-MIS normalization of the untargeted dataset summarized by internal standard; Table S5: RSD of targeted analytes before and after B-MIS normalization in example sample sets; and Table S6: Results from evaluating internal standards as analytes |
| S12 | Figure S1: Ion suppression of polar analytes via RP or HILIC chromatography |
| S13 | Figure S2: Total ion chromatograms in RP and HILIC chromatography |
| S14 | Figure S3: Evaluation of internal standards as analytes |
| S15 | Figure S4: Demonstration of how increasing the number of internal standards improves B-MIS normalization |
| S16 | Figure S5: B-MIS selection for targeted analytes in retention time vs m/z space |
| S17 | Figure S6: RSD before and after normalization in our sample sets |
| S17 | Figure S7: Extracted ion chromatograms of example mass features from repeat injections of the pooled environmental sample |
| S18 | Supplementary References |

Supplementary Methods

Chromatography

For reversed phase (RP) chromatography, a Waters Acquity UPLC HSS Cyano column (1.8 μm particle size, 2.1 mm x 100 mm) equipped with a Waters Acquity UPLC HSS Cyano guard column (1.8 μm particle size, 2.1 mm x 5 mm) was used with 0.1% formic acid in water (Solvent A) and 0.1% formic acid in acetonitrile (Solvent B) at a flow rate of 0.4 mL/min. The HSS Cyano column has previously been used for analysis of compounds in complex marine samples with robust reproducibility of peak size and shape.^{S1,S2} The column was held at 5% B for 2 minutes, ramped to 100% B over 16 minutes, held at 100% B for 2 minutes, and equilibrated at 5% B for 5 minutes (25 minutes total). The column was maintained at 35°C. The injection volume was 15 μL . When starting a batch, the column was equilibrated at the starting conditions for at least 10 minutes; several water blanks were run before and throughout a batch. After a batch, the column was flushed with 100% acetonitrile for 20 to 30 minutes.

For HILIC chromatography, a SeQuant ZIC-pHILIC column (5 mm particle size, 2.1 mm x 150 mm, from Millipore) was used with 10 mM ammonium carbonate in 85:15 acetonitrile to water (Solvent A) and 10 mM ammonium carbonate in 60:40 water to acetonitrile (Solvent B) at a flow rate of 0.15 mL/min. This column was compared with a Waters UPLC BEH amide and a Millipore cHILIC column; the pHILIC showed superior reproducibility and peak shapes. The column was held at 100% A for 2 minutes, ramped to 100% B over 18 minutes, held at 100% B for 5 minutes, and equilibrated at 100% A for 25 minutes (50 minutes total). In marine samples, a major salt peak elutes at approximately 23 minutes (Figure S2A). The column was maintained at 30°C. The injection volume was 2 μL for samples and standard mixes. When starting a batch, the column was equilibrated at the starting conditions for at least 30 minutes. To improve the performance of the HILIC column, we maintained the same injection volume, kept the instrument running water blanks between samples as necessary,

and injected standards in a representative matrix in addition to standards in water. After each batch, the column was flushed with 10 mM ammonium carbonate in 85:15 water to acetonitrile for 20 to 30 minutes.

Mass Spectrometry

For targeted metabolomics on the Waters Xevo TQ-S triple quadrupole (TQS) with electrospray ionization (ESI), the selected reaction monitoring (SRM) transitions were monitored over a two to three minute window around the retention time for the RP and a five to ten minute window around the retention time for the HILIC. For most metabolites, two SRM transitions were monitored, one for quantification and an additional for compound confirmation. Scheduling was set up to ensure at least 12 data points per peak were collected. MS parameters were as follows: capillary voltage of 0.5 kV for positive mode and 1.0 kV for negative mode, source temperature of 130°C, cone gas flow at 150 L/h and desolvation gas flow at 1000 L/h for both positive and negative modes. Desolvation temperature was 500°C or 600°C for HILIC and RP, respectively.

For untargeted analyses with HILIC, the capillary temperature was 320°C, the H-ESI spray voltage was 3.3 kV, and the auxiliary gas heater temperature was 100°C. The S-lens RF level was 65. Sheath gas, auxiliary gas, and sweep gas flow rates were maintained at 8, 3, and 0, respectively. Polarity switching was used with a scan range of 80 to 900 m/z and a resolution of 60,000. For RP analysis of samples, the capillary temperature was 320°C, the H-ESI spray voltage was 3.8 kV, and the auxiliary gas heater temperature was 100°C. The S-lens RF level was 65. Sheath gas, auxiliary gas, and sweep gas flow rates were maintained at 40, 10, and 1, respectively. For RP, a full scan method was used with a scan range of 90 to 900 m/z and a resolution of 120,000 in positive mode.

Sensitivity Analysis

We performed a sensitivity analysis to evaluate the differences in instrument performance between the QE and TQS, which emphasized the need for careful normalization (Table S3). We determined the limit of detection (LOD) and linear range for each compound on column, and ion suppression in matrix for both of our MS platforms. To do this we injected mixes of standards at 0.05, 0.5, 5, 50, 250, 500, and 1000 nM (for compounds monitored by RP) or 5, 10, 50, 250, 500, and 5000 nM (for compounds monitored by HILIC) at least three times. LOD was determined as the concentration at which a peak was observed in all three replicate injections after passing quality control (see section on quality control parameters). Five replicate injections of standards spiked in a matrix representative of a marine microbial community (our environmental samples pooled together, see next section) were used to calculate the relative standard deviation (RSD) of compound peak areas measured on both RP and HILIC columns on the TQS and QE platforms. To assess ion suppression, we injected standards in water, standards in the pooled environmental samples, and the pooled environmental samples without standards five times.

Sample Preparation

Diatom Cultures

Axenic cultures of the sea ice diatom *Nitzschia lecointei* were grown in 35 mL borosilicate tubes and artificial seawater media (Instant Ocean salts in 1 L MilliQ water with f/2 nutrients added) with a 12 hr light:dark cycle. Triplicate cultures were grown at the following conditions: -1°C and salinity of 32 g/kg, -1°C and salinity of 41 g/kg, 4°C and salinity of 32 g/kg, and 4°C and salinity of 41 g/kg. Cultures were monitored by fluorescence and harvested during exponential growth by gentle vacuum filtration onto 47 mm 0.2 μ m PTFE filters. Filters were stored at -80°C until extraction.

Environmental Samples

Environmental samples for metabolite analysis were collected at four depths (15, 45, 75, and 125 meters) north of the Hawaiian Islands at 24° 33.284' N, 156° 19.790' W on July 31, 2015. Water was sampled from Niskin bottles into acid-washed 20 L carboys. All samples were filtered onto 142 mm 0.2 μm Durapore filters using peristaltic pumps, polycarbonate filter holders, and Masterflex PharMed BPT tubing (Cole-Parmer). Filters were quenched in liquid nitrogen immediately after filtration and stored at -80°C until extraction. Each sample was 30-40 L filtered seawater; each filter was split into three equal parts for triplicate extractions.

Metabolite Extraction

All solvents were Fisher LC-MS Optima grade. Work was done on combusted aluminum foil whenever possible; tweezers and dissecting scissors used to cut filters were rinsed with water in 10% formic acid, methanol, and dichloromethane before use and between samples; glassware (including Pasteur pipettes) was combusted before use. A blank PTFE filter was extracted alongside environmental and diatom samples as a methodological blank.

We tested four extraction procedures on marine microbial samples: boiling ethanol, hot water, cold acetonitrile:methanol:water (as described in detail in Canelas *et al.* 2009), and modified Bligh-Dyer extraction.^{S3,S4} We found that the boiling ethanol method resulted in an overall lower yield, the acetonitrile:methanol:water extraction yielded smaller peaks for the phosphorylated compounds, and the boiling water method would not extract more lipid-like compounds. Therefore, we decided to use the modified Bligh-Dyer as described below.

Extraction solvents (dichloromethane (DCM) and 1:1 methanol (MeOH) to water) were prepared the day before and stored at -20°C overnight. Bead beating tubes (15 mL, PTFE Jensen Inert Products) were filled with approximately 3 mL of a mixture of 100 μm and 400 μm silica beads and samples were placed into tubes. Approximately 2 mL of both extraction solvents were added with isotope-labeled extraction internal standards (Table S1). The

samples were shaken on a FastPrep-24 Homogenizer (MP Biomedicals) for 30 seconds, then chilled at -20°C for 15 minutes. The homogenization step was repeated twice for a total of 30 minutes at -20°C over three bead beating cycles. The samples were centrifuged at 4,300 rpm for 2 minutes at 4°C to separate the aqueous and organic phases.

The aqueous phase was removed from the bead beating tube using a glass Pasteur pipette and transferred to a 50 mL glass centrifuge tube. The remaining DCM fraction was rinsed 3 times with ~ 2 mL of cold $\text{MeOH:H}_2\text{O}$ with a bead beating and centrifuge step after each rinse. Cold DCM (~ 2 mL) was added to the centrifuge tubes containing the combined aqueous rinses. The tubes were shaken by hand, and centrifuged at 1,800 rpm for 5 minutes at 4°C . The aqueous phase was removed and transferred to a clean 20 mL glass vial and dried under clean N_2 .

DCM remaining in the bead beating tubes was transferred to the glass centrifuge tubes and rinsed twice with 2 mL of DCM, vortexing for 10 seconds after each rinse. The glass centrifuge tubes containing the combined organic rinses were spun at 1,800 rpm for 5 minutes at 4°C . The organic phase was transferred to a clean 20 mL glass vial, and dried under clean N_2 .

The dried aqueous fractions were re-dissolved in 380 μL of water; the dried organic fractions were re-dissolved in 380 μL of 1:1 water:acetonitrile. Both received a 20 μL spike of isotope-labeled injection standards (Table S1). Isotope-labeled internal standards added before extraction yielded information about extraction efficiency and phase separation, while internal standards added after extraction monitored for injection errors. In the samples discussed here, and in other cases with high extraction reproducibility, these internal standards were interchangeable during B-MIS normalization because they both experienced obscuring variation.

Pooled samples were run multiple times throughout a sample run, at least three times at full strength and three times at half strength (diluted 1:1 with pure water) to train B-MIS normalization, as described in the main text. Standards in water and standards spiked into

an aliquot of the pooled sample were analyzed at the beginning and end of each sample run to set limits on the in-house quality control retention time and ion ratio ranges. Samples were analyzed within 12 or 72 hours of extraction for the HILIC and RP analyses, respectively, and stored at -80°C until analysis.

Standards

Most metabolite standards (isotope-labeled and unlabeled) were purchased from Sigma Aldrich, Santa Cruz Biotechnology, Millennium Enterprises, American Custom Chemical Corporation, and Cambridge Isotope Laboratories. Cinnamoyl-HSL was provided by the C. Harwood laboratory (University of Washington, Seattle, WA); DHPS (2,3-Dihydroxypropane-1-sulfonate) was provided by A. Bourdon and S. Champagna (University of Tennessee, Knoxville); N-acetyltaurine was provided by A. Cook and K. Denger (University of Konstanz, Germany); DMSP (Dimethylsulfoniopropionate) was provided by W. Whitman and M. Moran (University of Georgia). Full information about providers is included in Table S2.

Internal standard peaks in both targeted and untargeted data were integrated using Skyline for small molecules.^{S5} We inspected the peak shape and intensity of the internal standards and determined whether the signal was reliable and should be used for normalization. In the untargeted HILIC data for the environmental samples, D₄-succinic acid and ¹³C₂-acetyl CoA were removed because of these reasons. Thus these internal standards are not included for the environmental data in Tables S4 and S6 or Figure S3.

Quality control of integrated peaks

After integrations, we use an in-house quality control (QC) filter to ensure proper metabolite identification based on retention time, signal to noise ratio (> 4), absolute peak height, exact mass (< 5 ppm, for QE), and fragment ion ratio (for TQS). Acceptable retention times and ion ratios of a compound were set by the retention time of the standards (deviation ≤ 0.2

minutes of the standard run for RP, ≤ 0.4 minutes for HILIC). If the peak area of an analyte in a sample was not more than three times the peak area in the methodological blank, the data for that analyte were not used. Finally, peak height was used to flag peaks that were overloaded and thus not acceptable for relative quantitation. To filter low quality LC-MS features (pairs of m/z and retention time) in untargeted analyses, we discarded features from the first 0.5 minutes, after 17.8 minutes (RP), or after 25 minutes (HILIC) because features that elute at these times exhibit severe ion suppression. We applied additional filters to remove features that were detected in less than one quarter of the samples, and those with an average pooled signal less than three times the average methodological blank signal.

Supplementary Tables and Figures

Table S1: Isotope-labeled internal standards used in this study. The chromatography column, ion mode (z), injection concentration, and fraction used for analysis for each standard are provided. Note that the group of standards used for B-MIS normalization can be customized to suit the needs of individual studies; this is just an example set.

| | Standard | Fraction | Column | z | Inj Conc (nM) | Extracted m/z |
|-----------------------------|--|----------|--------|----|---------------|---------------|
| Extraction Standards | Indole 3 Acetic Acid, D ₅ | Aqueous | RP | +1 | 5000 | 181.1021 |
| | Arachidonic Acid, D ₈ | Organic | RP | +1 | 20000 | 313.2977 |
| | Tryptamine, D ₄ | Aqueous | RP | +1 | 2500 | 164.1246 |
| | L-Cysteic Acid, D ₃ | Aqueous | HILIC | -1 | 1000 | 171.0161 |
| | Sulfolactic Acid, ¹³ C ₃ | Aqueous | HILIC | -1 | 1000 | 171.9913 |
| | Sulfoacetic Acid, ¹³ C ₂ | Aqueous | HILIC | -1 | 2500 | 140.9774 |
| | Isethionic Acid, ¹³ C ₂ | Aqueous | HILIC | -1 | 500 | 126.9981 |
| | Taurine, D ₄ | Aqueous | HILIC | -1 | 1000 | 128.0325 |
| Injection Standards | L-Phenylalanine, D ₈ | Aqueous | RP | +1 | 100 | 174.1365 |
| | Riboflavin-dioxypyrimidine (Vitamin B ₂), ¹³ C ₄ , ¹⁵ N | Aqueous | RP | +1 | 100 | 383.1531 |
| | L-Tryptophan, D ₃ | Aqueous | RP | +1 | 100 | 208.1161 |
| | Pyridoxal, D ₃ | Aqueous | RP | +1 | 200 | 171.0843 |
| | Thiamine, ¹³ C ₃ | Aqueous | RP | +1 | 200 | 269.1252 |
| | Alpha-Tocopherol, D ₆ | Organic | RP | +1 | 5000 | 436.4182 |
| | Docosahexaenoic Acid (DHA), D ₅ | Organic | RP | +1 | 2000 | 334.2789 |
| | Eicosapentaenoic Acid (EPA), D ₅ | Organic | RP | +1 | 5000 | 308.2632 |
| | DL-Proline, D ₇ | Aqueous | HILIC | +1 | 1000 | 123.1145 |
| | DL-Valine, D ₈ | Aqueous | HILIC | +1 | 1000 | 126.1365 |
| | L-Methionine, D ₃ | Aqueous | HILIC | +1 | 1000 | 153.0722 |
| | DL-Histidine, ¹⁵ N | Aqueous | HILIC | +1 | 2500 | 157.0738 |
| | DL-Alanine, D ₃ | Aqueous | HILIC | +1 | 5000 | 93.0738 |
| | L-Isoleucine, ¹⁵ N | Aqueous | HILIC | +1 | 500 | 133.0989 |
| | Acetyl CoA, ¹³ C ₂ | Aqueous | HILIC | -1 | 1000 | 810.1252 |
| | Succinic Acid, D ₄ | Aqueous | HILIC | -1 | 10000 | 121.0444 |
| | Guanosine Monophosphate, ¹⁵ N ₅ | Aqueous | HILIC | -1 | 2000 | 367.0359 |
| | Adenosine Monophosphate, ¹⁵ N ₅ | Aqueous | HILIC | -1 | 5000 | 351.0411 |

Table S2: MS and LC conditions for each analyte in our targeted method. Column (RP or HILIC) and retention times (RT) for each compound, with extracted m/z (for QE) and transitions monitored (for TQS). Collision energy (CE) and cone voltage (CV) were tuned for the SRM analysis only, MS conditions for high-resolution reported in text. We did not run RP-QE under negative polarity, thus we did not collect high-resolution data for a few compounds, noted as nd in the m/z column. Compounds listed in bold are used as internal standards. Compounds with a † at the end of their name are monitored on both HILIC and RP columns. Compounds with a ‡ were added to our method after the sensitivity analysis was conducted, thus we report their MS and LC conditions here but they are not included in Table S3. This table is provided as a separate file.

Table S3: Full results from sensitivity analysis. Linear range is reported for compounds in water as fmol on column with the lower limit being the LOD (described in text) using raw areas. Relative standard deviation (RSD) calculated over 5 injections of the standard in representative matrix, reported as %. R^2 is the average R^2 of the 3 standard curves of each compound within the reported linear range. We did not run the RP-QE under negative polarity (noted as nd in the Lin Range, RF ratio, RSD, and R^2 columns). Occasionally, compounds exhibited ion suppression so intensely that we were no longer able to detect the compound in matrix (noted as dl). A number of compounds were added after the initial linear range analysis (noted as **); for these compounds we do not report the linear range or R^2 , but do supply an RF ratio and RSD. Compounds in Table S2 marked with a ‡ are not included in this table since they were added to the method after the sensitivity analysis was completed. This table is provided as a separate file.

Table S4: Results of B-MIS normalization of the untargeted dataset, summarized by internal standard. Fraction of A-MIS (acceptable-matched internal standard) is the fraction of quality mass features of the technical replicates that may have used each internal standard, while the fraction of B-MIS is the fraction of quality mass features that choose each internal standard. If the fraction of A-MIS is equal to the fraction of B-MIS, then every time the internal standard was considered acceptable, it is chosen. Users should consider adding more internal standards that are similar in chemistry to internal standards where the fraction of A-MIS is \approx fraction of B-MIS. This table is provided as a separate file.

Table S5: RSD of targeted analytes before and after B-MIS normalization in our two example sample sets (environmental and diatom cultures). For each analyte in each sample set, we have listed the acceptable internal standards for normalization. The best option (B-MIS) is listed first with other internal standards that meet the cut-off rules listed after. The 'Matched?' column is true if there is an acceptable internal standard in common between the sample sets. Analytes that were not detected or were overloaded are noted as: not detected in pooled. This table is provided as a separate file.

Table S6: Results from evaluating internal standards as analytes for the untargeted HILIC data from the environmental samples. 'RSD of Pooled Sample' is the RSD of multiple injections of the pooled sample at full and half strength normalized to the internal standard given in the 'Internal Standard used for Normalizing' column. 'RSD of Samples' is the RSD of the internal standard across the whole sample set, normalized to the internal standard given in the 'Internal Standard used for Normalizing' column. 'RSD of pooled with no normalization' is the same value as the injection volume normalized value in the 'RSD of Pooled Sample' column; this is used to calculate the 'percent change from raw' value. The 'acceptable normalization?' column is true if 'percent change from raw' is above 0.4 (see text for discussion of this cutoff value). Data are plotted in Figures 2 and S3.

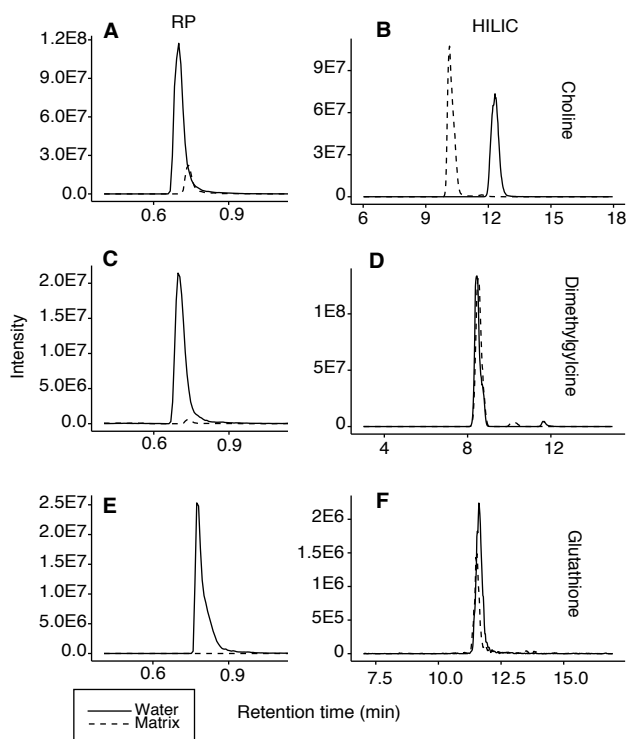


Figure S1: Ion suppression of polar analytes via RP or HILIC chromatography. Example chromatograms of standards of three compounds (choline (A, B), dimethylglycine (C, D), and glutathione (E, F)) run in water (solid line) or representative matrix (dashed line), on two different LC systems (RP (A, C, E) and HILIC (B, D, F)). These data were acquired on the triple quadrupole mass spectrometer.

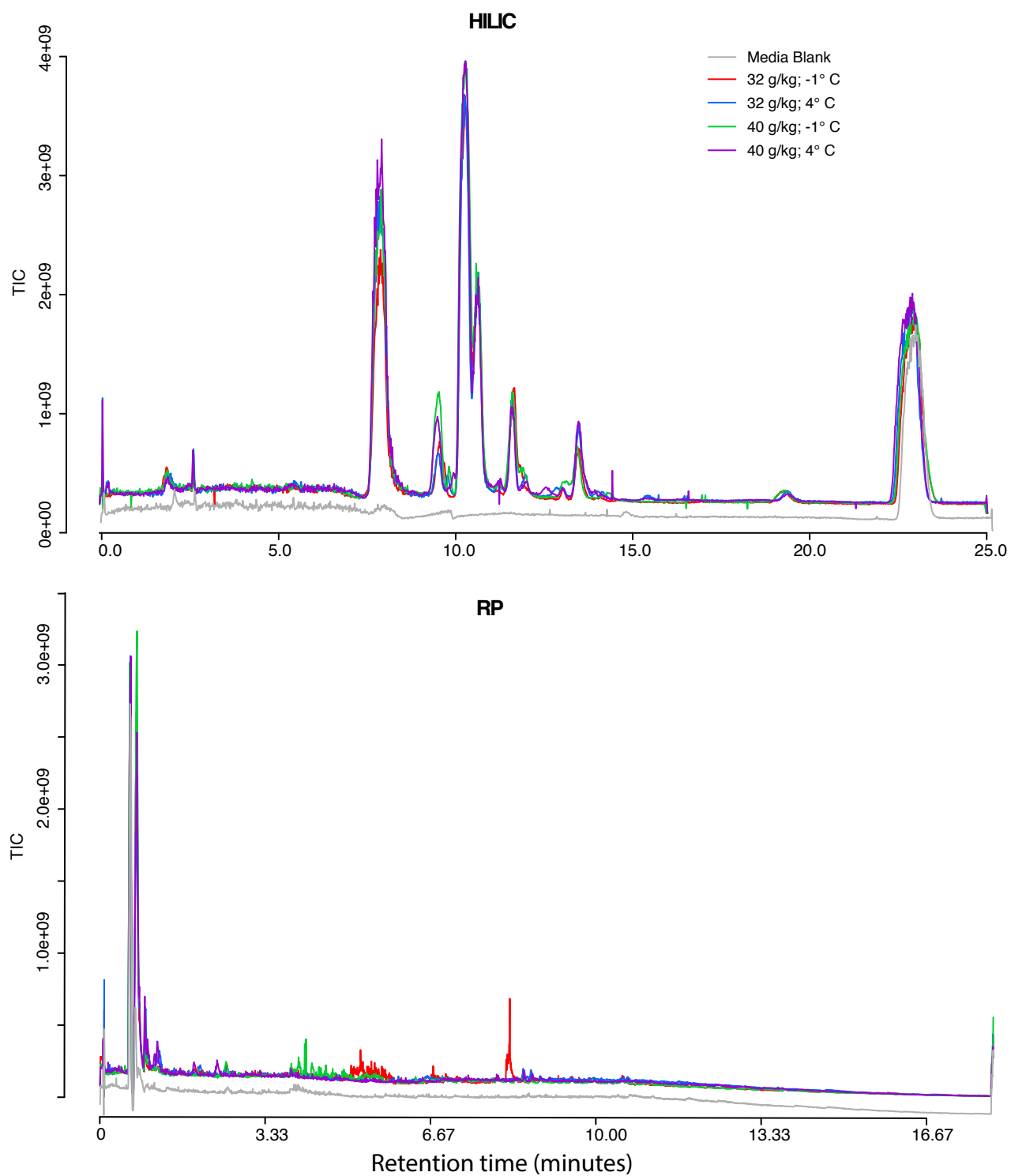


Figure S2: Total ion chromatograms from the QE for the aqueous fraction of four treatments of the diatoms and a methodological blank (offset on the y-axis by 1.25×10^8), run on the HILIC column (top), and on the RP column (bottom).

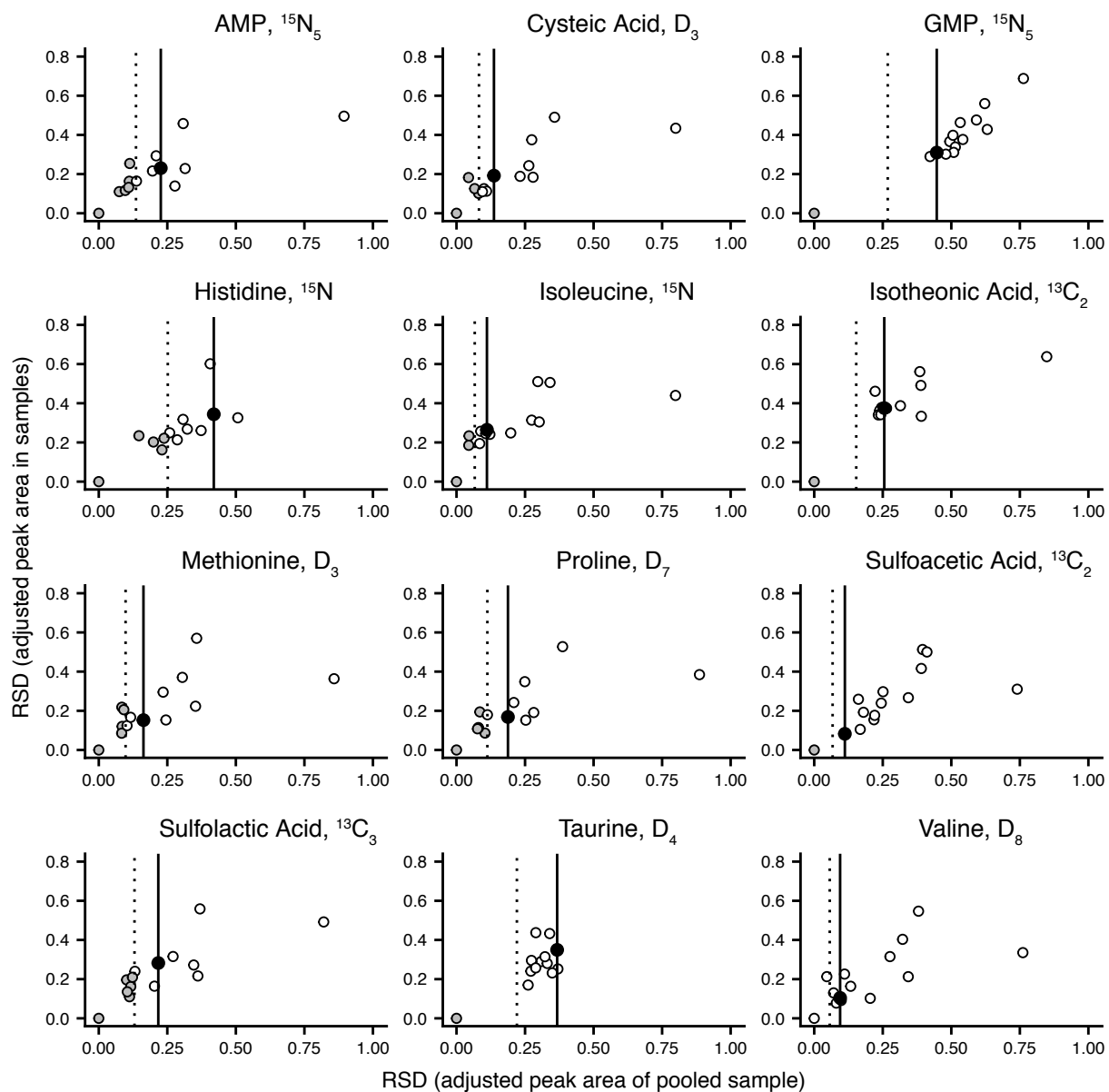


Figure S3: Results from evaluating internal standards as analytes for the untargeted HILIC data from the environmental sample set. Colors and lines are as in Figure 2. Note that Valine, D_8 had an original RSD < 0.1, therefore B-MIS was not applied (Scheme 1, all points white).

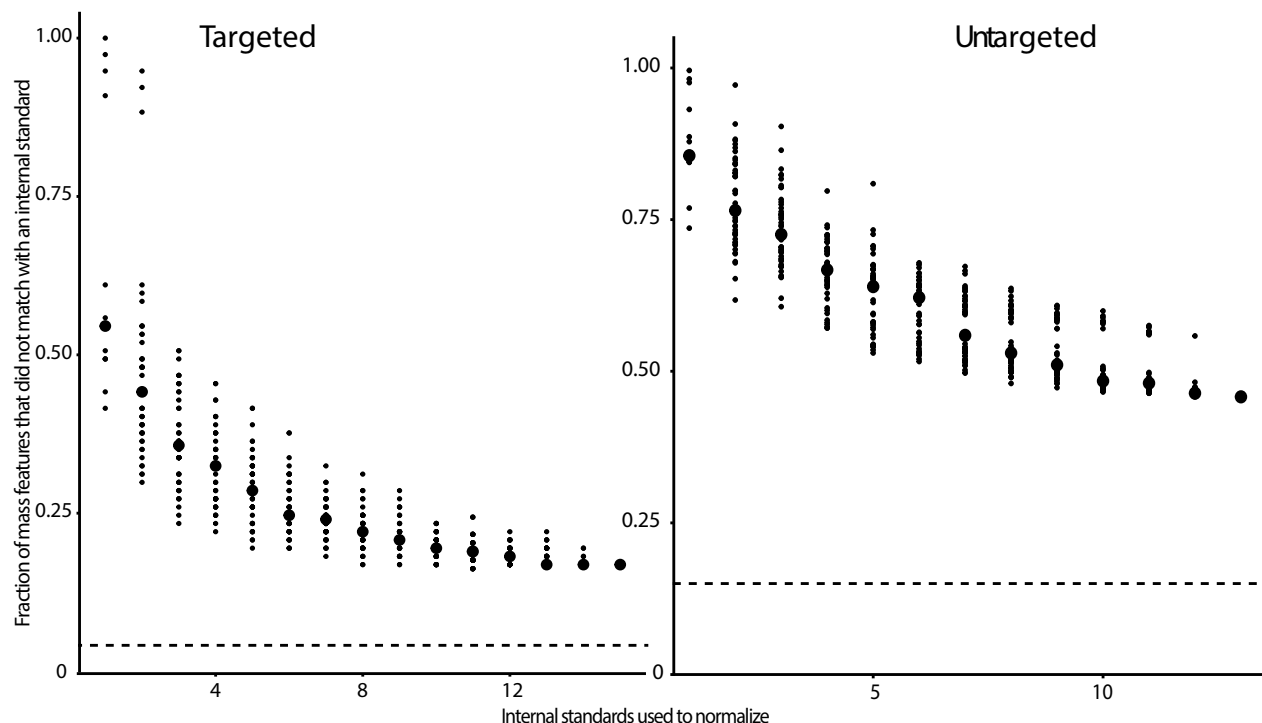


Figure S4: Number of internal standards included during B-MIS normalization vs the fraction of analytes (in targeted analyses, left) or mass features (in untargeted analyses, right) whose obscuring variation did not match with an acceptable internal standard. For each value along the x-axis, normalization was done either with all possible combinations of internal standards or with 50 possible combinations of internal standards (if there were more than 50 possible combinations). Each combination is represented in a small point; the median of the iterations is shown in large circle. The data used here are from the environmental HILIC samples. Only quality mass features (final RSD of the pooled < 0.2 after normalization by all included internal standards) were used in the untargeted analysis. Dashed lines indicate the 5% and 15% of mass features that had an original RSD of < 0.1 in targeted and untargeted analyses, respectively; B-MIS normalization does not attempt to normalize these mass features. For this untargeted data, 13 internal standards are used for B-MIS normalization because D₄-succinic acid and ¹³C₂-acetyl CoA do not ionize well enough at the concentrations injected in this matrix for robust detection with the QE. With all internal standards, over 54% of mass features in the untargeted data and 83% of compounds in the targeted data matched with an internal standard for normalization.

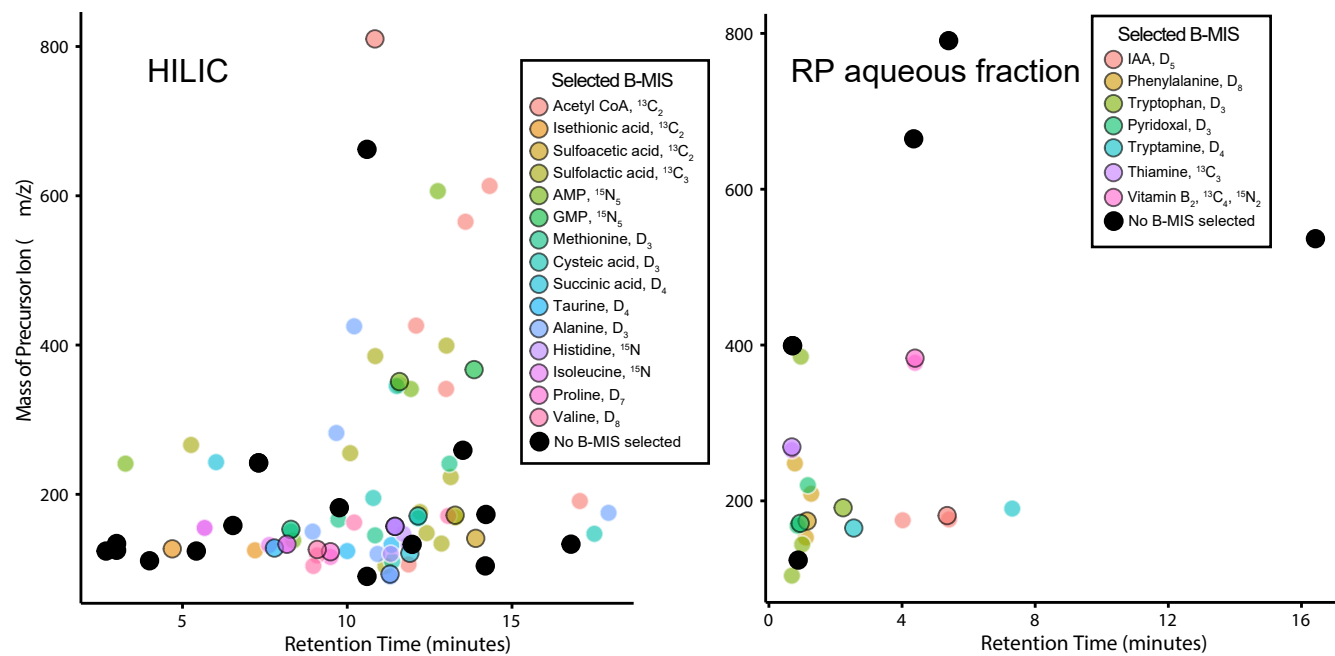


Figure S5: Retention time vs m/z of the measured analytes in the targeted environmental aqueous extracts. Color shows which internal standard was chosen for each analyte, with each internal standard outlined in black. Black markers are analytes whose obscuring variation was not improved by normalizing to any internal standard. Full results are given in Table S5.

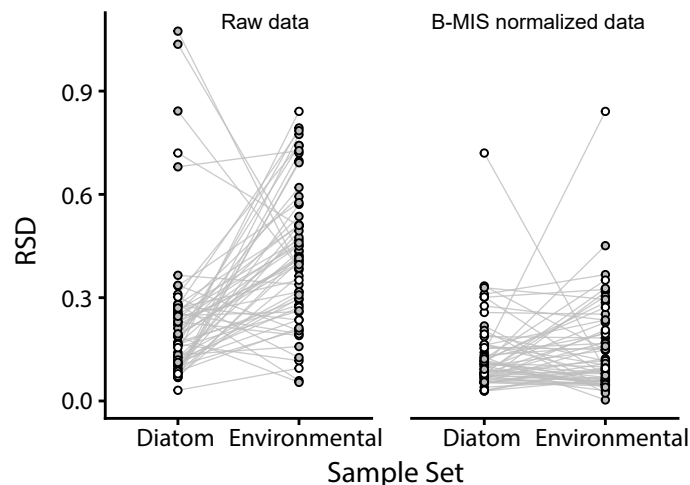


Figure S6: RSD of raw and B-MIS adjusted areas of replicate injections of pooled samples in the two matrices used in the targeted analysis of this study. The lines connect the same compound between both sample sets, demonstrating that the obscuring variation of each analyte changes with matrix type, but not in a predictable manner. Open circles are analytes that did not find an acceptable matched internal standard and are therefore not normalized (see Scheme 1), while closed circles are analytes that did use B-MIS normalization. After B-MIS normalization, the range of RSD is similar between the sample sets. Full results are given in Table S5.

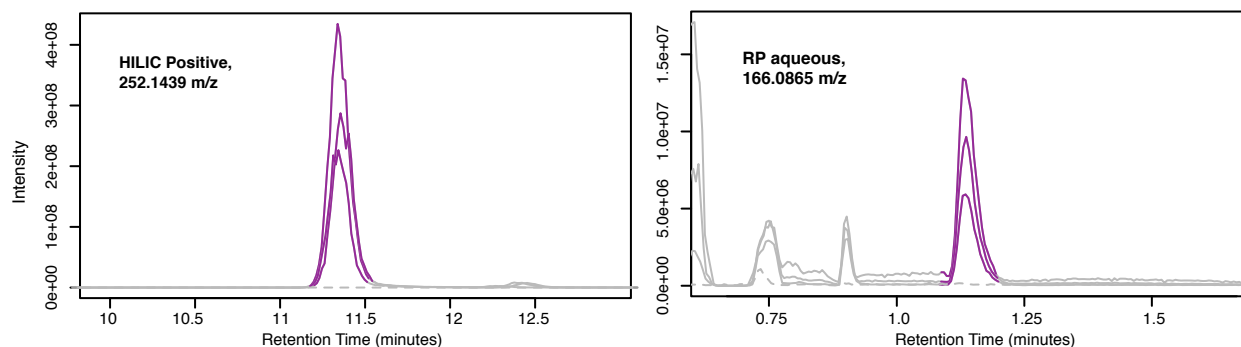


Figure S7: Extracted ion chromatograms of repeat injections of the pooled environmental sample acquired during the untargeted analysis of environmental samples. Two example mass features (from HILIC (left) and RP chromatography (right)) showed significant obscuring variation that could be minimized by B-MIS. Both of these mass features would have been excluded from downstream data analysis without B-MIS normalization if applying an RSD filter. XCMS peak integration shown in violet, methodological blank shown as dashed line.

Supplementary References

References

- (S1) 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.
- (S2) 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.
- (S3) Bligh, E. G.; Dyer, W. J. *Can. J. Biochem. Physiol.* **1959**, *37*, 911–917.
- (S4) Canelas, A.; ten Pierick, A.; Ras, C.; Seifar, R.; van Dam, J.; van Gulik, W.; Heijnen, J. *Anal. Chem.* **2009**, *81*, 7379–7389.
- (S5) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebner, D. C.; MacCoss, M. J. *Bioinformatics* **2010**, *26*, 966–968.