Identification of LC-MS metabolite compounds and Fragments using IROA (Isotopic Ratio Outlier Analysis)

Felice de Jong[1], Chris Beecher[1,2]
[1] IROA Technologies, Ann Arbor, MI, [2] University of Florida, Gainesville, FL

Overview

This study investigated the use of an IROA-based workflow which includes a matrix sample and IROA Discovery software together with an internal standard, to improve data quality, identify LC/MS metabolomic compounds, fragments and adducts.

Introduction

The standard workflow for unbiased metabolomics matches peaks against a library for identification and finds a large number of unknowns since fragments are rarely identified as such. Source variations yield fragmentation variation; therefore, spectral databases are of little use in the face of instrumentation variation, and data impurities stemming from artifacts, noise and ion-suppression. We have developed a general method for associating all fragments, adducts, dimers, etc., that also makes the structure elucidation of unknowns easier.

The IROA protocol incorporates stable-isotopes into metabolites, creating stable-isotopic Internal Standards (IS) for each and every metabolite measured so that specific alterations can be accurately measured and quantitated. Biochemically-complex IS are generated using growth media wherein all natural abundance ^{12}C compounds (amino acids, sugars etc.) are replaced with randomly and universally 95% ^{13}C-labelled compounds (95% ^{13}C media) so when populations of cells are incubated in such media, all biological components in the cells, including metabolites, carry unique ^{13}C signatures. When an IS, so created, is added to natural abundance samples and analyzed by MS, each metabolite peak carries a ready identifier of its origin, an enhanced M-1, M-2 etc. for the 95%^{13}C IS and natural abundance M+1 for the ^{12}C (Experimental) sample. Because of the presence of the IS all adducts and fragments may now be correctly identified even within areas of high co-elution, by the characteristics of the IROA peaks without the need of a fragmentation library.
ClusterFinder software identifies IS parent ions and their collective fragments, adducts, and provides structural confirmation of metabolites since the masses and ratios between the IS and it's NA analogues will be a unique determinant for each such cluster. Furthermore, for each fragment the number of carbons present and their monoisotopic masses provide accurate formulae which supports structure elucidation of the parent compound in an unknown.

Acknowledgements

Special thanks to the Southeast Center for Integrated Metabolomics (SECIM) for the LC/MS analysis and Cambridge Isotope Laboratories for the isotopic raw materials.

This poster was made possible by the University of Florida's Southeast Center for Integrated Metabolomics (SECIM) through grant number U24DK097209 from the National Institute of Health's Common Fund metabolomics program.

IROA-based workflow

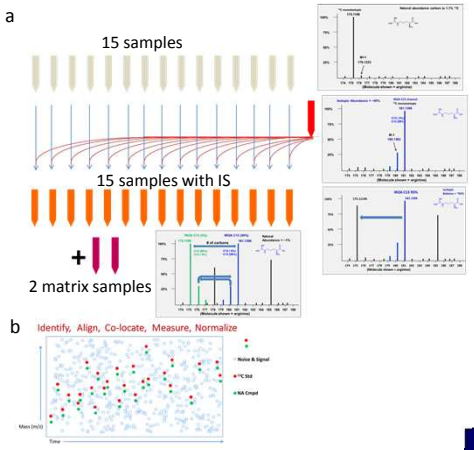


Figure 1 (above). The IROA workflow (a), use (b)

Alpha-aminoadipate

We find the following 8 ions for alpha-aminoadipic acid (a library lookup would find only 1):
- M+H-FA @ 116.071
- M+H-H2O @ 144.065
- M+H-NH3 @ 145.049
- M+H @ 162.076
- M+Na @ 184.058
- 2M+H @ 324.145
- 2M+Na @ 345.127
- 3M+H @ 484.214



Figure 2 (left). Example Dictionary entry for amino-adipic acid.

Workflow

The workflow consisted of (See Figure 1.):

1. The construction of an **authentic standard "dictionary" or library** for each chromatographic or MS data acquisition mode employed is accomplished by injecting authentic standards, and using the IROA Discovery™ software. Once completed the dictionary includes compounds, their fragments, adducts, dimers, trimers, *etc.*
2. The dictionaries generated above, were used to identify and name compounds, their fragments and adducts, *etc.* in two yeast extract (YE) IROA **matrix samples** labeled at both 95% and 5% ^{13}C. Note: The IROA Discovery software is also used to find natural abundance compounds in experimental samples.
3. Since the **matrix samples** are analyzed as part of the experimental sample-set and is run once for each state or mode of analysis it perfectly reflects the chromatographic and analytical qualities of the experimental samples. Therefore, these matrix samples may be used to create high-quality, concurrent mapping of 500+ compounds which are also **Internal standards (IS)** in each sample, likely to be found in each experimental sample, and serves as an alignment ladder for all non-IROA peaks.
4. In the experimental samples the IS is used to:
   - **a) Identify, align, co-locate, measure and normalize (Figure 1b) 500+ biochemicals in each experimental sample**; then by comparison all such samples are comparable through reference to the IS.
   - **b) Reduce variance and error:**
     - i. variance introduced during sample preparation
     - ii. variance introduced during sample analysis
     - iii. error introduced during data interpretation.
   - **c) Remove artifacts and noise, and eliminate ion suppression**
     - i. Experimental non-IROA paired peaks are authenticated by their location, relative to authenticated IROA IS peaks. Unpaired natural abundance artifacts and noise can be removed. The same degree of ionization enhancement or suppression due to the matrix will occur for both the target (experimental) analyte and the labeled IS.
   - **d) Identify fragments and adducts**
     - i. The dictionary was constructed on the same instrument, under very similar chromatographic conditions. When constructed all fragments, adducts, dimers, *etc.* were noted. Thus, **all of these will be named as such and not be considered unknowns.**

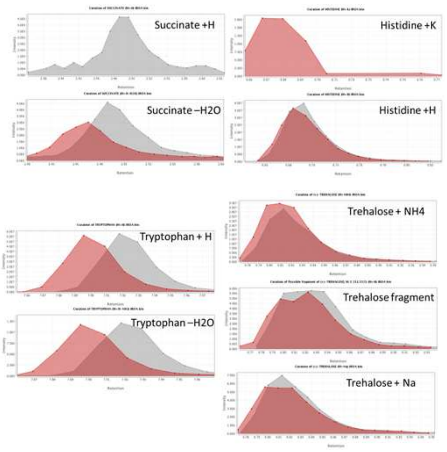


Figure 3 Compounds and named fragments that would otherwise be "Unknowns".

Experimental

Mixtures of meats, pork and beef, were made and homogenized in distilled water, and extracted using an 8:1:1 methanol:acetonitrile: acetone (MAA) protein precipitation method. An IROA-labeled yeast extract was added as a recovery Standard, and was the basis of the IROA-based workflow. The extracted samples were analyzed by reverse phase UHPLC-HRMS using an Ace Excel PFP-C 18 2.1x100 mm, 2.0 µm column. The UHPLC pump was a Dionex Ultimate 3000, and the mass spectrometer was a Thermo Fisher Q-Exactive Orbitrap. Solvent A was 0.1% formic acid in water, and solvent B was acetonitrile. The collected data were initially processed by Xcalibur Workstation software (version 3.0). Peak picking and compound identification was performed using IROA ClusterFinder™ software.

Conclusions

The Dictionary built in Step 1 contains between 1 and 8 entries per compound (See Figure 2). These entries, fragments, adducts, dimers, *etc.*, once put into the dictionary become identifiable peaks that would otherwise be classified as "Unknowns. (See figure 3)

The use of the IROA workflow assured that all named peaks were not artifacts which further improved the quality of the final dataset. (See Figure 1). The Matrix samples were easily mined for the exact location of all biological peaks in the IS. The IS in these experiments contained 500+ compounds. The samples were shown to contain 450 of these IS compounds. The absence of false peaks made the quality of the dataset extremely high. The dataset correctly predicted the concentrations of the mixed meats.

Summary

The use of the IROA workflow and dictionaries that were specifically compiled to include all adducts, fragments, dimers, *etc.* Created a dataset in which more peaks were named and which contained no artifactual peaks.

References

1. de Jong F, Beecher C, "Addressing the current bottlenecks of metabolomics: Isotopic Ratio Outlier Analysis (IROA®), an isotopic-labeling technique for accurate biochemical profiling", *Bioanalysis* 2012, 4(18), 2303-14.
2. Stupp GS, Clendinen CS, Ajredini R, Szewc MA, Garrett T, Menger RF, Yost RA, Beecher C, Edison AS. "Isotopic Ratio Outlier Analysis Global Metabolomics of Caenorhabditis elegans." *Analytical Chemistry* 2013 *85*(24), 11858-11865. DOI: 10.1021/ac4025413.