IROA TECHNOLOGIES
Correcting Variability. Revealing Biology.
Metabolomics You Can Trust.

Error Correction as the Basis of Reproducible Metabolomics

Evidence from three studies of the IROA TruQuant workflow

Executive summary

Reproducible metabolomics requires measurements distinguish biological differences from variances in analytical response and sample input; ideally from correcting for these losses. Three complementary studies evaluate the IROA TruQuant workflow at different levels: 1) Suppression Correction Study: suppression correction across analytical conditions where suppression was the dominant source of variance; 2) Batch Correction Study: correction across sequential batches where instrumental variance was the dominant source of variance, and; 3) Round-robin Study: reproducibility across eight laboratories where lab-lab differences (SOP, instrumentation, etc.) provided the variance. Together, they support a measurement strategy that uses chemically matched carbon-13 internal standards, reference-based feature filtering, and normalization selected for the biological question. [1–3]

Suppression correction restored expected sample-input responses for detectable, isotope-paired metabolites. A six-batch yeast study reported approximately 1% residual relative standard deviation (RSD) with core-compound normalization and demonstrated preservation of treatment-related patterns. An eight-laboratory study reproduced separation of known heterogeneous mixtures and showed strong cross-lab abundance correlations for metabolites verified with authentic standards. [1–3]

Study	Question tested	Principal contribution
Suppression Correction Study : Mahmud et al. [1] Nature Communications 2025	Can matched standards correct variable suppression?	Correction of analytical and sample variability across matrices and chromatographic conditions.
Batch Correction Study: Ghosh et al. [2] Analytical Chemistry 2026	Can batch correction preserve biological information?	Correction of variability across batches where batch effect was intentionally introduced. Enhanced Dual MSTUS normalization using core compounds. Means to test normalization methods to correctly group similar samples – “data fidelity test”.
Round-robin Study : Chatelaine et al. [3] Analytical Chemistry 2026	Can independent laboratories recover consistent conclusions?	Eight-lab mixture discrimination and assessment of abundance comparability.

The correction of mass spectral data is seen here as a two-step process involving first the correction of instrumental errors followed by a second step which corrects the samples for individual differences. These three studies demonstrate these corrections result in datasets that are significantly improved over the raw datasets.

How the TruQuant measurement system works

The IROA TruQuant Workflow (Figure 1) combines two yeast-derived standards (containing hundreds of compounds) with ClusterFinder data processing. The internal standard (IROA-IS) contains hundreds of metabolites universally labeled at approximately 95% ^{13}C producing a distinctive isotopolog pattern (C13); (B'). The long-term reference standard (IROA-LTRS) is a Standard Sample that combines the same 95% C-13 IS (B') with a chemically equivalent set of 5% ^{13}C isotopolog peaks (C12), in a 1:1 mixture (D'). Their characteristic isotopolog patterns provide information on carbon count and facilitate recognition of matched natural-abundance and labeled signals. [1,4]

As a Standard Sample, repeated LTRS injections track daily instrument performance as the basis of QA/QC and instrument maintenance.

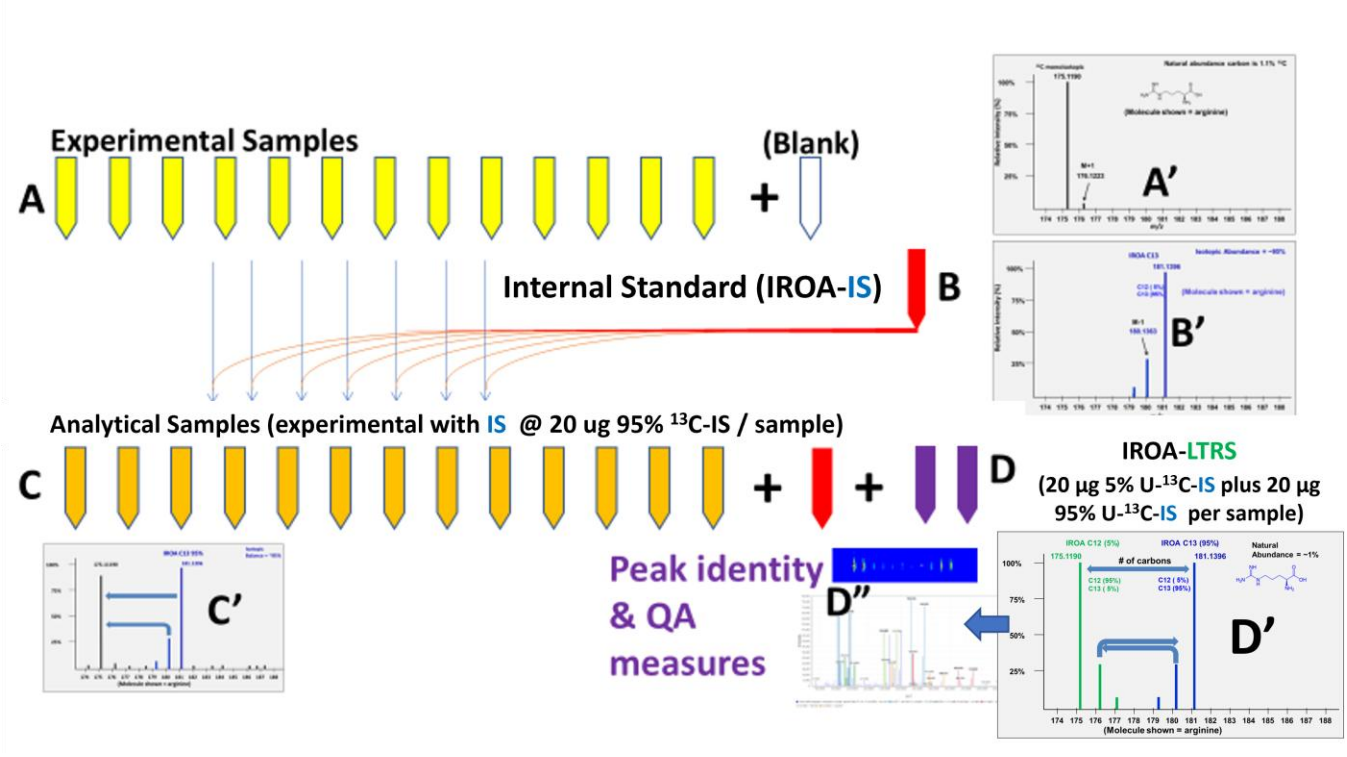


Figure 1 – The IROA Workflow

The LTRS establishes the analytical dictionary and instrument performance

From the analysis of the LTRS, ClusterFinder records the real-time retention time, mass, isotopic pattern, and fragmentation characteristics of specific reference features to build a daily RT library and as a measure of instrument performance. ClusterFinder then uses this RT library to search for corresponding features in experimental samples containing the IS. This semi-targeted approach uses the RT library for targeted search of authenticated peaks, and in addition, ClusterFinder searches for untargeted metabolites from user generated authentic standard libraries and public database entries such as PubChem. Identification confidence scores are generated based on peak attributes and conformity to referenced library standards. The detectable and usable coverage depends on the sample, extraction, chromatography, ionization, and instrument sensitivity. [1,4]

The isotope pattern strengthens formula assignment and rejects many unmatched artifacts, but a formula is not a unique structural identification. Isomers, adducts, and in-source fragments still require analytical resolution

and supporting evidence. Authentic standards provide an additional check on mass and retention time; suitable MS/MS data strengthen the assignment further. The Round-robin Study illustrates why these distinctions matter. [3]

Suppression correction acts on individual metabolites

A constant amount of IS is added to each experimental sample. Loss of IS signal provides a mechanism for correction of the loss. For a detectable analyte with a co-eluting labeled counterpart, the natural-abundance signal and IS signal experience a shared analytical environment. Their ratio therefore provides a basis for correcting signal loss. Both C12 and C13 isotopic envelopes must be summed to assure accurate quantitation. [1,3] Only the natural abundance (NA) compounds that are associated with an IS equivalent peak can be corrected for suppression.

Dual MSTUS acts on the sample as a whole

After suppression correction, Dual MSTUS (Mass Spectrometry Total Useful Signal) sums useful, paired signals in the NA and IS channels. The ratio of these sums defines a sample-specific normalization factor. As described by Mahmud et al. [1], the normalization factor (NF) is calculated as the suppression-corrected NA total divided by the suppression-corrected IS total, and all suppression-corrected analyte values are divided by NF in order to obtain their normalized values. The Batch Study demonstrates improved consistency by deriving the NF from compounds present in every sample “core compound” normalization. [2]

These steps address different sources of variation: the first is metabolite-specific; the second rescales a sample using aggregate information. It is important to note that while total useable signal is used in these correction methods, absolute quantitation is achievable when individual IS peaks are quantified. [1,4] IROA is currently determining the concentration of hundreds of compounds in the internal standard (IS) by using a complex natural abundance mixture, in which the concentration of every compound is known.

Study one establishes suppression correction and normalization

Mahmud et al., “Ion suppression correction and normalization for non-targeted metabolomics,” examines how matrix loading, chromatography, ionization polarity, and source condition affect measured metabolite signals. Plasma and urine extract aliquots of increasing size were dried and reconstituted with a constant amount of IS. The expected behavior was an increasing natural-abundance signal and a constant labeled-standard amount. Deviations from those expectations revealed suppression. [1]

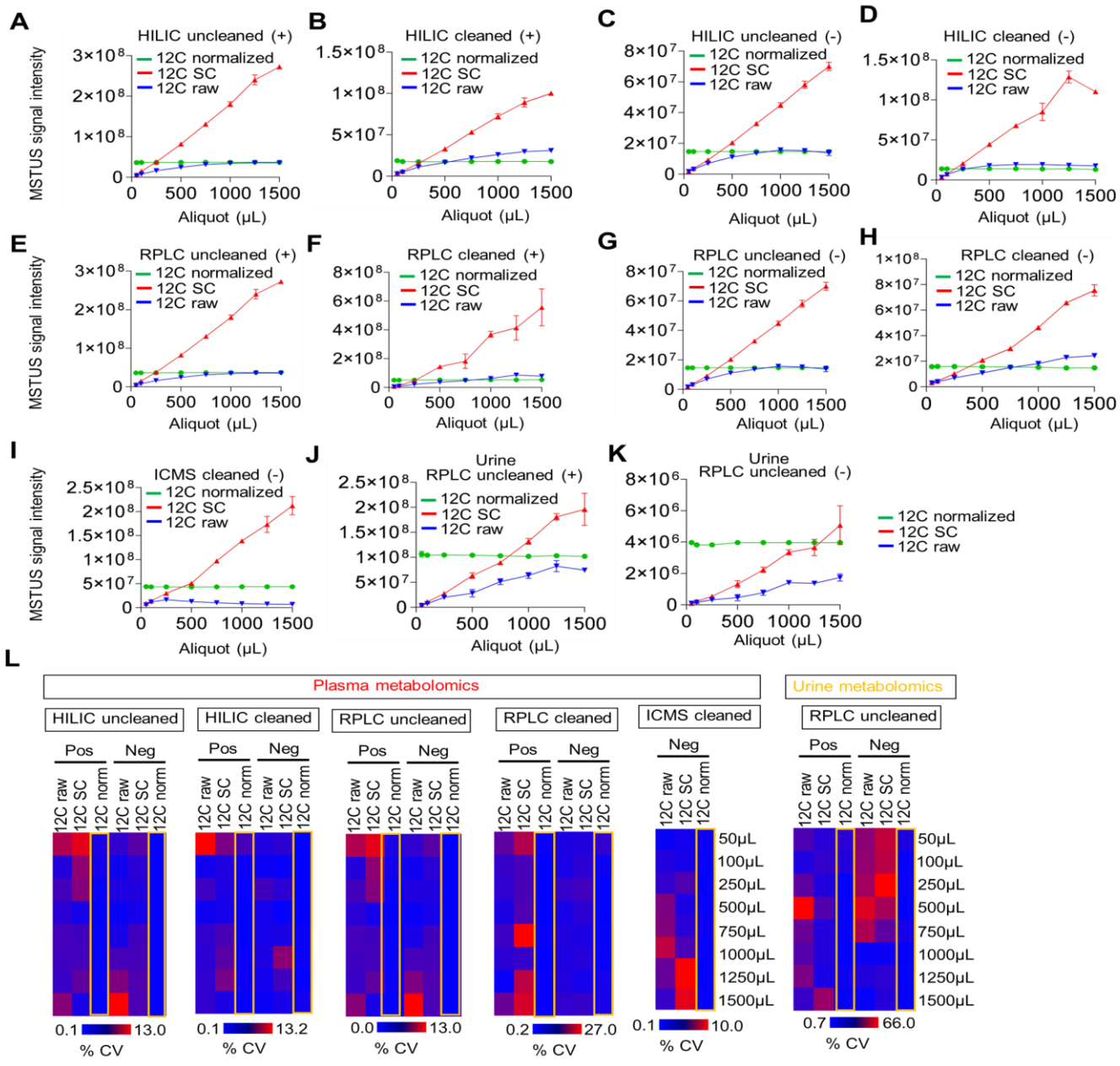


Figure 2 - Dual MSTUS normalization standardizes metabolomic data across a broad range of analytical conditions. Raw MSTUS-12C values (blue lines), normalized MSTUS-13C values (green lines), and suppression-corrected MSTUS-12C values (red lines) for indicated chromatographic systems and conditions in plasma. a, b HILIC positive mode, uncleaned and cleaned source; c, d HILIC negative mode, uncleaned and cleaned source; e, f RPLC (C18) positive mode, uncleaned and cleaned source; g, h RPLC (C18) negative mode, uncleaned and cleaned source; i ICMS negative mode, clean source. Dual MSTUS normalization for urine by RPLC uncleaned positive mode (j) and negative mode (k). l Percent coefficient of variation (%CV) for raw, suppression-corrected, and normalized data from uncleaned and cleaned source conditions across different sample matrices and chromatographic systems including IC, HILIC, and RPLC. Colors represent percent peak intensity as indicated by the color bar. Source data are provided as a Source Data file. Data are shown as mean ± s.d. (a–k).

Broad testing exposed substantial signal distortion

The study evaluated reversed-phase liquid chromatography (LC), hydrophilic interaction liquid chromatography (HILIC), and ion chromatography with mass spectrometry (MS), including positive and negative modes where applicable and clean and unclean source conditions. Suppression ranged from 1% to greater than 90% among detected metabolites under the conditions examined. These results show that suppression can vary widely and therefore peak height is often very loosely connected to absolute concentration. For example, in the Suppression Correction Study, phenylalanine showed relatively low suppression levels (8.3%) in positive-mode reversed-phase LC with a cleaned source and would therefore roughly correlate with concentration; whereas a compound such as pyroglutamylglycine which showed 97% suppression in negative-mode ion chromatography, peak height would be negatively correlated to concentration. With an average suppression level across all compounds of 40%, many raw MS peaks misrepresent their actual concentration. The IROA protocol restored the expected input-related response in these examples. [1]

Correction changed biological interpretation

The study also investigated urine and ovarian cancer cells. In ovarian cancer cell lines treated with L-asparaginase, Dual MSTUS normalization clarified changes in amino acid and peptide metabolism. In resistant OVCAR-4 cells, the apparent direction of peptide changes differed between analyses with and without the IROA protocol. The corrected findings motivated hypotheses involving peptide generation and drug resistance. [1]

What the study establishes

The work supports isotope-paired correction of inherent inaccuracies of MS analyses including variable analytical response and sample size differences across chromatographic systems, ion sources, and biological matrices. The IROA protocol corrects ion suppression across diverse analytical conditions and produces robust normalization of non-targeted metabolomic data. [1]

Study two tests correction across batches

Ghosh et al., “Achieving Effective Batch-to-Batch Error Correction through Suppression Correction and Dual MSTUS Normalization,” asks whether analytical variation across batches can be reduced while still retaining biological differences. *Saccharomyces cerevisiae* was exposed to four oxidants at five time points across six weekly batches. The experiment comprised 210 control and treatment samples, arranged as 35 samples from each of the six batches, plus 24 LTRS quality control (QC) injections, 4 per batch. [2]

The experiment deliberately introduced extreme batch effects

Batch 1 remained in the autosampler for two days. Between batches 4 and 5, the electrospray source and capillary were replaced with similar but different components, introducing a pronounced analytical discontinuity. The design challenged correction across both gradual changes and an abrupt instrument change. [2]

Core compounds improved normalization consistency

The initial “all compound” approach used paired metabolites found in each individual sample. Because the composition of that set could vary between samples, the authors then tested a “core compound” approach using compounds present in every sample to calculate normalization factors. For the reported sample-signal analysis, RSD decreased from 40% in raw data to 8% with the all-compound approach and to 1% with the core-compound approach (shown in Figure 3 below). The paper separately reports a QC metabolite median RSD of approximately 1% after IROA normalization. [2]

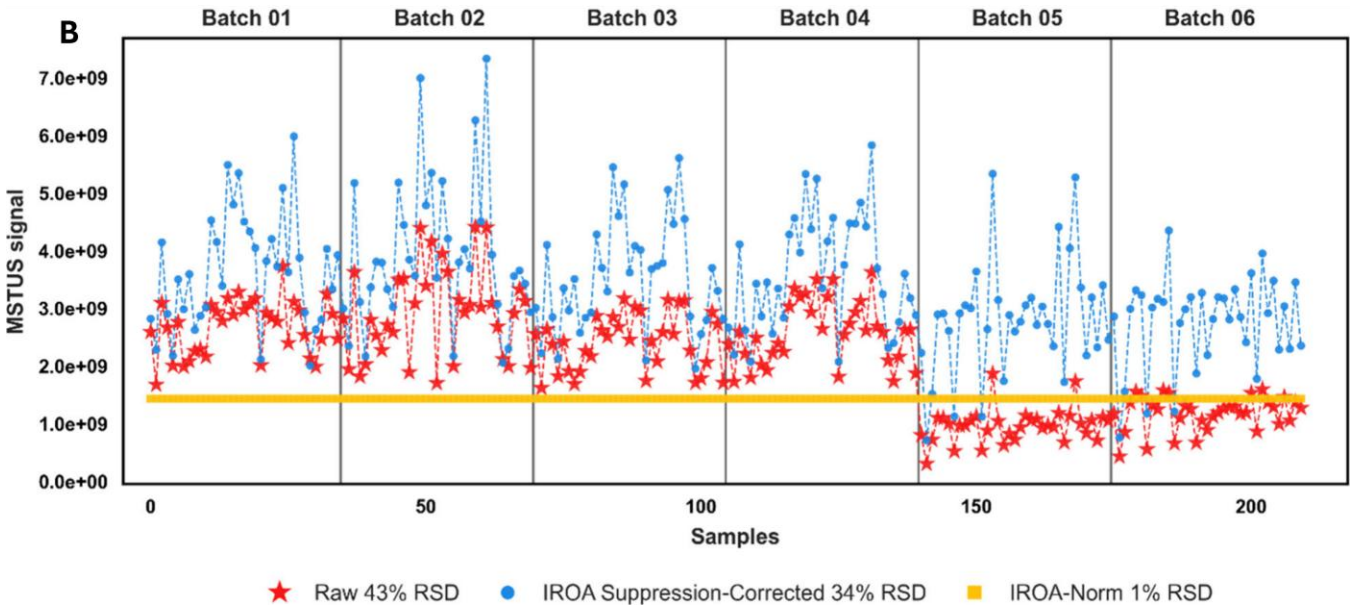


Figure 3 – Batch-batch correction

Preserving biology was part of the test

The study compared IROA normalization with TIGER, ComBat, CordBat, EigenMS, ANCOVA, and SERRF. Evaluation included QC behavior, principal component analysis (PCA), and a data-fidelity test asking whether comparable treatment and time-point samples grouped together across batches. Menadione responses at 3 and 20 minutes remained distinct after IROA normalization (Figure 4b below).

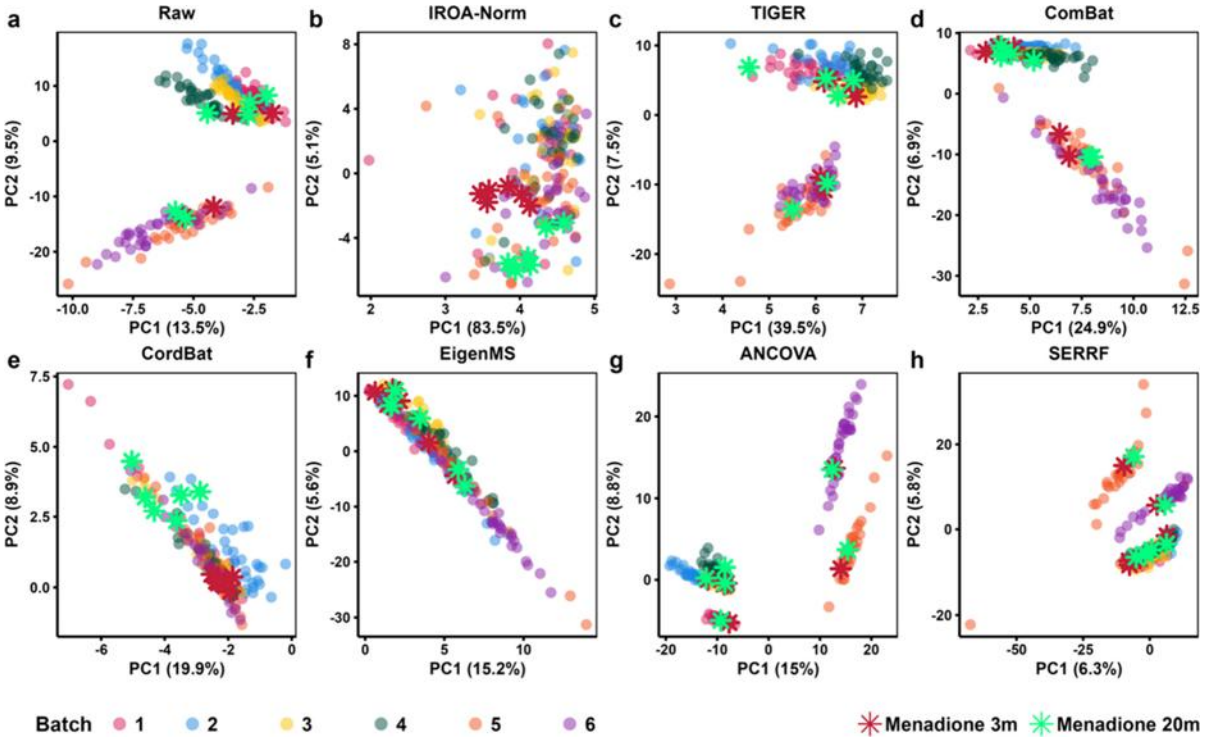


Figure 4 - PCA of batch raw data normalized by different normalization methods. PCA score plot of the raw (uncorrected) data (a) and PCA score plots of data normalized by IROA-Norm (b), TIGER (c), ComBat (d), CordBat (e), EigenMS (f), ANCOVA (g), and SERRF (h). Asterisks in the plot represent the overlay of experimental samples treated with menadione for two different time-points, 3 and 20 min (shown in cardinal red and green, respectively).

PERMANOVA analysis (p-values) of different normalization methods were conducted using either raw or suppression-corrected (SC) datasets (Figure below). PERMANOVA p-values for IROA-norm exhibited a median below 0.05, also indicating that the biological informational integrity was maintained using this method (5 c). CordBat and EigenMS showed a median near 0.1 (Figure 5 f,j) while the remaining normalization approaches had substantially higher values regardless of whether or not SC data was used (Figure 5 d,e,h,i,l,m,n,o). [2]

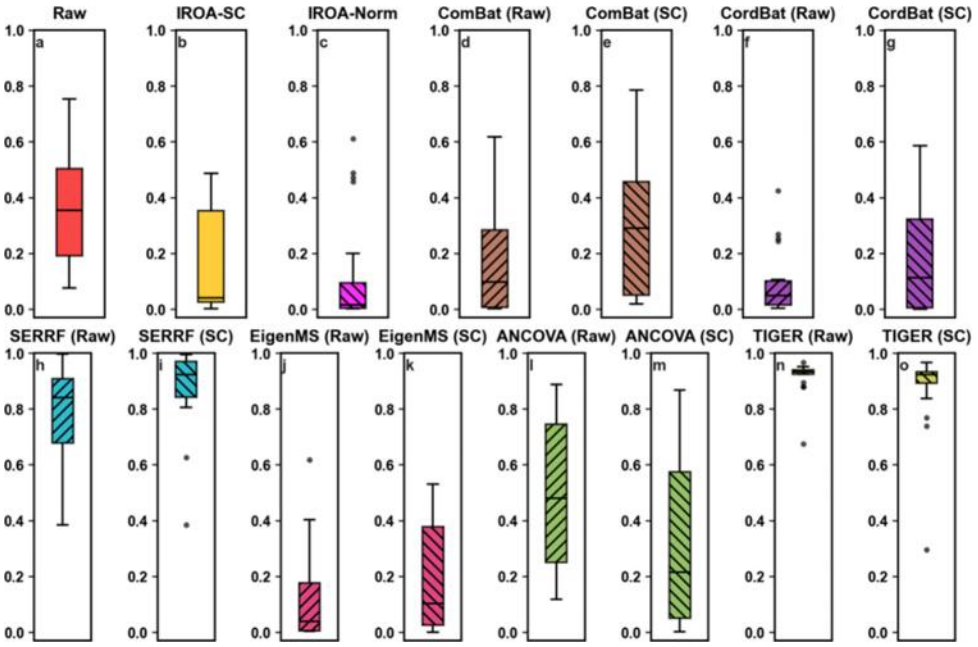


Figure 5 - Evaluation of p-value distribution of raw or suppression-corrected (SC) data across different normalization methods after performing PERMANOVA between groups (treatment vs control) across all samples. The x-axis shows the methods used to perform different normalization, and the y-axis shows 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.

The experimental design intentionally created severe batch effects using identical sample sets generated for each batch where the reproducibility of samples was maintained in order to examine post-correction data fidelity. As seen in Figure 4, post-correction fidelity is retained. [2]

What this adds to the first study

The first paper establishes the value of correcting analyte-specific response. The second adds longitudinal evidence of error correction across batches and shows why the composition of the normalization set matters. It also demonstrated that the IROA protocol preserves biological integrity by correcting variances, rather than massaging data typical of most other normalization techniques.

Study three tests reproducibility across laboratories

Chatelaine et al., “Metabolomic profiles reproducibly differentiate heterogeneous sample mixtures: a multi-lab ‘round robin’ study,” extends the evaluation to eight laboratories in the US and UK. Each received identical aliquots of 32 samples: duplicate sets of 16 preparations made from pure or mixed beef, chicken, and pork extracts at known relative amounts and at different concentrations. Sample preparation was centralized, and extracts were protein-normalized before distribution. [3]

Each laboratory used its own reversed-phase LC–MS protocols (SOPs) and analyzed both ionization polarities on their own instruments including Orbitrap and TOF/QTOF platforms from multiple manufacturers. Data conversion and processing were centralized, with laboratory-specific ClusterFinder tolerances. Thus, the experiment tested different acquisition systems within a shared sample and processing framework. [3]

Mixture structure was reproducible despite different coverage

Initial sample feature counts ranged from 560 to 1,727 across laboratories. Using suppression-corrected, normalized data, Non-negative Matrix Factorization (NMF) achieved an adjusted Rand index of 1 in seven laboratories and 0.84 in the eighth, indicating strong clustering across labs (Figure 6). The NMF results from every lab clearly defined the composition of each sample and normalized each to their effective equivalent concentration.

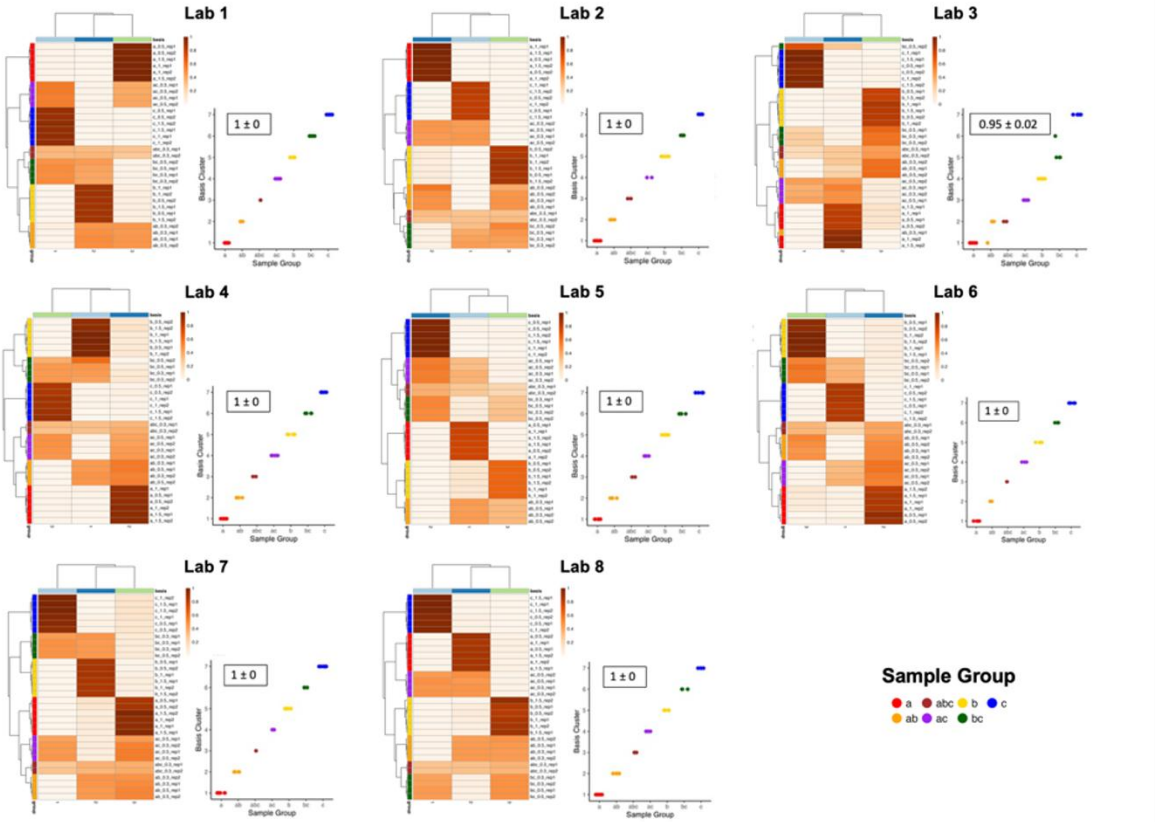


Figure 6 – Unsupervised Non-negative Matrix Factorization (NMF) demonstrates all labs correctly identified all sample types using the IROA protocol

Unsupervised random forest silhouette scores ranged from 0.33 to 0.73, with a reported mean of 0.58 ± 0.14. Pure samples generally occupied distinct extremes, with mixtures located between their component groups (Figure 7). [3]

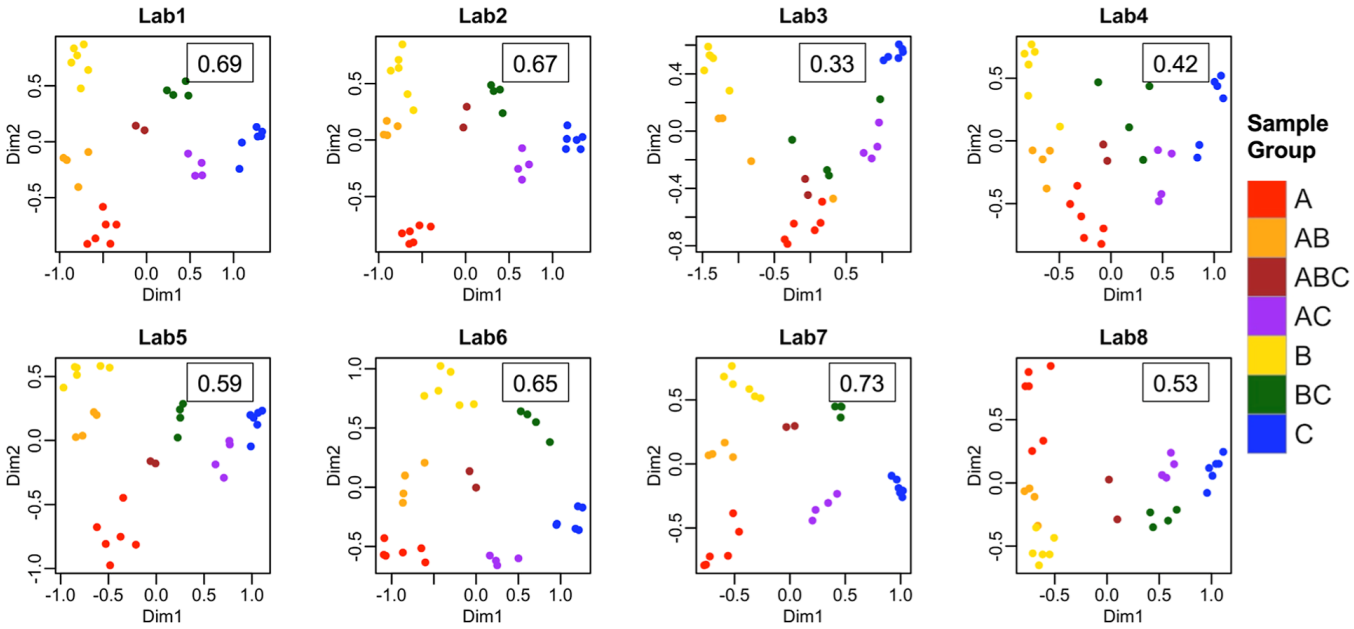


Figure 7 - Unsupervised random forest results from 8 independent laboratories demonstrate strong similarities

Identity verification strengthened abundance comparisons

Comparison using suppression-corrected data	Mean Pearson correlation
Authentic-standard-supported metabolites across all lab pairs	0.87 ± 0.12
Same supported subset excluding Lab 2	0.92 ± 0.06
31 metabolites matched by retained name in all labs	0.67 ± 0.31
350 metabolites matched by retained name in at least one lab pair	0.59 ± 0.34

The authentic-standard analysis concerned metabolites supported by standards in at least three laboratories. The broader name-matched groups are different subsets, so the table demonstrates an association with stronger identity support, not a controlled estimate of the gain from adding standards. It nevertheless highlights annotation as a major determinant of cross-lab comparability. [3]

Lab 2 had variable IS signals, consistent with unequal IS addition in some samples. Its mean carbon-13 coefficient of variation (CV) was 49%, compared with 11–23% in the other laboratories. Its mixture-clustering results remained strong, but abundance correlations were affected. This distinction shows why successful sample classification cannot substitute for evaluating standard delivery and quantitative comparability. [3]

What the three studies mean: data fidelity across time or space

The evidence supports a common principle: measurements become more interpretable when instrumental and sample variances are accounted for and corrected. The three studies reinforce each other, but their endpoints differ. [1–3]

The Suppression Correction Study demonstrated the ability to correct for ion losses, establishes the broad range of suppression values seen across all compounds, and establishes that the normalization of suppression corrected data is capable of recovering original biological signal despite significant losses from instrumentation and variances from the sample sources. This paper defines the basis of the IROA protocol.

The Batch Study demonstrates that the suppression correction and normalization when applied using the IROA protocol can correct raw data despite the introduction of severe instrumental modifications and do so on a single instrument over weeks of injections. Each batch in the Batch Study used samples that had an identical experimental design which enabled a test of “data fidelity” across batches. By initially correcting for technical losses and then normalizing sample-sample variances, the IROA protocol correctly grouped similar samples demonstrating that the workflow could successfully distinguish between technical and biological variation. Other commonly used normalization strategies tested in the study were not successful in retaining biological integrity.

The Round-robin Study demonstrates that using the same internal standard across different labs with different SOPs and instruments all labs got the same answer, they correctly identified the same sample compositions and the same biomarkers. It is clear that the ability to correct instrumental and sample-to-sample variances to a common internal standard is a mechanism to produce accurate and reproducible results.

If eight labs get the same results, would you trust the results? Of course you would.

If you were one of the eight, you wouldn’t need the other 7 results.

References and source notes

1. Mahmud I, Wei B, Veillon L, et al. Ion suppression correction and normalization for non-targeted metabolomics. Nature Communications. 2025;16:1347-61. doi:10.1038/s41467-025-56646-8.

2. Ghosh D, Kar J, de Jong FA, Beecher C, Shulaev V. Achieving Effective Batch-to-Batch Error Correction through Suppression Correction and Dual MSTUS Normalization. Analytical Chemistry. 2026;98:10966–10977. doi:10.1021/acs.analchem.5c05210.

3. Chatelaine H, Balcells C, Barnes S, et al. Metabolomic profiles reproducibly differentiate heterogeneous sample mixtures: a multi-lab “round robin” study. Analytical Chemistry. 2026. doi:10.1021/acs.analchem.6c01670

4. IROA Technologies LLC. TruQuant Yeast Extract Semi-targeted QC Workflow Kit User Manual. Cat. No. WORKFLOW. Document identifier 12192025IROAWORKFLOW. Supplied PDF. Operational background, particularly the workflow and pilot experiment sections and Appendix B.

Scope of this white paper

Prepared by IROA Technologies LLC from the three supplied published research papers and one supporting company documents. Quantitative findings are attributed to the research papers.

Disclosure

IROA Technologies develops and supplies the TruQuant standards and ClusterFinder software. IROA-affiliated authors participated in the underlying research. This white paper is a company-prepared synthesis of that evidence and its practical implications.