

Isotope Ratio Outlier Analysis (IROA) and Variable Window SWATH[®] Acquisition allows for Unambiguous Metabolite Identification

A Unique Labeling Technology (IROA) combined with TripleTOF[®] 6600 System for Untargeted Metabolomics

Chris Beecher^{1,2} & Felice A. de Jong² and Baljit K. Ubhi³

¹IROA Technologies, Boston, MA, USA, ²University Florida, Gainesville, FL, USA and ³SCIEX, Redwood City, CA, USA

Metabolomics focuses on the chemical processes central to cellular metabolism. Mass spectrometry and specifically data dependent workflows tend to be the choice for the measurement of these metabolites. IROA is an isotopic methodology in which all biological molecules are uniformly and randomly labeled to create informative isotopic patterns that are readily discriminated from artifacts (Figure 1). The IROA protocol generates an IROA Internal Standard providing specific molecular information so that the small biochemical molecules within biological samples may be easily and more accurately identified (Figure 2). Because of the uniform nature of the labeling, these patterns are revealed not only in the MS, but also in all fragments in any subsequent MS/MS. SWATH[®] acquisition, a data independent acquisition (DIA) workflow is well adopted in quantitative discovery proteomics, but still not commonly used in discovery metabolomics.

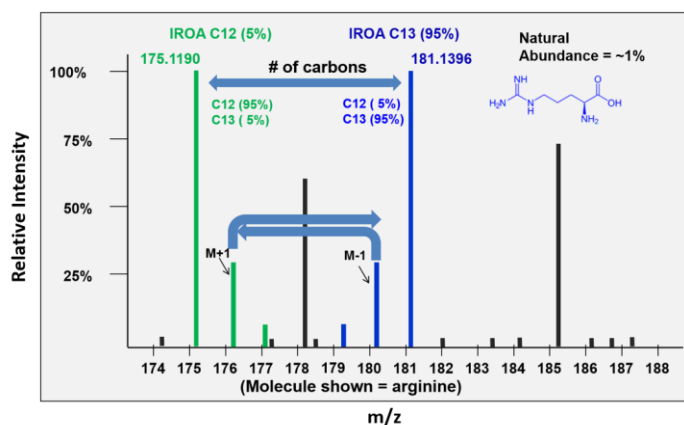


Figure 1: The Unique IROA Signature. The IROA peaks shown in this figure are for the 6-carbon molecule arginine. There are both the peaks from the experimental samples (shown in green) at 5% U-¹³C, and the Internal Standard or control samples (shown in blue) at 95% U-¹³C. Note the spacing of peaks is exactly a neutron mass apart, and the peaks are symmetrical. IROA peaks are discriminated from natural abundant artifactual peaks (shown in black). This system represents a triply redundant information system; the number of carbons is verified by the relative heights of the M+1 and M-1 and the distance between the monoisotopic peaks. The formula for the molecule is constrained by the number of carbons.

SWATH acquisition allows a user to collect MS and MS/MS of every detectable metabolite in their sample, thus creating a digital map of the metabolome.

Variable Window SWATH acquisition (an enhanced way of collecting MS/MS, using targeted mass windows in denser regions of the MS spectrum) allows for targeted specificity (Figure 4). The use of IROA with a SWATH variable window acquisition has allowed us to collect extremely advanced information for the biological components of a mixture, with significantly enhanced accurate identification and quantitation, clearly differentiating MS/MS IROA-SWATH peaks from artifacts. The MS/MS IROA pattern can be observed by varying the mass window overlap during SWATH acquisition which is unique for any data independent acquisition (DIA) approach. IROA can correctly assign formulae to all IROA peaks in both the MS and MS/MS scans. This is the first example whereby the correct formula is routinely found not only for the parent peak but also all fragments. Quantitation of any compound may be done at either the MS or MS/MS level.

Key Benefits of the IROA Technologies and SCIEX SWATH Acquisition Workflow

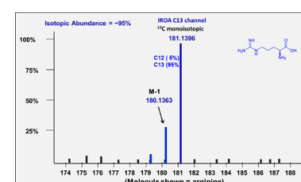
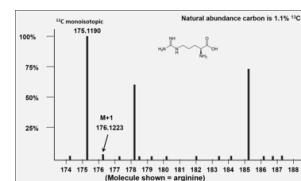
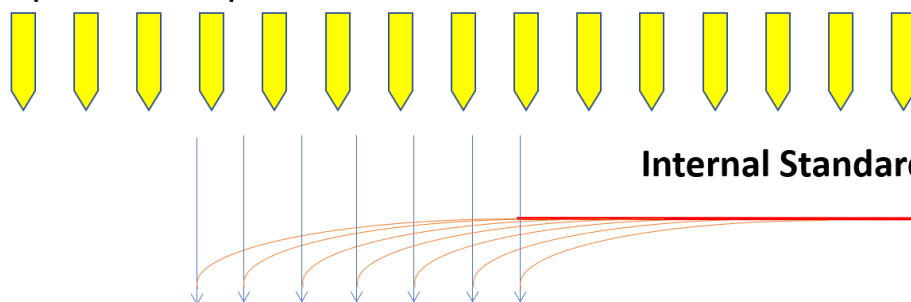
- Efficiently collect MS and MS/MS data in a single injection using the TripleTOF[®] System
- Unambiguously identify and accurately quantitate all detectable compounds of biological origin in every sample using the unique IROA pre-labeled Internal Standard Workflow Kit.
- Enhancement of data quality through data reduction by removal of background/noise
- Reduction of false discovery rate by excluding all non-biological peaks
- Ensures a high level of QA/QC for both ID and quantitation by monitoring a complete set of known metabolites every analysis
- Costs less than procuring individual metabolomics internal standards for every compound for use in every sample
- Automated software for data analysis using IROA Technologies ClusterFinder Software

Materials and Methods

The IROA Workflow Kit includes two (2) pre-labeled Standards. A 95% U-13C-labeled biochemically complex Internal Standard (IS) which contains 100's of biochemicals, each with an IROA isotopic pattern, was added to samples to accurately identify and quantitate complex mixtures without the need for baseline separation, and to overcome both sample-to-sample variances and ion suppression. A QA/QC Standard (Matrix) which contains both the IS and its balanced IROA 5% U-13C equivalent, was used to build a reference library of compounds and analyzed every 10 samples to create a Retention Index (RI) which provided an internal calibration for every sample and instrument. The identification of all IROA compounds and their fragments by ultra-high-resolution mass measurement made it possible to determine the empirical formula for all fragments. Data were collected from a TripleTOF[®] 6600 System in SWATH[®] acquisition using a variable window strategy in which defined windows of varying mass ranges were applied in areas of the chromatogram where there were many co-eluting ions. SWATH acquisition fragmentation of the IROA peaks completely differentiated fragments, and artifacts.

Chromatography: The reverse phase HPLC separation was performed using a Shimadzu LC System, operating at a flow rate

Experimental Samples



Experimental Samples with IS

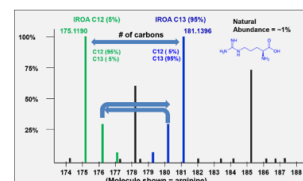
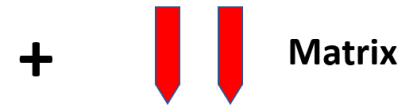
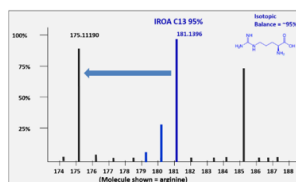


Figure 2: The IROA Workflow Protocol. The experimental samples are of natural abundance and the Internal Standard (IS) is labeled at 95% ¹³C. The IS contains hundreds of primary and secondary metabolites that may be used to quantitate and prove the identity of isotopomeric natural abundance compounds in the experimental samples. The heights and distances between IROA peaks are all mathematically calculable and IROA ClusterFinder Software employs algorithms to remove irrelevant data, identify and quantitate molecules of interest. The kit-supplied Matrix Standards are QA/QC samples that not only provide for cleaner faster identification, but also provide day-to-day QA/QC and inter-day conversion to compensate for instrumentation variance. SWATH acquisition provides unambiguous identity and the ability to assuredly quantitate at either the MS or MS/MS level.

of 350 $\mu\text{L}/\text{min}$. The column used was an Ace Excel C18-PFP column (100 x 1mm, 2 μm) from ACE, maintained at 30 $^{\circ}\text{C}$. A standard reverse phase gradient was used employing mobile phase A as 0.1% formic acid in water and mobile phase B as acetonitrile. The injection volume was 3 μL in positive ion mode and 5 μL in negative ion mode.

Mass Spectrometry: The data was collected using SWATH[®] acquisition on the TripleTOF[®] 6600 System. Using optimized source conditions, the MS mass range analyzed was 50-1000 m/z and the MS/MS was acquired with a mass range of 40-1000 m/z with a 25 msec accumulation time. The mass window overlap was set across the mass range to reflect the diversity of labeled metabolites. The collision energy was set to 35 V with a 15 V collision energy spread.

Data Processing: The IROA ClusterFinder software was adapted to automatically handle the SWATH[®] acquisition data by 1) finding the appropriate SWATH acquisition windows for every IROA peak, 2) examining the appropriate scans for their IROA fragments, 3) interpreting the IROA information for both MS and SWATH scans, and 4) determining the relationships between all IROA peaks. In order to assure that all IROA peaks would be captured, a specific SWATH acquisition window protocol was applied (See Figure 3). This windowing protocol assured that, unlike other MS/MS selection protocols, for every IROA peak there was an optimal fragmentation scan. Once ClusterFinder identifies an IROA peak it automatically retrieves the correct SWATH acquisition scan. This is an extremely efficient and accurate workflow. If all IROA fragments are sorted by descending mass, since they all have formulae, the full fragmentation history of the molecule is realized. This information is available for all molecules. It may be used support the identification of unknowns, adducts, fragments, polymeric species, etc.

Max # carbons below center	center	window	overlap	count	min	max	SWATH Window	Overlap
3	59	10	5	1	49	69	Window = 20	= 10
4	74	15	8	2	59	89	Window = 30	= 15
5	89	15	8	3	74	104		
7	119	30	15	4	89	149		
9	149	30	15	5	119	179		
13	209	60	30	6	149	269		
17	269	60	30	7	209	329		
27	389	120	60	8	269	509		
36	509	120	60	9	389	629		
54	749	240	120	10	509	989		

Figure 3: SWATH[®] Acquisition Variable Window example. Highlighting the window size in red arrows and the mass window overlap in green arrows

Uniquely IROA Labeled MS/MS Spectra

Applying a variable window SWATH acquisition strategy to the data analysis of an IROA Internal Standard (IS) spiked-sample made it possible to unambiguously identify and accurately quantify hundreds of detectable metabolites in a single unbiased metabolomics analysis using ClusterFinder Software. The IS contains 500+ well characterized metabolites, which migrated in an HPLC separation with their natural abundance isotopomers, and enabled both identification and standard quantitation for accurate measurement even in a non-baseline, “unbiased” metabolomics separation. Using traditional DIA, all compounds with the same retention time are fragmented without selection, minimal benefit is gained and the process is manually work-intensive.

SWATH acquisition subjects all ions within a precisely selected m/z window (see Figure 3) to fragmentation allowing specific precursor ions to be selected, making it easier to analyze fragmentation spectra. A corresponding spectral library of metabolites is however required for accurate identification. As is shown in Figure 4 variable window SWATH acquisition is a variant of SWATH acquisition in which the centers and widths of the SWATH acquisition windows are defined according to the minimum and maximum number of carbons in any metabolite that may appear in a given window.

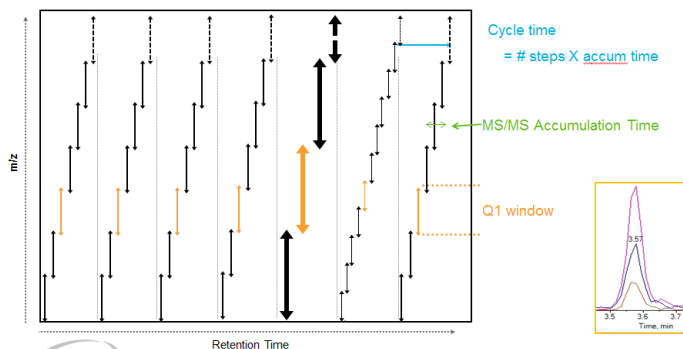


Figure 4: Investigating Variable Q1 Window Widths for SWATH Acquisition. During SWATH acquisition, wider width Q1 windows are stepped across the mass range, and high resolution MS/MS is acquired for a specific accumulation time (top). To achieve better specificity in complex matrices, smaller Q1 windows are desirable especially in the m/z dense regions where many peptide precursors are measured. The m/z density histograms constructed from the TOF MS data for the metabolome of interest (bottom, blue line) can be used to construct variable sized windows, where the density of precursors in each of the isolation windows is equalized across the m/z range.

Uniquely-labeled IROA metabolites are captured within a variable SWATH acquisition window are subjected to fragmentation (see Figure 5). IROA fragments and adducts will

also show IROA patterning and therefore the number of residual carbon atoms and the formulae for all fragments will be known. All artefactual (stray non-IROA) peaks captured from within the SWATH acquisition window will be identified as irrelevant and ignored. If two or more IROA peaks are captured within the window the ratios of the parent monoisotopomers will be inherited by all fragments allowing them to be easily sorted. The data may be quantitated based on MS or MS/MS peaks. Identity of the peaks is assured by the IROA patterns in both the parent, and fragments. In short, the combination of IROA and variable window SWATH acquisition provides a workflow in which a basic metabolomic-style system may be used for the accurate quantitation of several hundred metabolites in a single sample without the need for a baseline separation.

Data Reduction

Because biologically derived molecules (metabolites) and artifacts (non-biological contaminants) are completely distinguishable by their isotopic patterns the ability to remove non-biological data from further statistical analysis is assured; better data yields better results. It was recently demonstrated from a typical metabolomics dataset that 25,000 peaks were extracted and that the actual sample contained less than 1000 isotopically labeled (IROA) metabolites³. These additional non-biological peaks constitute false data which may become the basis for non-reproducible results. The IROA Technologies and SCIEX SWATH acquisition workflow assures that only biological data is considered.

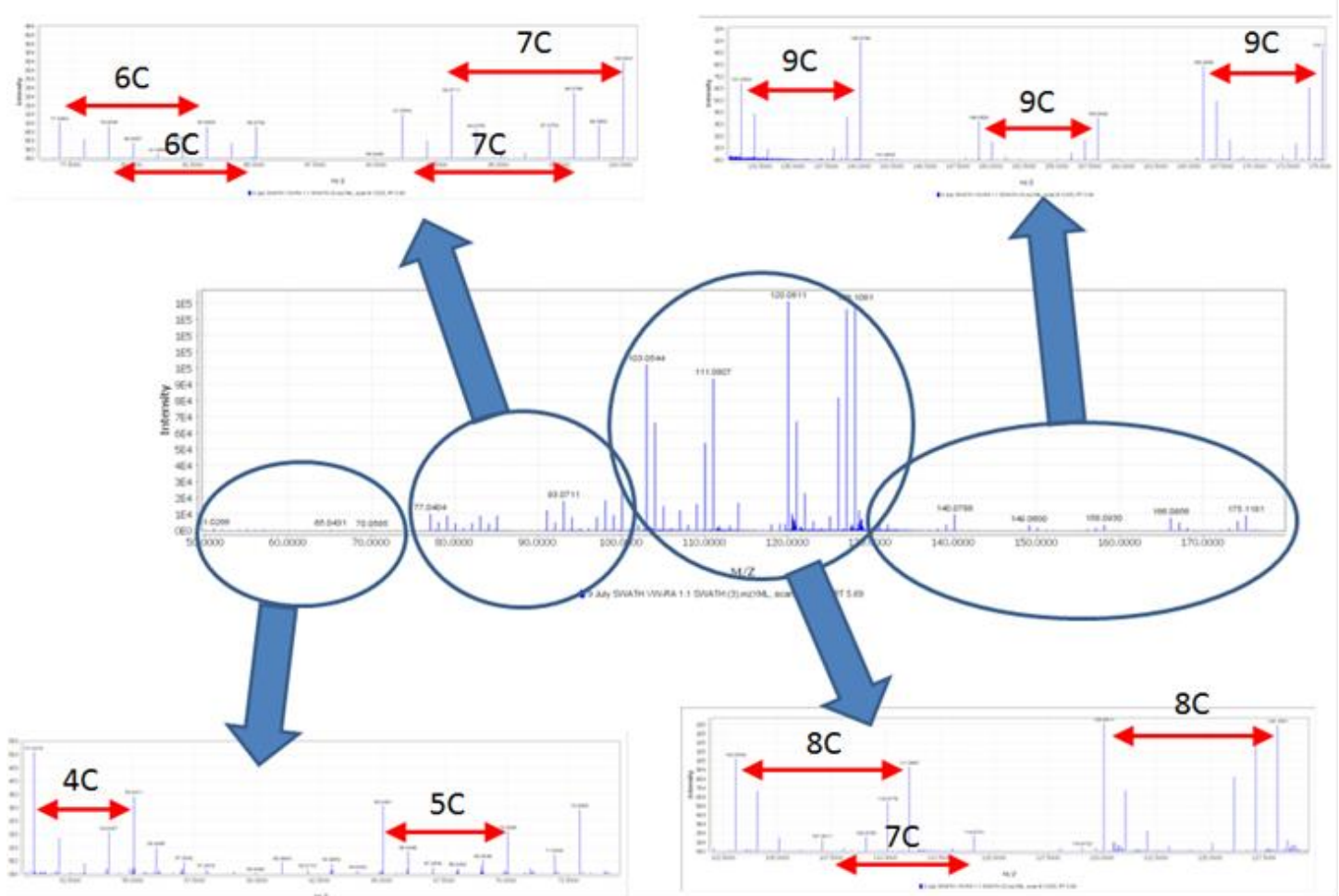


Figure 5: The IROA Technologies and SCIEX SWATH Acquisition Spectral Data. The fragmentation of phenylalanine is seen in the central SWATH acquisition scan (A). ClusterFinder detected all of the IROA peaks in this scan (B & C). Software was developed to automatically find, quantitate and identify all natural abundance peaks in experimental biological samples that corresponded to their known IROA isotopomers in the IROA-IS. The identification of compounds of unknown identity is simplified because all fragments are identified by their complete formula making the mode of fragmentation of the parent compound clear.

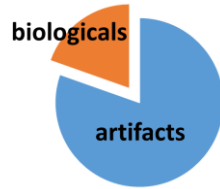


Figure 6: Remove non-biological Artifacts from your Data. Peak-picking yielded peaks of biological (Orange), and non-biological (Blue) origin. In this study, only 24% of the peaks found came from the biological sample being analyzed. The remainder was natural abundance and artifacts.

Use as a Daily QA/QC Check

The Matrix Standard is both used to support the identification,

quantitation, and interpretation of the experimental samples. As a Standard sample that is analyzed on a daily basis, any instrumental deviations may be understood, and appropriate corrections applied to normalize, and thereby integrate any collection of datasets. The Matrix is a Standard point of reference that can provide absolute identifiers for signal intensity and quality.

Cost Savings

It can be challenging when quantitating metabolites from an untargeted perspective. To apply absolute quantitation an internal standard for every metabolite would be needed, preferably heavy ¹³C –labeled, which has proved to be costly.



Figure 6: Data Review in ClusterFinder Software. The fragmentation of phenylalanine is seen in the central SWATH acquisition scan (A). ClusterFinder detected all of the IROA peaks in this scan (B & C). Note that for each peak the number of carbons in the fragment can be determined, and therefore the formula for the fragment is known (D). ClusterFinder Represents them both as (D) & (E). The actual fragmentation path (F) is easily determined from (E).

Even simple metabolite standard mixtures are far and few in between with many of them again being quite costly. In fact, available currently to date there are mainly amino acid standard mixtures which are heavy labeled and these can cost around 5 cents per metabolite for analysis, but only allowing quantitation of around 20 metabolites.

The IROA Internal Standard is comprised of over 500 uniformly-labeled metabolites in a mixture enabling reproducible accurate quantitation. The Workflow Kit costs less than 0.4 cents per metabolite for analysis which is over ten times less expensive than current standard mixtures available and offers more comprehensive quantitation.

Conclusions

The IROA Technologies and SCIEX SWATH acquisition workflow presented here is a powerful toolset for the assured identification of any detectable metabolite, the determination of its structure, via fully identified fragmentation, and the complete quantification of all of the components of a complex biological mixture. Because of the nature of the IROA and SWATH acquisition routines, these processes are readily automated and highly reproducible. This workflow allows a user to efficiently collect MS and MS/MS data in a single injection using the TripleTOF[®] System.

The enhancement of data quality is enabled through data reduction by removal of background/noise. Metabolites and their fragments may not be confused with artifacts, or noisy peaks. The fragmentation path attributable from the combination of IROA labeling and variable window SWATH acquisition reinforces the identity of the molecule and data quality.

The IROA Internal Standard ensures a high level of QA/QC for both identification and quantitation by monitoring a complete set of known metabolites every analysis.

This approach costs less than procuring individual metabolomics internal standards for every compound for use in every sample. And finally the data processing is automated using ClusterFinder Software allowing the user to efficiently move from data to results more swiftly.

AB Sciex is doing business as SCIEX.

© 2017 AB Sciex. For Research Use Only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.

Document number: RUO-MKT-02-0747-A

We see additional opportunity for improvement but already believe this system, in terms of assured quantitation and compound identification, will produce higher quality data than any other mass spectrometer-based system, and should be a useful adjunct in metabolomics and eventually clinical measurements.

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.
3. Mahieu NG, Patti GJ Systems-Level Annotation of a Metabolomics Data Set Reduces 25 000 Features to Fewer than 1000 Unique Metabolites. *Analytical Chemistry* 2017 Oct 3;89 (19): 10397-10406.