Measurement and Normalization System for targeted and untargeted MS analysis Identification, quantitation and QC of metabolites across time and instruments

Kit contents

Materials and tools for the analysis of 90 experimental samples

Unique fully-labeled S. Cerevisiae Yeast Extract*

- 3 vials of lyophilized IROA-IS; U-¹³C, 95%
- 3 vials of lyophilized IROA-LTRS (Long-Term Reference Standard); paired U-¹³C, 95% and 5%; mixed 1:1
- ClusterFinder[™] software (including library of IROA peaks in the LTRS and their physical characteristics)
- User manual

Storage: -80°C, protected from light

Summary of Benefits of the TruQuant Measurement System

- 1. Daily instrument QA/QC.
- 2. No false data. Reports only compounds of biological origin; excludes artifactual peaks.
- 3. Accurate compound formula ID for MS alone; <u>complete ID</u> with addition of SWATH, or IM.
- 4. **Fragments** have the IROA ratio pattern derived from their parent peaks and can be **identified** using the **"peak correlation"** ClusterFinder module.
- 5. Suppression-corrected measurements for significantly better quantitation.
- 6. Reproducibility. Sample normalization to a universal standard for complete comparability.
- ClusterFinder software solution builds libraries, IDs/quantitates compounds, corrects for ionsuppression and normalizes data.
- 8. Mass-spec vendor agnostic.
- 9. Economical. \$200 for 90 samples.

*Prepared from proprietary U-5%/U-95% ¹³C IROA-labeled glucoses specially produced for IROA Technologies by Cambridge Isotope Laboratories, Inc.

Standards included in the Kit

IROA Long-Term Reference Standard – IROA-LTRS (complex mixture of fully labeled 5% and 95% U-¹³C metabolites, mixed 1:1) is used to build a daily, verified-identity "dictionary" of Retention Time (RT), *m/z* and physical characteristics for all IROA peaks. Identity of isobaric compounds is verified using fragmentation and/or Collisional Cross Section (CCS) data. Provides complete daily QA/QC on instrument performance. IROA-LTRS data are collected randomly within every day's injection sequences.

IROA Internal Standard – IROA-IS (equivalent to the 95% U-¹³C component of IROA-LTRS) is added to analytical samples to identify and quantitate metabolites using the co-incidentally-run LTRS "dictionary". The chemical makeup and chromatographic behavior of the IROA-LTRS is identical to that of the IROA-IS; so the IROA-LTRS "dictionary" is completely applicable to the IROA-IS added to analytical samples. Isobaric compounds distinguished in the IROA-LTRS based on fragmentation or CCS are distinguished in the IROA-LTRS. Even if analyzed on different chromatographic systems, results can be equated using the dictionary because of comparability of the IROA-LTRS and IROA-IS, by querying the same mass and secondary physical characteristics across systems.

Why 5% and 95% U- ¹³C

- Mathematically calculable and readily identified by the IROA ClusterFinder software
- Unique IROA patterns discriminate peaks of biological origin from artifactual peaks allowing the removal of false data
- Fragments have the IROA ratio pattern derived from their parent peaks and can be identified using the "peak correlation" ClusterFinder module.

3 STEP PROCESS:

1. LC-MS analysis of experimental samples resolvated with IROA-IS (experimental sample + IROA IS = analytical sample) are randomly analyzed.

Within the random analytical injections the IROA-LTRS is injected approx. once for every 10 analytical samples, as a QA/QC Standard to account for fluctuations in mass and RT drift.

- 2. Generate "dictionary" of RT, m/z, formula and physical characteristics from the analysis of IROA-LTRS using ClusterFinder software to validly identify all compounds in the IROA-LTRS.
- 3. Use ClusterFinder and dictionary to identify and quantitate, suppression-correct and normalize all compounds in original experimental samples.

The IROA-LTRS

The IROA-LTRS, a Long-Term Reference Standard, is a fully labeled *S. cerevisiae* yeast cell extract IROA labeled at 5% and 95% U-¹³C and mixed 1:1. The result is a collection of over five hundred metabolites* that exhibit unique labeling patterns as represented by the peak pairs of the molecule arginine in Figure 1.



Figure 1. Example of an IROA-LTRS labeled molecule, shown here is arginine. The IROA-LTRS "U-shaped smile" pattern of peaks contains both the IROA-IS envelope (U-95% ¹³C peaks; M-1 etc.) and its mirror-image envelope (U-5% ¹³C peaks; M+1 etc.). The height of the M+1 and M-1 differ directly according to the number of carbons in a molecule; here 32% the height of their monoisotopic peaks, ¹²C and ¹³C, for a six-carbon molecule. This is true not only for the M+1 and M-1, but also the shape of the entire isotopic envelope is different for every number of carbons. The number of carbons in a biological molecule can be also determined by the distance between the two monoisotopic peaks. Since the relative height of the M+1, the relative height of M-1, and the distance between the monoisotopic peaks all provide confirmation of this fact, this results in a <u>triply redundant</u> quality control check point.

The monoisotopic IROA peaks (often the base peak) and the associated "carbon envelope" isotopic peaks can be clearly detected during MS analysis. In all IROA peak pairs the selection of the correct monoisotopic peaks is always confirmed by either the M+1 (shown in Figure 1 as green peaks and containing approximately 5% universal and random ¹³C; defining the C-12 envelope) or the M-1 (shown in Figure 1 as blue peaks and containing approximately 95% universal and random ¹³C; defining the C-12 envelope).

*IROA-LTRS contains a broad spectrum of metabolites including amino acids, organic acids, peptides, vitamins, carbohydrates, and coenzymes. Over five hundred metabolites have been characterized in the IROA-LTRS Standard using ClusterFinder software. The final number is expected to reach between 700-800 as different analytical approaches enable the identity of more compounds. The number of carbons in the molecule may be determined in <u>three ways</u>; by the mass difference between the C12 and C13 monoisotopic peaks, the height of the C13 M-1, and the height of the C12 M+1. In all IROA peak clusters or carbon envelopes these three values will be identical and therefore achieve triply redundancy with the added restriction that the peaks will exhibit symmetry (Figure 1). It is therefore very easy to computationally find all IROA peak pairs in the IROA LTRS and characterize them. For masses below 500 amu the isotopic mass plus the number of carbons identifies the molecular formula. Isobaric formulae are then distinguished using secondary MS scans collecting Collisional Cross Section (CSS), or fragmentation; thus all IROA-LTRS peaks have verified identities that correlate directly with their associated IS containing analytical samples.

The IROA-LTRS supports three major functions every day:

1) *Library Building Standard* – The IROA-LTRS (interdispersed in the sample set) and IROA-IS (spiked into samples) have the same concentration of compounds and should have extremely reproducible RTs and peak characteristics across all samples.

The IROA-LTRS is used to build a triply redundant dictionary (library) of RT, *m/z* and physical characteristics including fragmentation data. All peaks are named according to the IROA-LTRS database. The adducts and fragments for each compound are indexed individually. These compound IDs are loaded into ClusterFinder's internal databases as a separate editable IROA-LTRS database.

A library of IROA peaks in the LTRS and their physical characteristics is provided in the ClusterFinder[™] software and distributed with the TruQuant kits. This library is distributed without retention time information, however it is updateable by the user to assure that every compound in it will be reproducibly named according to the library.

2) **Validated Compound Identification** – The IROA-LTRS sample and ClusterFinder software and databases are used to create a <u>daily RT library of all the compounds</u> to be found in the IROA-IS so that their identification is reproducible and assured. The daily library is used as the basis of a targeted search of the IS in the analytical samples, and to quantify the natural abundance isotopologues of each compound. For each compound, once the IROA-IS is found in any sample, a value, even <LOD, will be returned; therefore, the targeted search should yield a non-sparse dataset.

All compound identities are validated in the IROA-LTRS where secondary physical characteristics such as fragmentation (ms/ms), ion mobility (ims/ms – CCS) may be collected to assure identity without lessening the quantitative aspects of the IS in the analytical sample.

3) **QC Standard** – As the name implies, the Long-Term References Standard (LTRS) is a Reference Standard and is always the same collection of compounds at the same concentrations. Although chemically complex it is well characterized so that on a daily basis its composition will provide insights into instrument performance: 1) The total of compounds seen is a measure of instrument sensitivity, 2) The retention times for familiar compounds are a measure of chromatographic performance, 3) The relative strength of many compounds and their known fragments is an important measure of in-source fragmentation, and 4) The total found signal for all IROA peaks is a measure of injection accuracy. The characteristics ClusterFinder sees today should compare with what it saw yesterday; after all, it is a Long-Term Reference Standard.

The IROA-LTRS is a pure IROA mixture with unique isotopic signatures. Following LC-MS, the paired compound IROA peaks (U-95% ¹³C or U-5% ¹³C) can be readily identified by the IROA ClusterFinder software. The IROA peaks represent actual compounds, fragments and adducts and can be discriminated from **unsigned artifacts and noise which can be removed from the data, eliminating false discoveries**. As a composite sample, sample-to-sample and analytical variance is removed and during MS analysis the identical compounds (labeled with either U-95% ¹³C or U-5% ¹³C) experience the same ionization efficiency and suppression.

Over 1000 peaks can be detected in the IROA-LTRS (see Figure 2D). The resulting IROA-LTRS dictionary of compounds is used to identify compounds in the IROA-IS saving time, effort and related costs.



Figure 2. ClusterFinder analysis of IROA-LTRS samples. A. Average isotopic pattern of adenosine 5'monophosphate IROA-IS; B. IROA peak chromatograms of adenosine 5'monophosphate IROA-IS; C. Identification panel for adenosine 5'monophosphate; D. IROA-IS peaks identified in the dataset.

The IROA Internal Standard (IROA-IS)

The **IROA-IS** is chemically identical to the U-95% ¹³C component of the IROA-LTRS. **The IROA-IS together** with ClusterFinder software are used to <u>co-locate</u>, <u>identify</u>, <u>and quantitate</u> 400 to 600 biochemicals in experimental samples, depending upon chromatographic mode(s) employed.</u>

The unique IROA labeling pattern (Figure 3A) again ensures that the monoisotopic peaks and the *carbon envelope* of the associated isotopic peaks (M-1 etc.) can be detected during LC-MS. The carbon envelope differentiates the IROA-IS from natural abundance peaks (Figure 3B) and is used to identify compounds of interest and exclude artifacts the may look otherwise similar.





The IROA-IS is a true Internal Standard and can be spiked into any natural abundance experimental sample (cells, tissue biopsy, plant material, blood, etc.) and all the IS peaks may be easily identified using the ClusterFinder software according to the presence of their characteristic M-1 peak and associated carbon envelope.

IROA-LTRS and analytical samples (experimental samples containing IROA-IS) are randomly interspersed into each sample set (e.g. such as there is one IROA-LTRS injection for approximately every 10 experimental/IS samples). Since the chemical makeup and chromatographic behavior of the IROA-LTRS sample is identical to the IROA-IS, it is possible to use the in-depth and informationally strong, triply redundant chemical information obtained from the IROA-LTRS sample and apply it to identify compounds in the experimental sample. The dictionary catalog of all peak pairs, their RT, number of carbons, and IM and fragmentation characteristics provides information where each of these same IROA peaks will be found in the experimental samples through the use of the IROA-IS. The experimental natural abundance peaks are easily located and quantitated as they will co-locate with their corresponding IROA peaks at a mass that is the mass of the IROA ¹³C monoisotopic peak less the number of carbons it contains times the mass of a neutron.



Figure 5. The IROA TruQuant IQQ measurement system is based on a well characterized Long-Term Reference Standard (D = LTRS) and a carefully matched Internal Standard (B = IS) to provide instrument and process QA/QC, validated compound identification and quantitation.

(A) Experimental samples (plasma, urine, cells, etc.) are mixed with (B) a complex 600+ compound Internal Standard (IS), fully labeled with U-95% ¹³C (B'). Experimental/IS samples are analyzed using LC-MS, injecting the LTRS (D) intermittently, approx. every 10 samples. The LTRS has the same concentration and is chemically identical to the IS but also contains a 1:1 mixture of fully labeled U-5% and 95% U-¹³C metabolites, producing a U-shaped pattern of carbon envelopes (D'). The relative height of the M+1, the relative height of M-1, and the distance between the monoisotopic peaks all provide confirmation of the number of carbons in each biological molecule resulting in a <u>triply redundant</u> quality control check point. The IROA peaks represent actual compounds, fragments and adducts and can be discriminated from unsigned artifacts and noise which can be removed from the data, eliminating false discoveries. As a composite sample, sample-to-sample and analytical variance is removed and during MS analysis the identical compounds (labeled with either U-95% ¹³C or U-5% ¹³C) experience the same ionization efficiency and suppression.

Over 1000 peaks can be detected in the IROA-LTRS. Following analysis, the resulting IROA-LTRS dictionary of compounds is used to identify compounds in the IROA-IS saving time, effort and related costs. The IROA-IS together with ClusterFinder software are used to <u>co-locate</u>, <u>identify</u>, <u>and quantitate</u> 400 to 600 biochemicals in experimental samples, depending upon chromatographic mode(s) employed. The characteristic for each compound in the IROA-IS allow the calculation of a suppression-corrected area for each compound.

Unambiguous ID and quantitation (ms2) in a single injection. Complete identification of compounds is achieved with the addition of IM or SWATH. The IROA IM peaks retain their patterns perfectly because all IROA isotopomers share the same CCS (D'). In IROA msms fragmentation, such as SWATH, the IROA peaks retain their patterns (D2') because wide windows are used. Since all fragments retain their IROA character, their formulae and the relationships between them (D3') are determinable.

The Process

A) Perform LC-MS analysis of the IROA-LTRS sample and experimental samples using IROA-IS

Aliquots of the IROA-LTRS are analyzed together with experimental samples spiked with IROA-IS, i.e. analyzed using positive and negative ion LC-MS modes, and other analytical modes such as HILIC and orthogonal, second-stage analyses such as an Ion Mobility, SWATH fragmentation etc. that are employed in the laboratory.

See **IROA-LTRS** and **IROA-IS** sections below for instructions on how to prepare materials for analysis.

IROA-LTRS

- The IROA-LTRS is injected periodically i.e. approximately 10 sample intervals and is separately analyzed.
- The IROA-LTRS is used to build a "**dictionary**" of RT, *m/z* and physical characteristics including fragmentation data.
- The IROA-LTRS is used as a **QC Standard to account for fluctuations in chromatography**, **mass**, **or retention time drift**, **source**, **or instrumentation malfunctions**.
- IROA-LTRS is always the same and the catalog of all IROA peaks found in each daily IROA-LTRS analysis provides a way to quantitate the performance characteristics for the instrumentation for every day's analysis and a mechanism for correcting any instrumental error.

The IROA-LTRS is used for identifying all of the compounds. Since there is no need to accurately quantify the compounds in the IROA-LTRS, relative quantitation is sufficient, fewer scans are needed to establish quantitative accuracy, and these scans may be better used to support compound identity verification. Thus, the IROA-LTRS may be used with alternating scans devoted to collection of secondary compound identification information, such as fragmentation, Ion Mobility, etc., while in the analysis of the analytical samples keeping all scans as primary will assure better peak shape and, thus, better quantitation.

To Prepare the IROA-LTRS Standard:

- Add 40µL injection solvent to the vial containing IROA-LTRS. Pipette up and down to ensure solubilization of IROA-LTRS material. Keep on ice until use.
- Use 3µL for LC-MS analysis. Note: Some Users prefer to add 2ul for positive mode and 4ul in negative mode.

These IROA-LTRS analyses assure that the maximum number of compounds in the IROA-IS can be found, and that daily performance quality expectations are met. This procedure creates a daily library or map of the IROA-IS (Figure 6). Individual compounds can be readily identifiable (Figure 7 shows examples of IROA-LTRS derived amino acids).



Noise & Signal

• 13C Std

Figure 6. The IROA-LTRS (mixture of the IROA-IS and its U-5% mirror image is mapped using LC-MS to assure that the maximum number of compounds in the IROA-IS can be found



Figure 7. Examples of IROA-LTRS derived amino acid compounds after LC-MS. The compounds show the typical U-shaped pattern (mixture of the IROA-IS, U-95% ¹³C, and its U-5% ¹³C mirror image). The distance between the two monoisotopic peaks clearly indicate the number of carbons in the molecule.

IROA-IS

• The IROA-IS is equivalent to the 95% U-¹³C component of the IROA-LTRS only and is used as an Internal Standard for it's natural abundance counterparts. It provides enough information

for complete identification and quantitation of samples without the need for chromatographic base-line correction. Since the chemical makeup, concentration, and chromatographic behavior of the IROA-LTRS sample is identical to the IROA-IS, it is possible to use the in-depth and informationally strong, triply redundant chemical information obtained from the IROA-LTRS sample collected in the dictionary and apply it to identify compounds in the experimental sample.

- The IROA-IS is used to find and align all the peaks.
 - The alignment of all samples ensures reduced variability in day-to-day measurements.
 - The characteristics for each compound in the IROA-IS allows the calculation of a suppression-corrected area for each compound
 - The normalization of the total area under the curve (AUC) for all natural abundance suppression-corrected peaks to the total AUC of their corresponding IROA-IS peaks is a "Dual MSTUS*" algorithm that allows for normalization of samples not only within a single day's run, but since the IROA-IS is equivalent every day; all normalizations are to a true Standard, i.e. that is the same every day.
- **Suppression correction and normalization** are implemented in ClusterFinder using the qualities inherent in the IS peaks). The user is always presented with raw, suppression corrected, and normalized so the best choice may always be available according to any experimental design.
- Normalization overcomes sample-to-sample variation.
- The IROA-IS can be used to build a Retention Index to track "unknown" compounds not identified in the IROA-IS.

To prepare experimental samples with IROA-IS

For use as an injection standard

The IROA-IS is incorporated into the re-solvation solvent. In this case there is no question that the IROA-IS is in exactly the same concentration in every sample and will be directly comparable to other samples prepared the same way. The material in each amber vial is expected to be suitable for 30 samples.

- 1) Re-solvation solvent is either the initial chromatographic solvent or dH2O. (Where the chromatographic solvent contains acid dH2O is preferable to reduce degradation prior to injection, but either will work).
- 2) Add 1.2 ml re-solvation solvent to the amber vial, mix, and then vortex. This is the Internal Standard solution (IROA-IS).
- 3) Hold the IROA-IS on ice until you are ready to re-solvate your samples.
- 4) To each dried sample add 40 ul of the IROA-IS, then vortex.
- 5) If you have a filtration step, then filter before analysis after re-solvation.

*MSTUS described in "Normalization strategies for metabonomic analysis of urine samples" Warrack et. al

A Note on Normalization

We all know that sample amounts can be highly variable due to fluxuations in sample preparation and delivery. A "pre-prep" normalization should be performed on each sample prior to crash. Our recommendation is that a protein reading be done on each sample homogenate, and equal amounts of protein be delivered as the "prep sample". While this may not always be appropriate it is always better to collect all the information one can on every sample and attempt a prep-normalization.

The above protocol delivers approximately 2 micrograms each of U-95% ¹³C and U-5% ¹³C (IROA-LTRS) and 2 micrograms U-95% ¹³C (IS) per sample which should be appropriate for most MS platforms, however we recommend that a "calibration" step is done ahead of time, i.e. test different amounts of your standard prep with the same amount of IROA-IS to figure out how much to balance with the IROA-IS. Please see "Recommendations for Pilot Experiments" in the Appendix.

- B. Build Dictionary using IROA-LTRS data files. Identify compounds, fragments and adducts using ClusterFinder software.
 - Upload the IROA-LTRS LC-MS data files to IROA ClusterFinder Software. ClusterFinder uses algorithms to build, evaluate, edit and export the result of untargeted IROA search. Refer to ClusterFinder User Manual and instructional videos, provided separately.
 - If you have not yet created a dictionary specific to your method connect the relevant IROA-LTRS library, likely IROA-LTRS-pos, or IROA-LTRS-neg. These libraries are devoid of retention times but are a good starting point. If you have created method specific libraries (see last step) please attach that instead as it will contain your previously recorded retention times.
 - Compounds, fragments and adducts are identified in minutes using algorithms that search for the specific IROA signatures (See Figure 8).



Figure 8. Example of IROA-LTRS-derived compound L-Phenylalanine identified using ClusterFinder software. The compound shows the IROA-LTRS U-shaped pattern (mixture of the IROA-IS, U-95% ¹³C, and its U-5% ¹³C mirror image). The distance between the two monoisotopic peaks clearly indicates that the molecule contains 9 carbons.

- Curate the data
 - Examine spectrum bins and their constituent IROA peaks.
 - Look at the detected features in the context of underlying raw data and manually correct the results in cases where the automatic algorithm has failed.
 - The curation process takes approximately an hour to complete.
- Build a targeted library (.cflib file) from each method or LC-MS mode. Each library will be used to perform a targeted search of the clinical/experimental IROA-IS-containing files run coincidently to the IROA-LTRS.
- Create a method specific dictionary as a ClusterFinder database:
 - Disconnect the IROA-LTRS dictionary from your search,
 - Create a new library (database), for instance, myHILICLTRSPos, and attach it to the method of analysis,
 - Select all bins in your curated method,
 - Right-click the selected bins, and
 - Select "add to Library"
 - The new library will now have retention times, spectra, and all related information, and is likely the dictionary to connect in future work.
- Export the resultant dictionary in comma separated value (CSV) format or compound exchange format (CEF) for storage or transport to other ClusterFinder installations.

- C. Identify compounds, fragments and adducts in the experimental samples using ClusterFinder software and the IROA-LTRS dictionary.
 - Import LC-MS data files to IROA ClusterFinder Software.
 - Run a targeted analysis using the targeted library (.cflib file) that was created following the curation of the IROA-LTRS files.
 - Compounds, fragments and adducts are identified in minutes using algorithms that search for the specific IROA signatures.
 - Curate the data
 - Examine spectrum bins and their constituent IROA peaks.
 - Look at the detected features in the context of underlying raw data and manually correct the results in cases where the automatic algorithm has failed.
 - The curation process takes approximately an hour to complete.
 - Export the results as tab separated value (tsv) files.
 - The quantitative output of targeted analysis, together with experiment design may serve as input for the statistical and biochemical interpretation. The User can employ their favorite statistical package.

APPENDIX: Recommendations for Pilot Experiments

The IROA TruQuant IQQ WorkFlow Kit contains enough materials for 3 sets of experiments. Consider what type of experiment you want to do? For example, do you have cell types (transformed or untransformed) available or biopsies that might be interesting?

If this your first use of the TruQuant IQQ method, your **first experiment** should be an experimental sample IROA-IS calibration experiment (see below). The purpose of the calibration experiment is to optimize the amount of experimental sample that you will use with the kit (the quantity of your experimental sample is optimized against the quantity of IROA Internal Standard used in the method). You only need to do this once for any sample type or SOP.

Internal Standard (IS)-Calibration

The calibration of IS to your SOP simply requires the generation of a single large portion of prepared "prepped" sample (whatever kind of sample type you plan to use for your experiment, i.e. plasma, cells or tissue), approximately 20X the normal sample size routinely used is sufficient.

- 1. Pool and filter prepped sample to create a single homogeneous sample.
- Deliver differing aliquots of the pooled sample, ranging from half of the normal amount delivered to 3 times the amount normally to Eppendorf capped sample tubes. (For example: 0.5X, 0.75X, 1X, 1.5X and 3X.) Make triplicate samples.
- 3. Dry aliquots under a gentle nitrogen stream.
- 4. Add 1.2 ml of deionized H2O to a single vial of IS. Thoroughly mixed to ensure a single homogeneous solution. (Note: It should go into solution freely). Keep on ice until ready to use.
- 5. Use a 40 ul aliquot of the IS to resolvate each of the dried samples. Upon addition of IS, mix thoroughly.
- 6. Resolvate a single vial of IROA LTRS with 40 ul of deionized H2O, thoroughly mix to ensure complete solvation. Keep on ice until ready to use.
- 7. Analyze samples using the chromatographic method you would normally use. The analysis may be repeated in multiple modes, i.e. positive reverse phase, negative reverse phase, positive HILIC, or negative HILIC.
- 8. Use the ClusterFinder software to find and identify all IROA peaks in the LTRS in an unbiased analysis.
- 9. Export the compounds found in the previous step to run a targeted analysis of the IS containing samples. Use the ClusterFinder software to export all of the quantitative data from the IS containing samples for statistical analysis. Export all of the compound identification and QA/QC data from the LTRS.
- 10. Analyze the samples to identify the amount of the prepped samples that yields an overall mass spectral signal that is equal to the overall mass spectral signal of the IS. This is the amount of sample that will most accurately be measured using the IS in the future, i.e. well balanced by the standard 40 ul of IS.