analytical chemistry

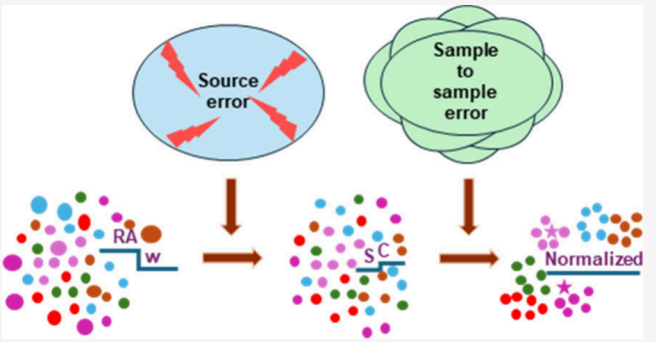
pubs.acs.org/ac
Article

Achieving Effective Batch-to-Batch Error Correction through Suppression Correction and Dual MSTUS Normalization

Debasish Ghosh, Janmejoy Kar, Felice A. de Jong, Chris Beecher, and Vladimir Shulaev*

Cite This: *Anal. Chem.* 2026, 98, 10966−10977
Read Online

ABSTRACT: Batch effect corrections are necessary to remove instrumentally introduced variance but should retain biological variance. Most common batch correction techniques are based on analysis of quality control (QC) samples, sample amplitude normalization, or other statistical and normalization techniques. Many of these approaches correct for the totality of the MS signal and do not differentiate between analytical and biological variances. In this study, *Saccharomyces cerevisiae* was treated with four different oxidants at five different time points. Biologically replicated sample sets were analyzed across six batches where the batch effect was intentionally introduced. We used the Isotopic Ratio Outlier Analysis (IROA) workflow method to first correct suppressed raw MS data for source related errors and then subsequently applied an enhanced normalization technique. These results were compared to those of other widely used normalization methods. Importantly, for validation of our approach, we conducted careful statistical evaluation and biological validation to ensure that the correction had effectively removed batch effects without losing biological signal. This workflow method could successfully distinguish between technical and biological variation and reduce the relative standard deviation (RSD) error to 1% compared to raw data, a clear indication that this method can be used to reliably enhance data quality. The experimental design afforded us with a means to test the ability of each normalization method to correctly group similar samples in a "data fidelity test" that may have utility for future studies.



INTRODUCTION

Technological advances in analytical methods such as liquid chromatography−mass spectrometry, data analytics, and bioinformatic tools have rapidly advanced the ability to make and interpret a wide range of metabolites necessary and useful for the clinical metabolomics community.[1] Currently there are several major approaches to metabolomics, namely targeted analysis, metabolite profiling, and metabolic fingerprinting.[2−4] While almost all clinical measurements are targeted, most metabolomic investigations start with nontargeted, or "unbiased", analyses, which is considered a good means of hypothesis generation and the first step in a larger series of "discovery" investigations. Unfortunately, the techniques employed in nontargeted analyses largely encounter pitfalls that are often forgotten, misunderstood, or neglected. The two most serious of these misconceptions are that (1) ion losses, often called "suppression" or the "matrix effect", are likely consistent across all samples for a given sample type and are therefore ignored, and (2) the samples are similar enough in size and form that normalization is either simple or unnecessary.

The major portion of suppression is ion suppression, which is a well characterized,[5−7] though still poorly understood phenomenon having to do with the variance of the efficiency of ionization of any molecule. It was first described in the early 1990s[8] and is considered to be present to some extend in almost all ionization sources, including matrix-assisted laser desorption/ionization (MALDI),[9] desorption electrospray ionization (DESI)/MALDI,[10] electrospray ionization (ESI), and atmospheric pressure chemical ionization (APCI),[11] and all postsource MS processes, including MS/MS.[12] In general, suppression results in reduced detection capability, precision, and accuracy as an effect of the complex function of molecular structure that involves all of its physical parameters, i.e., acidity/basicity, polarity/aromaticity, hydrophobicity/lipophilicity, electronic and physical structures, the concentration of the compound itself,[5] the nature of the source, elution solvents, gas temperature, and the general chemical environment and complexity of the effluent in which the compound elutes. Methods to reduce the chemical complexity of the samples,

Received: August 26, 2025
Revised: March 20, 2026
Accepted: March 24, 2026
Published: April 7, 2026

https://doi.org/10.1021/acs.analchem.5c05210
Anal. Chem. 2026, 98, 10966−10977

such as during sample preparation by SPE or by baseline chromatographic separation, or similarly reduction of the total concentration of a compound, for instance by use of nanospray sources, may reduce suppression for some compounds, but do not completely remove suppression.[13] In addition to ion suppression, other losses or sources of variance that are common in current MS sources include in-source fragmentation, inefficient ion transport, injector variances, etc. Even though they arise through different processes and are not common to all ion sources, the correction for these losses has been a long-considered goal.

The most common way to correct for the effects of suppression is through the use of internal standards[14] or stable isotope dilution−mass spectrometry.[15] If the number of compounds being measured is small, individual aliquots of purified stable isotope labeled internal standard (SIL-IS) represent the optimal solution; however, as the number of compounds that needs to be measured increases, the SIL-IS may not always be available and is often very expensive.

Therefore, techniques such as Mass Isotopomer Ratio Analysis of U-^{13}C Labeled Extracts (MIRACLE)[16,17] based on isotopically enriched biological mixtures were developed. Other attempts at using SIL-IS have tried using a single representative compound as a standard for other compounds of that class, however, as was previously noted the immediate source environment is a critical component of suppression and the likelihood that the standard compound is suffering the same level of suppression as the analyte is difficult to ensure if they are in different environments.[18] Where the SIL-IS is an extract of an organism, such as *Escherichia coli*, the analysis becomes a targeted analysis for all of the compounds represented in the extract. In MS detection, both the standard and the analyte are always suppressed to the same extent if they are coeluting,[19] and under these circumstances, the ratio of the analyte to the standard allows the calculation of the concentration of the analyte as long as the concentration of the standard is known. However, it should be noted that a deuterium labeled internal standard will almost always alter the chromatographic behavior of a compound,[20] causing a difference of retention time and hence differential matrix effects,[21] and for these reasons the use of heavy stable carbon isotopes, such as ^{13}C, is preferred. The difficulty of techniques such as MIRACLE, where the isotopic standards are synthesized or biosynthesized using ≥99% ^{13}C, is that the resulting isotopic standards are difficult to identify from other natural abundance and noise peaks. Using specific isotopic probabilities for each carbon, generally 5% U−^{13}C and/or 95% U−^{13}C, the IROA protocols[19,22−26] create isotopomeric patterns in the mass spectrum of each compound that are unique for each formula, and most importantly are both easily identified within mass spectral scans and will be rarely mimicked by random noise. They perform exactly like any isotopic standard and, aside from the complexity of their isotopolog clusters, share almost all the physical chemical properties of their natural abundance counterparts, except mass.

In a recent publication, Mahmud and coauthors[19] successfully demonstrated effective suppression correction and normalization using an IROA-based stable isotope-labeled internal standard (IROA-IS) plus novel companion algorithms (IROA Workflow) across chromatographic systems, ion sources and biological matrices. Considering the values of both the IROA-IS and natural abundance samples, the authors used a modified version of the MS Total Useful Signal (MSTUS) algorithm, a Dual-MSTUS algorithm, to normalize all suppression-corrected natural abundances and markedly increase accuracy and precision.

While suppression-correction may correct for MS in-source losses, it cannot correct for sample-to-sample (STS) variances. When quantitating metabolites, it is critical to reduce or eliminate variations due to differences in the sample's sizes, dilutions, or biologically dissimilarity. Despite the criticality of this step, there is not a commonly used method for data normalization, and most methods yield different results because they are based on different principles. There are many statistical and software approaches for batch effect correction (normalization). These commonly used methods include Technical variation elImination with ensemble learninG architEctuRe (TIGER),[27] ComBat,[28] Concordance-based Batch effect correction (CordBat),[29] EigenMS,[30] MS Total Useful Signal (MSTUS),[31] analysis of covariance (ANCOVA),[32] and systemic error removal using random forest (SERRF).[33] It is important to note that these statistical- and software-based approaches do not actually correct the suppressed values in the raw data. While normalization methods that do not consider the effects of suppression yield better interpretable data sets than using un-normalized data, a more accurate data set normalization can be achieved by correcting the variances of the suppressed data before normalization.[25]

The consequences of not correcting the variances associated with these two key sources of MS data error, both suppression and normalization, are numerous and can be potentially devastating to clinical metabolomic outcomes. Ion suppression causes a decrease of the analyte signal, reducing detection capability affecting ion ratio, linearity, and quantification. Without proper suppression correction and normalization, data sets may be skewed by high-variance noise leading to false positives (or negatives) and the underestimation of analyte concentration. Importantly, since many clinical metabolomic studies rely on the analysis of 1000s of samples, sometimes over weeks or months of time, batch effects arising from variations in experimental conditions, instrumentation, and sample handling across different batches will impact the quality of data, leading not only to compromised data analysis but also to data that will not be comparable lab-to-lab and may not even be comparable day-to-day within a given lab. With adequate normalization and reduction of instrumental and technical variation, not only will statistical power be improved, but the quality, reproducibility, and reliability of data will be significantly enhanced to support clinical biomarker validity.

In this study we applied the IROA Workflow to demonstrate batch-to-batch error correction by first correcting the raw MS data for all source errors and then applied a variation of the Dual-MSTUS normalization mechanism to the suppression-corrected data set. We compared the normalization performance of the IROA Workflow with other normalization methods and then tested the ability of each normalization method to correctly group similar samples in a "data fidelity test".

MATERIALS AND METHODS

Yeast Strains and Culture Media

Yeast strain BY4743 ([4741/4742] MATa/MATalpha his3delta1/his3delta1 leu2delta0/leu2delta0 lys2delta0/+met15delta0/+ura3delta0/ura3delta0) was obtained from American Type Culture Collection (ATCC, Manassas, VA, USA, Catalog #201390). Cultures

https://doi.org/10.1021/acs.analchem.5c05210
Anal. Chem. 2026, 98, 10966−10977

were kept in long-term storage and frozen at −80 °C in glycerol stocks. Work cultures were kept at 4 °C in Yeast Extract Peptone−Dextrose (YPD) Medium (yeast extract 0.1% (w/v), peptone 0.5% (w/v), and dextrose 2% (w/v)) agar plates.

Cell Growth, Oxidative Stress Conditions, and Metabolite Extraction

Saccharomyces cerevisiae cells were initially batch grown overnight at 30 °C, pH 6.0, 150 rpm, in minimal medium (25 mL) containing Difco yeast nitrogen base without amino acids, 2% (w/v) dextrose supplemented with uracil 20 mg/L, L-leucine 60 mg/L, and L-histidine 20 mg/L until midexponential phase was reached, based on measuring optical density measurement at 600 nm ($OD_{600} \approx 1.5$). The inoculum from overnight culture was used to inoculate minimal media (25 mL) containing Difco yeast nitrogen base without amino acids, 2% (w/v) dextrose supplemented with uracil 20 mg/L, L-leucine 60 mg/L, and L-histidine 20 mg/L. The cultures were grown at 30 °C, pH 6.0, 150 rpm. Conditions of oxidative stress were achieved by using previously described protocol[34] at the mid-exponential phase by the addition of 1.2 mM H_2O_2, cumene hydroperoxide (CHP) 190 μM, menadione 150 μM, or diamide 1.5 mM at the mid-exponential phase ($OD_{600} \approx 1.5$). Control samples were prepared by adding the same volume of solvent used to prepare different oxidants. Six biological replicates were used for each condition. Post oxidative stress treatment samples were collected at different time points: 3, 6, 9, 12, and 20 min. Samples were quenched according to the protocol in which Spura and coauthors[35] described using a quenching solution containing 0.8% NaCl (*w*/*v*) in aqueous 40% ethanol (*v*/*v*) solution using a dry ice−ethanol bath kept at −40 °C. Metabolite extractions were performed from 5 mg of dry yeast cells using a two-step extraction process with 2:2:1 Acetonitrile:Methanol:H_2O (v/v) extraction solvent. Yeast cells were broken using 100 μL of 0.5 mm zirconia beads in a Retsch Mixer Mill MM 300 (Retsch GmbH, Haan, Germany) at 30 Hz for 30 s followed by rest on dry ice for 1 min. Combined extractions were dried in a vacuum concentrator and resuspended in 30 μL of IROA IS.

Internal Standard Preparation and Calibration Experiment

The IROA Workflow kit was obtained from IROA Technologies. The kit contains two different types of standard samples, IROA Long-Term Reference Standard (IROA-LTRS), a 1:1 mixture of 5% U−^{13}C and 95% U−^{13}C labeled metabolites and IROA internal standard (IROA-IS), 95% U−^{13}C labeled metabolites. IROA-LTRS and IROA-IS were prepared according to manufacturer guidelines by adding 40 μL and 900 μL aliquots of 2% methanol (v/v), respectively (for workflow and discussion, see Figure S1, ref 19, and the IROA Basic Concepts document in Supporting Information 1 for a full discussion of the IROA protocol methods, theory, and the algorithms applied). Before processing the experimental samples, a calibration experiment was performed to determine the amount of experimental sample needed to establish an approximate 1:1 ratio of experimental metabolites to IROA internal standard metabolites for optimal quantitation.

Batch Effect Experiment

Metabolites were extracted in six batches, where each batch contained one replicate from each individual group, including both controls and treatments.

On different weeks batches were processed, and samples were immediately analyzed by LC-MS. Batch 1 was intentionally left inside the autosampler for 2 days to articulate the batch effect. Batches 2−4 were processed continuously over several weeks. After analyzing batch 4, the ESI source and capillary were replaced with a similar but different ESI source and capillary in order to introduce source error, and batches 5 and 6 were processed using the replaced ESI source and capillary. An IROA-LTRS sample was injected every 10 samples to establish a quality control (QC) set for each batch.

Metabolite Analysis by High-Performance Liquid Chromatography Mass Spectrometry

Extracted metabolites were analyzed by a hybrid linear ion trap (LTQ) Orbitrap Velos Pro mass spectrometer (Thermo Scientific, San Jose, CA, USA) coupled with an Accela Open autosampler and Accela 1250 UHPLC pump. A reverse-phase ACE Excel 2 C18-PFP column (150 mm × 2.1 mm) 2.0 μm was used for metabolite separation at a 350 μL/min flow rate using a 30 min gradient. The gradient system consisted of solvent A, 0.1% formic acid in water, and solvent B, 0.1% formic acid in methanol. Five microliters of the extracted metabolites were injected into the system and eluted with a gradient starting from 100% solvent A for 0−3 min, then 3−13 min of linear gradient up to 20% Solvent A, 13−16 min 20% Solvent A, and finally 16−16.5 min linear gradient from 20 to 100% Solvent A. The column was reconditioned using 100% Solvent A for 13.5 min before starting the next gradient cycle. The column temperature was maintained at 55 °C throughout the gradient cycle. The LTQ Orbitrap Velos Pro equipped with an electrospray ionization (ESI) source was used to carry out data acquisition where data was acquired in both positive and negative ionization mode at 15,000 resolution and mass range 50−1000 *m*/*z*. Nitrogen was used as a sheath gas. The source voltage was 4.5 kV and the heated capillary temperature was 250 °C.

Data Analysis

Raw files were converted to mzXML format using MSConvert from the ProteoWizard Suite (version 3.0.24164 64 bit, proteowizard.sourceforge.io) by enabling the vendor specified peak picking function. Converted data were processed using IROA ClusterFinder software (Version 4.2.97). Prior to actual sample processing, LTRS samples were processed to build a library which was later used for semitargeted analysis of batch experimental samples. Processed raw data obtained from ClusterFinder software was used to compare other normalization methods such as ComBat, EigenMS, ANCOVA, CordBat, and TIGER. Raw data were uploaded to the MetaboAnalyst 6.0[36] platform to perform ANCOVA, ComBat, and EigenMS batch effect correction. CordBat and TIGER batch effect correction were performed using the published R script from the literature. In order to apply SERRF, which requires a minimum number of QC samples per batch, we reorganized the raw data into three batches by reconstructing batches 1 and 2, 3 and 4, and 5 and 6 into three individual batches. SERRF normalization was performed using the temporal web-based platform created by SERRF authors. Post-normalized data obtained from all the analysis were processed in R Studio (version 4.4.1) (cran.r-project.org) for statistical analysis and visualization. Permutational Multivariate Analysis of Variance (PERMANOVA) was conducted utilizing the Bray−Curtis dissimilarity measure to evaluate differences between groups (treatment vs control). The analysis was performed with 999 permutations to compute the statistical significance. Results obtained from PERMANOVA were visualized using boxplots, where separate plots were generated for pseudo-F statistics and *p*-values.

Data Fidelity Analysis

The experimental design was established so that each batch consisted of a collection of similar (analogous) biological samples. Since biologically similar samples were present in every batch, it was reasoned that the similar samples should be able to be grouped close to one another after normalization in a standard clustering technique if the normalization did not disrupt the data fidelity. Two sets of samples representing very different biological outcomes (early- and late-stage toxicity for the same toxin) were selected randomly. All normalization approaches were compared using unsupervised PCA performed using either raw or suppression-corrected (SC) data sets. The homogeneity of sample distribution and dispersion was examined within the first three principal components (PC1 vs PC2 and PC2 vs PC3).

https://doi.org/10.1021/acs.analchem.5c05210
Anal. Chem. 2026, 98, 10966−10977

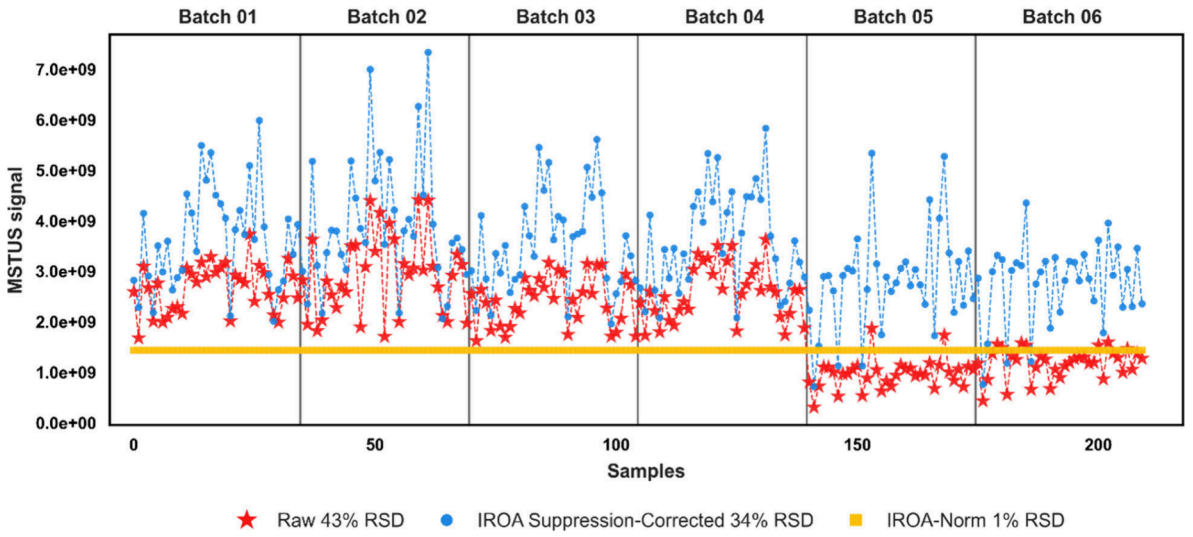


Figure 1. Core compound-based normalization. Suppression-correction and Dual-MSTUS normalization were applied to reduce batch effects in the sample data by using common compounds present in all samples. Red asterisks indicate the summed MSTUS signal of all the metabolites from raw data in each sample; blue circles indicate the suppression-corrected MSTUS value of each sample; and yellow squares indicate the summed MSTUS signal of all the metabolites from normalized data.

RESULTS AND DISCUSSION

S. cerevisiae cells were treated with four different oxidants that are known to induce oxidative stress. We processed 210 samples (35 per batch) in six batches over several weeks with each batch providing a complete analyzable set of samples. All samples for each individual batch were independently generated and can be expected to exhibit normal experimental noise, i.e., duplicate samples were purposefully not generated as the information from them would be gleaned from the consistent use of the LTRS sample. The first four batches were analyzed with at least one or more intervening week(s). Between the fourth and fifth batches the ESI source and capillary were replaced with a similar but different ESI source and capillary to simulate common instrument downtime issues with long-running experiments. Samples were analyzed during this extensive time period, and a very strong batch effect was observed between the early and later otherwise identical batches, as well as within both the early and later batch-sets. We then assessed the performance of the IROA Workflow for the ability to normalize the samples and correct for batch effects across multiple batches of samples (see Figure 1).

We followed the IROA Workflow protocol for sample preparation (see Figure S1, ref 19, and the IROA Basic Concepts document in Supporting Information 1 for a full discussion of the IROA protocol methods, theory, and the algorithms applied), LC/MS-based metabolite profiling, library creation, data normalization, and accurate quantification of endogenous metabolites.

As defined in the workflow, a constant amount of an internal standard (IS) was added to every experimental sample, and a Long-Term Reference Standard (LTRS) was analyzed approximately every 10 injections. In addition to providing continuous and real time instrument performance QA/QC, the LTRS was used to build a reference library of RT and MS-MS characteristics for each peak, which provided the basis of verifiable ^{12}C peak identification in experimental samples, relative to the LTRS. The IS provided a mechanism to (1) remove artifactual data, (2) correct ion suppression, and (3) perform a sample-to-sample normalization. The removal of artifactual data was based on the removal of all of the peaks in the experimental samples that were not matched to an IS peak, all of which are of biological origin, universally and randomly labeled at 95% ^{13}C, and added to each experimental sample at a constant concentration. Accordingly, there were two isotopically differentiable populations of peaks (experimental peaks (^{12}C) and IS peaks (^{13}C)), and since the IS signal suffered the same variances as the experimental signal, i.e., source, sample input, matrix, in-source fragmentation, chromatographic conditions, etc., the loss of ^{13}C signals in each compound in each sample was measured (Figure S2) and used to correct for the loss of signal in the corresponding ^{12}C signals. Figure S2 illustrates this in a calibration experiment in which the samples represent a gradient with known increasing concentrations of sample extract in each injection that are analyzed against a constant concentration of IS that was added to each sample. The blue diamonds show the MSTUS value of the suppression corrected data set, and the yellow dots show the MSTUS values in the raw data set.

After correcting suppression losses, a Dual-MSTUS algorithm was used to normalize the samples and remove the batch effect in the data. In the original MSTUS normalization protocol,[31] only known peaks were used in an effort to remove "signal noise", and the normalization value was selected in an arbitrary fashion. Here, we used an improved IROA-based "Dual MSTUS" normalization that uses only IS accompanied peaks to determine two MSTUS values (experimental and IS), namely, the sum of the total area under the curve (AUC, suppression-corrected) for all ^{12}C peaks and their matched ^{13}C peaks.

For the calculations, initially we followed the previously published routine of selecting all the metabolites that were matched with an isotopically known metabolite present in each individual sample for normalization.[19] This normalization (Figure S3) was effective and yielded a reduction in sample-to-sample variance that was lower than the sample-to-sample variance of the samples in any individual batch and brought all batches to be comparable. We refer to this approach using all

https://doi.org/10.1021/acs.analchem.5c05210
Anal. Chem. 2026, 98, 10966−10977

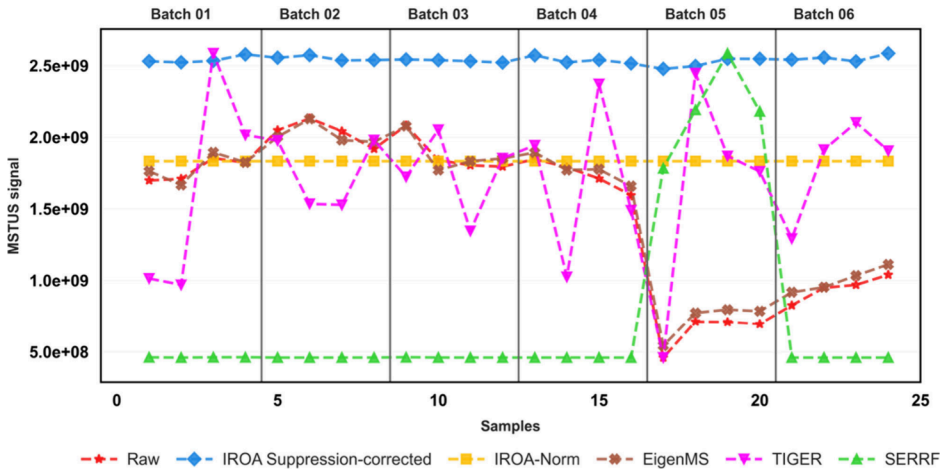


Figure 2. Summed values of QC samples after performing different normalization methods. The *x*-axis shows the injection order of QC samples, and the *y*-axis shows the summed MSTUS signal from individual QC samples. Each dot in the plot represents the summed value of all of the metabolites present in one QC sample.

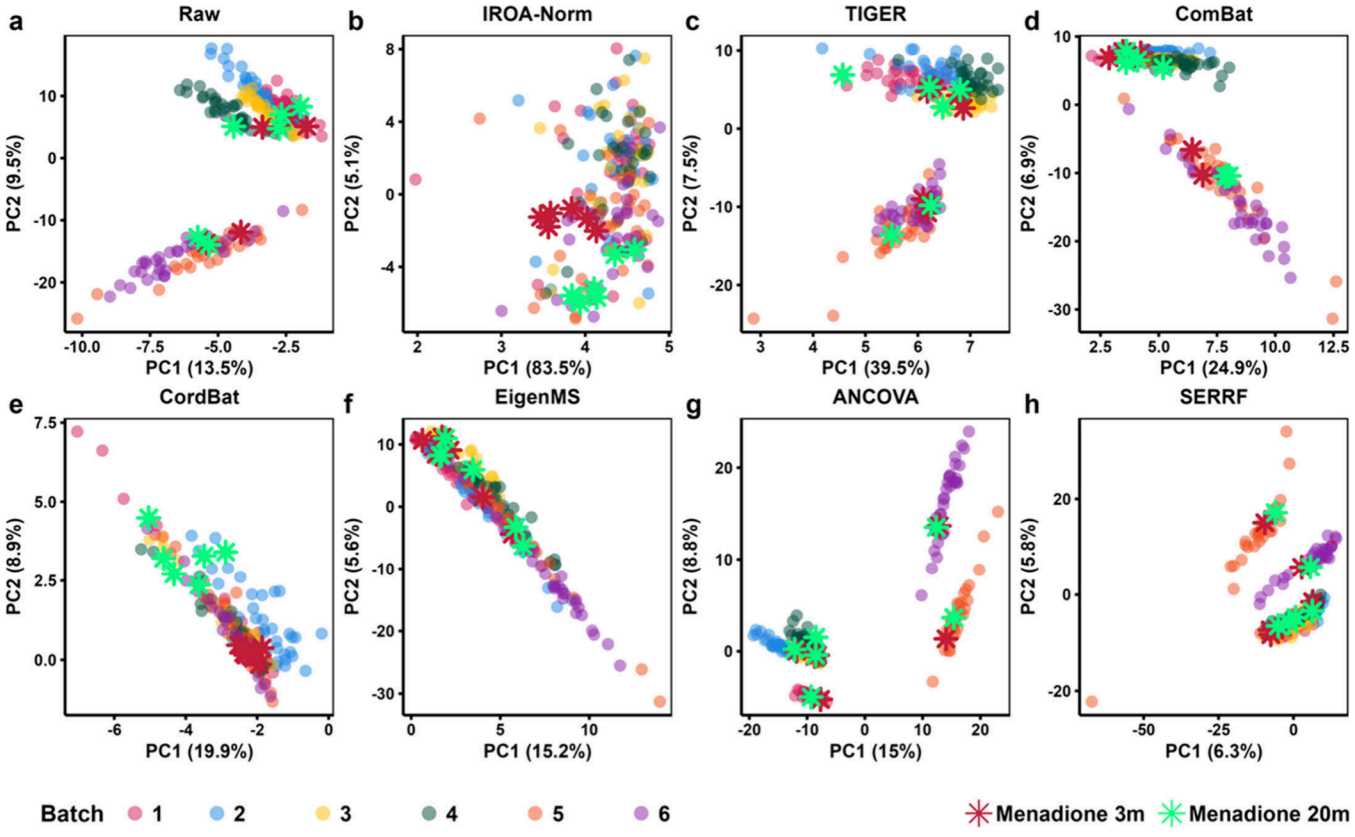


Figure 3. PCA of batch raw data normalized by different normalization methods. PCA score plot of the raw (uncorrected) data (a) and PCA score plots of data normalized by IROA-Norm (b), TIGER (c), ComBat (d), CordBat (e), EigenMS (f), ANCOVA (g), and SERRF (h). Asterisks in the plot represent the overlay of experimental samples treated with menadione for two different time-points, 3 and 20 min (shown in cardinal red and green, respectively).

peaks that have both an IS and natural abundance peak as the "all compound" method.

"All Compound" vs "Core Compound" for Batch Effect Correction (Normalization)

While the "all compound" method produced a reasonable result for batch effect correction (normalization), we realized that the number and composition of peaks would be different sample-to-sample in large batch-to-batch type experiments, so we decided to explore an additional approach we called the "core compound" method (Figure 1). In the "core compound" method only compounds present in all of the samples are used to generate the normalization factor used to normalize all the samples (based on the same collection of compounds in each sample). Our result with the "all compound" normalization (Figure S4) indicates that there is still some random scatteredness within batches 1 to 4 and batches 5 to 6, despite reducing the error in the raw data from 40% to 8%.

https://doi.org/10.1021/acs.analchem.5c05210
Anal. Chem. 2026, 98, 10966−10977

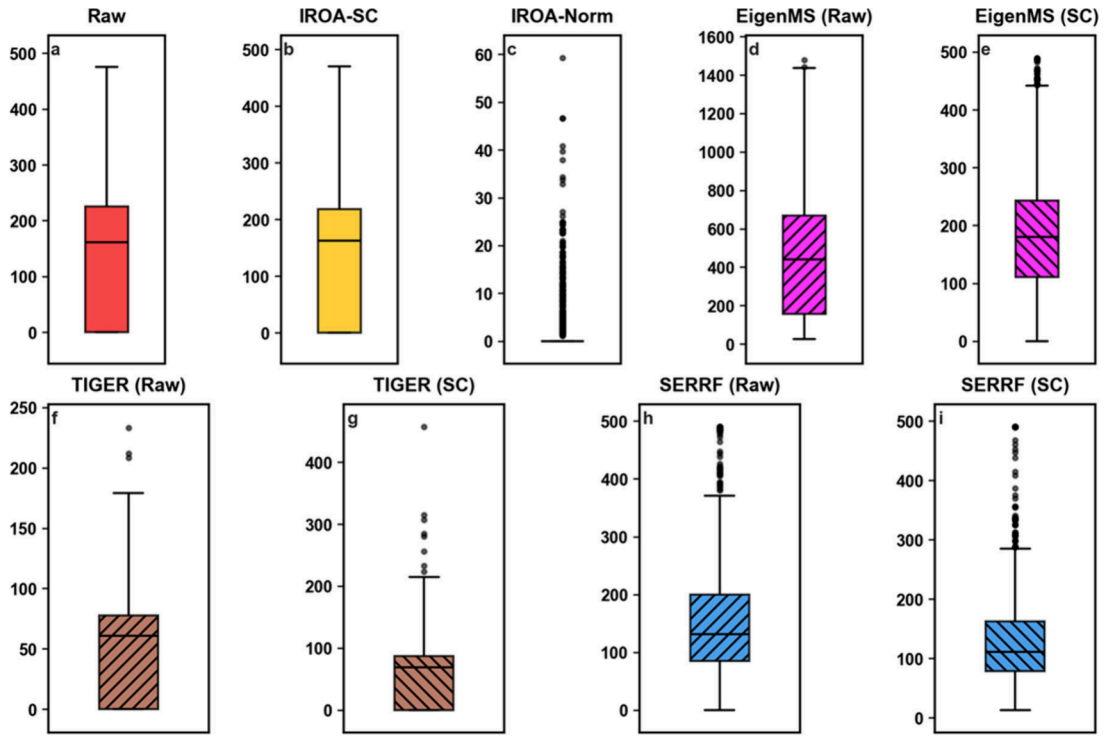


Figure 4. Boxplots showing the RSD distribution of compounds in QC samples by different normalization methods using either raw or suppression-corrected (SC) data. The results were obtained after analyzing QC-RSD of each metabolite from all the samples across different normalization methods. The *x*-axis shows the methods used to perform normalization of QC samples, and the *y*-axis shows the deviation of data points from the median. In the box plot, the black line inside the box represents the median distribution of the data, and the upper and lower box edges represent the deviation from median. The results show that for all of the normalization methods, except TIGER, the median RSD improved using SC data instead of raw data. The median RSD for EigenMS improved using SC data over raw data; however, there was no improvement relative to the raw (uncorrected) data. IROA-Norm performed well with a majority of compounds within a 1% median window. All values are calculated based on the standard deviation from the mean.

Presumably, the remaining variance can be explained by the fact that samples in the data set are biologically different, and the "all compound" approach increased the deviation during normalization because it contained too many compounds that were sample specific. In contrast, our "core compound" method normalization (Figure 1) not only shows much tighter clustering both within and across all of the batches but also reduced the RSD error from 8% to 1%.

In both cases, we believe that the MSTUS algorithm was successful in reducing RSD error because it is being applied to suppression-corrected data which already has technical variance reduced. This suppression-correction is based on a constant concentration of the IS that is itself also subjected to the same correction. These two error correction methods, suppression correction and normalization, remove a significant source of variance inherent in the raw data across all batches.

Performance Evaluation of Different Normalization Methods

To evaluate the performance of the IROA Dual-MSTUS normalization of suppression-corrected data, we compared it to normalization procedures that employ statistical batch averaging approaches (Figure S5), including TIGER,[27] ComBat,[28] CordBat,[29] EigenMS,[30] ANCOVA,[32] and SERRF.[33] Raw data show systematic drift over injection order and across all the batches. SERRF demonstrated improved normalization performance following the introduction of batch error; however, samples from batch 5 deviated substantially, possibly due to batch reconstruction. Methods such as EigenMS and IROA Normalization (IROA-Norm) display much tighter clustering and more consistent baselines, effectively reducing the batch-to-batch variation. ComBat and ANCOVA provide good normalization but still exhibit moderate spread, not effectively reducing the batch-to-batch variation. CordBat and TIGER exhibit similar patterns with partial improvement compared with raw data.

Summed MSTUS Values of QCs (LTRS) across Batches

The metabolites across six batches of QC (LTRS) samples were evaluated by using the summed value of their AUCs (MSTUS values). A total of 1253 metabolites from 24 QC samples (4 per batch) were used to evaluate the normalization efficiency (Figure 2). Our results indicate that there was a massive drift observed in the raw sample MSTUS signal after the 16th QC sample in batch 4, and these batch effects were significantly corrected following the IROA-Norm.

SERRF demonstrated improved normalization performance until the source change, and TIGER showed a slight improvement in reducing batch effect for batches 5 and 6 when compared to the raw data. The EigenMS normalization method showed a similar pattern compared to that of raw data and did not significantly improve the normalization of QC data. No significant drift was observed in the suppression-corrected and IROA normalized QC samples, and suppression-corrected data were shown to greatly improve the batch-to-batch variation compared to raw data, suggesting that suppression-correction is necessary to reduce data variances and plays a key role in the success of normalization.

https://doi.org/10.1021/acs.analchem.5c05210
Anal. Chem. 2026, 98, 10966−10977

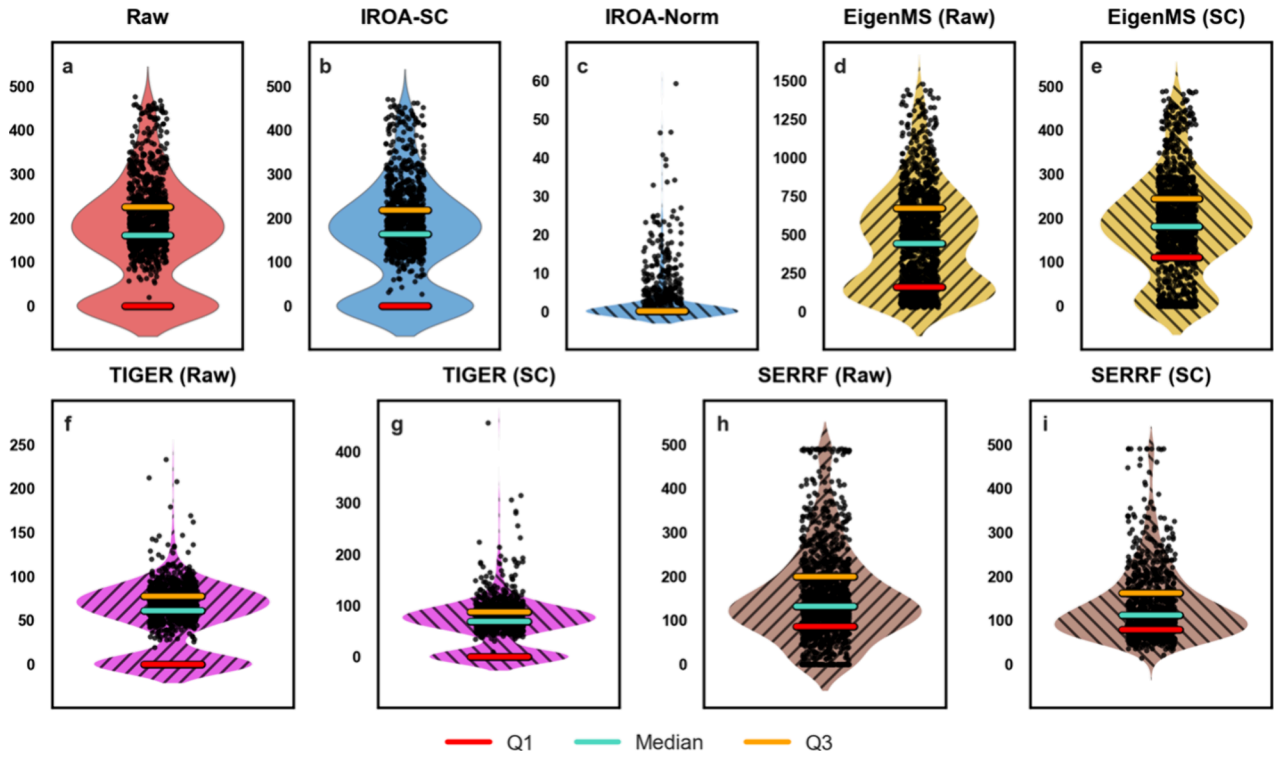


Figure 5. RSD distribution of QC samples after performing different normalization methods using either raw or suppression-corrected (SC) data. The violin plot combines a box plot and a kernel density plot, providing both summary statistics and a data distribution shape. The interquartile range (IQR) spans from the first quartile (Q1, red) to the third quartile (Q3, orange) and contains the middle 50% of the data. The horizontal line inside the violin plot indicates the median (Q2, aqua). The bottom and top of the violin mark the minimum (Q0) and maximum (Q4) values observed in the data, respectively. Surrounding the violin plot is the kernel density estimation, where the width of the violin reflects the frequency of observations: wider sections indicate a higher data concentration. Results were obtained after analyzing the QC-RSD of each metabolite from all of the samples through different normalization methods. Each dot in the violin plot represents the corresponding value of one metabolite across all of the samples. The results show that the IROA-Norm data yield a 1% RSD, outperforming other methods. RSD is calculated based on the standard deviation from the mean.

Evaluation of Batch Effect Correction (Normalization) Efficiency

For further evaluation of the normalization methods for batch effect correction, we performed Principal Component Analysis (PCA) to measure the proportion of variance that was captured by each method (Figure 3). The percentage of variance explained by each principal component is calculated by dividing its eigenvalue by the sum of all eigenvalues and converting the number into a percentage. This percentage indicates how much of the original data's variability is captured by each component, with higher percentages reflecting stronger underlying patterns. A total of 1253 metabolites were identified from the analysis of 24 LTRS (a QC "standard sample" which is similar to the "pooled" samples in other methods). These compounds were then identified in the 90 untreated control samples and 120 experimental samples distributed across 6 batches. The complete 234 sample data sets were analyzed by PCA and plotted to observe the relationships of samples according to each principal component.

We performed a comparison of raw and suppression-corrected (SC) data normalized with different normalization methods. Figure 3 shows the PC1/PC2 plot of raw (raw) data normalized by different methods. The distributed variances of the original (unnormalized) raw experimental data across the first two PCs are 13.5% and 9.5%, respectively (total = 23%). The PCA of the IROA Normalized data set (IROA-Norm, Figure 3b) demonstrates the best normalization by capturing most of the variance in the first two PCs 83.5% and 5.1%, respectively (total = 88.6%). The second highest variance captured in the first two PCs is achieved by TIGER (Figure 3c), 39.1% and 7.5%, respectively (total = 46.6%). As we consistently observed, TIGER's PC1 clearly separated the LTRS (QC) samples from experimental samples but was unable to overcome the variances between the early and later batches in PC2, and they remained as two distinct groups. SERRF (Figure 3h) captured 6.3% in PC1 and 5.8% in PC2 (total = 12.1). As described in the Methods, the batch structure was reconstructed to meet SERRF requirements in this analysis, and the percentages of PC1 and PC2 may reflect the effects of this modification. The PCs of EigenMS (Figure 3f), capturing 15.2% in PC1 and 5.6% in PC2 (total = 20.8), similarly demonstrate the separation of the QC and experimental samples in PC1, and the experimental samples were collected along the PC2 axis. Other non-QC based methods such as ANCOVA, CordBat, and ComBat (Figure 3d,e,g) performed slightly better than the QC based method EigenMS but not as well as TIGER and IROA, capturing only 15%, 19.9%, and 24.9% in PC1 and 8.8%, 8.9%, and 6.9% in PC2, respectively (totals = 23.8%, 28.8%, and 31.8%). See Figures S6−S9 for the PCA using either batch raw data or suppression-corrected (SC) data samples (with and without LTRS (QC) samples) normalized by different normalization

https://doi.org/10.1021/acs.analchem.5c05210
Anal. Chem. 2026, 98, 10966−10977

methods (S6, raw, no LTRS; S7, raw with LTRS; S8, SC, no LTRS; S9, SC with LTRS).

Assessment of QC Metabolites Using RSD

In this study, some of the normalization methods we compared require the presence of QC samples while others do not (i.e., ComBat, EigenMS, and CordBat). The IROA LTRS (QC) standard samples (identical in composition and amount) provide another good way to assess the performance of the normalization method. The QC-RSD of each metabolite from all samples was analyzed and compared across different normalization methods (Figure 4). We observed a median value of over 150% RSD in the raw data (Figure 4a) that was drastically reduced to 1% using IROA-Norm (Figure 4b). The TIGER normalization median was approximately 60%, although a high deviation of up to 250% was observed for some data points (Figure 4c). The median RSD of the SERRF-normalized QC samples was approximately 78%, although certain data points exhibited deviations of up to 500% (Figure 4h). The EigenMS median RSD at 440% was higher than that of the raw data (Figure 4d). This is likely because the EigenMS approach focuses on removing the bias based on experimental treatment or group differences. To conclude, the IROA-Norm data method outperformed all other normalization methods that we tested, lowering the QC RSD to 1% compared to raw data despite very significant batch effects due to extensive source modifications.

RSD Distribution of Metabolites across All QC Batches

The representation of QC-RSD (Figure 4) shows only the distribution of data from its median value for the control samples but does not represent the overall distribution of the sample RSD data among different batches. In order to understand the overall spread of the data, we analyzed the longitudinal distribution of the raw and suppression-corrected (SC) data using violin plots (Figure 5). In the raw data (Figure 5a) two separate clusters were observed, one cluster distributed in the lower quantile Q1 and another in Q3. The second cluster captured the dispersed data points above the median range, suggesting systemic variation across batches. The tightest cluster and longitudinal distribution of the data were seen using IROA-Norm (Figure 5b). Although two main clusters were also generated with both raw and SC data using the TIGER method (Figure 5f,g), the longitudinal data distribution and RSD percentage decreased compared to raw data. The first cluster, predominantly in Q1 showed less systemic variation, while in the second cluster systemic variation was still observed and little difference was evident comparing raw and SC data. We also observed similar but more systemic variation in the EigenMS normalized raw and SC data sets (Figure 5d,e), which may be explained by the same phenomenon we observed and described in the violin plot section. SERRF (Figure 5h,i) exhibited a similar clustering pattern, distributing the QC data into two distinct clusters for both raw and SC data; however, the median values were substantially lower than those of EigenMS. Notable improvement was seen using SC data in both EigenMS and SERRF methods. The longitudinal distribution of data analysis across all different methods shows that IROA-Norm data distribution achieves significantly tighter clusters and lower % RSD compared to other QC-based normalization methods.

Preservation of Biological Significance

Normalization is a double-edged sword and should be handled cautiously as it can move the direction of the data in unintentional ways. The main purpose of normalization is not only to reduce the Sample-to-Sample (STS) variation but also to keep biological importance of individual data points preserved and meaningful. To further evaluate the normalization methods, it is instructive to return to the PCA analyses previously discussed. As is shown in Figure S7 (raw with QC) and Figure S9 (SC with QC), PC1 in all cases separated the QC samples from the experimental samples. To establish the degree to which each method preserves information integrity, we examined two principal components. The QC samples were removed to better visualize the sample distributions; see Figures S6 (raw) and S8 (SC). As described in the experimental discussion, there were four treatments (with three controls) and five time points in each batch. It would be expected that the same treatment/time sample should show batch to batch consistency if the method maintains informational integrity; therefore, two same-oxidant, nonzero time points were selected for comparison. Figure 3 shows the PCA (PC1 vs PC2) comparison of data normalized by different methods where each different experimental sample is represented by a dot. The samples for time periods 3 and 20 min (3m and 20m, respectively) treated with the oxidant/toxin menadione are highlighted with red or green asterisks, respectively. In the raw data (Figure 3a), the two time points are represented as separated samples in each batch, but as expected, all of the batches are distinctly different. For TIGER, ComBat, ANCOVA, and SERRF (Figure 3c,d,g,h), this same pattern is seen (i.e., the samples are more tightly associated post normalization); however, they all still appear in distinctly separated batches, i.e., the batch effects observed appear to be the greatest sources of variance, not their biological variance. In EigenMS (Figure 3f), the different batches are no longer distinct but are merged into a single linear structure; however, the two menadione time periods do not show distinct separation. Both IROA-Norm (Figure 3b) and CordBat (Figure 3e) distinctively cluster the menadione 3m and 20m samples. The IROA-Norm PCA PC1 and PC2 represent 83.5% and 5.1%, respectively, of the total variance remaining after normalization, while CordBat PC1 and PC2 represent 19.9% and 8.9%, respectively, of the remaining variance. This would suggest that there was significantly more variance remaining in the CordBat post normalization than with IROA-Norm normalization.

In Figure S6B, the results from PC2 and PC3 are examined and show a similar response compared to the PC1 and PC2 analyses. For the IROA-Norm and CordBat data (Figure S6B(b,e)), we observe that menadione 3m and 20m samples separately clustered, i.e., the 3m and 20m samples across all samples each collocate, suggesting that their biological informational integrity is maintained in all three principal components after normalization. In the PC2 by PC3 plot of the ComBat and EigenMS normalized data (Figure S6B(d,f)), the replicate groups of menadione 3m and 20m could not be discriminated in any of the three PCs. In the remaining methods, the PC2 by PC3 plots do not show any collocation nor do they batch correct effectively. We observed identical phenomena, namely, the coexistence of 3m and 20m samples, in the TIGER, ComBat, ANCOVA, and SERRF analyses (Figure S6B(c,d,g,h)). Notably, in Figure S8, PCA of suppression-corrected data, the PC1 vs PC2 plot (Figure

https://doi.org/10.1021/acs.analchem.5c05210
Anal. Chem. 2026, 98, 10966−10977

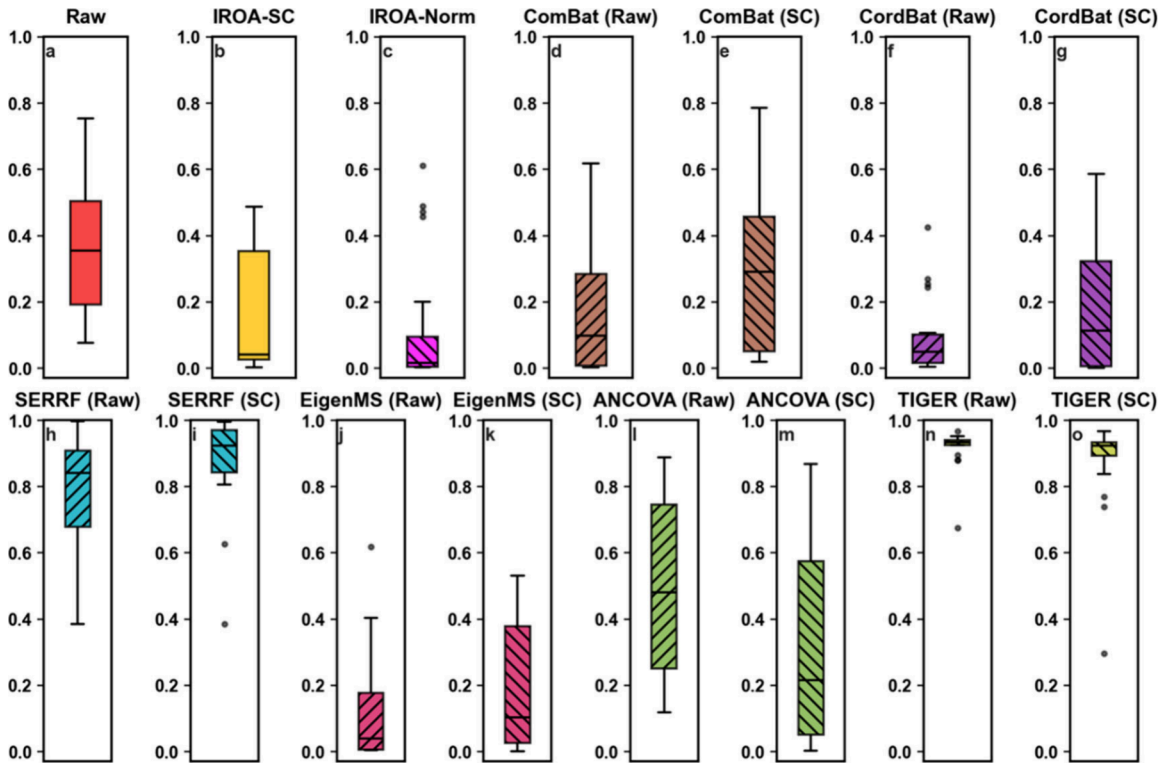


Figure 6. PERMANOVA analysis (p-value) of different normalization methods using either raw or suppression-corrected (SC) data sets. Evaluation of p-value distribution of raw or suppression-corrected (SC) data across different normalization methods after performing PERMANOVA between groups (treatment vs control) across all samples. The x-axis shows the methods used to perform different normalization, and the y-axis shows the calculated distribution of p-values. The black line inside the box represents the median distribution of the data, and the upper and lower box edges represent the 2nd and 3rd quartiles, respectively.

S8A(h)) shows that SERFF appears to separately cluster the menadione 3m and 20m samples effectively (capturing a total of 28.3% variability), but this same separation is not observed in the SERFF PC2 vs PC3 plot (Figure S8B(h)).

In order to establish the degree to which each method preserves biological information integrity in addition to graphical Data Fidelity Analysis, we performed Permutational Multivariate Analysis of Variance (PERMANOVA). PERMANOVA was conducted utilizing the Bray−Curtis dissimilarity measure to evaluate differences between groups (treatment vs control). The analysis was performed with 999 permutations to compute statistical significance. Table S1 shows the results of a quantitative ablation analysis comparing Median RSD, PCA, and PERMANOVA metrics of all normalization methods using either raw or suppression-corrected data sets to delineate the individual contributions of the suppression-correction step vs the Dual MSTUS normalization.

Figure 6 graphically represents PERMANOVA results obtained using boxplots. The raw p-value distribution (Figure 6a) indicated a higher median significance around 0.4, whereas IROA-Norm p-values (Figure 6b) exhibited a median below 0.05, suggesting that their biological informational integrity is maintained. Among other normalization techniques, CordBat and EigenMS exhibited median p-values less than 0.1 (Figure 6f,j), while ComBat showed a median near 0.1 (Figure 6d). In contrast, TIGER, ANCOVA, and SERRF normalization approaches had substantially higher median p-values of 0.94, 0.5, and 0.84, respectively (Figure 6n,l,h).

In PERMANOVA, the pseudo-F statistic is also determined, representing the ratio of between-group to within-group variance (Figure S10). Raw data exhibited a lower pseudo-F close to 1 (Figure S10a) in comparison to IROA-Norm data (Figure S10c), which showed the highest pseudo-F value, approximately 5. CordBat and EigenMS normalization of raw data resulted in pseudo-F values of approximately 2 and 3 (Figure S10d,j), with ComBat exceeding 2 (Figure S10d). TIGER, ANCOVA, and SERRF normalization methods demonstrated notably lower pseudo-F values of 0.15, 0.80, and 0.55, respectively (Figure S10n,l,h). With the exception of IROA-Norm, which represented the largest variance between treatment vs control, the PERMANOVA p and pseudo-F values underlying the variances between treatment vs control showed little or no improvement using suppression-corrected data across the remaining normalization approaches. Discrepancies observed across different normalization methods may reflect an emphasis on batch effect correction based on mathematical interaction without sufficient consideration of the underlying biological variation among samples.

Our ablation analysis suggests that the underlying theoretical differences between the normalization techniques are substantial (see Table S1). In some cases, as in the case of SERRF, the analysis improved with the use of SC data over raw data. For SERRF, the median RSD values were reduced, the discriminating strength of PCA was increased, and data fidelity was improved; however, the PERMANOVA (p and F values) results were effectively unchanged. The results from the other normalization methods using SC data were even more marginal. Neither SERRF nor any of these techniques achieved the level of consistent normalization that we report here with IROA-Norm. It is possible that future techniques will be formulated that will rely on SC data and achieve better results, but our analysis so far suggests that the current normalization

https://doi.org/10.1021/acs.analchem.5c05210
Anal. Chem. 2026, 98, 10966−10977

techniques rely on theoretical underpinnings that apply more to raw than SC data.

LIMITATIONS

The IROA Workflow, although highly effective, does have limitations. The Workflow uses two standards, the QC standard (LTRS) and the Internal Standard (IS), both of which are chemically identical but isotopically different. The internal standard is derived from a water-soluble yeast extract containing a broad range of 100s of polar primary metabolites but does not contain nonpolar or species-specific metabolites. Development of additional internal standards containing more diverse metabolic mixtures will increase the applicability of the approach.

Computationally, the values seen in the LTRS are used to report instrument performance and correct for two major sources of error, instrument and sample preparation. The two sources of error are necessarily corrected sequentially. In the first step, each compound is corrected individually for ion losses anywhere from the point of introduction of the IS to the detection of the ions (by determining the loss of signal in the IS in each natural abundance/IS pair). In the second step, the samples are corrected for their physical differences in their original pre-extraction makeup, a sample-to-sample normalization (using a Dual MSTUS algorithm).

This approach uses only peaks that have both an IS and natural abundance peak, which ensures that each correction step (suppression-correction and normalization) uses paired metabolites and not artifactual data. Metabolites that are not present in the IS cannot be suppression-corrected and will still reflect suppressed signals and contain more errors and therefore are not directly comparable to metabolites that have been subjected to both ion-suppression correction and normalization. However, improvement in biological and statistical significance has been demonstrated using dual MSTUS normalization without suppression correction in various biological samples, including urine, plasma, and cancer cell lines.[19]

CONCLUSIONS

In the original implementation of the MTUS protocol,[31] artifactual data was removed in an effort to clean the data set and remove "signal noise", but the normalization value was determined in an arbitrary fashion. The IROA Workflow protocol provides a simple way to isolate only peaks of biological interest by considering only the peaks used in the analysis that are matched to a complex IS of biological origin. Additional benefits are leveraged based on the presence of the total IS signal added to each experimental sample. First, the IS provides a mechanism to correct individual compound ion losses that occur (1) in the injector, (2) due to ionization inefficiencies, or (3) due to other source-based losses (this is the basis of suppression-correction). Second, the IS provides a nonarbitrary value that is used for calculating a sample-based normalization factor unique to each sample (the total IS signal). Since each sample contains two isotopically differentiable populations of peaks (Experimental and IS), and because the IS signal suffers the same variances as the Experimental signal, i.e., suppression, injection fragmentation, etc., the protocol first corrects the raw data for all instrumentation based losses, and then the "Dual MSTUS" normalization provides a sample specific normalization factor based on the error corrected data that adjusts for sample preparation based variances. Furthermore, this error correction approach allows the adjustment to normalized values to be universally applied to each sample without losing the suppression-corrected biological variability that is inherent in the sample. In this study, we also present the benefit of calculating the normalization factor not on all compounds identified in a data set but instead only on all compounds present in every individual sample, yielding a better normalization across all batches. Unlike most normalization procedures that employ statistical batch averaging approaches or other model-based techniques, this method corrects for the sources of error in both sample preparation and instrumentation before the normalization is performed. This "core compound" normalization is not only more specific than the original method but is less sensitive to error. Most normalization procedures that employ statistical batch averaging approaches normalize by adjusting data that contain substantial sample variance without removal of the biological variances; i.e., the variance still exists in the normalized data. By correcting for technical losses (variance) first and then normalizing for sample-to-sample differences separately, data normalization with low RSDs and biological integrity can be achieved. We expect that this approach could be a valuable method for future large studies involving multiple batches and possibly even clinical trials. Also in this paper, we have demonstrated a technique that can be used to determine the retention of biological integrity post normalization. These techniques could be of great utility for other researchers.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.5c05210.

IROA conceptualization and TrueQuant workflow for IS and sample preparation (PDF)

Additional figures illustrating semitargeted workflow, all v/s core compounds-based normalization, and the statistical evaluation of different normalization methods (PDF)

R script used to perform different normalization (PDF)

AUTHOR INFORMATION

Corresponding Author

Vladimir Shulaev − *Department of Biological Sciences, College of Science, University of North Texas, Denton, Texas 76203, United States; Advanced Environmental Research Institute, University of North Texas, Denton, Texas 76203, United States;* orcid.org/0000-0001-7306-6340; Email: shulaev@unt.edu

Authors

Debasish Ghosh − *Department of Biological Sciences, College of Science, University of North Texas, Denton, Texas 76203, United States; Advanced Environmental Research Institute, University of North Texas, Denton, Texas 76203, United States*

Janmejoy Kar − *Group O, Roanoke, Texas 76262, United States*

Felice A. de Jong − *IROA Technologies LLC, Chapel Hill, North Carolina 27517, United States*

Chris Beecher − *IROA Technologies LLC, Chapel Hill, North Carolina 27517, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.5c05210

Notes

The authors declare the following competing financial interest(s): The authors D.G., J.K., and V.S. declare no competing interest. C.B. and F.d.J. are employed by IROA Technologies.

REFERENCES

(1) Misra, B. B. New software tools, databases, and resources in metabolomics: updates from 2020. *Metabolomics* **2021**, *17* (5), 49.
(2) Fiehn, O. Metabolomics - the link between genotypes and phenotypes. *Plant Molecular Biology* **2002**, *48* (1), 155−171.
(3) Roberts, L. D.; Souza, A. L.; Gerszten, R. E.; Clish, C. B. Targeted metabolomics. *Curr. Protoc Mol. Biol.* **2012**, 1−24.
(4) Shulaev, V. Metabolomics technology and bioinformatics. *Brief Bioinform* **2006**, *7* (2), 128−139.
(5) Annesley, T. M. Ion suppression in mass spectrometry. *Clin Chem.* **2003**, *49* (7), 1041−1044.
(6) Jessome, L. L.; Volmer, D. A. Ion suppression: A major concern in mass spectrometry. *LCGC North America* **2006**, *24*, 498−510.
(7) Trufelli, H.; Palma, P.; Famiglini, G.; Cappiello, A. An overview of matrix effects in liquid chromatography-mass spectrometry. *Mass Spectrom. Rev.* **2011**, *30* (3), 491−509.
(8) Ikonomou, M. G.; Blades, A. T.; Kebarle, P. Electrospray mass spectrometry of methanol and water solutions suppression of electric discharge with SF6 gas. *J. Am. Soc. Mass Spectrom.* **1991**, *2* (6), 497−505.
(9) Knochenmuss, R.; Stortelder, A.; Breuker, K.; Zenobi, R. Secondary ion−molecule reactions in matrix-assisted laser desorption/ionization. *J. Mass Spectrom* **2000**, *35* (11), 1237−1245.
(10) Taylor, A. J.; Dexter, A.; Bunch, J. Exploring Ion Suppression in Mass Spectrometry Imaging of a Heterogeneous Tissue. *Anal. Chem.* **2018**, *90* (9), 5637−5645.
(11) Souverain, S.; Rudaz, S.; Veuthey, J. L. Matrix effect in LC-ESI-MS and LC-APCI-MS with off-line and on-line extraction procedures. *J. Chromatogr A* **2004**, *1058* (1−2), 61−66.
(12) Freitas, L. G.; Gotz, C. W.; Ruff, M.; Singer, H. P.; Muller, S. R. Quantification of the new triketone herbicides, sulcotrione and mesotrione, and other important herbicides and metabolites, at the ng/l level in surface waters using liquid chromatography-tandem mass spectrometry. *J. Chromatogr A* **2004**, *1028* (2), 277−286.
(13) Vats, P.; Verma, S. M.; Khuroo, A.; Monif, T. Effects of Sample Processing Techniques and Instrument Model on Matrix Effect in LC-MS/MS for a Model Drug. *Journal of Liquid Chromatography & Related Technologies* **2015**, *38* (13), 1315−1323.
(14) Baillie, T. A. The use of stable isotopes in pharmacological research. *Pharmacol Rev.* **1981**, *33* (2), 81−132.
(15) de Leenheer, A. P.; Thienpont, L. M. Applications of isotope dilution-mass spectrometry in clinical chemistry, pharmacokinetics, and toxicology. *Mass Spectrom. Rev.* **1992**, *11* (4), 249−307.
(16) Mashego, M. R.; Wu, L.; Van Dam, J. C.; Ras, C.; Vinke, J. L.; Van Winden, W. A.; Van Gulik, W. M.; Heijnen, J. J. MIRACLE: mass isotopomer ratio analysis of U-13C-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites. *Biotechnol. Bioeng.* **2004**, *85* (6), 620−628.
(17) Wu, L.; Mashego, M. R.; van Dam, J. C.; Proell, A. M.; Vinke, J. L.; Ras, C.; van Winden, W. A.; van Gulik, W. M.; Heijnen, J. J. Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly 13C-labeled cell extracts as internal standards. *Anal. Biochem.* **2005**, *336* (2), 164−171.
(18) Hewavitharana, A. K.; Abu Kassim, N. S.; Shaw, P. N. Standard addition with internal standardisation as an alternative to using stable isotope labelled internal standards to correct for matrix effects-Comparison and validation using liquid chromatography- tandem mass spectrometric assay of vitamin D. *J. Chromatogr A* **2018**, *1553*, 101−107.
(19) Mahmud, I.; Wei, B.; Veillon, L.; Tan, L.; Martinez, S.; Tran, B.; Raskind, A.; de Jong, F.; Liu, Y.; Ding, J.; et al. Ion suppression correction and normalization for non-targeted metabolomics. *Nat. Commun.* **2025**, *16* (1), 1347.
(20) Stokvis, E.; Rosing, H.; Beijnen, J. H. Stable isotopically labeled internal standards in quantitative bioanalysis using liquid chromatography/mass spectrometry: necessity or not? *Rapid Commun. Mass Spectrom.* **2005**, *19* (3), 401−407.
(21) Wang, S.; Cyronak, M.; Yang, E. Does a stable isotopically labeled internal standard always correct analyte response? A matrix effect study on a LC/MS/MS method for the determination of carvedilol enantiomers in human plasma. *J. Pharm. Biomed Anal* **2007**, *43* (2), 701−707.
(22) Clendinen, C. S.; Stupp, G. S.; Ajredini, R.; Lee-McMullen, B.; Beecher, C.; Edison, A. S. An overview of methods using (13)C for improved compound identification in metabolomics and natural products. *Front Plant Sci.* **2015**, *6*, 611.
(23) Qiu, Y.; Moir, R.; Willis, I.; Beecher, C.; Tsai, Y. H.; Garrett, T. J.; Yost, R. A.; Kurland, I. J. Isotopic Ratio Outlier Analysis of the S. cerevisiae Metabolome Using Accurate Mass Gas Chromatography/Time-of-Flight Mass Spectrometry: A New Method for Discovery. *Anal. Chem.* **2016**, *88* (5), 2747−2754.
(24) Stupp, G. S.; Clendinen, C. S.; Ajredini, R.; Szewc, M. A.; Garrett, T.; Menger, R. F.; Yost, R. A.; Beecher, C.; Edison, A. S. Isotopic ratio outlier analysis global metabolomics of Caenorhabditis elegans. *Anal. Chem.* **2013**, *85* (24), 11858−11865.
(25) Beecher, C.; de Jong, F. A. Isotopic Ratio Outlier Analysis (IROA) for Quantitative Analysis. In *Metabolomics: Methods and Protocols*; Bhattacharya, S. K., Ed.; Springer US, 2025; pp 69−89.
(26) Lippa, K. A.; Aristizabal-Henao, J. J.; Beger, R. D.; Bowden, J. A.; Broeckling, C.; Beecher, C.; Clay Davis, W.; Dunn, W. B.; Flores, R.; Goodacre, R.; et al. Reference materials for MS-based untargeted metabolomics and lipidomics: a review by the metabolomics quality assurance and quality control consortium (mQACC). *Metabolomics* **2022**, *18* (4), 24.
(27) Han, S.; Huang, J.; Foppiano, F.; Prehn, C.; Adamski, J.; Suhre, K.; Li, Y.; Matullo, G.; Schliess, F.; Gieger, C. TIGER: technical variation elimination for metabolomics data using ensemble learning architecture. *Briefings in Bioinformatics* **2022**, *23*, No. bbab535.
(28) Muller, C.; Schillert, A.; Rothemeier, C.; Tregouet, D. A.; Proust, C.; Binder, H.; Pfeiffer, N.; Beutel, M.; Lackner, K. J.; Schnabel, R. B.; et al. Removing Batch Effects from Longitudinal Gene Expression - Quantile Normalization Plus ComBat as Best Approach for Microarray Transcriptome Data. *PLoS One* **2016**, *11* (6), No. e0156594.
(29) Guo, F.; Lin, G.; Dong, L.; Cheng, K.-K.; Deng, L.; Xu, X.; Raftery, D.; Dong, J. Concordance-Based Batch Effect Correction for Large-Scale Metabolomics. *Anal. Chem.* **2023**, *95* (18), 7220−7228.
(30) Karpievitch, Y. V.; Nikolic, S. B.; Wilson, R.; Sharman, J. E.; Edwards, L. M. Metabolomics data normalization with EigenMS. *PLoS One* **2014**, *9* (12), No. e116221.
(31) Warrack, B. M.; Hnatyshyn, S.; Ott, K.-H.; Reily, M. D.; Sanders, M.; Zhang, H.; Drexler, D. M. Normalization strategies for metabonomic analysis of urine samples. *Journal of Chromatography B* **2009**, *877* (5), 547−552.
(32) Khammar, A.; Yarahmadi, M.; Madadizadeh, F. What Is Analysis of Covariance (ANCOVA) and How to Correctly Report Its Results in Medical Research? *Iranian journal of public health* **2020**, *49* (5), 1016−1017.
(33) Fan, S.; Kind, T.; Cajka, T.; Hazen, S. L.; Tang, W. H. W.; Kaddurah-Daouk, R.; Irvin, M. R.; Arnett, D. K.; Barupal, D. K.; Fiehn, O. Systematic Error Removal Using Random Forest for Normalizing Large-Scale Untargeted Lipidomics Data. *Anal. Chem.* **2019**, *91* (5), 3590−3596.
(34) Pandey, P.; Zaman, K.; Prokai, L.; Shulaev, V. Comparative Proteomics Analysis Reveals Unique Early Signaling Response of

Saccharomyces cerevisiae to Oxidants with Different Mechanism of Action. *Int. J. Mol. Sci.* **2021**, *22* (1), 167.
(35) Spura, J.; Christian Reimer, L.; Wieloch, P.; Schreiber, K.; Buchinger, S.; Schomburg, D. A method for enzyme quenching in microbial metabolome analysis successfully applied to gram-positive and gram-negative bacteria and yeast. *Anal. Biochem.* **2009**, *394* (2), 192−201.
(36) Pang, Z.; Lu, Y.; Zhou, G.; Hui, F.; Xu, L.; Viau, C.; Spigelman, A. F.; MacDonald, P. E.; Wishart, D. S.; Li, S.; et al. MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation. *Nucleic Acids Res.* **2024**, *52* (W1), W398−W406.

https://doi.org/10.1021/acs.analchem.5c05210
Anal. Chem. 2026, 98, 10966−10977