Li Liang (Orcid ID: 0000-0002-9347-2108)

(Manuscript submitted to *Journal of Mass Spectrometry*, Special Issue on AOMSC2020)

High Tolerance to Instrument Drifts by Differential Chemical Isotope Labeling LC-MS: A Case Study of the Effect of LC Leak in Long-Term Sample Runs on Quantitative Metabolome Analysis

(Short Title: Overcoming Instrument Drifts in LC-MS Metabolomics)

Deying Chen[1], Shuang Zhao[2], Wei Han[2], Elvis Lo[2], Xiaoling Su[1], Liang Li[1,2*] and Lanjuan Li[1*]

[1]State Key Laboratory and Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, the First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310003, China

[2]Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jms.4589

*Corresponding authors. E-mail: Liang.Li@ualberta.ca or ljli@zju.edu.cn

Abstract

Metabolomics study of a biological system often involves the analysis of many comparative samples over a period of several days or weeks. This process of long-term sample runs can encounter unexpected instrument drifts such as small leaks in LC-MS, degradation of column performance and MS signal intensity change. A robust analytical method should ideally tolerate these instrumental drifts as much as possible. In this work, we report a case study to demonstrate the high tolerance of differential chemical isotope labeling (CIL) LC-MS method for quantitative metabolome analysis. In a study of using a rat model to examine the metabolome changes during rheumatoid arthritis (RA) disease development and treatment, over 468 samples were analyzed over a period of 15 days in three batches. During the sample runs, a small leak in LC was discovered after a batch of analyses was completed. Re-analysis of these samples was not an option as sample amounts were limited. To overcome the problem caused by the small leak, we applied a method of retention time correction to the LC-MS data to align peak pairs from different runs with different degrees of leak, followed by peak ratio calculation and analysis. Herein we illustrate that using ^{12}C-/^{13}C-peak pair intensity values in CIL LC-MS as a measurement of concentration changes in different samples could tolerate the signal drifts, while using the absolute intensity values (i.e., ^{12}C-peak as in conventional LC-MS) was not as reliable. We hope that the case study illustrated

and the method of overcoming the small-leak-caused signal drifts can be helpful to others who may encounter this kind of situation in long-term sample runs.

Keywords: Chemical isotope labeling, LC-MS, retention time correction, peak intensity normalization, metabolomics

1. Introduction

Metabolome analysis is a key step in metabolomics study of a biological system such as in discovering potential metabolite biomarkers of a disease. To increase the statistical power and better account for the contributions of confounding factors, many comparative samples are collected and used for metabolome comparison. The entire process of running these samples by techniques such as liquid chromatography mass spectrometry (LC-MS) may take days or weeks to complete.[1,2] During this process, unexpected instrument-related problems can occur. A major instrument failure such as malfunction of a power supply to a component of LC or MS would interrupt the sample running process during a sample run. At most, one sample would be wasted, which may not significantly affect the outcome of the whole study, as multiple samples per group are often used and missing one sample data point may be tolerable. However, if a minor instrument drift occurs, the problem may not be noticeable during the run of a whole batch of samples. After discovering the problem, a fraction of data collected have already been affected. Re-running these affected samples may not be an option in many real world metabolomics studies, due to limited amounts of samples, limited instrumental time or other logistics issues.

There are several causes that can result in minor instrument drifts.[3,4] MS signal intensity change during a long-term run can be minimized by carrying out regular maintenance of the MS instrument, including cleaning electrospray tip and MS interface.

Moreover, using isotope-analog standards as a reference, minor MS signal drifts can be overcome.[5] For comprehensive metabolome profiling, chemical isotope labeling (CIL) can be used to produce the isotope reference standards.[6–15] For example, ^{13}C-reagent-labeled pooled sample can serve as a reference for ^{12}C-reagent-labeled individual sample.[16] The mixture of ^{12}C-labeled sample and ^{13}C-labeled pool is analyzed by LC-MS. The intensity ratio of ^{12}C- and ^{13}C-peak is used as a measure of relative concentration difference. Minor MS signal intensity drift would not affect the peak ratio values significantly.

From our experiences, the most frequently encountered problems are related to LC performance. LC column may be degraded after many sample runs, resulting in changes in chromatographic peak shapes.[17] Small leaks in LC may occur due to the raise of backpressure in LC from column blockage or due to imperfect tightening of fittings in LC and tubing connections. Small leaks can result in not only retention time shifts, but also changes in peak shapes. These changes can potentially affect the metabolome results. A robust analytical method should be able to tolerate these changes as much as possible so that the affected data can still be processed and analyzed as part of the whole metabolome dataset. In this report, we wish to document a case where we encountered a small leak in LC in a long-term metabolome analysis project. We discovered this problem after one batch of sample runs have been completed. A large fraction of these runs (144 out of a total of 468 runs in this batch) were affected. Using retention time correction, we found that the affected data could be re-stored, demonstrating the resilience of the CIL LC-MS method for minor instrument drifts. While the applied methods such as retention time correction have been documented in the literature, the novelty of this present work lies in the description of these methods within the context of dealing with small leaks during the long-term analysis of many samples, as well as illustrates the possibility of using the correction method to re-store the integrity of affected data, which may help others to address similar problems.

2. Materials and Methods

2.1. Chemicals and Reagents

All the chemicals and reagents, unless otherwise stated, were purchased from Sigma-Aldrich. The light-chain reagent, ^{12}C-dansyl chloride (Dns-Cl), was purchased from Sigma-Aldrich and the heavy-chain reagent, ^{13}C-Dns-Cl, was obtained from www.mcid.chem.ualberta.ca. LC-MS grade water, methanol, and acetonitrile were purchased from ThermoFisher Scientific.

2.2. Plasma Samples and Sample Processing

The rat model used for metabolomics study of RA has been described previously.[18] The samples collected from the rats consisted of 7 groups, including a group of normal controls (n=96), RA and other treated rats. In total, there were 468 samples which were processed and analyzed in three batches. For each batch, the analysis took 5 days.

Each sample was labeled using ^{12}C-dansylation after protein precipitation and the resultant labeled sample was analyzed by LC-UV to measure the total concentration of labeled metabolites for sample amount normalization.[18,19] Based on the total concentration, an appropriate volume of an individual unlabeled sample was taken to mix with an equal mole amount of other unlabeled samples to generate a pooled sample which was then labeled by ^{13}C-dansylation. An equal mole amount of the ^{12}C-labeled individual sample and the ^{13}C-labeled pooled sample was mixed for LC-TOF-MS analysis.

2.3. LC-MS Analysis

An Agilent 1290 UPLC system with an Waters ACQUITY UPLC BEH C_{18} column (2.1 mm × 10 cm, 1.7 μm particle size, 130 Å pore size) connected to an Agilent electrospray ionization (ESI) time-of-flight mass spectrometer (Model 6230, Agilent, Palo Alto, CA) was used for LC-MS. For LC-MS, LC solvent A was 0.1% (v/v) FA in water, and solvent B was 0.1% (v/v) FA in ACN. The gradient was: 0 min 10 %B, 0.3 min 10 %B, 1 min 20 %B, 20 min 80%B, 23 min 98%B, 25 min 100%B, 30 min 100%B, 30.5 min 10%B, 35 min 10%B. The flow rate was 300 μL/min. A quality control (QC) sample was prepared by mixing an equal mole amount of the ^{12}C-labeled and ^{13}C-labeled pooled samples and injected into LC-MS after every 10 sample runs. A retention time (RT) calibrant, which was a mixture of standards with known RTs (see Table 1), was injected into LC-MS every 10 sample runs as well.

2.4. Data Processing and Analysis

LC-MS data were first exported to .csv file with ExportMHDatafile (Agilent, Palo Alto, CA) software. The exported data were uploaded to IsoMS Pro 1.2.3 (Nova Medical Testing Inc.) for data analysis, which was performed sequentially in three modules: data processing, data cleansing and metabolite identification.[20] The workflow of data processing consists of peak pair picking from the raw files, alignment of peak pair ratios among all the samples, and zero-filling, which retrieves missing values by searching through the raw files.[21] In this step, all redundant pairs (e.g., those of adduct ions and dimers) and noise signal (shown as singlet peak) were filtered out. Only information of peak pairs, which represent metabolites, was extracted. The peak intensity ratio was also calculated for each peak pair. The data cleansing module excludes peak pairs with too many missing values. Peak pairs that present in at least 80% of samples in any group were retained. The remaining missing values were then estimated by a ratio-based imputation algorithm. An integrated table listing the

ratio values of peak pairs among samples was generated (i.e., peak ratio table). A table only listed ^{12}C-peak intensities was also generated for comparison (i.e., peak intensity table).

Retention time (RT) correction[22] and post-data-acquisition normalization[23] were performed in the software to improve the data quality during data processing. Supplemental Note S1 describes the mathematical procedures for RT correction and signal normalization. The retention times of chromatographic peaks were corrected using the nearest RT calibrants (see Results and Discussion). In this way, all experimental retention times were converted into normalized retention times, which can be used for peak pair alignment and metabolite identification.[24] In post-data-acquisition normalization, either peak ratio or peak intensity values were used for normalization (see Results and Discussion).

Metabolite identification was carried out using the three-tier metabolite identification approach.[20,25] In tier 1, peak pairs were searched against a labeled metabolite library (CIL Library) based on accurate mass and retention time. The CIL Library (i.e., dansyl amines and phenols) contains 711 experimental entries, including metabolites and dipeptides. In tier 2, linked identity library (LI Library) was used for identification of the remaining peak pairs. LI Library includes metabolic-pathway-related metabolites (more than 7,000 entries extracted from the KEGG database), providing high-confidence putative identification results based on accurate mass and predicted retention time matches. In tier 3, the remaining peak pairs were searched, based on accurate mass match, against the MyCompoundID (MCID) library composed of 8,021 known human endogenous metabolites (zero-reaction library) and their predicted metabolic products from one metabolic reaction (375,809 compounds) (one-reaction library) and two metabolic reactions (10,583,901 compounds)(two-reaction library).[26]

3. Results and Discussion

For the analysis of rat plasma samples collected from normal control, RA, sham, and four treatment groups, samples were prepared together, but analyzed by LC-MS in three batches. The first and second batches were from the normal control, RA, sham samples collected in weeks 0 to 6 and weeks 8 to 14 weeks, respectively, while the third batch was from the treatment groups collected in weeks 8 to 14. QC samples were run every 10 sample runs. In total, 468 samples and 46 QC samples were analyzed. To investigate the effect of LC leak on quantitative metabolome analysis, we focused on the metabolome datasets generated from the normal control rats in batches 1 and 2. LC leak was discovered after the first batch data acquisition was completed. The second batch was run under the normal condition without LC leak. The normal control rats were under no influence of surgery or drug treatment between batch 1 and batch 2 and thus were representative of growing rats with their metabolome affected by aging alone.

Figures 1A and 1B show the overlaid LC chromatograms obtained from all the QC runs in batch 1 and batch 2, respectively. The expanded chromatograms are shown in Figures 1C and 1D. LC leak encountered in batch 1 runs caused retention time (RT) shifts of up to 60 s for many of the chromatographic peaks (Figure 1A and Table 1). In batch 2 runs, all peaks were overlapped very well, indicating little or no RT shift (Figure 1B, Figure 1D and Table 1). In addition, the chromatographic peak intensities were also affected by the LC leak (Figure 1C), indicating that the shape of chromatographic peaks have changed.

To illustrate that the intensity changes were not uniform, Figure 2 shows a series of extracted ion chromatograms (EICs) of a representative metabolite, dansyl labeled proline, from QC runs in batch 1. The chromatographic peak area was reduced due to a decreasing amount of samples being injected into the column, indicating that the LC leak became larger

and larger as time went by. The peak was gradually shifted to a larger retention time, suggesting a reduction in flow rate caused by the leak.

Figure 3 shows the expanded mass molecular ion regions of representative mass spectra. The peak pair was from dansyl labeled proline. For the unaffected chromatograms in batch 2, the peak intensity values did not change significantly (Figures 3C and 3D). In contrast, the intensity changed from one run to another (Figures 3A and 3B), which was consistent with the result of chromatographic peak intensity changes caused by LC leak. However, the peak ratio values in Figure 3 did not change much. This is not surprising as the ^{12}C-labeled metabolite and its counterpart, ^{13}C-labeled metabolite, experienced the same degree of LC leak effect on their absolute intensity. For a specific metabolite, the ^{12}C-labeled-peak intensity from a sample and the ^{13}C-labeled-peak intensity of the same metabolite from a pool would change in absolute intensity in the same manner (i.e., either linearly reduced or nonlinearly reduced from no leak to a leak). As a result, their ratio (i.e., peak ratio value) would not change.

The above results illustrate that small leaks in LC can cause RT and intensity changes in LC-MS data. On RT shifts among different sample and QC runs, an RT correction method can be implemented using a set of RT calibrants. The calibrants are a set of standards with known RT values spaced intermittently across the entire LC elution time and are run external to the sample or QC runs. With the RT correction algorithm, experimental RTs from samples and QCs in batch 1 and batch 2 were all calibrated to the same standards. Figures 4A and 4B show the overlaid chromatograms of batch 1 and batch 2 QC runs after applying the RT calibration method, respectively. Figures 4C and 4D are the corresponding expanded chromatograms. It is clear that RT shifts caused by LC leak were effectively corrected. RT correction is critical for proper alignment of MS peaks of the same metabolites from different samples based on RT values and accurately measured masses. It should be noted that since

experimental RTs were calibrated to the RT calibrants, the RT values were changed and RT windows were stretched (i.e., from 0-24 minutes in Figure 1A-B to 0-30 minutes in Figure 4A-B). The corrected sample RTs can also be used in metabolite identification with accurate mass and RT matches as the library RTs were also normalized to the same calibrants.

However, changes in chromatographic peak intensity or shape caused by sample leak cannot be readily corrected for. These changes can have a significant effect on quantitative results in metabolome analysis. We examined how the LC and MS signal changes affected the metabolome data analysis of the normal control samples in our rat model study. We note that the main objective of metabolome analysis is to determine the metabolomic differences of samples of different groups. In the normal rats, by taking plasma samples periodically over an extended period of time, longitudinal metabolomic changes can be monitored. In our study, samples were collected every two weeks for 8 time points, forming a total of 8 longitudinal groups. In metabolome analysis of these different groups, two types of data affect the overall quality of group separation: intra-group data separation and inter-group data separation. Ideally, we want to have a very small intra-group separation so that we can reveal subtle inter-group differences caused by a biological process. The extent of intra-group separation depends on the homogeneity of the samples within the group which is determined by the intrinsic biological variations. It also depends on the technical variations in quantitative metabolome analysis. We want to use an accurate quantitative method to minimize the technical variations. In this work, we compared the datasets obtained from the 8 groups of samples to determine the effect of LC leak on metabolome analysis using two types of measured quantitative parameters. One was to use the peak ratio value (i.e., ^{12}C-labeled-peak-intensity/^{13}C-labeled-peak-intensity) determined from a peak pair. Another one was to use the absolute peak intensity value of the ^{12}C-labeled peak within a peak pair. The latter

mimicked the conventional or label-free LC-MS approach where the absolute intensity of a peak is used for relative quantification of metabolites in different samples.

Because batch 2 dataset was collected under the normal condition without LC leak, we first examined this batch to determine the extent of intra-group and inter-group separations. Figures 5A and 5B show the PCA plots of the peak ratio values of samples collected in weeks 8-14 with and without QC samples. The four groups (weeks 8, 10, 12 and 14) were separated well and the intra-group data points were clustered together. The extent of intra-group and inter-group clustering or separation within a PCA plot (Figure 5B) can be seen in the stick plot in Figure 5C. In the stick plot, each horizontal line represents the Euclidean distance between the centroid of a specific group and that of the baseline group (in this case, the 08-Week group), and each vertical stick, centered at the horizontal line, reflects the Euclidean distance between an individual sample within the group and the centroid.[27] The number in parentheses is the average value of the sample-to-centroid distance within the group. A smaller average distance indicates a better clustering. In Figure 5C, the average distance ranges from 8.53 to 13.39 for the four groups with an average of 10.48±2.08.

Peak intensity normalization after data acquisition can be helpful to reduce signal variations during the data acquisition process. For peak ratio normalization, we first calculated the normalization factor by taking the ratio between the sum of the intensities of ^{12}C-labeled peaks and the sum of the intensities of ^{13}C-labeled peaks of peak pairs found in each sample, and then divided a peak pair ratio in an individual sample by this normalization factor. Using this approach, we can correct for the imperfect 1:1 mixing of a ^{12}C-labeled sample and a ^{13}C-labeled-pool. Ideally, after LC-UV sample normalization, all the ^{12}C-labeled individual samples would have the same amount when mixing with the same ^{13}C-labeled-pool. However, LC-UV quantification error and/or aliquoting error during mixing can cause minor variation in the amount of ^{12}C-labeled individual sample in a 1:1 mixture. As

part of the standard operating procedure (SOP) for CIL LC-MS, we performed peak ratio normalization to correct for imperfect 1:1 mixing, if any. Using the normalized peak ratio values, we produced the PCA plots that are shown in Figure 5D with QC and Figure 5E without QC. The scattering stick plot corresponding to the PCA plot in Figure 5E is shown in Figure 5F. In Figure 5F, the average distance ranges from 7.75 to 11.87 for the four groups with an average of 10.59±1.93. The average intra-sample distances as a whole in Figure 5C and Figure 5F are not significantly different, indicating that peak ratio normalization did not change the intra-sample clustering.

For the QC samples shown in Figure 5A and Figure 5D, it appears that their intra-group clustering is different. After normalization, the QC samples are more tightly clustered. Because the same QC sample was used with a fixed 1:1 mixing, peak ratio normalization should not affect the QC data clustering. The observed change in QC clustering can be attributed to the changes in the distribution of the four groups of samples with respect to the QC samples. The inter-group separations can affect the QC data distribution in the PCA plot. Overall, peak ratio normalization only improves the group separation slightly, suggesting that the quality of LC-UV quantification and ^{12}C- and ^{13}C-aliquot mixing were very good in this case.

Similarly, we generated the PCA plots of the samples and QCs using the absolute peak intensity values with and without peak intensity normalization that are shown in Figure 5G-5J. The PCA plots without QCs are shown in Figure 5H and Figure 5K and their corresponding scattering stick plots are shown in Figure 5I and Figure 5L, respectively. Normalization was done by dividing the individual ^{12}C-labeled-peak intensity in a sample with the sum of the ^{12}C-labeled peak intensities in each sample. These plots show more scattered data points, compared to the plots produced from the peak ratio values. For example, in Figure 5E, individual groups were clustered together tightly. However, in Figure

5K, the data points from week 8 group and week 12 group had wider distributions, particularly in the component 1 axis. The intra-group average distance shown in Figure 5I ranges from 6.24 to 21.55 with an average of 12.49±6.47, compared to a range of 7.37 to 31.33 with an average of 17.40±10.25 in Figure 5L. Comparing inter-group separation, the distance between 14W group and 08W group is slightly increased after normalization, while the distance between 10W group and 08W group is slightly decreased. The 12W and 08W group separation does not change much. Overall, peak intensity normalization did not affect the PCA plots significantly.

Next, we examined the dataset produced from the batch 1 samples and QCs where some of the runs were collected with small LC leaks. The problem of small leak was readily detectable by examining the RT shifts of the QC samples. After RT correction, we aligned the same peak pairs from different runs with different degrees of leak. We used the RT-corrected peak pairs for statistical analysis of the batch 1 samples and QCs. The PCA plots shown in Figure 6A and Figure 6B were obtained using the peak ratio values with and without QCs. To quantify the extent of data scattering within a group and between groups, Figure 6C shows the stick plot corresponding to Figure 6B. The PCA plots and stick plot after peak ratio normalization are shown in Figure 6D-F. In Figure 6C, the average intra-sample distance of the four sample groups ranges from 7.75 to 27.07 with an average of 13.02±9.41, compared to the range of 9.88 to 17.64 with an average of 13.46±3.22 in Figure 6F. The data points in week 0 were scattered (Figure 6B). However, after peak ratio normalization, these points were more closely distributed (Figure 6E and Figure 6F). The distance between 0W group and 02W group or 04W group is similar in Figure 6C and Figure 6F, while the distance between 0W group and 06W group is larger in Figure 6C than in Figure 6F. Overall, the peak ratio normalization does not affect the group clustering

significantly, similar to the observation found in batch 2 samples. The four groups in batch 1 had similar intra-group clustering (Figure 6E) to those shown in Figure 5E from batch 2.

In comparison, using peak intensity values (Figure 6G-K), the intra-group separation of the 0W group was greater without peak intensity normalization (Figure 6H vs. Figure 6K and Figure 6I vs. Figure 6L). However, for other three groups, normalization worsens the intra-group clustering slightly. The average intra-group distance ranges from 11.43 to 27.47 with an average of 17.62±7.19 in Figure 6I, compared to the range of 13.09 to 22.70 with an average of 19.17±4.63. The inter-group separation does not change much. Overall, the group clustering was not affect significantly by signal normalization.

We combined the 8 groups from batch 1 and batch 2 to produce the PCA plots shown in Figure 7. By comparing Figure 7B produced from the normalized peak ratio values and Figure 7D produced from the normalized peak intensity values, we can clearly see that the intra-group data points were more closely clustered together if we used the peak ratio as a relative quantification parameter. The inter-group separations among the 8 groups were more clearly seen in Figure 7B, compared to these shown Figure 7D. The data scattering stick plots shown in Figure 7E and Figure 7F provide more quantitative comparisons of the intra-group and inter-group separations. These comparisons illustrate the advantage of using peak ratio over peak intensity for quantitative metabolome analysis.

It should be noted that the QC samples in batch 1 and batch 2 were clustered together in the PCA plots shown in Figure 7A and Figure 7C. The QC sample clustering was not used for detecting LC leak. As indicated earlier, in processing the batch 1 data, RT correction was first applied after discovering the LC leak problem. After RT correction, the batch 1 and batch 2 data with and without LC leak could be merged.

To further examine the performance of signal normalization on improving data quality, we calculated the relative standard deviations (RSD) of the peak pair ratios of

individual metabolites and the absolute intensities of the ^{12}C-labeled peaks detected in each group's before and after the signal normalization for the 14 weeks (see Supplemental Table S1). Figure 8 shows the box plots to plot the distribution of these RSD values. For example, 0W_Before_Ratio denotes RSDs from week 0 samples' peak ratio values before the signal normalization. The peak pair ratio data have lower RSDs with a median of 21% and an interquartile range (IQR) from 14% (first quartile) to 29% (third quartile). In comparison, the average of the medians for the absolute intensities of the ^{12}C-labeled peaks is 34% with the average first quartile value of 23% and the average third quartile value of 50%. Using the peak ratios, the RSDs were slightly reduced after normalization in most sample groups. However, on the whole, the RSDs of absolute intensities did not change much after normalization. These conclusions are consistent with those drawn from the scattering stick plot analysis of the PCA plots.

We note that, while the QC data point in PCA analysis can be used to demonstrate the analytical reproducibility, it will not tell the analytical accuracy. In this work, our focus is on how small LC leak affects the quantification of metabolites. If we use the biological replicate samples, instead of QC samples, we can tell how analytical accuracy can affect the PCA plots. The reason is that the metabolite compositions of biological samples within a group are similar, but not the same (the QC samples have the same composition). Because these samples have different metabolite compositions, each metabolite would have different matrix and ion suppression effects in LC-MS analysis. To determine the small compositional differences among the samples of the same group, an accurate method would produce similar metabolome data with small differences (i.e., these intra-group data will be clustered tightly). However, if a method is not accurate, the small differences cannot be accurately measured. With large errors in measuring each metabolite, the intra-group differences will become larger, resulting in scattering of the samples in PCA.

LC leak or any signal intensity change during data acquisition would affect the analytical accuracy. Therefore, in this work, we examined biological replicates in PCA to determine if the metabolome analysis accuracy has been compromised or not by LC leak. We have a total of 8 time points or 8 groups of biological replicates. The rats grew normally from week 0 to week 14 and thus we expect that the pattern of clustering among the biological replicates or intra-group distribution should be similar for all the 8 groups. Using the peak ratio values after RT correction for the leak effect, in combination with peak ratio normalization, we saw similar patterns of the clustering for week 0 to week 6 groups in Figure 6D and for week 8 to week 14 groups in Figure 5D. The merged data from the 8 groups shown in Figure 7B also show similar intra-group distribution patterns. This is not the case for the 8 groups of samples using peak intensity values for quantification. For example, in Figure 7D, week 10 samples are more scattered than the others.

Taken together, the results shown in Figures 7 indicate that the normalized peak ratio values after RT correction from the batch 1 samples were of high quality. These data could still be used for quantitative comparison of metabolome changes for biological studies (e.g., observing aging dependent changes of some metabolites), without the need of re-analyzing the affected samples. Thus, the CIL LC-MS method is very robust and can have high tolerance to small variations in the running conditions, including a small leak. This is important for long-term sample runs where some unexpected instrument drifts cannot be avoided.

Conclusions

We report a study of how a small LC leak during sample runs might affect the metabolome analysis results and how we could correct for the changes in order to use the affected data within a large dataset to avoid re-running the samples. A retention time

calibration method could be applied to correct the RT variations among different sample runs caused by the LC leak. The chromatographic and mass spectral peak intensity changes could be accounted for by using peak ratio values determined in differential CIL LC-MS, instead of using peak intensity. While we could not determine the limit of LC leak this method could tolerate to, the example shown in this study has had a relatively large fraction of the overall samples being affected by LC leak. In our experience of other studies while we encountered small LC leaks, we did not have these many affected samples. Frequent checking of the LC-MS running conditions to discover any problems promptly is highly recommended. However, when unexpected instrument drifts are found, the affected data should be carefully analyzed to see if any correction can be applied to re-store the integrity of the data.

Acknowledgements

This work was supported by National Natural Science Foundation of China (No. 21804118), Public Welfare Technology Application Research Plan Project of Zhejiang Province in China (LGC20B050012 and 2018C37059), Independent Project Fund of the State Key Laboratory for Diagnosis and Treatment of Infectious Disease (SKL DTID) and General Medical Science and Technology Plan Project of Zhejiang Province (No.2018KY366). This work was also supported by a visiting professorship to L. Li by Zhejiang University K. P. Chao's Hi-Tech Foundation for Scholars and Scientists. Additional support was provided by the Natural Sciences and Engineering Research Council of Canada, Canada Research Chairs, Canada Foundation for Innovations, Genome Canada and Alberta Innovates.

Conflict of Interest

The authors have no conflicts of interest in this article.

Supporting Information

Supplemental Table S1 listing RSD values for the box plots shown in Figure 8. Supplemental Note S1A describing the mathematical procedure for retention time correction. Supplemental Note S1B for the mathematical procedure for post-data acquisition normalization.

Table 1. List of retention time calibrants for correcting the RT shifts caused by LC leak.

	RTs in Batch 1* (Mean ± SD, min)	Range in Batch 1 (Max-Min, min)	RTs in Batch 2* (Mean ± SD, min)	Range in Batch 2 (Max-Min, min)
Dns-Arginine	3.01 ± 0.18	0.53	2.6 ± 0.02	0.1
Dns-Serine	3.82 ± 0.3	0.88	3.28 ± 0.02	0.08
Dns-Glutamic acid	4.04 ± 0.31	0.95	3.46 ± 0.03	0.1
Dns-Threonine	4.62 ± 0.34	1.03	3.98 ± 0.03	0.09
Dns-Glycine	4.83 ± 0.35	1.05	4.18 ± 0.03	0.09
Dns-Alanine	5.64 ± 0.37	1.14	4.93 ± 0.03	0.09
Dns-Proline	7.46 ± 0.36	1.1	6.7 ± 0.04	0.12
Dns-Methionine	7.92 ± 0.34	1.07	7.15 ± 0.03	0.09
Dns-Phenylalanine	9.34 ± 0.36	1.1	8.57 ± 0.03	0.09
Dns-Cystine	10.53 ± 0.32	0.98	9.86 ± 0.04	0.12
Dns-Lysine	12.16 ± 0.29	0.88	11.49 ± 0.02	0.09
Dns-Tyrosine	15.1 ± 0.19	0.6	14.56 ± 0.02	0.05

* Calibrants were run in batch 1 and batch 2 (n=21).

References

1. Lewis MR, Pearce JTM, Spagou K, et al. Development and Application of Ultra-Performance Liquid Chromatography-TOF MS for Precision Large Scale Urinary Metabolic Phenotyping. *Anal Chem*. 2016;88(18):9004-9013. doi:10.1021/acs.analchem.6b01481
2. Broadhurst D, Goodacre R, Reinke SN, et al. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics*. 2018;14(6). doi:10.1007/s11306-018-1367-3
3. Rodríguez-Coira J, Delgado-Dolset M, Obeso D, et al. Troubleshooting in Large-Scale LC-ToF-MS Metabolomics Analysis: Solving Complex Issues in Big Cohorts. *Metabolites*. 2019;9(11):247. doi:10.3390/metabo9110247
4. Sánchez-Illana Á, Piñeiro-Ramos JD, Sanjuan-Herráez JD, Vento M, Quintás G, Kuligowski J. Evaluation of batch effect elimination using quality control replicates in LC-MS metabolite profiling. *Anal Chim Acta*. 2018;1019:38-48. doi:10.1016/j.aca.2018.02.053
5. Wang M, Wang C, Han X. Selection of internal standards for accurate quantification of complex lipid species in biological extracts by electrospray ionization mass spectrometry—What, how and why? *Mass Spectrom Rev*. 2017;36(6):693-714. doi:10.1002/mas.21492
6. Han W, Li L. Chemical isotope labeling LC-MS for human blood metabolome analysis. In: *Methods in Molecular Biology*. Vol 1730. Humana Press Inc.; 2018:213-225. doi:10.1007/978-1-4939-7592-1_14
7. Leng JP, Wang HY, Zhang L, Zhang J, Wang H, Guo YL. A highly sensitive isotope-coded derivatization method and its application for the mass spectrometric analysis of analytes containing the carboxyl group. *Anal Chim Acta*. 2013;758:114-121. doi:10.1016/j.aca.2012.11.008
8. Tayyari F, Gowda GAN, Gu HW, Raftery D. N-15-Cholamine-A Smart Isotope Tag for Combining NMR- and MS-Based Metabolite Profiling. *Anal Chem*. 2013;85(18):8715-8721. doi:10.1021/ac401712a
9. Wong JMT, Malec PA, Mabrouk OS, Ro J, Dus M, Kennedy RT. Benzoyl chloride derivatization with liquid chromatography-mass spectrometry for targeted metabolomics of neurochemicals in biological samples. *J Chromatogr A*. 2016;1446:78-90. doi:10.1016/j.chroma.2016.04.006
10. Chu JM, Qi CB, Huang YQ, et al. Metal Oxide-Based Selective Enrichment Combined with Stable Isotope Labeling-Mass Spectrometry Analysis for Profiling of Ribose Conjugates. *Anal Chem*. 2015;87(14):7364-7372. doi:10.1021/acs.analchem.5b01614
11. Yuan W, Edwards JL, Li SW. Global profiling of carbonyl metabolites with a photo-cleavable isobaric labeling affinity tag. *Chem Commun*. 2013;49(94):11080-11082. doi:10.1039/c3cc45956j
12. Dai WD, Huang Q, Yin PY, et al. Comprehensive and Highly Sensitive Urinary Steroid Hormone Profiling Method Based on Stable Isotope-Labeling Liquid Chromatography Mass Spectrometry. *Anal Chem*. 2012;84(23):10245-10251. doi:10.1021/ac301984t
13. Yu Y, Li G, Wu D, et al. Thiol radical-based chemical isotope labelling for sterols quantitation through high performance liquid chromatography-tandem mass spectrometry analysis. *Anal Chim Acta*. 2020;1097:110-119. doi:10.1016/j.aca.2019.11.007

14. Jia S, Xu T, Huan T, et al. Chemical Isotope Labeling Exposome (CIL-EXPOSOME): One High-Throughput Platform for Human Urinary Global Exposome Characterization. *Environ Sci Technol*. 2019;53(9):5445-5453. doi:10.1021/acs.est.9b00285
15. Hao L, Zhu Y, Wei P, et al. Metandem: An online software tool for mass spectrometry-based isobaric labeling metabolomics. *Anal Chim Acta*. 2019;1088:99-106. doi:10.1016/j.aca.2019.08.046
16. Guo K, Li L. Differential 12C-/13C-isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. *Anal Chem*. 2009;81(10):3919-3932.
17. Dolan JW. Column protection: Three easy steps. *LC-GC North Am*. 2014;32(12):916-920.
18. Chen D, Su X, Wang N, et al. Chemical isotope labeling LC-MS for monitoring disease progression and treatment in animal models: Plasma metabolomics study of osteoarthritis rat model. *Sci Rep*. 2017;7. doi:10.1038/srep40543
19. Wu Y, Li L. Determination of Total Concentration of Chemically Labeled Metabolites as a Means of Metabolome Sample Normalization and Sample Loading Optimization in Mass Spectrometry-Based Metabolomics. *Anal Chem*. 2012;84(24):10723-10731. doi:10.1021/ac3025625
20. Zhao S, Li H, Han W, Chan W, Li L. Metabolomic Coverage of Chemical-Group-Submetabolome Analysis: Group Classification and 4-Channel Chemical Isotope Labeling LC-MS. *Anal Chem*. August 2019:acs.analchem.9b03431. doi:10.1021/acs.analchem.9b03431
21. Zhou R, Tseng C-L, Huan T, Li L. IsoMS: Automated Processing of LC-MS Data Generated by a Chemical Isotope Labeling Metabolomics Platform. *Anal Chem*. 2014;86(10):4675-4679. doi:10.1021/ac5009089
22. Li Y, Li L. Retention Time Shift Analysis and Correction in Chemical Isotope Labeling LC/MS for Metabolome Analysis. *Rapid Commun Mass Spectrom*. November 2019. doi:10.1002/rcm.8643
23. Wu Y, Li L. Sample normalization methods in quantitative metabolomics. *J Chromatogr A*. 2015;1430. doi:10.1016/j.chroma.2015.12.007
24. Huan T, Wu Y, Tang C, Lin G, Li L. DnsID in MyCompoundID for rapid identification of dansylated amine- and phenol-containing metabolites in LC-MS-based metabolomics. *Anal Chem*. 2015;87(19):9838-9845. doi:10.1021/acs.analchem.5b02282
25. Dahabiyeh LA, Malkawi AK, Wang X, et al. Dexamethasone-Induced Perturbations in Tissue Metabolomics Revealed by Chemical Isotope Labeling LC-MS analysis. *Metabolites*. 2020;10(2):42. doi:10.3390/metabo10020042
26. Li L, Li R, Zhou J, et al. MyCompoundID: Using an Evidence-Based Metabolome Library for Metabolite Identification. *Anal Chem*. 2013;85(6):3401-3408. doi:10.1021/ac400099b
27. Goodpaster AM, Kennedy MA. Quantification and statistical significance analysis of group separation in NMR-based metabonomics studies. *Chemom Intell Lab Syst*. 2011;109(2):162-170. doi:10.1016/j.chemolab.2011.08.009

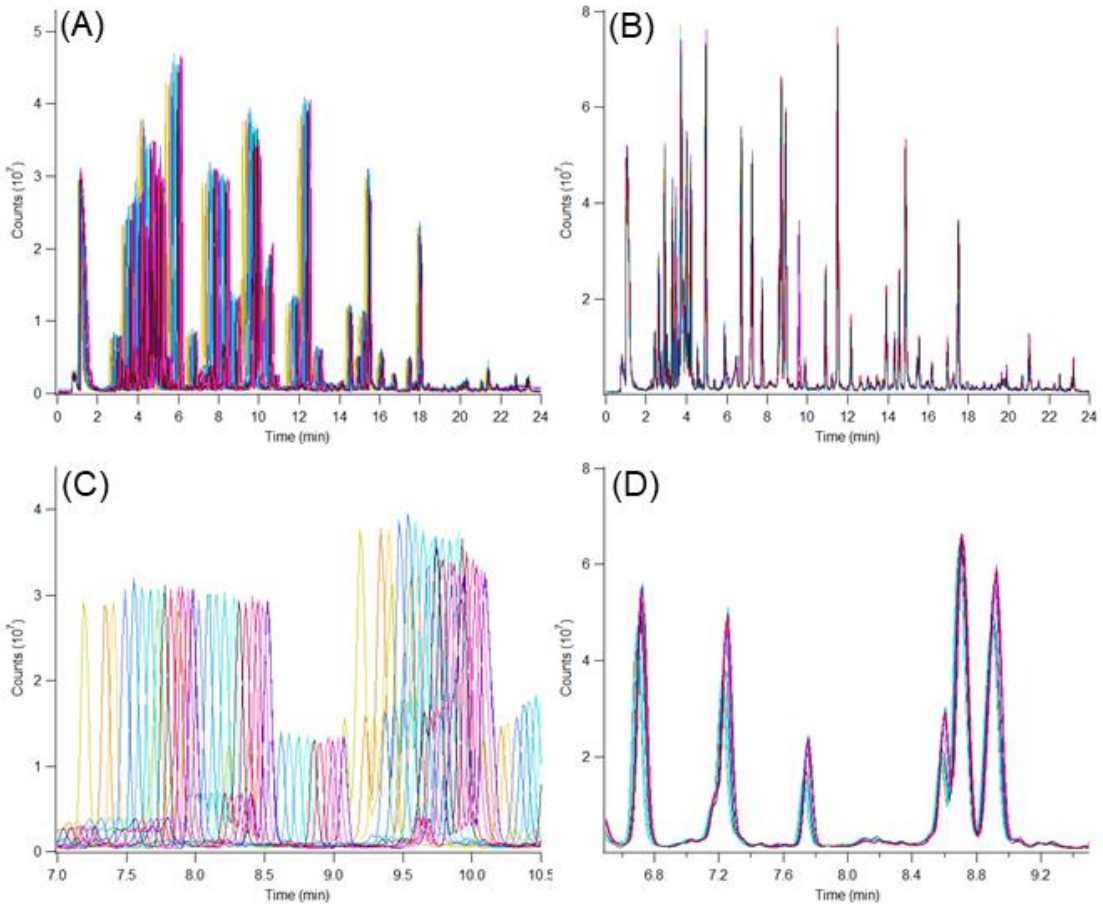

Figure 1. Total ion chromatograms of (A) 15 quality control (QC) samples run in batch 1 (144 rat plasma samples) and (B) 15 QC samples run in batch 2 (144 rat plasma samples). (C) Expanded chromatograms of (A) showing the retention time shifts as well as peak intensity changes. (D) Expanded chromatograms of (B) showing the well overlapped chromatographic peaks.

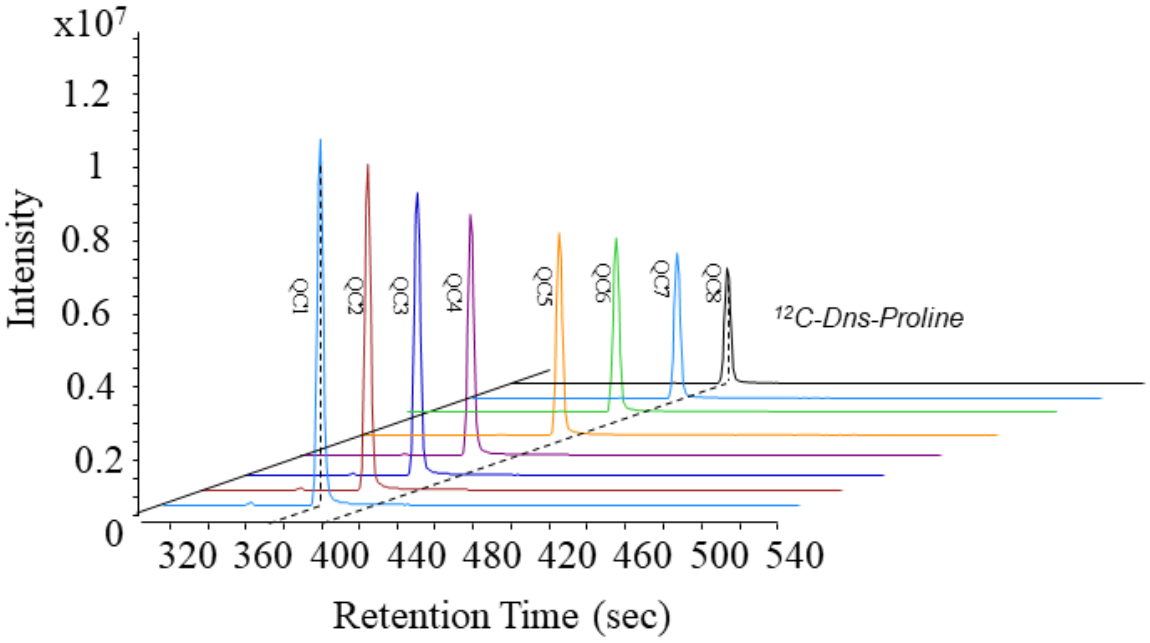

Figure 2. Extracted ion chromatograms of dansyl labeled proline detected in QC sample runs in batch 1.

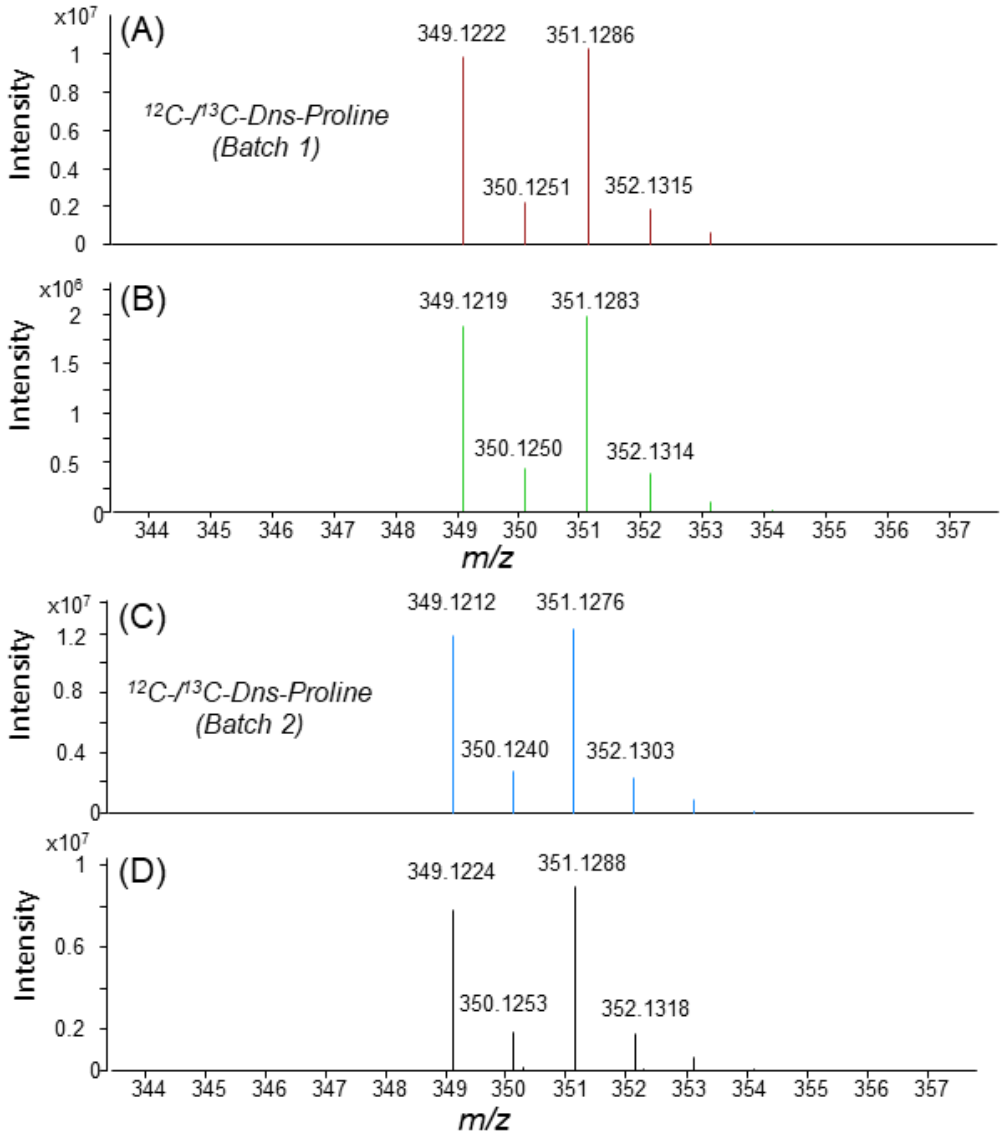

Figure 3. Expanded mass spectra of dansyl labeled proline detected from (A) a QC sample in batch 1, (B) a different QC sample in batch 1, (C) a QC sample in batch 2, and (D) a different QC sample in batch 2.

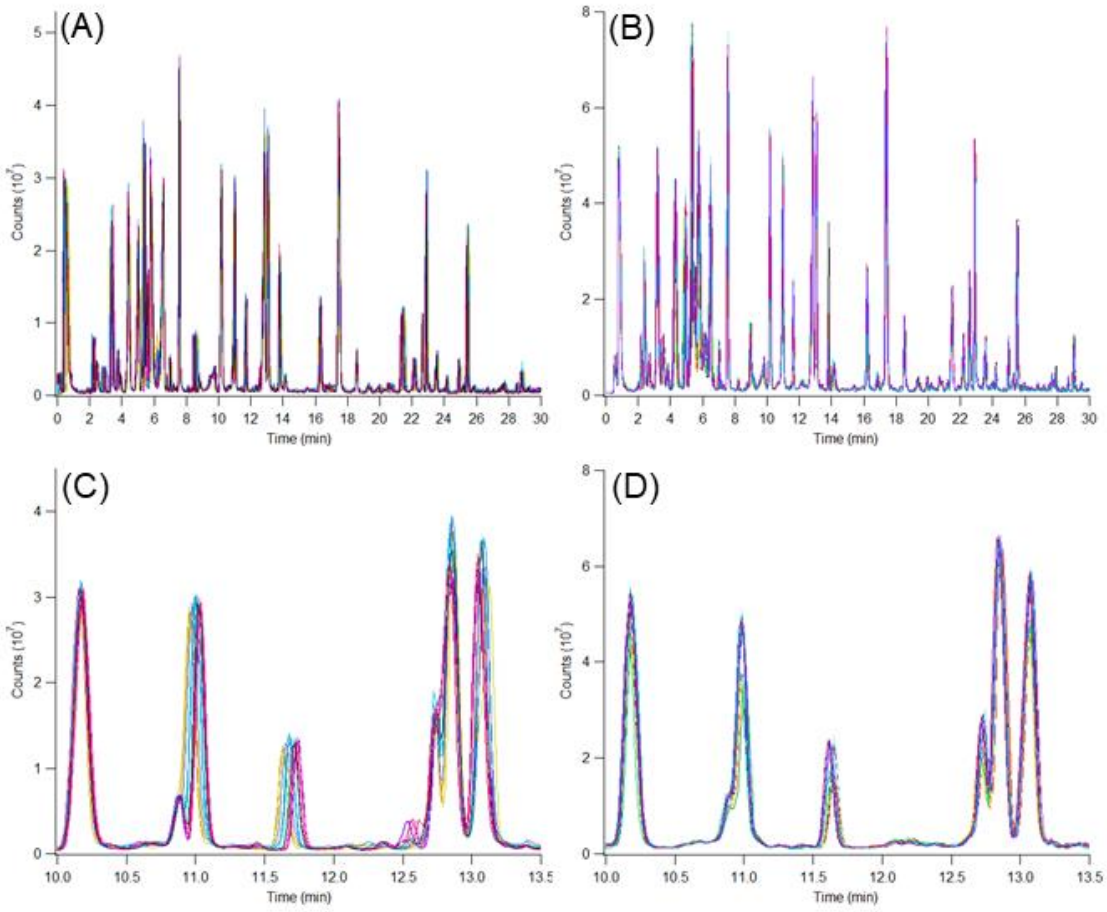

Figure 4. Total ion chromatograms of (A) the same 15 QC samples run in batch 1 and (B) 15 QC samples run in batch 2 after retention time correction. (C) Expanded chromatograms of (A) showing the well overlapped chromatographic peaks after RT correction. (D) Expanded chromatograms of (B).

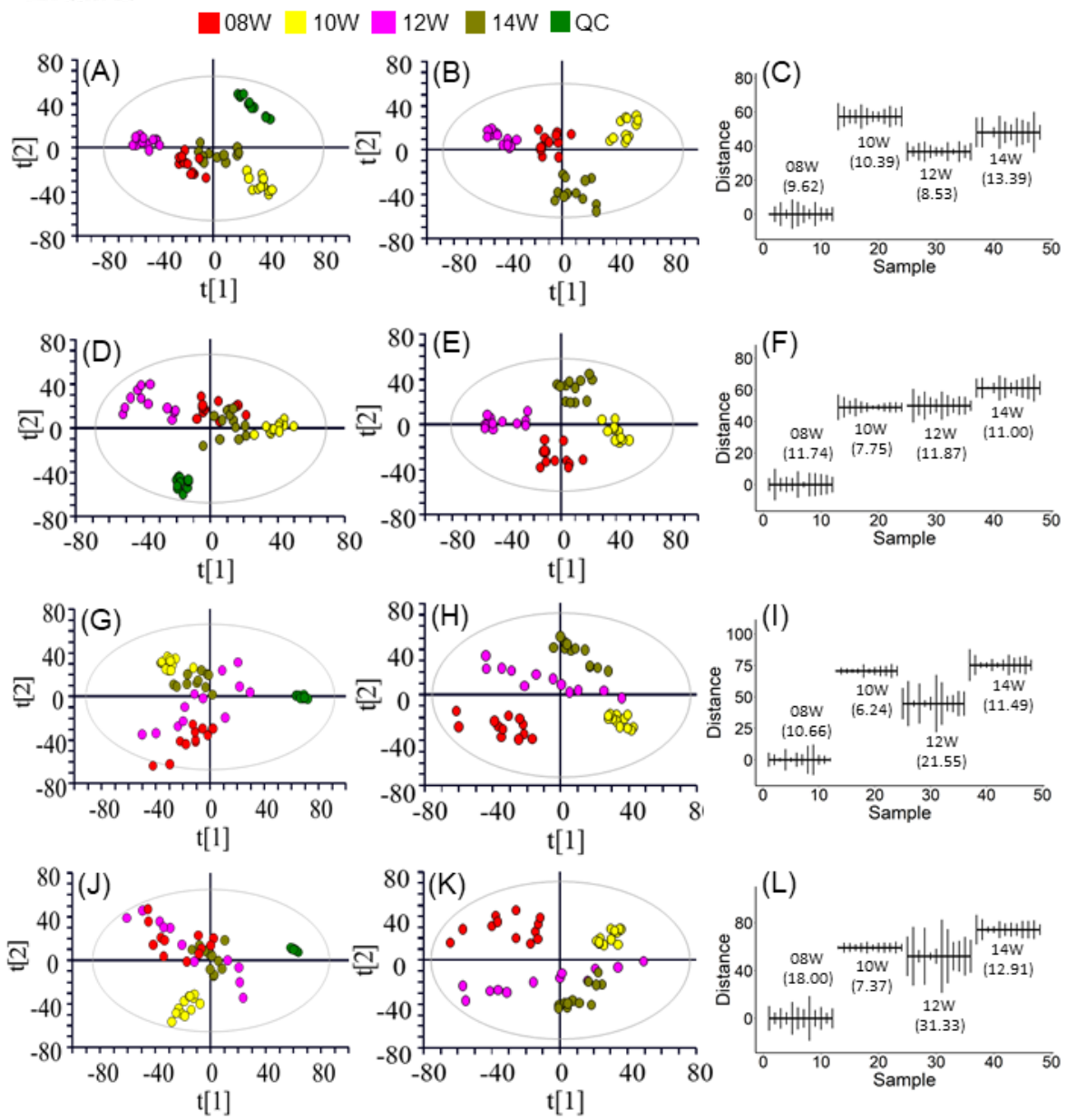

Figure 5. PCA plots and scattering stick plots of the metabolome dataset generated from CIL LC-MS runs of the normal control rat plasma samples and their QC samples in batch 2. Metabolome data were analyzed using peak ratio values without peak ratio normalization for (A) samples plus QCs and (B) samples only. (C) Scattering stick plot for (B). Metabolome data were analyzing using peak ratio values with peak ratio normalization for (D) samples plus QCs and (E) samples only. (F) Scattering stick plot for (E). Metabolome data were analyzed using peak intensity without normalization for (G) samples plus QCs and (H) samples only. (I) Scattering stick plot for (H). Metabolome data were analyzed using peak intensity with normalization for (J) samples plus QCs and (K) samples only. (L) Scattering stick plot for (K).

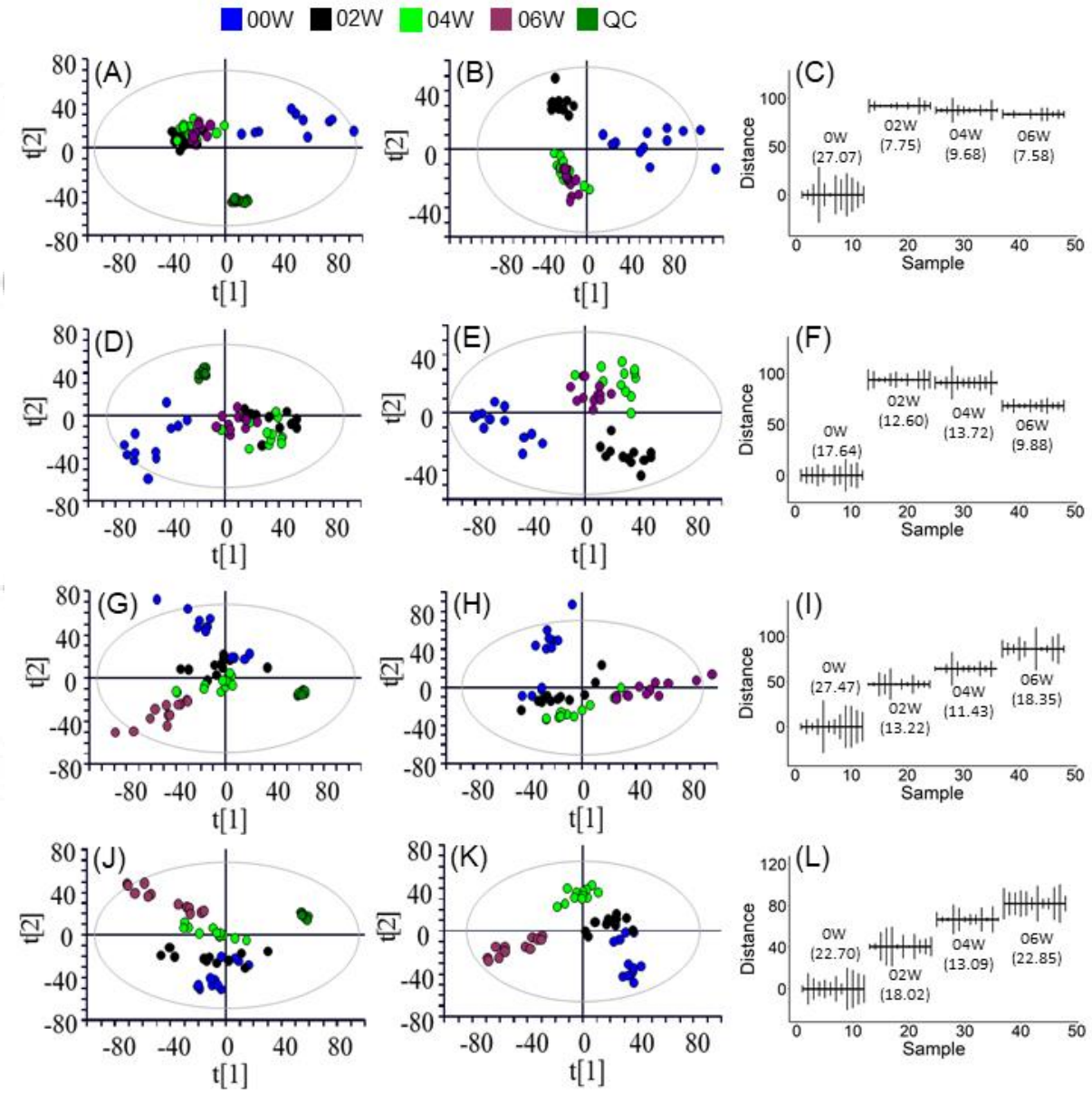

Figure 6. PCA plots and scattering stick plots of the metabolome dataset generated from CIL LC-MS runs of the normal control rat plasma samples and their QC samples in batch 1. Metabolome data were analyzed using peak ratio values without peak ratio normalization for (A) samples plus QCs and (B) samples only. (C) Scattering stick plot for (B). Metabolome data were analyzing using peak ratio values with peak ratio normalization for (D) samples plus QCs and (E) samples only. (F) Scattering stick plot for (E). Metabolome data were analyzed using peak intensity without normalization for (G) samples plus QCs and (H) samples only. (I) Scattering stick plot for (H). Metabolome data were analyzed using peak intensity with normalization for (J) samples plus QCs and (K) samples only. (L) Scattering stick plot for (K).

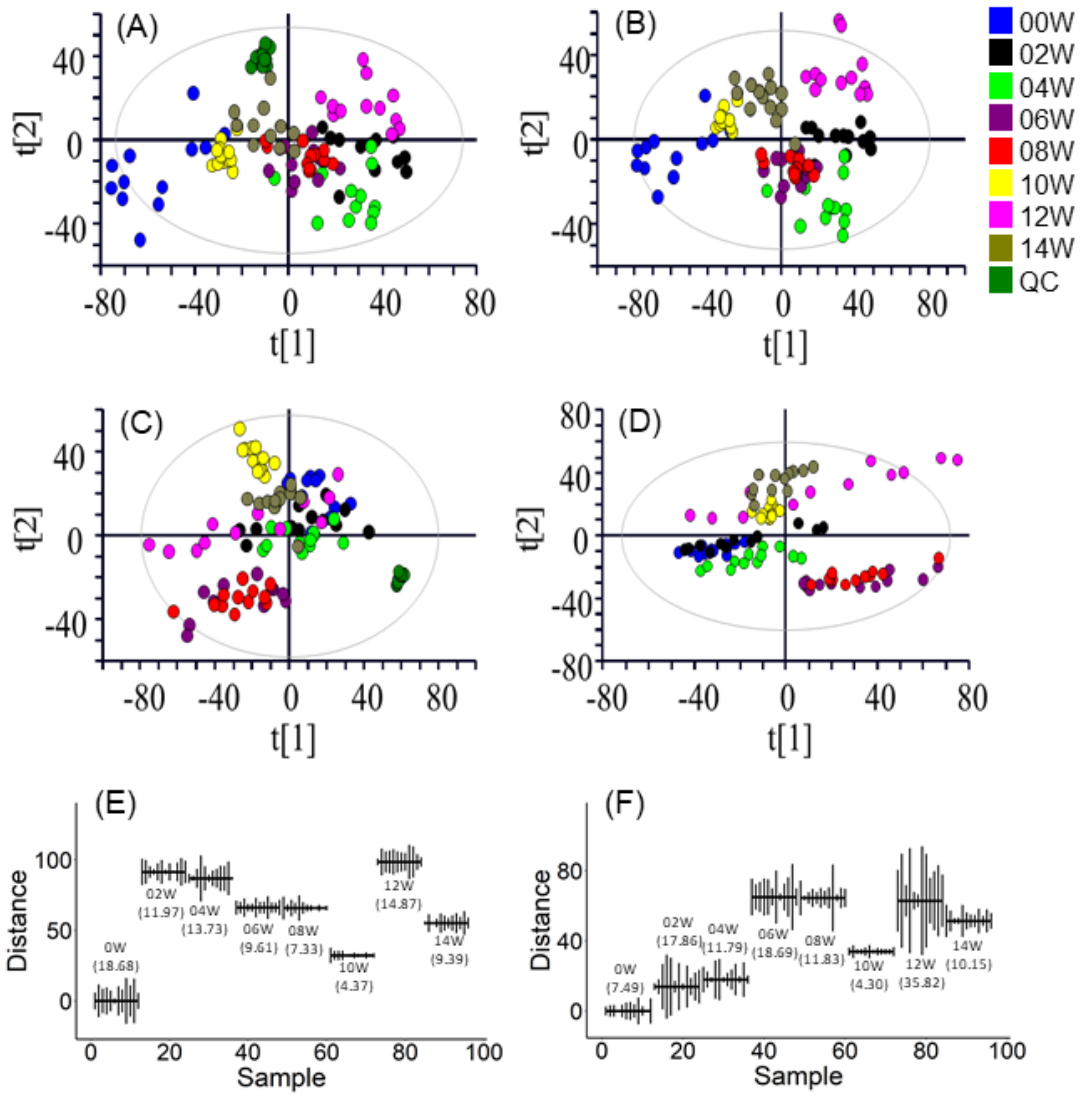

Figure 7. PCA plots and scattering plots of the combined metabolome dataset generated from CIL LC-MS runs of the normal control rat plasma samples and their QC samples in batch 1 and batch 2. Metabolome data were analyzing using peak ratio values with peak ratio normalization for (A) samples plus QCs and (B) samples only. Metabolome data were analyzed using peak intensity with normalization for (C) samples plus QCs and (D) samples only. (E) Scattering stick plot for (B). (F) Scattering stick plot for (D).

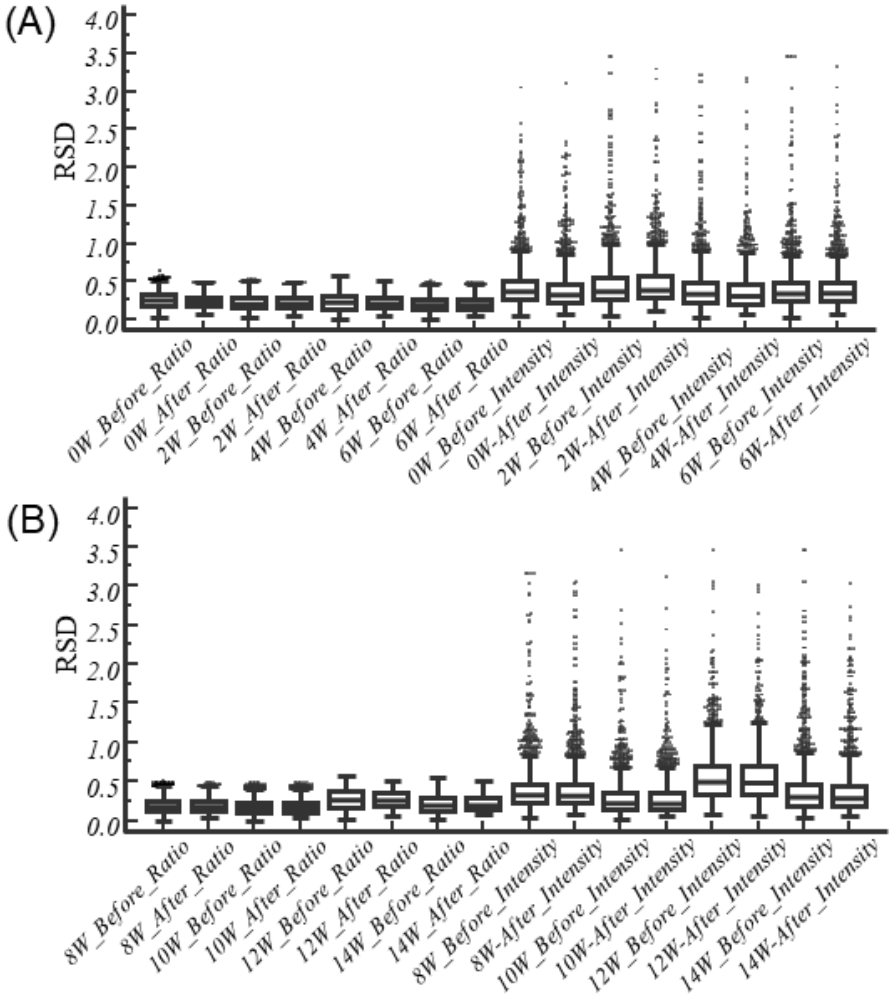

Figure 8. Boxplots of RSD distributions. For the y-axis label, 0W means week 0 samples; before and after mean before and after signal normalization; ratio means using peak ratios for quantification; intensity means using peak intensity for quantification.