ACS Partner Journal

Journal of the American Society for Mass Spectrometry

pubs.acs.org/jasms

Research Article

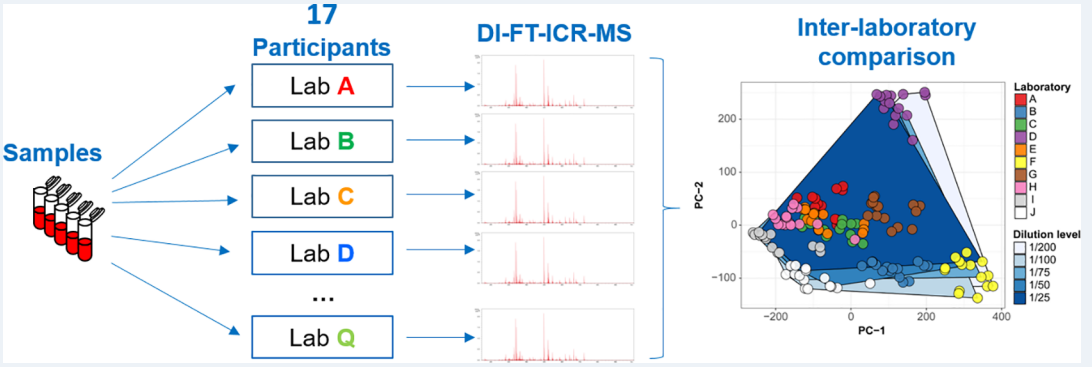
Large-Scale Interlaboratory DI-FT-ICR MS Comparability Study Employing Various Systems

Sara Forcisi,* Franco Moritz,* Christopher J. Thompson,* Basem Kanawati, Jenny Uhl, Carlos Afonso, Chantal D. Bader, Aiko Barsch, Berin A. Boughton, Rosalie K. Chu, Justine Ferey, Francisco Fernandez-Lima, Céline Guéguen, Dimitri Heintz, Mario Gomez-Hernandez, Kyoung-Soon Jang, Nikolas Kessler, Vaughn Mangal, Rolf Müller, Ryo Nakabayashi, Edith Nicol, Simone Nicolardi, Magnus Palmblad, Ljiljana Paša-Tolić, Jacob Porter, Isabelle Schmitz-Afonso, Jong Bok Seo, Eduardo Sommella, Yuri E. M. van der Burgt, Claire Villette, Matthias Witt, Ashley Wittrig, Jeremy J. Wolff, Michael L. Easterling, Frank H. Laukien, and Philippe Schmitt-Kopplin*

Cite This: *J. Am. Soc. Mass Spectrom.* 2022, 33, 2203−2214

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Ultrahigh resolution mass spectrometry (UHR-MS) coupled with direct infusion (DI) electrospray ionization offers a fast solution for accurate untargeted profiling. Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometers have been shown to produce a wealth of insights into complex chemical systems because they enable unambiguous molecular formula assignment even if the vast majority of signals is of unknown identity. Interlaboratory comparisons are required to apply this type of instrumentation in quality control (for food industry or pharmaceuticals), large-scale environmental studies, or clinical diagnostics. Extended comparisons employing different FT-ICR MS instruments with qualitative direct infusion analysis are scarce since the majority of detected compounds cannot be quantified. The extent to which observations can be reproduced by different laboratories remains unknown. We set up a preliminary study which encompassed a set of 17 laboratories around the globe, diverse in instrumental characteristics and applications, to analyze the same sets of extracts from commercially available standard human blood plasma and Standard Reference Material (SRM) for blood plasma (SRM1950), which were delivered at different dilutions or spiked with different concentrations of pesticides. The aim of this study was to assess the extent to which the outputs of differently tuned FT-ICR mass spectrometers, with different technical specifications, are comparable for setting the frames of a future DI-FT-ICR MS ring trial. We concluded that a cluster of five laboratories, with diverse instrumental characteristics, showed comparable and representative performance across all experiments, setting a reference to be used in a future ring trial on blood plasma.

INTRODUCTION

The study of complex chemical systems requires instrumentation that captures their chemical space.[1] Mass spectrometry is probably the most versatile among the techniques at the disposal of an analytical chemist. It provides high sensitivity and a multitude of means for molecular characterization.

There are two poles toward chemical characterization in mass spectrometry. At one pole there is quantitative hyphenation of separation techniques to tandem mass spectrometry, which provides structural information on the most abundant ions produced from a sample. The chemical information on the less

Received: March 18, 2022
Revised: August 12, 2022
Accepted: August 23, 2022
Published: November 13, 2022

https://doi.org/10.1021/jasms.2c00082
J. Am. Soc. Mass Spectrom. 2022, 33, 2203−2214

abundant proportion of ions is lost oftentimes as these signals do not produce abundances of fragment ions that suffice for annotation. On the other end, fast profiling using direct infusion electrospray ionization with ultrahigh resolution mass spectrometry (DI-ESI-UHR) allows for qualitative and nonquantitative analyses and unambiguous assignment of molecular formulas to those MS signals that escape annotation by tandem MS.

Fourier transform mass spectrometers such as FT-ICR MS and Orbitrap offer highly resolved, accurate, and precise determination of mass-to-charge ratios (m/z), state of the art prerequisites for the characterization of complex mixtures.[2] FT mass spectrometers, in general, have a capacity for very sensitive characterizations of MS features, as analyte ions can be trapped for prolonged periods of time. Within the family of FT instruments, FT ion cyclotron resonance (FT-ICR) MS enables the detection of thousands of peaks in complex matrices at a time, with lower parts-per-billion mass accuracy and even a resolution as high as 2,400,000 at 400 m/z (21T instrumentation).[3] Its power to characterize the compositional space of a multitude of complex chemical systems is well acknowledged within disciplines such as metabolomics, petroleomics, foodomics, lipidomics, microbiome analysis, natural organic matter (NOM), and dissolved organic matter (DOM) analysis.[3−8]

Several works describe the application of DI-MS in answering specific questions by looking at the whole MS profile, eventually integrating orthogonal techniques such as LC-MS or GC-MS for deeper isomeric elucidation.[8−11] However, broad applications require strict interlaboratory comparability.[12,13] To date, interlaboratory comparability in untargeted metabolomics was mainly tested considering LC-MS or GC-MS, DI-stable isotope dilution MS, and NMR techniques.[14−17] Each laboratory employing FT-ICR MS has an interest to achieve results that are the most representative for the analytical matrix under inspection (IUPAC project 2016-015-2-600[15,14]). To observe the same distinctive features, other laboratories must be able to observe identical and unique analytical patterns, as the same sample set is analyzed.

FT-ICR MS data were investigated by Kirwan et al.[16] to study the experimental reproducibility of a large multibatch metabolomics study of mammalian cardiac tissue extracts acquired by means of nanoinfusion FT-ICR MS in one laboratory. They developed a batch correction algorithm based on cubic spline interpolation across quality control (QC) samples.[17] Intralab reproducibility was also examined in NOM[14,18] and DOM[19] investigations. Although it was proven that batch effects and systematic errors within DI-MS data can be controlled using appropriate study designs on one instrument,[16] other resources indicate that the overlap of detected signals between two laboratories can be lower than 25%.[20] Assessments on the comparability of untargeted DI-MS data produced by different laboratories are under-represented in the present literature.

How can data generated by different laboratories be compared when they used nonquantitative DI-FT-ICR MS?

We set up a preliminary study which encompassed a set of laboratories with a high variability in terms of instrumental characteristics and routine applications. The same sample set and a standard operating procedure (see Supporting Information) were sent to 17 different laboratories worldwide, which have expertise in diverse application areas. The aim of this study was to assess the extent to which the outputs of differently tuned FT-ICR mass spectrometers, with different technical specifications, are comparable for setting the frames of a future DI-MS ring trial.

Three experiments were set up to evaluate typical effects observed when acquiring DI-ESI-FT-ICR mass spectra of standard human blood plasma SPE eluates.

Experiment X is intended to capture how different instruments data (generated by FT-ICR MS of different build) are comparable as matrix effects vary due to variation of whole matrix concentration. We found that matrix effects can generally be reproduced across different laboratories and that appropriate data normalization can produce interlaboratory coefficients of variation below 20% for those laboratories that co-detected the most signals.

Experiment Y simulates strong concentration changes in specified sets of analytes at constant matrix concentration. This is particularly interesting, as it is not clear how many signals native to an analytical matrix are sacrificed due to suppression with internal standards. We found that at least 48% of blood plasma signals get repressed by pesticide spiking and that the effects were very comparable between laboratories with coefficients of variation below 20% across laboratories.

Experiment Z intended to capture how many FT-ICR MS scans need to be accumulated in order to provide a sufficient description of the analytical matrix. This point is of high relevance for routine analysis (e.g., clinical cohort analysis[21]) and planning of future ring trials. The National Institute of Standards and Technology (NIST) Standard Reference Material (SRM) for blood plasma (SRM 1950)[22] was used for this experiment as several hundreds of individual metabolites are quantified and certified for this reference matrix. We found that the accumulation of merely 50 scans, which is performed within 60−90 s, provides 75% of the information acquired after 300 scans with respect to those metabolites certified for this reference material.

A cluster of five laboratories, with diverse instrumental characteristics, achieved the highest comparability across all experiments. Participants of a future ring trial on blood plasma profiling could tune their instruments against the median spectral intensities produced by this cluster of laboratories within this preliminary study.

EXPERIMENTAL SECTION

Aim of Study. We developed a study design that could be completed by a single person within 8 working hours, at any partner laboratory.

C18-SPE eluates of standard human plasma were prepared at different dilution levels (experiment X) to capture the matrix effect, a major concern in DI-MS analysis. A second batch of the same C18-SPE standard human plasma eluate was spiked with different concentrations of pesticides (experiment Y) to compare both matrix effects derived from internal standards and the upper bounds of dynamic ranges. C18-SPE eluate of a second human plasma standard recognized as reference material (SRM 1950)[22] was acquired at different scan numbers. Concentration levels of several hundreds of metabolites are available for SRM 1950. Following the presence of MS signals potentially related to the set of quantified compounds over the course of different scanning times was scheduled to obtain a rough estimate of the minimum number of scans required to produce meaningful data.

Study Design. Seventeen partner laboratories across four continents were recruited for the comparability study (the

https://doi.org/10.1021/jasms.2c00082
J. Am. Soc. Mass Spectrom. 2022, 33, 2203−2214

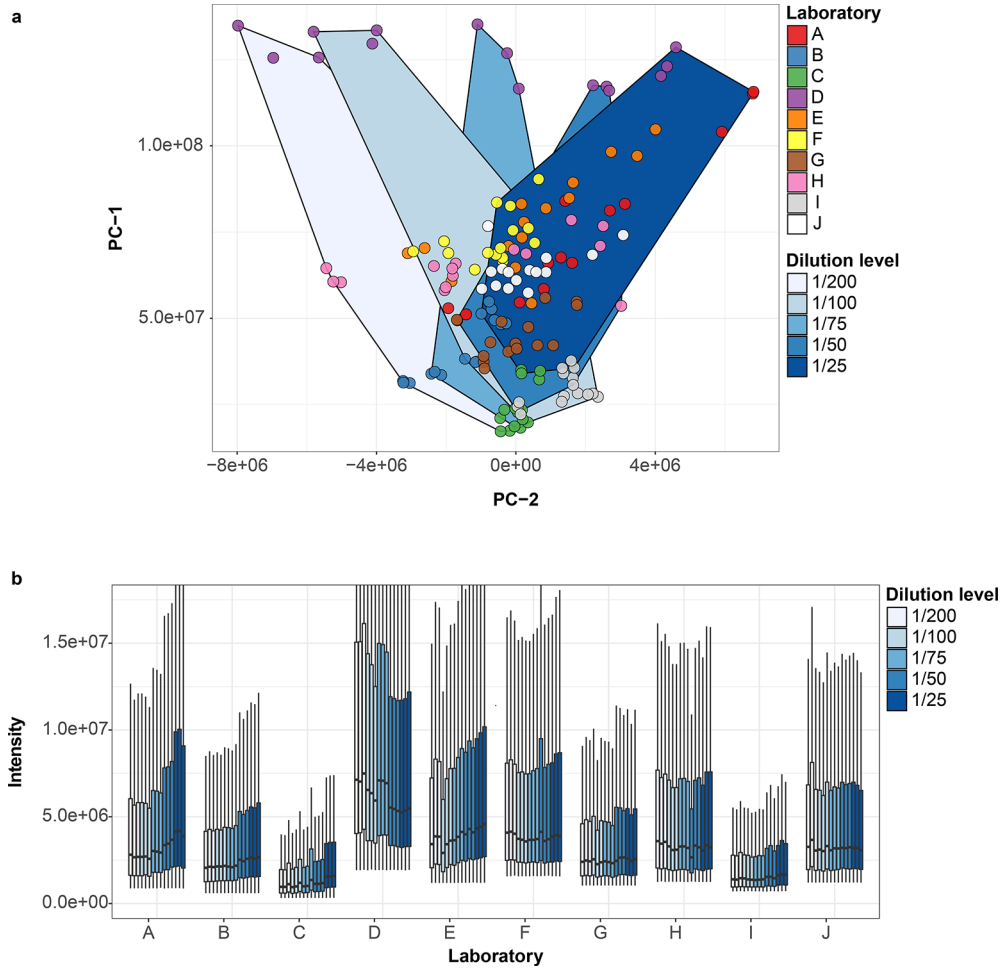


Figure 1. (a) Multivariate dynamic ranges, experiment X. Median-centered PCA on 5th to 95th percentiles of spectral intensities involving 150 spectra from 10 laboratories. The first positive component (ordinate) covers approximately 99.9% of median-centered covariance and is proportional to each laboratory's spectral signal magnitudes. The second PC (abscissa) covers approximately 0.1% of median-centered covariance and covers the experimental effect. (b) Univariate dynamic ranges, experiment X. Box plots showing the spread of distribution and magnitude of spectral intensities. Note the correspondence of box plot magnitudes to the above PC1. Both the magnitude of differences between laboratories and the outlying behavior of laboratory D are confirmed. The PCA plot is more informative in terms of lab-to-lab comparisons of intensity distributions.

laboratories identity is kept anonymous, they are classified by alphabetic letters). A kit for the analysis, standard operating procedure (Supporting Information, SI-10) were sent to each participant. Three main experiments (X,Y,Z) were set up:

(X) A C18-SPE eluate of standard human blood plasma (P9523, Sigma-Aldrich; HIV, hepatitis B, and hepatitis C, none detected) was delivered in triplicates of five different dilutions with dilution ratios ranging from 1/25 to 1/200 (v/v). Number of scans accumulated: 100 purpose: comparison of matrix effects derived from matrix dilution.

(Y) A C18-SPE eluate of standard human blood plasma (P9523, Sigma-Aldrich; HIV, hepatitis B, and hepatitis C, none detected) was delivered in triplicates at a constant concentration (1/50), spiked with five different concentrations of a mixture of 85 pesticides (LC/MS pesticide standard kit, mix 5, 6 and 7; Agilent Technologies) ranging from 1 to 20 ppb for all 85 compounds. Number of scans accumulated: 100 purpose: comparison of matrix effects derived from internal standards.

(Z) A C18-SPE eluate of standard human blood plasma (NIST SRM 1950 plasma, Sigma-Aldrich; HIV, hepatitis B, and hepatitis C, none detected) delivered as single sample at a constant concentration (1/50) and acquired in triplicate accumulating 50 scans, 100 scans, and 300 scans. Purpose: comparison to the P9523 standard (experiment X) and investigation of the sufficient number of scans for routine profiling.

Sample Preparation and Analysis. *Sample Preparation.* A pool of citrated plasma sample (P9523, Sigma-Aldrich) was extracted by OMIX C18 solid-phase extraction (SPE) pipet tips (Agilent Technologies, USA) in combination with a liquid-handling system epMotion (Eppendorf, Germany). Five milliliters of plasma was diluted (1:1) with 2% phosphoric acid, vortex mixed for 30 s, and transferred into a 96 well-plate reservoir. 96 Omix C18 100 μL tips (Agilent, A57003100) were placed in epMotion 96 trays and loaded with 100 μL of the acidified plasma. The extraction followed the protocol described in Forcisi et al.[7] The same procedure was applied for the preparation of SRM 1950 blood plasma (Sigma-Aldrich). All of the eluates were pooled, vortex-mixed, and split into aliquots following the study design.

Sample Analysis. The 17 laboratories have different FT-ICR MS instruments varying in their magnetic field strength (encompassing 7, 9.4, 12, and 15 T systems) and their ICR

https://doi.org/10.1021/jasms.2c00082
J. Am. Soc. Mass Spectrom. 2022, 33, 2203−2214

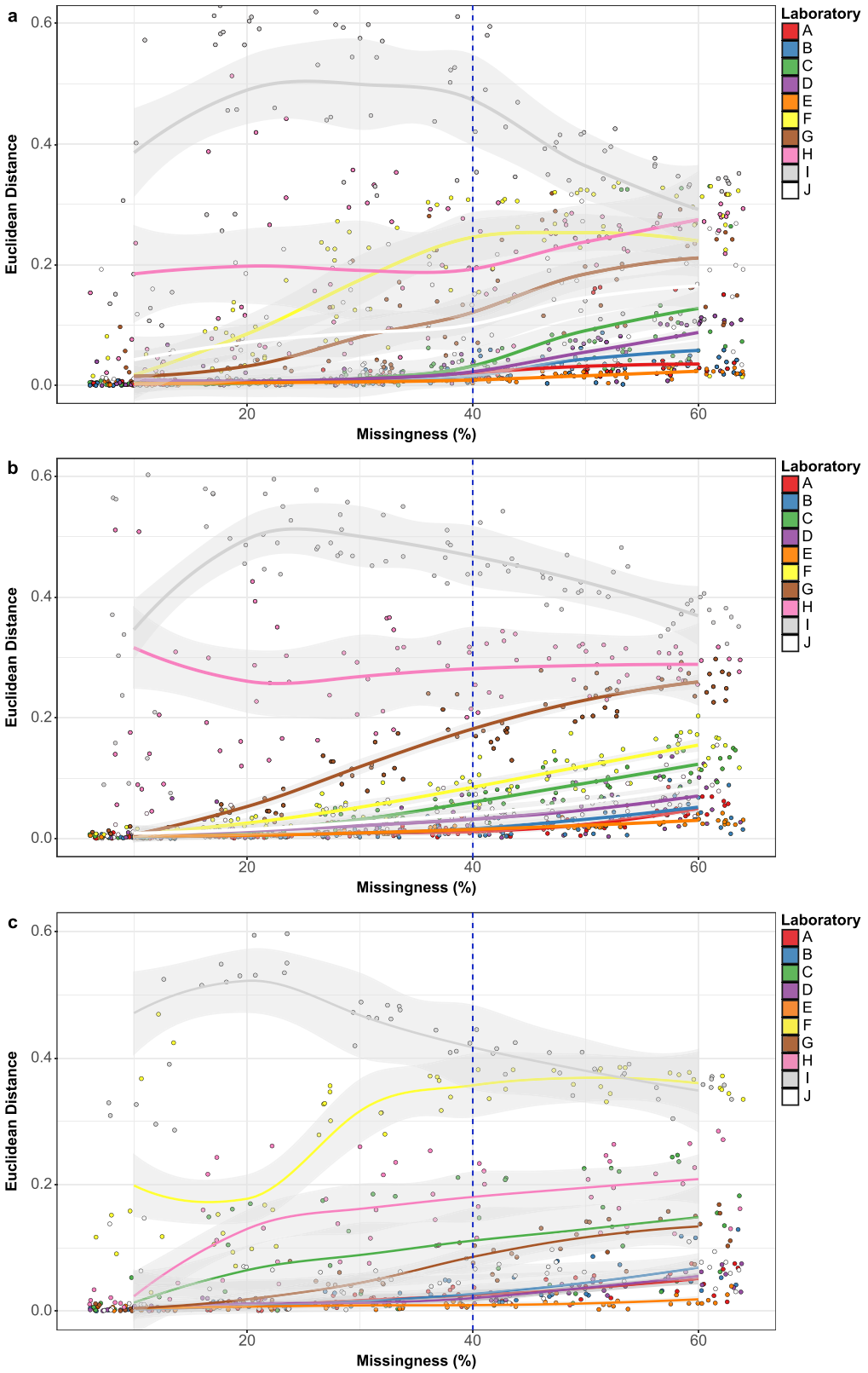


Figure 2. Summary of co-presence analysis at different degrees of missingness. See the formation of CPCs (laboratories that remain tightly associated despite increasing overall missingness) in the Supporting Information, SI-4. Vertical lines at 40% missingness indicate the order of laboratory removal from top to bottom as performed in Figure 3. (a) Normalized Euclidean distance of every laboratory's scores computed on CPA relative to the mean scores of the CPC {A,B,C,D,E} of experiment X. PCAs were computed per missingness level (expressed in %) and dilution. (b) Normalized Euclidean distance of every laboratory's scores computed on CPA relative to the mean scores of the CPC {A,B,C,D,E,F,J} of experiment Y. PCAs were computed per missingness level (expressed in %) and dilution. (c) Normalized Euclidean distance of every laboratory's scores computed on CPA relative to the mean scores of the CPC {A,B,D,E} of experiment Z. PCAs were computed per missingness level (expressed in %) and dilution.

https://doi.org/10.1021/jasms.2c00082
J. Am. Soc. Mass Spectrom. 2022, 33, 2203−2214

cell designs (Infinity or ParaCell). Their field of expertise and application was also very diverse, ranging from the fields of small molecules or metabolomics on different body fluids and biological matrices to proteomics, environmental chemistry or petroleomics, and additionally different combinations of ion source usage (ESI versus MALDI). Each laboratory received a kit containing the same samples to be measured in triplicate. A standard operating procedure (SOP) was established in Munich and distributed to each laboratory participant (see Supporting Information, SI-10). A centralized collection of the acquired data was organized to study the interlaboratory comparability. Each participant in the interlaboratory study was kept anonymous. The part of the study presented here consisted of 10 samples analyzed in triplicate with additional quality controls and blanks resulting in 713 spectra to be calibrated and processed accordingly in files of 4 Mega word (MW; measure of time domain transient length in FT-ICR MS) size containing on average 5094 ± 1168 m/z peaks over a mass range from 150 to 1000 m/z.

Spectral Calibration and Alignment. The authors used a two-step calibration scheme starting by external calibration on-site to remove the influence of local mass error drifts prior measurement, followed by internal calibration using any suitable tool to correct sample-specific space-charge effects. All instruments were calibrated on arginine clusters externally prior measurement. Data from all laboratories were peak-picked using Bruker Compass Data Analysis 4.4 at a signal-to-noise threshold (S/N) ≥ 4. Peak-picked MS lists (m/z, intensity, resolving power) were exported to tab delimited text files. All spectra were calibrated internally against a list of calibration m/z values designed for the study. The calibration list contained 285 m/z peaks composed of 153 theoretical pesticide m/z values as well as 132 human plasma m/z values whose formulas were assigned through an in-house mass-difference-based algorithm (NetCalc)[23] and further validated by isotopic fine structure. All m/z values, known from blood plasma and the Agilent pesticide mix, were excluded if they were within 5 ppm proximity to another calibrant. This way, the same calibration list could be used for the spectra generated within experiments X, Y, and Z. Mass spectra were calibrated using kernel-based calibration[24] (available upon request). Mass spectra calibrated with a standard deviation of mass measurement error exceeding 300 ppb were excluded. Gibbs peaks were removed on the basis of resolution following Kanawati et al.[25] Multiply charged features were removed on the basis of mass defect regions (width = 20 ppm) that were covered by the Pubchem database. Absolute feature intensity cutoffs were adjusted manually using absolute mass defect (AMD) plots.[1] Intensity thresholds were adjusted to be the minimum intensity that kept the region 0.1 < AMD < 0.9 at 150 < m/z < 200 empty. Spectra were aligned into an MS feature versus observation matrix using a moving alignment error window of 0.5 ppm width.

Data Processing and Statistics. All computations were performed excluding missing data marked as "NaN" (not a number), except stated otherwise. MS features were kept for further analyses if they met the following criteria after locating triplicate measurements with at least two nonzero detections: (1) Keep triplicates that occur at least twice in at least three laboratories. (2) Count frequency of missing features (missingness [%]) per dilution (across laboratories) and keep features whose missingness is <40% in at least one dilution. The resulting data distributions can be viewed in Figure 1. Details on scaling, imputation, visualization, principal component analysis (PCA), linear regression models, and comparison of signal magnitudes are detailed in Supporting Information, SI-1.

Normalization. Here, no normalization, probabilistic quotient normalization[26] (PQN), and smoothed quotient correction (SQC) were compared. PQN is commonly used for the correction of median linear shifts of spectral intensities. SQC was devised here because mass spectrometers from different generations can be equipped with varying ion optics and hardware combined with diverse instrumental settings that cause instrument specific biases. As an example, the instrument of laboratory A has poor ion transmission at m/z < 200, and varying time-of-flight in the hexapole or ion varying accumulation time in the ICR cell will change sensitivity toward lower or higher m/z. Such differences in instrumental characteristics cannot be adjusted with a monoparametric data correction method such as PQN. SQC uses smoothed quotients for data correction: compute a median spectrum R across all spectra T (excluding NaN's) and compute Q = R/T as performed in PQN. Smooth the quotients Q of each spectrum t in T with an appropriate smoother. Here, smoothed quotients SQ were computed following the matlab function SQ = smooth data (Q, "sgolay", "Window", "omitnan"), with "sgolay" being the Savitzky-Golay smoother, "Window" specifying the number of features passed to the smoother at each spectral position, and "omitnan" specifying that NaN's are to be omitted by the smoother. To avoid smoothing experimental characteristics, a window size of 50% of each spectrum's nonzero entries was used.

Co-presence Analysis (CPA). The filtered MS feature matrix was transformed into a binary data matrix with nonzero entries and NaN's replaced by ones and zeroes, respectively. PCAs were performed at different levels of missingness to locate co-presence clusters (CPCs). CPCs were determined visually in X, Y, and Z experiments. Here, CPA means co-presence PCA on binary data. Mean positions of CPCs on the first and second PCs were computed, and each laboratories' Euclidean distances to the CPC's center were computed (jointly normalized to the range [0,1]). Diverging and converging behavior as a function of missingness was visualized (Figure 2 and Figure S11).

Univariate Reproducibility. Coefficients of Variance (CV). Intra-CVs were computed on triplicates of MS features. Inter-CVs were computed on the features' triplicate means across all laboratories within the same level in terms of spiking or dilution. As such, inter-CVs report the relative standard error of the means across laboratories.

Multivariate Comparability. Each laboratory's dilution/spiking data (mean centered per laboratory) were used as generators for PCA models (scores and loadings of the first PC were stored for each laboratory). Loadings of each model generator were applied to compute "pseudoscores" given the data of the nongenerator laboratories. Scores and pseudoscores were transformed to the domain of dilution/spiking levels. That is, scores were normalized to the Euclidean length of the series {−2,−2,−2,−1,−1,−1,0,0,0,1,1,1,2,2,2} of centered dilution/spiking levels. Addition of the constant "3" shifted the resulting scores into the same domain as the sequence of dilution levels {1,1,1,2,2,2,3,3,3,4,4,4,5,5,5}. Normalized scores and pseudoscores were subjected to linear regression. Further, linear regression models of each laboratory's generator scores against dilution/spiking levels were generated. All models were visualized in a heatmap of laboratories versus laboratories and dilution/spiking levels.

https://doi.org/10.1021/jasms.2c00082
J. Am. Soc. Mass Spectrom. 2022, 33, 2203−2214

RESULTS AND DISCUSSION

Comparability of Laboratory Performances. *Raw Data.* The data of seven out of 17 laboratories had to be excluded due to asymmetric or split peak shapes, strong contamination or absence of matrix-specific MS feature patterns (Supporting Information, SI-2). The only criterion for the inclusion/exclusion of a laboratory was that the spectra were calibratable (standard deviation of mass measurement error <300 ppb). The criteria were not known to participating laboratories in advance. MS peak asymmetry and splitting are usually caused by the gas pressure in the ICR cell increasing from e^{-10} mbar toward e^{-9} mbar. Drifts of excitation powers in the ICR cell as well as ICR-cell overloading all result in aberrant shapes and trajectories of ion clouds within the ICR cell and the magnetic field. The new dynamically harmonized cell provided in new instrumentation enables better control over some of the underlying phenomena.[27] Strong peak shifts (several ppm) can occur at unstable electrospray as fluctuating ionization efficiency leads to different ion densities in the ICR cell distorting the mass-error distribution. Data matrices of plasma dilution series (experiment X), spiked plasma series (experiment Y), and varying depth of acquisition (experiment Z) were computed (see Experimental Section). The resulting matrices X and Y comprised 7974, 8918, and 8163 features over 150, 150, and 90 samples, respectively. A first assessment of laboratory comparability was performed using the filtered raw data. Figure 1 visualizes a PCA on a matrix composed of the 5th to 95th percentiles of the intensities in each spectrum of experiment X (see Experimental Section).

The first PC's magnitude is proportional to spectral signal strength, while the second PC reflects the general effect of the dilution. PCs 1 and 2 covered 99.9 and 0.10% of covariance. The scatter plot of both scores shows how overall signal magnitudes decrease as a function of dilution level (increasing matrix concentration). Therefore, increasing matrix concentrations lead to increasing signal magnitudes in all laboratories except for laboratory D. Laboratories F, G, I, and J do not show strong responses to matrix concentration in overall signal magnitudes in Figure 1a,b. Laboratory H shows a flat but monotone increase on Figure 1a and no response in Figure 1b. The behavior of these laboratories allows for two different hypotheses: (1) the analytical systems are in suppression already in the lowest matrix concentration; (2) the analytical system has strong contaminations or background that are suppressed with increasing matrix concentration. The Multivariate Comparability section will deliver which of the two hypotheses is correct. Figure 1 displays the high diversity of dynamic ranges in terms of signal magnitudes and responses to experimental intervention. The PCA's on matrices Y and Z are not shown since the results are similar.

Univariate Comparability. Different laboratories may detect diverse sets of MS features and produce different MS traces, depending on instrument parametrization, purpose of instrument use and contaminations. Here, two strategies are followed to select the laboratories for comparison: (1) analysis of CPCs on binary data and (2) removal of laboratory-specific biases by means of smoothed quotient correction (detailed rationale in the Supporting Information, SI-3).

Co-presence Analysis. While it is common to exclude features at more than 10% missingness, it is important to first identify what spectra detect the same features. The data from all experiments were transformed into binary data matrices with ones and zeros indicating feature presence and absence to compute co-detection tables (here, "co-detection" means the joint detection of a feature in at least two laboratories). Table S1 shows that laboratories A−E co-detected 4952, 5155, and 5522 MS features on average in experiments X, Y, and Z, respectively. Correspondingly, laboratories F−J co-detected 2761, 3556, and 3225 MS features on average. The overlap of laboratories A−E among each other computes to 84, 83, and 91% in X, Y, and Z, respectively. These laboratories' average overlap with laboratories F−J amounts to 62, 69, and 69%. Globally, laboratory E co-detected most features with the other laboratories followed by laboratory B. Laboratories F, H, and I have the least feature counts, sharing the least number of features with the other laboratories. More detailed insights on aberrant co-detection can be obtained by performing a PCA on binarized data. The scores of a principal component analysis on such a binary data matrix indicate feature frequencies within CPCs. Here, CPAs were performed within the dilution levels for experiment X, within spiking levels for experiment Y and within levels of scan number for experiment Z.

CPA scores were generated at varying levels of missingness, and their normalized Euclidean distance to the most robust CPC was computed (Figures S6, S7, and S8). Figure 2 shows the convolutions of co-presence structure along with increasing missingness for experiments X, Y, and Z. Laboratory I had the most aberrant co-presence structure across all levels of missingness, closely followed by laboratories H and F. Laboratories A, B, C, D, and E showed consistently similar peak detection across all experiments and levels of missingness. Those laboratories are expected to show the best intercomparability. Any clustering was independent of magnetic field strength or type of analyzer cell.

Global Co-detection between Experiments X, Y, and Z. Table S1 contains a comparison of the overlap of MS features between all three experiments computed at 40% missingness. Two thirds of all MS features detected within experiments X and Z overlapped, while one-third was found to be specific for either X or Z. Comparing the spiking experiment Y to the matrices of X and Z revealed that 48% (3815 features) and 57% (4668) of all MS features originally contained in X and Z cannot be detected when spiking with pesticides. Spiking a mixture of 85 pesticides produced 4759 (versus X) and 5423 (versus Z) features that were never detected in blood plasma extracts. Future ring trials for routine analyses will have to assess the type and concentration of internal standards thoroughly as was done in Chekmeneva et al.[28]

Normalization and Univariate Comparability. Here, the major measure to assess interlaboratory comparability is the coefficient of variation computed on co-detected feature intensities of the laboratories to be compared. Figure 3 shows the results of this computation at different levels of laboratory exclusion on the abscissa. Inter-CVs in all plots of Figure 3 decrease as the most dissimilar laboratories are removed from computation following the vertical dashed lines of Figure 2 from top to bottom. This result indicates that CPA captures the similarity between detection patterns well. The exception to that observation was the convolution of inter-CVs on raw data in experiment Y. Pesticide spiking emphasizes the differences in the instruments' dynamic ranges and this difference appears to be stronger than the effect of MS feature co-detection. Different strategies of spectral normalization (PQN and SQC) and feature scaling (UVC and L2) were tested on the data available, knowing that there are a multitude of effects influencing the comparability of mass spectra between different batches and

https://doi.org/10.1021/jasms.2c00082
J. Am. Soc. Mass Spectrom. 2022, 33, 2203−2214

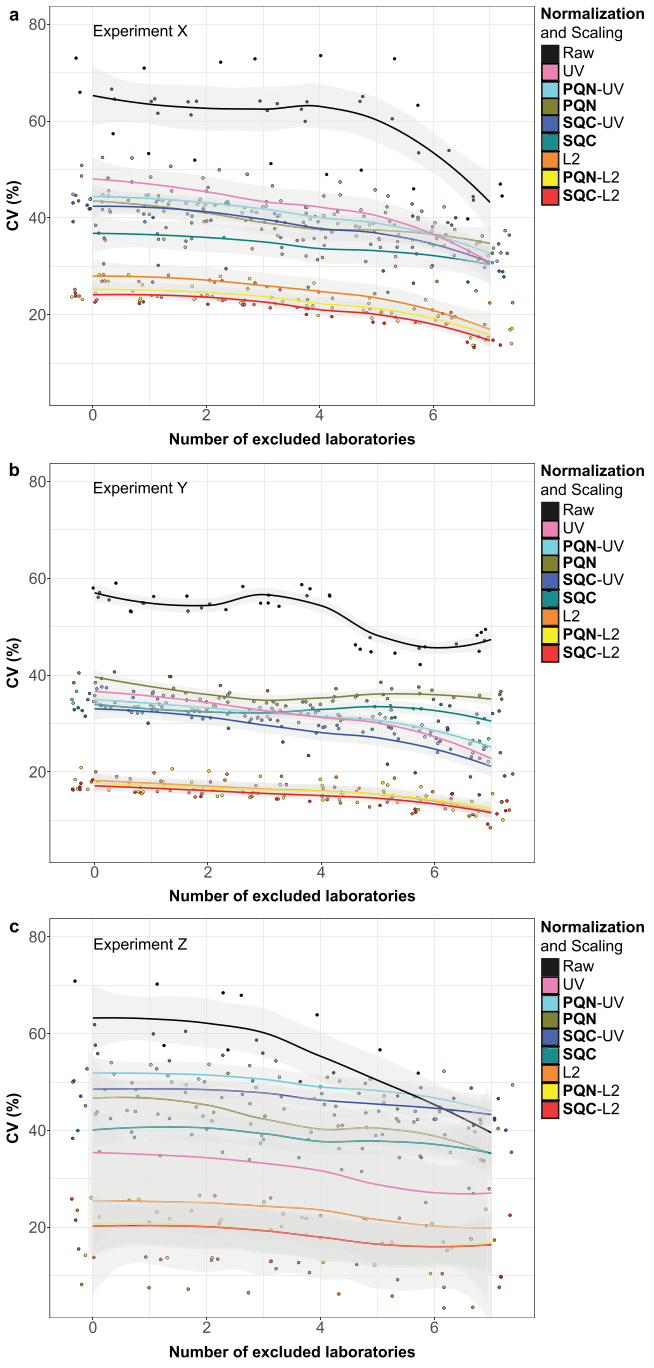


Figure 3. Convolution of CVs as a function of increasing co-detection of MS features on experiments X, Y, and Z. Inter- and intra-CVs were computed across the entire data set iteratively removing laboratories following Figure 2 at 40% missingness, beginning with the uppermost (most aberrant) laboratory, with the members of CPCs being the last to be removed. Inter- and intra-CVs of the following data treatments are compared: raw (Raw), raw with laboratory Euclidean (L2), raw with laboratory UV scaling (UV), PQN on raw data (PQN), SQC on raw data (SQC), PQN followed by L2 (PQN_L2), PQN followed by UV (PQN_UV), SQC followed by L2 (SQC_L2), and SQC followed by UV (SQC_UV). The effects of both SQC and Euclidean-norm affect inter-CVs majorly, while intra-CVs do not change due to Euclidean normalization (of features) and improve only slightly as SQC is applied (a zoom-in toward intra-CVs is provided in the Supporting Information, SI-6).

laboratories. Normalization methods act on the distribution of signal intensities in a spectrum, while scaling methods act on the magnitude of an individual MS feature across spectra. We found that feature scaling across all laboratories at once distorts intensity distributions. Here, scaling was used on each laboratory's data individually. Figure 3 shows that the application of the L2 norm (Euclidean norm) was always superior to UV scaling. L2 exercises the greatest impact in terms of inter-CV improvement relative to raw data among individual correction methods tested. At the same time, any combination involving SQC was superior to PQN except for experiment Z. Here, PQN-L2 and SQC-L2 performed the best resulting in exactly the same results while both PQN-UV and SQC-UV performed worse than raw data. A likely reason for this behavior is that the denominator in the computation of standard deviation (itself the denominator in UV scaling) is corrected for degrees of freedom, which exercises a large effect in small sample sizes. The removal of aberrant laboratories had the least noticeable effect on inter-CVs in terms of absolute removal of error. The improvement of inter-CVs was the weakest in experiment Z when data correction was used. No matrix composition was modified in experiment Z and the single sources of variation were instrument specific biases and number of scans. While Figure 3 displays inter-CVs only, it is to be noted that any combination of scaling and normalization methods did not alter intra-CV significantly (Supporting Information, SI-5).

The results substantiate that L2 normalization should be used for feature scaling and that SQC corrects interlaboratory variation majorly and that intra-CV magnitude is not an artifact of data treatment.

Multivariate Comparability. Multivariate comparability was assessed in a cross validation scheme for experiments X and Y. A PCA model was built for each laboratory. Each laboratory's original model was then applied on the data of all other laboratories, generating validation scores that were scaled to same Euclidean length. R^2 values of regression analyses between a laboratory's original model score (rows in Figure 4) and the validation scores plus dilution or spiking levels (columns in Figure 4) were visualized in a heat map. Figure 4 shows the cross-validation map computed on the experiment X (matrix effect) at 40% missingness per dilution level and laboratory L2 norm applied on features. The interesting case of laboratory H shows that applying its model to the data other laboratories generated high R^2 values ($R^2 > 0.8$). A linear regression of the H-scores against dilution levels showed good linearity in the scores ($R^2 = 0.94$). At the same time, the data of H did not perform well using the other laboratories' PCA models. H is one of the two laboratories with a strongly aberrant missingness structure in experiments X, Y, and Z (Figure 2). The result can be interpreted as follows: Those features that follow matrix dilution levels in H show the same trends in the other laboratories. However, these features receive small loadings in the PCAs of other laboratories. This insight is supported by the scatter plots corresponding to Figure 4 (Figure S14). Recall the case of laboratory D, which showed decreasing signal magnitudes with increasing sample concentration. Figure 4 implies that the dependency of signal magnitudes on matrix concentration was a sign of increasing suppression strong background ions (Figure S16). The models of laboratory D produced an average $R^2 \approx 0.7$ when applied to the data of other laboratories. Likewise, the loadings of other laboratories produced an average $R^2 \approx 0.75$ when applied on the data of D. The cross-validation map of experiment Y (Figure S13) shows perfect R^2 values for all

https://doi.org/10.1021/jasms.2c00082
J. Am. Soc. Mass Spectrom. 2022, 33, 2203−2214

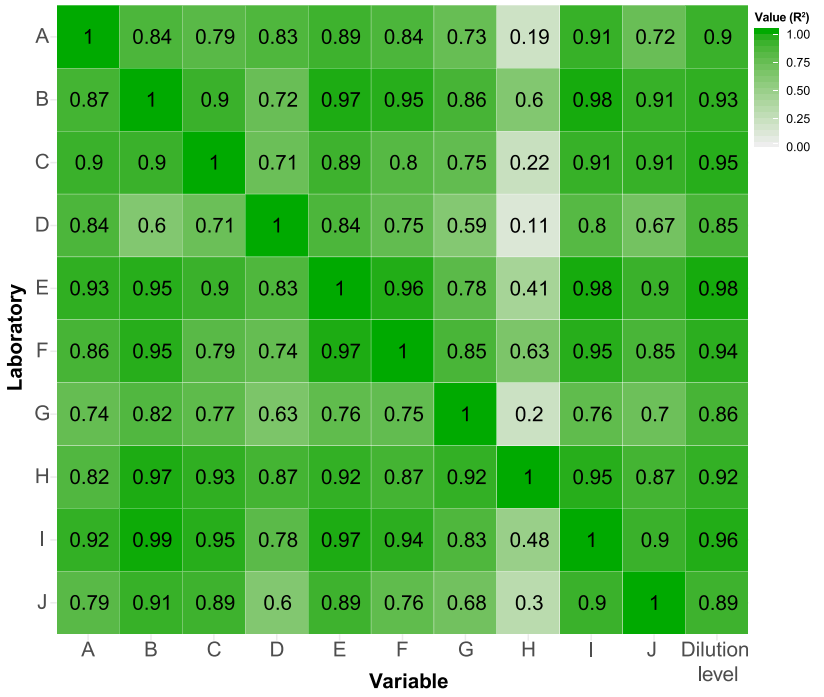


Figure 4. PCA cross-validation map for experiment X.

combinations reflecting excellent reproduction the effects pesticide spiking exercised on the analytical matrix. The scatter-plot map for the cross-validation of experiment X shows that the FT-ICR analyzers were challenged by large ion abundances at spiking levels 15 and 20 ppb (Figure S15).

Spectral normalization did not have significant effects on the cross-validation experiment. Univariate comparability as measured by inter-CVs did not appear to be of major importance for multivariate comparability. One possible explanation would derive from the central limit theorem: PCA scores represent a weighted mean of the hundreds or thousands of features' univariate signal magnitudes.

Evaluation of Scanning Depth. High-throughput routine analyses are required to be fast and robust and need to detect features of relevance. Experiment Z was performed using the NIST SRM 1950 standard for which quantified metabolite identities and concentrations are available. We manually extracted the names and identifiers for 682 small molecules, amounting to 430 individual compounds and 319 unique molecular formulas from Simón-Manso et al. All 319 molecular formulas transformed into seven fundamental ion types where possible: $[M + H]^+$, $[M + CH_4O + H]^+$, $[M + Na]^+$, $[M + H_2O + H]^+$, $[M + CHO_2Na + H]^+$, $[M - NH_3 + H]^+$, and $[M - H_2O + H]^+$. All ion types were combined combinatorically to build homodimers and homotrimers of different adduct types. The expanded list encompassed 26,908 positive ionization mode m/z values to be searched in the Z experiment. The raw data matrices of experiments X and Z were fused at 1 ppm error and matched against the built SRM 1950 metabolic feature collection. Mass matching at a 0.5 and 1 ppm search window size resulted in 1111 and 1345 hits against the fused XZ matrix.

While laboratory B showed significantly more hits than all other laboratories, only laboratories F and I detected significantly fewer putatively annotated MS features (Figure 5a). Setting the number of detected features after 300 scans to 100% for each laboratory, 76 and 89% of all SRM 1950 features are detected after 50 and 100 scans, respectively. SRM 1950 annotated MS features showed significantly higher detection frequencies compared to detection frequencies in the corresponding data sets (Figure 5b). Experiment Z showed consistently higher detection frequencies compared to experiment X, which is partially due to longer scanning times and no variation of matrix composition in experiment Z. Figure 5 and experiment Z imply that (i) features detected within this study are representative for the analytical matrix and that it is more likely to detect metabolic features that are already known; (ii) the accumulation of merely 50 scans, which is performed within 60−90 s, provides a read out that is already 75% complete (compared to acquisition of 300 scans, which takes approximately 10 min depending on instrumental settings). The detection of isotopic fine structures naturally requires longer scanning times.[27]

CONCLUSION

This preliminary interlaboratory study was set up to comprehend whether different FT-ICR mass spectrometers across the globe, at their routine performance, have the aptitude to detect the same signals generated on the same sample sets. The subsequent question was whether feature intensities were comparable and what data correction technique could be applied to minimize interlaboratory CVs. The global aim was to collect the adequate experience and knowledge for setting up a future DI-FT-ICR MS ring trial. The variability in laboratory performance across the FT-ICR community around the world was found to be significant, likely due to diverse scientific scopes, some of which may not fall into the field of analyzing small molecules.

The data provided by seven out of 17 laboratories had to be excluded because either MS-peak deformation, lack of signals to be calibrated or strong contaminations hindered calibration to below 300 ppb of mass error standard deviation. The remaining laboratories exhibited strong variability in terms of signal

https://doi.org/10.1021/jasms.2c00082
J. Am. Soc. Mass Spectrom. 2022, 33, 2203−2214

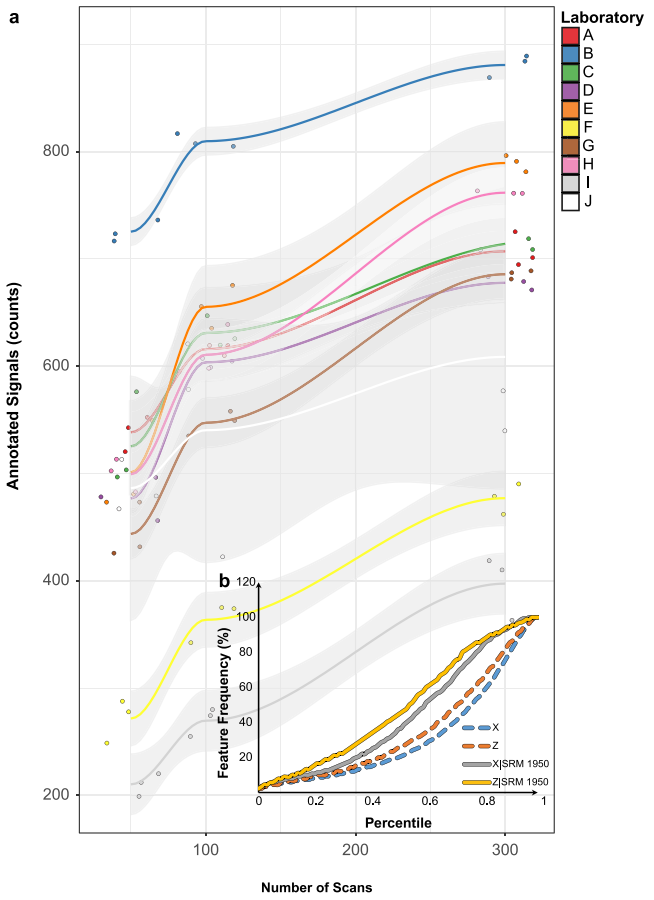


Figure 5. Putatively annotated signals in SRM 1950 blood plasma per scanning number. (a) Number of MS features putatively annotated against SRM 1950 metabolites per number of accumulated scans for each participant. (b) Comparison of the MS feature frequency (percentage of valid triplicate detection across all laboratories) between the MS features co-detected in X and Z alone against the frequency of those features with SRM 1950 annotations within either experiment.

magnitudes at smaller or larger m/z values and number of valid triplicate features detected. Co-detection analyses across three different experiments showed that laboratories A, B, C, D, and E co-detect up to 90% their valid triplicate signals. Laboratories B and E were found to be the most representative for all laboratories. Globally, 67 and 69% of the MS features detected in experiments X and Z overlapped, indicating a good representation of MS features typical for blood plasma SPE extracts. In turn, pesticide spiking caused the loss of almost half of the MS features detected in X, substantiating that the type and concentration of authentic standards have to be evaluated carefully when standard addition is performed in DI-MS. Comparisons of interlaboratory CVs consistently showed that smoothed quotient correction followed by scaling MS features on the L2 norm within each laboratory individually resulted in median CVs between 10 and 20%. Multivariate cross-validation on experiments X and Y showed that multivariate comparability of experimental effects was acceptable not necessarily depending on univariate comparability.

In effect, an appropriate strategy toward interlaboratory comparability for untargeted DI-FT-ICR MS would encompass the following steps: (1) Optimize instrumental parameters to meet the detection pattern of the laboratories that showed best co-detection (e.g., B and E). (2) Define a study design that is representative of the analytical task and maintain a constant batch size. (3) Always use the exact same amount of quality control samples that are randomly distributed in the batch so that L2 normalization can be performed. (4) Perform SQC toward the median of QCs and use the L2 norm of the QCs to scale the samples in question. We suggest the above points as one step toward applications of untargeted diagnostic UHR-MS profiling for fields of application such as quality control in food industries, pharmaceuticals, or clinical phenome centers.

Finally, experiment Z, using the SRM 1950 blood plasma standard, showed that 50 scans (60−90 s scanning time) were found to be sufficient to detect 75% of all potential SRM 1950 metabolites detected at 300 scans (8−10 min scanning time). These results suggest that FT-ICR mass spectrometers can be used for routine high-throughput measurements. Follow-up studies could encompass ring trials with cloned instruments (7Ts, 9Ts, 12Ts, Infinity versus ParaCell), clusters of laboratories of similar scope and expertise (proteomics, lipidomics, petroleomics, etc.), and more elaborate study designs (e.g., including a clinical study).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jasms.2c00082.

Experimental information, supporting concepts and figures (PDF)

Table S1 (XLSX)

■ AUTHOR INFORMATION

Corresponding Authors

Sara Forcisi − *Research Unit Analytical BioGeoChemistry, Helmholtz Zentrum München, 85764 Neuherberg, Germany; German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany;* orcid.org/0000-0003-1976-7976; Email: sara.forcisi@tum.de

Franco Moritz − *Research Unit Analytical BioGeoChemistry, Helmholtz Zentrum München, 85764 Neuherberg, Germany;* orcid.org/0000-0002-1167-6191; Email: franco.fmoritz@helmholtz-muenchen.de

Christopher J. Thompson − *Bruker Daltonics Inc., Billerica, Massachusetts 01821, United States;* orcid.org/0000-0002-3022-3710; Email: chris.thompson@brightspec.com

Philippe Schmitt-Kopplin − *Research Unit Analytical BioGeoChemistry, Helmholtz Zentrum München, 85764 Neuherberg, Germany; German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany; Analytical Food Chemistry, Technical University of Munich, 85354 Freising, Germany*; Email: schmitt-kopplin@helmholtz-muenchen.de

Authors

Basem Kanawati − *Research Unit Analytical BioGeoChemistry, Helmholtz Zentrum München, 85764 Neuherberg, Germany*

Jenny Uhl − *Research Unit Analytical BioGeoChemistry, Helmholtz Zentrum München, 85764 Neuherberg, Germany*

Carlos Afonso − *COBRA, UMR 6014 et FR 3038, INSA de Rouen, CNRS, IRCOF, Normandie Université, Université de Rouen, 76130 Cedex Mont Saint Aignan, France;* orcid.org/0000-0002-2406-5664

Chantal D. Bader − *Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection*

https://doi.org/10.1021/jasms.2c00082
J. Am. Soc. Mass Spectrom. 2022, 33, 2203−2214

Research, Saarland University Campus, 66123 Saarbrücken, Germany and Department of Pharmacy, Saarland University, 66123 Saarbrücken, Germany

Aiko Barsch − *Bruker Daltonik GmbH, 28359 Bremen, Germany*

Berin A. Boughton − *Metabolomics Australia, School of BioSciences, University of Melbourne, Melbourne, Victoria 3010, Australia;* Present Address: Australian National Phenome Centre, Murdoch University, Harry Perkins Institute of Medical Research, 5 Robin Warren Drive, 6150 Murdoch, Western Australia; orcid.org/0000-0001-6342-9814

Rosalie K. Chu − *Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99352, United States*

Justine Ferey − *COBRA, UMR 6014 et FR 3038, INSA de Rouen, CNRS, IRCOF, Normandie Université, Université de Rouen, 76130 Cedex Mont Saint Aignan, France;* Present Address: Metatoul-AXIOM Platform, MetaboHUB, Toxalim, INRAE, Toulouse, France. Toxalim, Toulouse University, INRAE, ENVT, INP-Purpan, UPS, Toulouse, France.

Francisco Fernandez-Lima − *Department of Chemistry and Biochemistry, Florida International University, Miami, Florida 33199, United States; Biomolecular Sciences Institute, Florida International University, Miami, Florida 33199, United States;* orcid.org/0000-0002-1283-4390

Céline Guéguen − *Chemistry Department, Trent University, Peterborough, ON K9J 7B8, Canada;* Present Address: Département de chimie, Université de Sherbrooke, 2500 Boulevard de l'université, Sherbrooke, Québec, J1K 2R1, Canada.

Dimitri Heintz − *Plant Imaging and Mass Spectrometry (PIMS), Institut de Biologie Moléculaire des Plantes, CNRS, Université de Strasbourg, 67084 Strasbourg, France*

Mario Gomez-Hernandez − *Department of Chemistry and Biochemistry, Florida International University, Miami, Florida 33199, United States; Biomolecular Sciences Institute, Florida International University, Miami, Florida 33199, United States*

Kyoung-Soon Jang − *Bio-Chemical Analysis Team, Korea Basic Science Institute, Cheongju 28119, South Korea*

Nikolas Kessler − *Bruker Daltonik GmbH, 28359 Bremen, Germany*

Vaughn Mangal − *Chemistry Department, Trent University, Peterborough, ON K9J 7B8, Canada*

Rolf Müller − *Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research, Saarland University Campus, 66123 Saarbrücken, Germany and Department of Pharmacy, Saarland University, 66123 Saarbrücken, Germany;* orcid.org/0000-0002-1042-5665

Ryo Nakabayashi − *Metabolomics Research Group, RIKEN Center for Sustainable Resource Science, Yokohama 230-0045, Japan;* orcid.org/0000-0002-8674-0928

Edith Nicol − *Laboratoire de Chimie Moléculaire (LCM), CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau, France;* orcid.org/0000-0001-8791-9949

Simone Nicolardi − *Center for Proteomics and Metabolomics, Leiden University Medical Center Leiden, 2333 ZC Leiden, The Netherlands;* orcid.org/0000-0001-8393-1625

Magnus Palmblad − *Center for Proteomics and Metabolomics, Leiden University Medical Center Leiden, 2333 ZC Leiden, The Netherlands;* orcid.org/0000-0002-5865-8994

Ljiljana Paša-Tolić − *Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99352, United States*

Jacob Porter − *Department of Chemistry and Biochemistry, Florida International University, Miami, Florida 33199, United States; Biomolecular Sciences Institute, Florida International University, Miami, Florida 33199, United States*

Isabelle Schmitz-Afonso − *COBRA, UMR 6014 et FR 3038, INSA de Rouen, CNRS, IRCOF, Normandie Université, Université de Rouen, 76130 Cedex Mont Saint Aignan, France*

Jong Bok Seo − *Seoul Center, Korea Basic Science Institute, Seongbuk-Gu 02841 Seoul, South Korea*

Eduardo Sommella − *Department of Pharmacy, University of Salerno, 84084 Fisciano, SA, Italy*

Yuri E. M. van der Burgt − *Center for Proteomics and Metabolomics, Leiden University Medical Center Leiden, 2333 ZC Leiden, The Netherlands;* orcid.org/0000-0003-0556-5564

Claire Villette − *Plant Imaging and Mass Spectrometry (PIMS), Institut de Biologie Moléculaire des Plantes, CNRS, Université de Strasbourg, 67084 Strasbourg, France*

Matthias Witt − *Bruker Daltonik GmbH, 28359 Bremen, Germany*

Ashley Wittrig − *ExxonMobil Research and Engineering Company, Clinton, New Jersey 08869, United States;* Present Address: ExxonMobil Chemical Company, 22777 Springwoods Village Parkway, Spring, TX 77389.; orcid.org/0000-0002-2712-2003

Jeremy J. Wolff − *Bruker Daltonics Inc., Billerica, Massachusetts 01821, United States*

Michael L. Easterling − *Bruker Daltonics Inc., Billerica, Massachusetts 01821, United States*

Frank H. Laukien − *Bruker Daltonics Inc., Billerica, Massachusetts 01821, United States; Department of Chemistry & Chemical Biology, Cambridge, Harvard University, Cambridge, Massachusetts 02138, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/jasms.2c00082

Author Contributions

Conception, study design and logistics: S.F., F.M., C.J.T., B.K., J.U., A.B., M.L.E., F.L., P.S.-K. On-site coordination of the experiments: S.F., F.M., B.K., J.U., C.A., B.A.B., F.F.-L., C.G., D.H., K.-S.J., R.M., R.N., E.N., M.P., L.P.-T., J.B.S., E.S., M.W., A.W., J.W., P.S.-K. Preparation and distribution of samples: S.F., F.M., J.U. Development, distribution and revision of standard operating procedure (SOP): S.F., F.M., C.J.T., B.K., J.U., S.N., M.P., Y.E.M.vdB., M.W. Data acquisition: S.F., F.M., B.K., C.D.B., B.A.B., R.K.C., J.F., F.F.-L., M.G.-H., K.-S.J., V.M., R.N., E.N., S.N., M.P., J.P., I.S.-A., J.B.S., E.S., Y.E.M.vdB., C.V., M.W., A.W., J.W. Data analysis, interpretation and drafting of the manuscript: S.F., F.M., N.K., C.J.T. Critical revision of the manuscript: S.F., F.M., C.J.T., B.K., J.U., C.D.B., B.A.B., F.F.-L., C.G., K.-S.J., E.N., E.S., C.V., M.W., P.S.-K. S.F., F.M., and C.J.T. contributed equally.

Notes

The authors declare the following competing financial interest(s): A.B., N.K., and M.W. are employees of Bruker Daltonik GmbH, F.H.L. is President and CEO of the Bruker Corporation, and C.J.T. and J.W. were employees of Bruker Daltonics Inc. which manufactures and sells mass spectrometers and software used in this study.

https://doi.org/10.1021/jasms.2c00082
J. Am. Soc. Mass Spectrom. 2022, 33, 2203−2214

■ ACKNOWLEDGMENTS

This research was supported by the German Center for Diabetes Research (DZD; Grants G-501900-482 and G-501901-020), the European Regional Development Fund (ERDF) No. HN0001343, the European Union's Horizon 2020 Research Infrastructures program (Grant Agreement 731077), the Région Normandie, the Laboratoire d'Excellence (LabEx) SynOrg (ANR-11-LABX-0029), and the national FT-ICR network (FR 3624 CNRS). A portion of the research was performed using EMSL (grid.436923.9), a DOE Office of Science User Facility sponsored by the Biological and Environmental Research program.

■ REFERENCES

(1) Schmitt-Kopplin, P.; Hemmler, D.; Moritz, F.; Gougeon, R. D.; Lucio, M.; Meringer, M.; Müller, C.; Harir, M.; Hertkorn, N. Systems Chemical Analytics: Introduction to the Challenges of Chemical Complexity Analysis. *Faraday Discuss.* **2019**, *218* (0), 9−28.

(2) Moritz, F.; Hemmler, D.; Kanawati, B.; Schnitzler, J.-P.; Schmitt-Kopplin, P. Mass Differences in Metabolome Analyses of Untargeted Direct Infusion Ultra-High Resolution MS Data. Fundamentals and Applications of Fourier Transform Mass Spectrometry. *Fundam. Appl. Fourier Transform Mass Spectrom.* **2019**, 357−405.

(3) Smith, D. F.; Podgorski, D. C.; Rodgers, R. P.; Blakney, G. T.; Hendrickson, C. L. 21 T FT-ICR Mass Spectrometer for Ultrahigh-Resolution Analysis of Complex Organic Mixtures. *Anal. Chem.* **2018**, *90* (3), 2041−2047.

(4) Hertkorn, N.; Frommberger, M.; Witt, M.; Koch, B. P.; Schmitt-Kopplin, P.; Perdue, E. M. Natural Organic Matter and the Event Horizon of Mass Spectrometry. *Anal. Chem.* **2008**, *80* (23), 8908−8919.

(5) Roullier-Gall, C.; David, V.; Hemmler, D.; Schmitt-Kopplin, P.; Alexandre, H. Exploring Yeast Interactions through Metabolic Profiling. *Sci. Rep.* **2020**, *10* (1), 6073.

(6) Valle, J.; Harir, M.; Gonsior, M.; Enrich-Prast, A.; Schmitt-Kopplin, P.; Bastviken, D.; Hertkorn, N. Molecular Differences between Water Column and Sediment Pore Water SPE-DOM in Ten Swedish Boreal Lakes. *Water Res.* **2020**, *170*, 115320.

(7) Forcisi, S.; Moritz, F.; Lucio, M.; Lehmann, R.; Stefan, N.; Schmitt-Kopplin, P. Solutions for Low and High Accuracy Mass Spectrometric Data Matching: A Data-Driven Annotation Strategy in Nontargeted Metabolomics. *Anal. Chem.* **2015**, *87* (17), 8917−8924.

(8) González-Domínguez, R.; Sayago, A.; Fernández-Recamales, Á. Direct Infusion Mass Spectrometry for Metabolomic Phenotyping of Diseases. *Bioanalysis* **2017**, *9* (1), 131−148.

(9) González-Domínguez, R. Metabolomic Fingerprinting of Serum Samples by Direct Infusion Mass Spectrometry. *Proceedings of the 1st International Electronic Conference on Metabolomics*; November 1−30, 2016; DOI: 10.3390/iecm-1-c001.

(10) González-Domínguez, R. Metabolomic Fingerprinting of Blood Samples by Direct Infusion Mass Spectrometry: Application in Alzheimer's Disease Research. *Journal of Analysis and Testing* **2017**, *1*, 213−222.

(11) Chekmeneva, E.; Dos Santos Correia, G.; Gómez-Romero, M.; Stamler, J.; Chan, Q.; Elliott, P.; Nicholson, J. K.; Holmes, E. Ultra-Performance Liquid Chromatography-High-Resolution Mass Spectrometry and Direct Infusion-High-Resolution Mass Spectrometry for Combined Exploratory and Targeted Metabolic Profiling of Human Urine. *J. Proteome Res.* **2018**, *17* (10), 3492−3502.

(12) Bowden, J. A.; Heckert, A.; Ulmer, C. Z.; Jones, C. M.; Koelmel, J. P.; Abdullah, L.; Ahonen, L.; Alnouti, Y.; Armando, A. M.; Asara, J. M.; Bamba, T.; Barr, J. R.; Bergquist, J.; Borchers, C. H.; Brandsma, J.; Breitkopf, S. B.; Cajka, T.; Cazenave-Gassiot, A.; Checa, A.; Cinel, M. A.; Colas, R. A.; Cremers, S.; Dennis, E. A.; Evans, J. E.; Fauland, A.; Fiehn, O.; Gardner, M. S.; Garrett, T. J.; Gotlinger, K. H.; Han, J.; Huang, Y.; Neo, A. H.; Hyotylainen, T.; Izumi, Y.; Jiang, H.; Jiang, H.; Jiang, J.; Kachman, M.; Kiyonami, R.; Klavins, K.; Klose, C.; Kofeler, H. C.; Kolmert, J.; Koal, T.; Koster, G.; Kuklenyik, Z.; Kurland, I. J.; Leadley, M.; Lin, K.; Maddipati, K. R.; McDougall, D.; Meikle, P. J.; Mellett, N. A.; Monnin, C.; Moseley, M. A.; Nandakumar, R.; Oresic, M.; Patterson, R.; Peake, D.; Pierce, J. S.; Post, M.; Postle, A. D.; Pugh, R.; Qiu, Y.; Quehenberger, O.; Ramrup, P.; Rees, J.; Rembiesa, B.; Reynaud, D.; Roth, M. R.; Sales, S.; Schuhmann, K.; Schwartzman, M. L.; Serhan, C. N.; Shevchenko, A.; Somerville, S. E.; St. John-Williams, L.; Surma, M. A.; Takeda, H.; Thakare, R.; Thompson, J. W.; Torta, F.; Triebl, A.; Trotzmuller, M.; Ubhayasekera, S. J. K.; Vuckovic, D.; Weir, J. M.; Welti, R.; Wenk, M. R.; Wheelock, C. E.; Yao, L.; Yuan, M.; Zhao, X. H.; Zhou, S. Harmonizing Lipidomics: NIST Interlaboratory Comparison Exercise for Lipidomics Using SRM 1950-Metabolites in Frozen Human Plasma. *J. Lipid Res.* **2017**, *58* (12), 2275−2288.

(13) Siskos, A. P.; Jain, P.; Römisch-Margl, W.; Bennett, M.; Achaintre, D.; Asad, Y.; Marney, L.; Richardson, L.; Koulman, A.; Griffin, J. L.; Raynaud, F.; Scalbert, A.; Adamski, J.; Prehn, C.; Keun, H. C. Interlaboratory Reproducibility of a Targeted Metabolomics Platform for Analysis of Human Serum and Plasma. *Anal. Chem.* **2017**, *89* (1), 656−665.

(14) Zherebker, A.; Kim, S.; Schmitt-Kopplin, P.; Spencer, R. G. M.; Lechtenfeld, O.; Podgorski, D. C.; Hertkorn, N.; Harir, M.; Nurfajin, N.; Koch, B.; Nikolaev, E. N.; Shirshin, E. A.; Berezin, S. A.; Kats, D. S.; Rukhovich, G. D.; Perminova, I. V. Interlaboratory Comparison of Humic Substances Compositional Space as Measured by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (IUPAC Technical Report). *J. Macromol. Sci., Part A: Pure Appl. Chem.* **2020**, *92* (9), 1447−1467.

(15) Perminova, I. V. From Green Chemistry and Nature-like Technologies towards Ecoadaptive Chemistry and Technology. *J. Macromol. Sci., Part A: Pure Appl. Chem.* **2019**, *91* (5), 851−864.

(16) Kirwan, J. A.; Weber, R. J. M.; Broadhurst, D. I.; Viant, M. R. Direct Infusion Mass Spectrometry Metabolomics Dataset: A Benchmark for Data Processing and Quality Control. *Sci. Data* **2014**, *1*, 140012.

(17) Kirwan, J. A.; Broadhurst, D. I.; Davidson, R. L.; Viant, M. R. Characterising and Correcting Batch Variation in an Automated Direct Infusion Mass Spectrometry (DIMS) Metabolomics Workflow. *Anal. Bioanal. Chem.* **2013**, *405* (15), 5147−5157.

(18) Kido Soule, M. C.; Longnecker, K.; Giovannoni, S. J.; Kujawinski, E. B. Impact of Instrument and Experiment Parameters on Reproducibility of Ultrahigh Resolution ESI FT-ICR Mass Spectra of Natural Organic Matter. *Org. Geochem.* **2010**, *41*, 725−733.

(19) Hawkes, J. A.; D'Andrilli, J.; Agar, J. N.; Barrow, M. P.; Berg, S. M.; Catalán, N.; Chen, H.; Chu, R. K.; Cole, R. B.; Dittmar, T.; Gavard, R.; Gleixner, G.; Hatcher, P. G.; He, C.; Hess, N. J.; Hutchins, R. H. S.; Ijaz, A.; Jones, H. E.; Kew, W.; Khaksari, M.; Palacio Lozano, D. C.; Lv, J.; Mazzoleni, L. R.; Noriega-Ortega, B. E.; Osterholz, H.; Radoman, N.; Remucal, C. K.; Schmitt, N. D.; Schum, S. K.; Shi, Q.; Simon, C.; Singer, G.; Sleighter, R. L.; Stubbins, A.; Thomas, M. J.; Tolic, N.; Zhang, S.; Zito, P.; Podgorski, D. C. An International Laboratory Comparison of Dissolved Organic Matter Composition by High Resolution Mass Spectrometry: Are We Getting the Same Answer? *Limnol. Oceanogr.: Methods* **2020**, *18* (6), 235−258.

(20) Clark, T. N.; Houriet, J.; Vidar, W. S.; Kellogg, J. J.; Todd, D. A.; Cech, N. B.; Linington, R. G. Interlaboratory Comparison of Untargeted Mass Spectrometry Data Uncovers Underlying Causes for Variability. *J. Nat. Prod.* **2021**, *84* (3), 824−835.

(21) Laber, S.; Forcisi, S.; Bentley, L.; Petzold, J.; Moritz, F.; Smirnov, K. S.; Al Sadat, L.; Williamson, I.; Strobel, S.; Agnew, T.; Sengupta, S.; Nicol, T.; Grallert, H.; Heier, M.; Honecker, J.; Mianne, J.; Teboul, L.; Dumbell, R.; Long, H.; Simon, M.; Lindgren, C.; Bickmore, W. A.; Hauner, H.; Schmitt-Kopplin, P.; Claussnitzer, M.; Cox, R. D. Linking the FTO Obesity rs1421085 Variant Circuitry to Cellular, Metabolic, and Organismal Phenotypes in Vivo. *Sci. Adv.* **2021**, *7*, eabg0108.

(22) Simon-Manso, Y.; Lowenthal, M. S.; Kilpatrick, L. E.; Sampson, M. L.; Telu, K. H.; Rudnick, P. A.; Mallard, W. G.; Bearden, D. W.; Schock, T. B.; Tchekhovskoi, D. V.; Blonder, N.; Yan, X.; Liang, Y.; Zheng, Y.; Wallace, W. E.; Neta, P.; Phinney, K. W.; Remaley, A. T.; Stein, S. E. Metabolite Profiling of a NIST Standard Reference Material

for Human Plasma (SRM 1950): GC-MSLC-MSNMR, and Clinical Laboratory Analyses, Libraries, and Web-Based Resources. *Anal. Chem.* **2013**, *85*, 11725−11731.
(23) Tziotis, D.; Hertkorn, N.; Schmitt-Kopplin, P. Kendrick-Analogous Network Visualisation of Ion Cyclotron Resonance Fourier Transform Mass Spectra: Improved Options for the Assignment of Elemental Compositions and the Classification of Organic Molecular Complexity. *Eur. J. Mass Spectrom.* **2011**, *17* (4), 415−421.
(24) Smirnov, K. S.; Forcisi, S.; Moritz, F.; Lucio, M.; Schmitt-Kopplin, P. Mass Difference Maps and Their Application for the Recalibration of Mass Spectrometric Data in Nontargeted Metabolomics. *Anal. Chem.* **2019**, *91* (5), 3350−3358.
(25) Kanawati, B.; Bader, T. M.; Wanczek, K.-P.; Li, Y.; Schmitt-Kopplin, P. Fourier Transform (FT)-Artifacts and Power-Function Resolution Filter in Fourier Transform Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2017**, *31* (19), 1607−1615.
(26) Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal. Chem.* **2006**, *78* (13), 4281−4290.
(27) Thompson, C. J.; Witt, M.; Forcisi, S.; Moritz, F.; Kessler, N.; Laukien, F. H.; Schmitt-Kopplin, P. An Enhanced Isotopic Fine Structure Method for Exact Mass Analysis in Discovery Metabolomics: FIA-CASI-FTMS. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (10), 2025−2034.
(28) Chekmeneva, E.; Dos Santos Correia, G.; Chan, Q.; Wijeyesekera, A.; Tin, A.; Young, J. H.; Elliott, P.; Nicholson, J. K.; Holmes, E. Optimization and Application of Direct Infusion Nanoelectrospray HRMS Method for Large-Scale Urinary Metabolic Phenotyping in Molecular Epidemiology. *J. Proteome Res.* **2017**, *16* (4), 1646−1658.

https://doi.org/10.1021/jasms.2c00082
J. Am. Soc. Mass Spectrom. 2022, 33, 2203−2214