Talanta 286 (2025) 127514

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Continuum of non-targeted data for long term study of complex samples generated by direct infusion ultra-high resolution mass spectrometry

Susanne E. Woodward [a,1], Stefan A. Pieczonka [b,c,1], Jasmine Hertzog [b,c], Richard Haydock [a], Mary J. Thomas [a], Chloé Roullier-Gall [b,c], Ciaran O'Flynn [a], Jenny Uhl [b,c], Michael Rychlik [b], Philippe Schmitt-Kopplin [b,c,**], James W. Marshall [a,*]

[a] Waltham Petcare Science Institute, Mars Petcare, Freeby Lane, Waltham on the Wolds, Melton Mowbray, LE14 4RT, United Kingdom
[b] Technical University of Munich, Analytical Food Chemistry, Maximus-von-Imhof Forum 2, 85453, Freising, Germany
[c] Helmholtz Munich, Analytical BioGeoChemistry, Ingolstädter Landstr 1, 85764, Neuherberg, Germany

ARTICLE INFO

Handling editor: A Campiglia

Keywords:
FT-ICR-MS
Non-targeted
Complex mixtures
Data matrix
Continuum
Food analysis
-Omics

ABSTRACT

Non-targeted chemical analysis is a powerful tool for exploration of the unknown chemistry of complex matrices such as food, biological, geochemical, environmental and even extra-terrestrial samples. It allows researchers to ask open, unbiased questions about their system chemistry. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS) offers these options and has been widely used to study complex mixtures, with its unmatched mass resolution enabling direct infusion methods and eliminating the challenges of chromatographic alignment in large-scale longitudinal projects. In this article, we demonstrate the use of FT-ICR-MS for generating a dataset for hundreds of complex samples with diverse chemistries over a period of 6 years and 32 batches, allowing confident comparison of data between samples from different batches. The resulting chemical database will be continuously expanded in future and retrospectively interrogated to test new hypotheses utilizing data of past projects and new knowledge of coming projects. We discuss the experimental setup and how other researchers could apply the same approaches, which is relevant for wide ranging applications and research fields.

1. Introduction

The chemistry of biological systems and the environment is complex and diverse; non-targeted chemical analysis is a powerful way for researchers to explore the unknown chemistry of complex systems. Over the last decades, non-targeted metabolomics, petroleomics and related approaches have been used to study samples from diverse fields including metabolism, the composition of natural organic matter, environmental samples, extra-terrestrial samples and foods [1–4]. The unique value of non-targeted analysis is that it does not presume the composition of the sample, allowing researchers to explore the key aspects of their system's chemistry as opposed to merely testing the significance of certain known or expected molecules or chemical classes. Conducting such hypothesis-generating studies often requires an extensive sample set and a long-term measurement effort that exceeds individual batches. Comparing non-targeted data across studies or over time is challenging, even within datasets generated by the same laboratory. Ultra-high-resolution methods such as Direct Infusion Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (DI-FT-ICR-MS) have been extensively used to study complex mixtures, with no prior separation [1,3,5–12]. Ultra-high resolution and parts-per-billion (ppb) mass accuracy means FT-ICR-MS is particularly well suited for determination of the molecular composition of complex biological and geochemical sample matrices. The absolute mass of molecules, determined by FT-ICR-MS, can be used to achieve alignment between samples and produce continuous data sets [4,13–18]. Although the potential to align FT-ICR-MS data on large sample sets has been well described, the intentional joining of FT-ICR-MS data from many discrete sample sets of the same type, into a continuum of complex chemical data has not been demonstrated. The high resolving power and mass accuracy of FT-ICR on the half electron mass scale enables the formal assignment of exact *m/z* signals to chemical formulae and if the mass resolution allows,

* Corresponding author. Waltham Petcare Science Institute, Mars Petcare, Freeby Lane, Waltham on the Wolds, Melton Mowbray, LE14 4RT, United Kingdom.
** Corresponding author. Helmholtz Munich, Analytical BioGeoChemistry, Ingolstädter Landstr 1, 85764, Neuherberg, Germany.
E-mail address: james.marshall@effem.com (J.W. Marshall).
[1] Equally contributed to the work.

https://doi.org/10.1016/j.talanta.2024.127514
Received 18 November 2024; Received in revised form 20 December 2024; Accepted 31 December 2024
Available online 1 January 2025

analysis of isotopic fine structure [19,20]. Each sample's spectrum is a representation or 'fingerprint' of the chemical composition with relative abundance for each mass signal - a summary of the individual chemistry of the sample, annotated at a molecular formula level.

In this non-targeted study, using pet food as a commercially relevant example of a highly complex mixture, over six years, we built a DI-FT-ICR-MS data set of the diverse chemistry of 384 samples from a range of manufacturers, brands and flavours. The generated data set shows the extent to which alignment of complex, non-targeted, chemical data, acquired longitudinally, is possible in well controlled studies which use FT-ICR-MS demonstrating the possibility to monitor complex chemistry of many types of samples longitudinally using our approach. Pet food is a biologically derived material containing mineral additives, and which has undergone thermal processing which leads to complex chemical reactions, and these qualities make pet food an inherently complex and challenging matrix with a wide dynamic range and significant potential for adverse ionization interferences. Samples of pet food contain diverse chemistry including, but not limited to, sugars, amino acids, peptides, fatty acids and derivatives, nuceotides, and Maillard reaction products. Findings from the study of pet food would be expected to be transferable to other types of samples and research fields.

2. Experimental section

2.1. FT-ICR-MS

High resolution mass spectra were acquired on a FT-ICR-MS (SolariX™, Bruker Daltonics GmbH, Bremen, Germany) equipped with a 12 T superconducting magnet (Magnex Scientific Inc., Yarnton, GB) and an APOLLO II electrospray ionization (ESI) source (Bruker Daltonics GmbH, Bremen, Germany) operated in negative mode. Further detailed instrumental settings can be found in Supplementary Information 1. During instrument optimization for analyzing pet food, negative mode ESI yielded the widest range of analytes for non-targeted analysis. Further detailed instrumental settings can be found in Supplementary Information 1. Negative mode ESI has been used to demonstrate the longitudinal aspect of the methodology and capabilities of the instrument regarding quality control and data processing. Other ionization and operating modes are available and may be more appropriate for other sample types and research questions.

Sample extracts were diluted 1:100 in pure methanol (LC-MS grade, CHROMASOLV™, Sigma-Aldrich) prior to injection and introduced into the ESI source at a flow rate of 120 μL h^{-1}. The 4 Megawords data sets were acquired within a detection range of *m/z* 90–1000. A total of 400 scans were acquired before the data were zero-filled once and sine apodised [21]. The resolving power consistently maintained its value at around 500,000 at *m/z* 400.

The pet food data was acquired in 32 batches of up to 60 samples run in triplicate. In each batch of samples, the analysis order was randomised so that the replicates of each sample were randomly distributed throughout the sequence.

2.2. Data processing and analysis

Raw spectra were post-processed by Compass DataAnalysis 4.2 (Bruker Daltonics, Bremen, Germany) and the RStudio (Posit PBC, Boston MA, USA) programming environment.

Internal accurate mass calibration of data was performed using an in-house list of 3000 recurrent compositions commonly found in FT-ICR-MS spectra of the samples. Linear calibration, with mass error below 0.1 ppm was achieved in the range of the calibration list (*m/z* 90–600) and extrapolated to the whole mass range. Space charge effects were compensated by subsequent Kernel density-based calibration [22]. Acquired masses were filtered to remove peaks below a signal-to-noise ratio of 6 and an absolute intensity threshold of 2,000,000. Exported *m/z* features were filtered with respect to FT-side loop artefacts [23,24]. Heavy isotopologues were filtered based on their specific mass difference and expected intensities on single spectrum basis. Aligned filtered mass lists were aligned within a window of ±1 ppm in a matrix containing averaged *m/z* values and corresponding peak intensities. After using Matrix Generator in-house software (maximum error 1.0 ppm), statistical analysis was carried out on aligned raw data peaks across all samples. For statistical analysis, data pre-treatment included zero-filling, zscore- normalization, and scaling.

The data matrix was processed to assign formulae using an in-house developed software tool (NetCalc) [25], annotate known metabolites, and remove spectra signals not be explained by combinations of elements C, H, N, O, P, S and Cl given $z = -1$ [21,26]. Only molecular formulae detected in two of three replicates of one sample were included in the exported matrix [6].

Exported *m/z* signals processed from each batch were grouped, to form features, using k-means clustering. A second pass was applied to wider feature groups containing multiple separate compounds, providing the new group widths were at least 0.2 ppm with at least 5 % of the larger group *m/z* values. These stopping criteria were identified by manually inspecting a series of *m/z* grouping results using a range of values to identify the most plausible groups. Following this process 44,712 features or *m/z* groups were identified. For each sample, peak intensities were normalized to the sample total intensity resulting in relative intensity values. Relative intensities of each feature per sample were used to form a matrix, which contained a single row per sample with the average relative intensity across the three replicates for each feature as columns. Due to the method by which this matrix was created, new batches of data can easily be added by allocating each detected *m/z* the appropriate feature group and adding new rows to the matrix. Approximately 1.0–5.5 % of *m/z's* in a typical new batch do not align to a feature group and are discarded to keep a stable matrix.

2.3. QC material preparation

All QC products were prepared as follows: 10 × 100 g pouches from a single factory production were homogenized (Robot-Coupe R2, Robot-Coupe Ltd, Isleworth, UK), and 20 g aliquots of homogenized material were frozen (−80 °C) to stabilize the sample until thawed for extraction. All batches of QCs were prepared from a sub-set of the same original 87 single manufacturing batch of pet food pouches. QC1 is a real meat pieces, beef variety and QC2 is a real fish pieces, tuna variety. Three batches of each QC homogenate were prepared over the course of the extended experimental timeframe (February 2015 (B1), June 2016 (B2), September 2018 (B3)). Material prepared in B1 and B2 was within the two years Best Before End date (BBE) whereas B3 was prepared after expiration of the BBE date. Analysis batches 17–23 include QCs from preparation batches 2 and 3, as cross-over checks were run to establish continuity of QC material after expiration date.

2.4. Sample extraction

All samples were extracted and analysed in triplicate. All product samples were analysed blinded, a coded table categorizing the samples by manufacturer, brand and other parameters is included in Supplementary Information 2. During method optimization, the following extraction procedure was determined to be the most robust and repeatable, which is important for long term research studies. The extraction was found to be suitable for a wide variety of pet food samples and formats, and able to provide a very diverse extract that captured chemicals from many compound classes of interest such as fatty acids, nucleotides, amino acids, peptides, sugars and Maillard reaction products [21]. A visual representation of the chemistry of pet food is demonstrated in Van-Krevelen diagrams of the common analytes found in the QC samples used in this study (Supplementary information 3). Samples and QCs (5 g) were homogenized with methanol: water (1:1; 20 mL, CHROMASOLV™, for HPLC, = 99.9 %, Honeywell Riedel-de Haën)

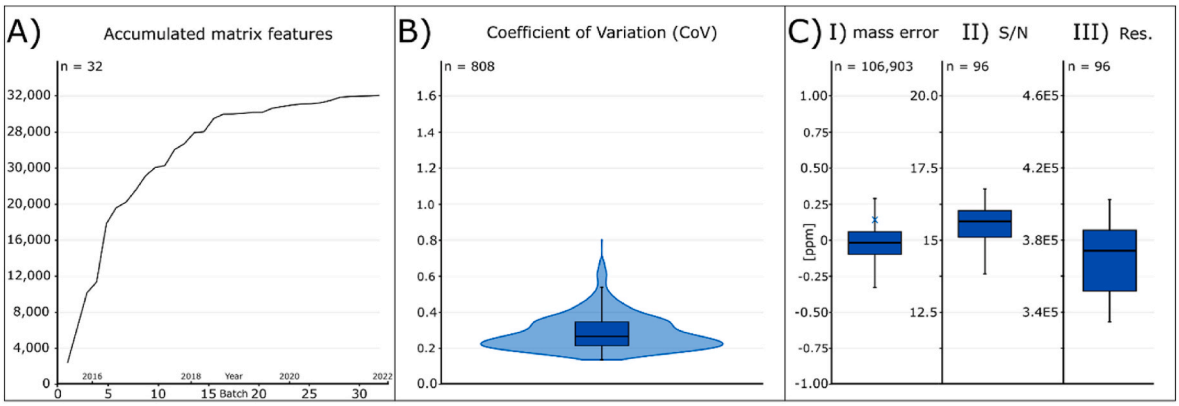

Fig. 1. The quality evaluation of the overall matrix was based on triplicate measurements of QC samples taken at the beginning of each of the 23 batches over a 6-year period. The accumulated features of the overall matrix show a saturated plateau phase over time (A). The coefficient of variation for common annotated features displays a uniform distribution, with only a few compounds exhibiting higher variability (B). The distribution of annotation mass errors and a very low average mass error of 0.13 ppm (highlighted with an 'x') confirm consistent mass accuracy (C–I). The median signal-to-noise ratio (C-II) and resolving power values (C-III) further validate the stable spectral quality.

and hexane (10 mL, CHROMASOLV™, for HPLC, ≥97.0 %, Honeywell Riedel-de Haën). The non-soluble lipid fraction (hexane phase) was separated by centrifugation (11,000 g). An aliquot (4 mL) of the methanol-water phase was dried (SpeedVac, Thermo Fisher Scientific, Loughborough, UK). Samples were reconstituted in HPLC grade water (2 mL), centrifuged and filtered through a SFCA syringe filter (0.2 µm; Thermo Fisher Scientific) then sealed and stored at −80 °C prior to FT-ICR-MS analysis [21]. Gold standard criteria were defined for extract storage temperature, storage duration, and shipping conditions. These measures allowed as much experimental control as possible and were applied to all samples analysed in this study. All extracts were stored at −80 °C for a maximum of 12 weeks prior to being shipped overnight on dry ice (temperature monitored) for analysis.

2.5. Statistical data analysis and data visualization

Coefficients of Variation (CVs) for each signal of the batch matrix were used as a data integrity check prior to integration of new batches to the longitudinal data matrix. Principal Component Analysis (PCA) was applied to reduce the dimensions of the resulting data set. The corresponding score plot of the two most influential Principal Components was utilized to visualize the diversity and common patterns of all 384 commercial pet foods. A second PCA was applied solely to QC samples to evaluate the consistency of the measurements over 6 years and identify potential sources of variation. Box plots were utilized to visualize intensities of the most significant signals that contributed to the differentiation between QC clusters (PC1) and QC homogenate batches (PC2). Significance of intensity differences was assessed by a one-sided *t*-Test ($\alpha = 0.05$, $p < 0.001$). The fifth percentile of molecular compositions most significant for respective QC clusters were visualized in a van Krevelen (VK) diagram. Plotting H/C versus O/C atomic ratios enables tentative classification of resolved metabolite signals and indicates molecular classes present [27]. Statistical analysis was performed using RStudio (Version 1.4.1103, Boston, USA) and SIMCA 13.0.3.0 (Umetrics, Umeå, Sweden).

2.6. Evaluation of the data matrix

The overall matrix was evaluated based on the total number of features accumulating with the increasing batch number, the coefficient of variation (CoV) of the intensities of annotated features found in 90 % of the QC triplicates measured at the beginning of each batch (n = 808), the annotation error of all QC features (n = 106,903), the median signal-to-noise ratios of the spectra (n = 96), and their median resolving power.

3. Results and discussion

Three hundred and eighty-four commercial pet foods were analysed over six years in 32 batches of up to 60 samples. Two (different) commercially available pet foods were used as quality control (QC) samples. The QC samples were analysed alongside the samples of each analysis batch to check the reproducibility of the analytical process, from sample preparation to data analysis. Data from the QC samples provided a reference for quality assurance of the new data.

For the generation of a continuum of non-targeted, multivariate chemical data, as with the FT-ICR-MS data in this study, the quality of each batch of new data must be controlled and evaluated to ensure interoperability with old data. In this study, consistently high-quality interoperable data was achieved over 6 years by carefully controlling the analysis methodology, which was developed to be robust and repeatable, from the extraction of the sample through to the acquisition, cleaning and alignment of the FT-ICR-MS data, and the data analysis.

3.1. Batch quality assurance

Non-targeted analysis with FT-ICR-MS generates data which shows the unique chemistry present in each sample. Our approach used the composition of QC1 and QC2 as a reference to quality assure each new batch of samples; if the QC sample data in new batches was typical and true to type, the new sample data could be trusted by extrapolation. First the QC samples were assessed for variability across features as a data integrity check [21]. Coefficients of variation (CV) for each mass detected in the three replicate extracts of each QC sample were calculated after each batch of new data was acquired. Typically, the CV of both QC samples in each batch were below 0.3 for more than 97 % of features. A CV over the run time of each batch was calculated based on the first Q1 replicate 1 (R1), re-injected every 12 samples throughout the sequence. Spectra are routinely consistent throughout, confirming repeatability over time of the FT-ICR-MS analysis (cā 50 h). Typically, the CV of Q1 (R1) over time was below 0.3 for more than 95 % of features.

3.2. Evaluation of the data matrix

Between 2015 and 2021 32 new batches of samples were analysed and the new data joined into a continuous compositional data matrix (Fig. 1). Based on the triplicate QC measurements at the start of each batch (n = 32 x 3), the quality of the overall matrix and its annotated features was assessed. The coefficient of variation (CoV) of the annotated features present in at least 90 % of the QC samples after exclusion

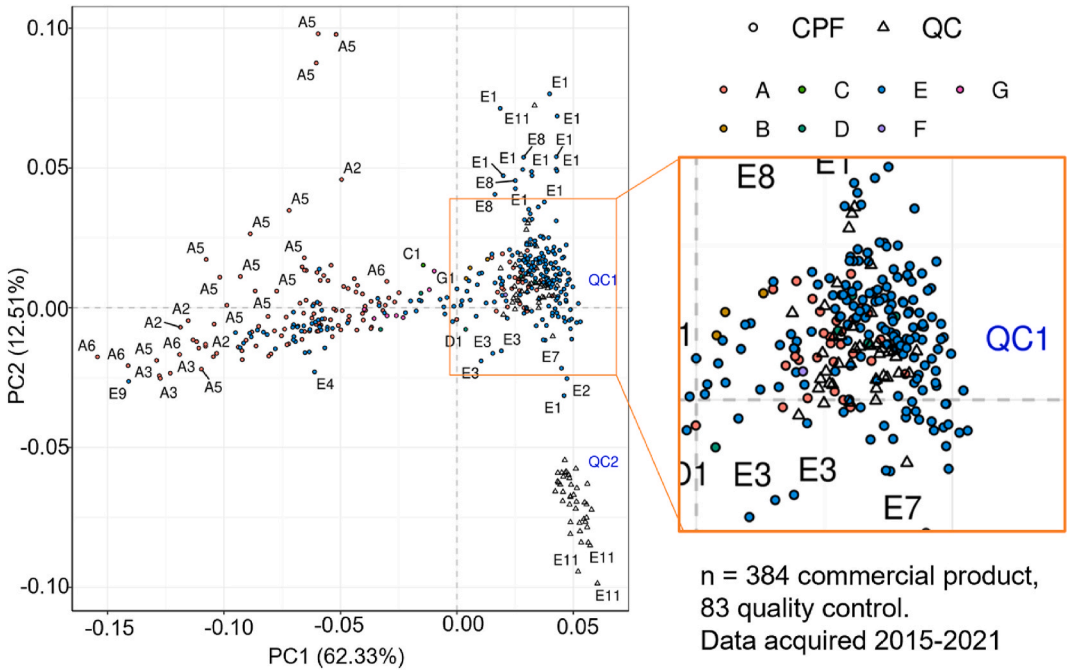

Fig. 2. PCA of 384 commercial pet food products over a six-year period with corresponding 83 QC samples analysed by DI-FT-ICR-MS. Commercial products (circular points) are labelled by letter (manufacturer) and number (brand). QC samples (triangles) are annotated as type, QC1 or QC2 and group separately. The two most represented manufacturers (A and E) dominate the separation across PC1, with other manufacturers distributed throughout the plot. Highlighted section enlarged to show QC1 data.

of outliers was 0.39 ± 0.14. Remaining intensity fluctuations of a few more variable compounds were mitigated by the triplicate measurement of all samples. The distribution of mass errors for annotated features across all measurements (total = 106,903), with an average of 0.02 ± 0.2 ppm, confirms the absence of systematic annotation errors. The absolute error values, averaging 0.13 ± 0.15 ppm, demonstrate exceptional mass accuracy over a span of 6 years. The median signal-to-noise ratio (S/N) of the measurements (n = 96) of 15.91 ± 1.44, and the median resolving power of 371,362 ± 20,369 further confirms the controlled consistent spectral quality.

3.3. *Interrogation of the combined data*

Analysis of complex samples using FT-ICR-MS generates complex, qualitative chemical data (for context the number of features detected in at least 10 % of all QC measurements in this data set are 5498 and 5283 (QC1 and QC2 respectively)). When Principal Components Analysis (PCA) is performed on the combined data (384 commercial pet foods and 83 repeated measurements of QC samples acquired in 32 batches over 6 years, 2015–2021) similar products cluster together (Fig. 2). The two QC samples have different chemistry and therefore cluster in separate areas of the PCA plot. Common features detected across QC1 (in at least 40 of 41 samples) and QC2 (39 of 40 samples) (plotted in VK diagrams [27] Supplementary Information 3), had an average absolute mass error of <0.05 ppm, demonstrating exceptional alignment of the data over six years. The compositional differences between QCs are predominantly evident in chemical class areas of lipids, peptides, carbohydrates, and amino sugars (such as Maillard reaction products). The pet food samples from the two most represented manufacturers A and E sit in separate areas of the PCA plot, separated on Principal Component 1. The extent of variability of the complex samples over the six year period is apparent in sample category A5 (manufacturer A brand 5 (Supplementary Information 2)), which is broadly distributed across the PCA plot relative to the well clustered QC samples (Fig. 2). The distribution of A5 is driven by diversity in the category, which includes products designed to meet the needs of 3 life-stages (having different macronutrient profiles), nine varieties and is made across four different manufacturing sites.

The data presented was acquired over a six-year time frame and several potential sources of variability needed to be managed during the experimental timeframe, for instance, staff turnover and hardware upgrades. A table summarizing leading factors which had the potential to introduce variability into the data is included in (Supplementary Information 4). Diligent control of the methodologies used for sample preparation, analysis, data processing and staff training mitigated the impact of these changes on the results. While the chemical composition of large numbers of samples have been interrogated using DI-FT-ICR-MS before [28], this study compares the composition of samples acquired over a 6-year timeframe. The QC material itself was re-prepared (from the same original batches of pet food) twice during the six-year experimental time frame. The data from analysis of the QC samples shows small differences in the composition of the QC material between the preparations, likely due to sample age at time of preparation; different people preparing the QC material or slight random variation between the single servings selected and pooled to produce each QC (10 single serve portions were combined and homogenized to create aliquots to use as QC material, each time the QC was prepared). When PCA is carried out on the QC samples on their own (Supplementary Information 5), 74 % of the variation is explained by the first two components. PC1 accounts for 62.33 % of the variability, which primarily relates to the compositional difference between the two QCs (QC1, QC2). PC2 represents 11.71 % of the variation and represents the differences within the same QCs (QC1 and QC2) over the different batches. When the loadings of the PCA of the QC samples are extracted and the intensities of the signals with the highest contribution to the second PC are evaluated, the variability of the signals between the different preparations of QC material is small (Fig. 3).

The formulae of the compounds whose signals drive most of the variability of the QC material over time are consistent with sugars and nucleotides (Fig. 3). It is possible that randomly sampling ten servings of each of these pet foods leads to slight differences in the sugar and nucleotide compounds observed due to the natural variability of food

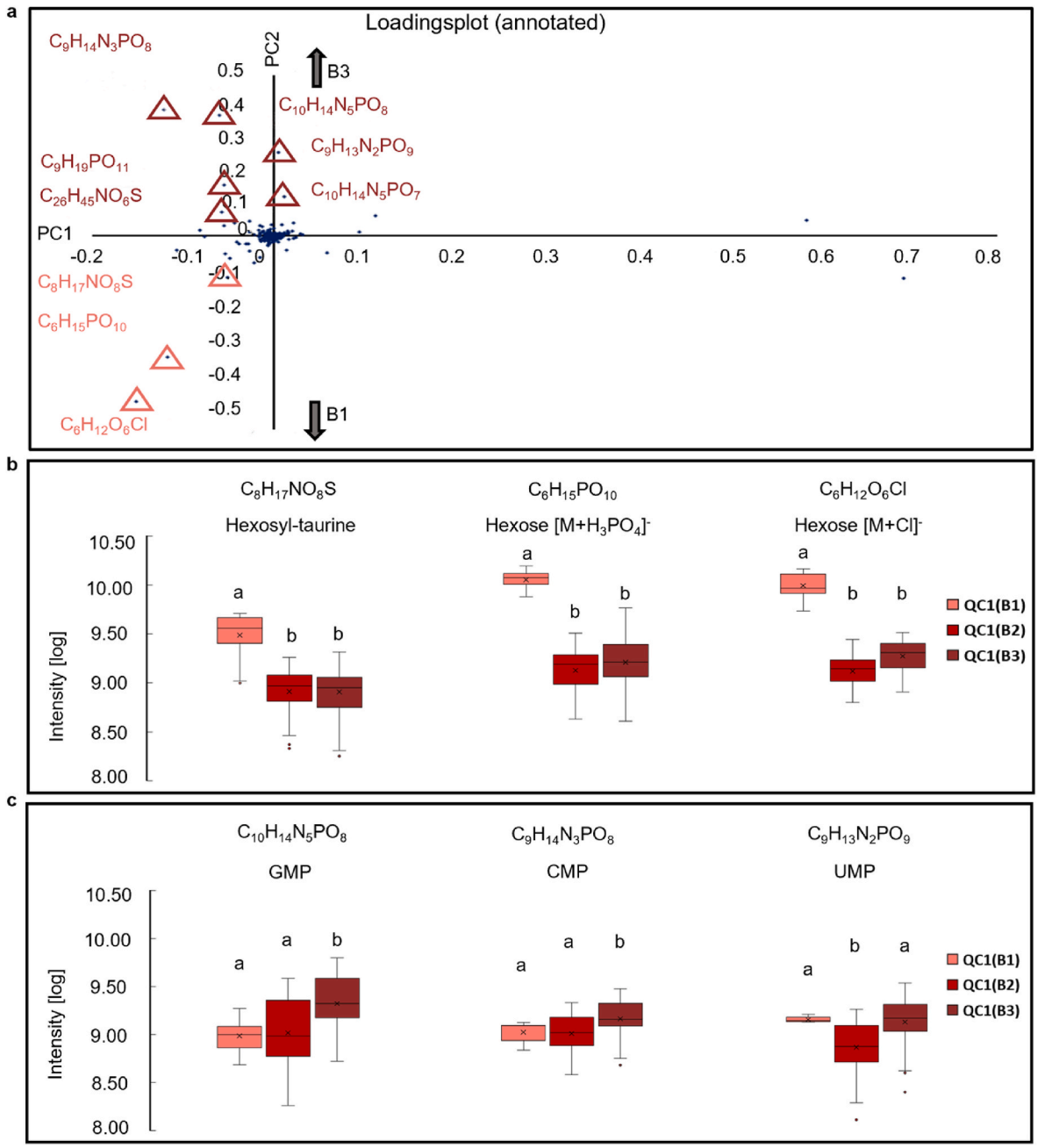

Fig. 3. Analytes driving separation between batch preparations of QC samples; a) Loading plot with highest contributing analytes annotated (early to late preparation pink-red). b) and c) box plots of log intensity of the three analytes with the greatest loadings for the three preparations of QC1 [B1, B2, B3]. Statistical significance ($p \leq 0.05$) denoted by a, ab, b.

ingredients which causes a variability between single servings. On the other hand changes in sugar compounds such as hexose $[M+Cl]^-$, hexose $[M + H_2PO_4]^-$, hexosyl-taurine could be due to chemical reactions such as the Maillard reaction [6,29] which may occur in the QC material stored at room temperature between preparations of QC material. Similarly, degradation of the food matrix could result in more abundant nucleotides such as cytosine monophosphate (CMP), guanosine monophosphate (GMP) and uridine monophosphate (UMP) which increase in signal intensity in later preparations of QC. Whatever the reason for the variability observed in the QC samples overtime, the variability is low, especially within the context of the real variability of the commercial pet food samples.

4. Conclusion

The consistency of the DI-FT-ICR-MS analysis allows data from individual experiments to be transformed into a continuum of complex chemical data. Such a continuum of data can be increased over time and can be interrogated as many times and in as many ways as necessary to investigate new theories and test hypotheses. Repetitive analysis of two QC samples, consistently extracted and analysed with each batch over new samples over six years, shows it is possible to achieve consistent long-term non-targeted measurement of hundreds of complex samples using DI-FT-ICR-MS.

The data points corresponding to each of the QCs form their own consistent cluster when PCA is carried out on the entire combined data set (Fig. 2). Benefiting from tightly controlled extraction and

experimental methodologies, the data showed no significant variance attributable to changes in hardware or staff members preparing or analysing samples. Actual disparities in compositional information, encompassing thousands of mass signals, can be discerned in long-term studies, for example between QC samples, and between manufacturers (Fig. 2). Minor variability within QC samples could be directly attributed to differences between batches prepared on different occasions. Future studies would be improved by the preparation of larger volumes of QC samples from the outset (once frozen, the material remains stable under controlled storage conditions at low temperature).

When beginning a program of work aiming to generate a continuum of data using any methodology, it is essential that the critical methodological and instrumental parameters are optimized as a primary step. Although FT-ICR-MS is relatively holistic, different compositional classes and functional groups tend to exhibit a preference for specific ionization techniques and polarities; and various instrumental parameters, collectively exert a profound influence on the output data. Given the impact of instrumental set up and matrix effects, a data continuum must be generated from interoperable data for each sample type such that the individual samples are in comparable matrices. Once all parameters are carefully optimized and standardized, studies could be transferred to a new or different FT-ICR-MS instrument and future data could be longitudinally joined into the continuum.

Through this experiment, our carefully controlled, quality assured data could be longitudinally joined to demonstrate that improvements to the methodology, (which are likely as analytical instrumentation evolves) were at no detriment to the interoperability of the data. With the absolute constancy of molecular masses over time, careful control of high-resolution mass spectrometry parameters affords longitudinal interoperability. This approach could be applied to a wide range of fields of study and types of samples. Non-targeted data is an excellent means for refinement in research when acquired with future use and long-term value in mind, particularly when creating a data continuum to complement with other omics approaches. Small volume samples that require ethical sourcing can be rendered more valuable in the long-term.

CRediT authorship contribution statement

Susanne E. Woodward: Writing – review & editing, Writing – original draft, Visualization, Validation, Investigation. **Stefan A. Pieczonka:** Writing – review & editing, Visualization, Validation, Methodology, Investigation, Formal analysis. **Jasmine Hertzog:** Writing – review & editing, Formal analysis. **Richard Haydock:** Writing – review & editing, Visualization, Validation, Software, Investigation, Formal analysis, Data curation. **Mary J. Thomas:** Writing – review & editing. **Chloé Roullier-Gall:** Formal analysis, Writing – review & editing, Methodology. **Ciaran O'Flynn:** Data curation, Writing – review & editing. **Jenny Uhl:** Writing – review & editing, Formal analysis. **Michael Rychlik:** Project administration. **Philippe Schmitt-Kopplin:** Writing – review & editing, Methodology, Investigation, Data curation, Conceptualization. **James W. Marshall:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization, Data curation.

Funding source

This work was funded by Mars Petcare UK. Sample extraction was carried out at Mars Petcare UK, sample acquisition at Technical University of Munich/Helmholtz Institute and data analysis performed at both Mars Petcare UK and Technical University of Munich/Helmholtz Institute. Mars authors listed are employees of Mars Petcare UK, a manufacturer of pet foods. All other authors are or were, Technical University of Munich/Helmholtz Institute employees or students at the time of contributing to the study.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Amanda Jackson, for her contribution in the early stages of extraction method development, and Daniel Hemmler for his contributions to data curation during this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.talanta.2024.127514.

Data availability

The data that has been used is confidential.

References

[1] M.J. Thomas, et al., Petroleomic depth profiling of Staten Island salt marsh soil: 2ω detection FTICR MS offers a new solution for the analysis of environmental contaminants, Sci. Total Environ. 662 (2019) 852–862.
[2] K.S. Smirnov, et al., Challenges of metabolomics in human gut microbiota research, International Journal of Medical Microbiology 306 (5) (2016) 266–279.
[3] M. Rychlik, et al., Chapter 21 - foodomics assessed by Fourier transform mass spectrometry, in: B. Kanawati, P. Schmitt-Kopplin (Eds.), Fundamentals and Applications of Fourier Transform Mass Spectrometry, Elsevier, 2019, pp. 651–677.
[4] A. Aharoni, et al., Nontargeted metabolome analysis by use of fourier transform ion cyclotron mass spectrometry, OMICS 6 (3) (2002) 217–234.
[5] S. Nicolardi, et al., Developments in FTICR-MS and its potential for body fluid signatures, Int. J. Mol. Sci. 16 (11) (2015) 27133–27144.
[6] D. Hemmler, et al., Evolution of complex maillard chemical reactions, resolved in time, Sci. Rep. 7 (1) (2017) 3227.
[7] M.J. Thomas, et al., Comprehensive analysis of multiple asphaltene fractions combining statistical analyses and novel visualization tools, Fuel 291 (2021) 120132.
[8] D.C.P. Lozano, et al., Petroleomics: tools, challenges, and developments, Annu. Rev. Anal. Chem. 13 (1) (2020) 405–430.
[9] C. Junot, F. Fenaille, Chapter 11 - metabolomics using Fourier transform mass spectrometry, in: B. Kanawati, P. Schmitt-Kopplin (Eds.), Fundamentals and Applications of Fourier Transform Mass Spectrometry, Elsevier, 2019, pp. 325–356.
[10] J. Hertzog, V. Carré, F. Aubriet, Chapter 22 - contribution of Fourier transform mass spectrometry to bio-oil study, in: B. Kanawati, P. Schmitt-Kopplin (Eds.), Fundamentals and Applications of Fourier Transform Mass Spectrometry, Elsevier, 2019, pp. 679–733.
[11] S.A. Pieczonka, et al., Hidden in its color: a molecular-level analysis of the beer's Maillard reaction network, Food Chem. 361 (2021) 130112.
[12] S.A. Pieczonka, et al., On the trail of the German purity law: distinguishing the metabolic signatures of wheat, corn and rice in beer, Front. Chem. 9 (2021).
[13] E.B. Kujawinski, Electrospray ionization fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS): characterization of complex environmental mixtures, Environ. Forensics 3 (3) (2002) 207–216.
[14] A.G. Marshall, et al., Petroleomics: a test bed for ultra-high-resolution Fourier transform ion cyclotron resonance mass spectrometry, Eur. J. Mass Spectrom. 16 (3) (2010) 367–371.
[15] K.O. Nagornov, et al., High-resolution Fourier transform ion cyclotron resonance mass spectrometry with increased throughput for biomolecular analysis, Anal. Chem. 86 (18) (2014) 9020–9028.
[16] A.G. Marshall, C.L. Hendrickson, G.S. Jackson, Fourier transform ion cyclotron resonance mass spectrometry: a primer, Mass Spectrom. Rev. 17 (1) (1998) 1–35.
[17] Y. Cho, et al., Developments in FT-ICR MS instrumentation, ionization techniques, and data interpretation methods for petroleomics, Mass Spectrom. Rev. 34 (2) (2015) 248–263.
[18] M.L. Easterling, J.N. Agar, Chapter 3 - fundamentals, strengths, and future directions for Fourier transform ion cyclotron resonance mass spectrometry, in: B. Kanawati, P. Schmitt-Kopplin (Eds.), Fundamentals and Applications of Fourier Transform Mass Spectrometry, Elsevier, 2019, pp. 63–88.
[19] M.B. Comisarow, A.G. Marshall, The early development of Fourier transform ion cyclotron resonance (FT-ICR) spectroscopy, J. Mass Spectrom. 31 (6) (1996) 581–585.
[20] M.P. Barrow, W.I. Burkitt, P.J. Derrick, Principles of Fourier transform ion cyclotron resonance mass spectrometry and its application in structural biology, Analyst 130 (1) (2005) 18–28.

[21] J.W. Marshall, et al., Monitoring chemical changes during food sterilisation using ultrahigh resolution mass spectrometry, Food Chem. 242 (2018) 316–322.
[22] K.S. Smirnov, et al., Mass difference maps and their application for the recalibration of mass spectrometric data in nontargeted metabolomics, Anal. Chem. 91 (5) (2019) 3350–3358.
[23] B. Kanawati, et al., Fourier transform (FT)-artifacts and power-function resolution filter in Fourier transform mass spectrometry, Rapid Commun. Mass Spectrom. 31 (19) (2017) 1607–1615.
[24] B.M. Ruddy, et al., Elemental composition validation from stored waveform inverse Fourier transform (SWIFT) isolation FT-ICR MS isotopic fine structure, J. Am. Soc. Mass Spectrom. 24 (10) (2013) 1608–1611.
[25] D. Tziotis, N. Hertkorn, P. Schmitt-Kopplin, 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. 17 (4) (2011) 415–421.
[26] C. Roullier-Gall, et al., A grape and wine chemodiversity comparison of different appellations in Burgundy: vintage vs terroir effects, Food Chem. 152 (2014) 100–107.
[27] P. Schmitt-Kopplin, et al., Systems chemical analytics: introduction to the challenges of chemical complexity analysis, Faraday Discuss 218 (0) (2019) 9–28.
[28] S.A. Pieczonka, et al., On the trail of the German purity law: distinguishing the metabolic signatures of wheat, corn and rice in beer, Front. Chem. 9 (2021) 715372.
[29] M.S. Feather, Some aspects of the chemistry of nonenzymatic browning (the Maillard reaction), in: Chemical Changes in Food during Processing, Springer, 1985, pp. 289–303.