MSTUS sample-to-sample Normalization has an interesting anomaly, yielding better normalizations

Chris Beecher[2]; Debasish Ghosh[1]; Felice de Jong[2]; Vladimir Shulaev[1]

[1]Department of Biological Sciences, College of Science, University of North Texas, Denton, Texas; [2]IROA Technologies LLC, Chapel Hill, North Carolina

INTRODUCTION

Previously we demonstrated a "Dual-MSTUS normalization" mechanism for batch-to-batch error correction by adding the same concentration of an Internal Standard (IROA-IS) into each experimental sample to create two isotopically differentiable population of peaks so we could define the value to be normalized. Now we demonstrate an improvement to this normalization that further educes sample variance. In any MSTUS algorithm it is assumed that the more compounds that are included the better. We now show that variances are further reduced if we use a subset of all compounds seen (the current algorithm), namely only those compounds found in all samples (Core compounds). The current belief is to include compounds that may appear in only a small portion of samples. However, even though they are real compounds their presence in the subset of samples alters their normalization. The rationale and consequences of this will be discussed.

METHODS

Saccharomyces cerevisiae were treated with 4 different oxidants known to be involved in oxidative stress. Samples collected from the entire experiment were processed in batches on several different days across several weeks. After four batches, the electrospray ionization source and capillary was removed, cleaned and retuned to articulate very strong batch effects before processing further batches. Data obtained from an Orbitrap Velos Pro mass-spectrometer was used to optimize our Dual-MSTUS normalization method. A Dual-MSTUS normalization corrected for the batch effect errors easily, but sample-to-sample (STS) variances could be reduced further by using only the "Core" peaks (seen in all samples), rather than the current standard of "All" peaks seen in each sample. An IROA-IS was included in every sample.

WORKFLOW

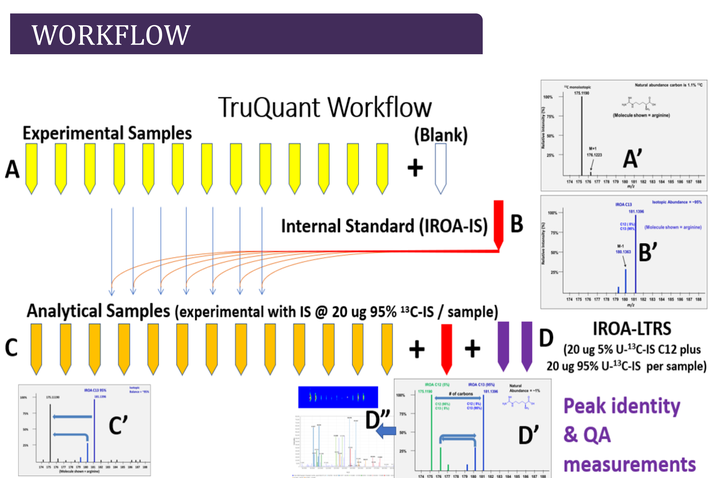

Figure 1. IROA TruQuant Yeast Extract Semi-Targeted Workflow

RESULTS

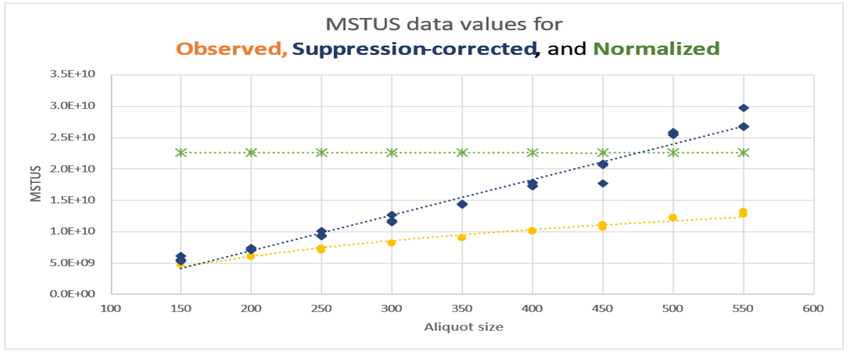

Figure 2. Suppression-correction and normalization for IS related peaks. The intersection of the suppression corrected MSTUS values (dark blue values), and the MSTUS C13 values for normalized data (green values) represents the aliquot size that will have optimal quantitation.

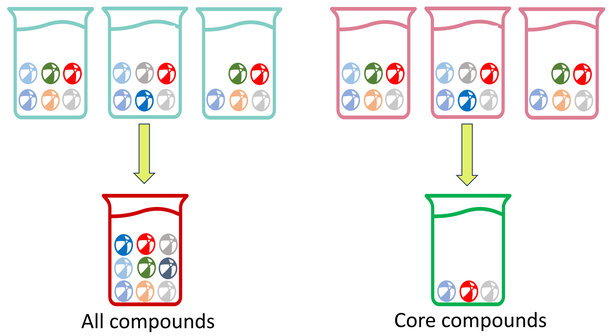

Figure 3. Conceptualization of All vs. Core compounds used in normalization. All and Core compounds used in normalization are represented as circles in red and green buckets, respectively.

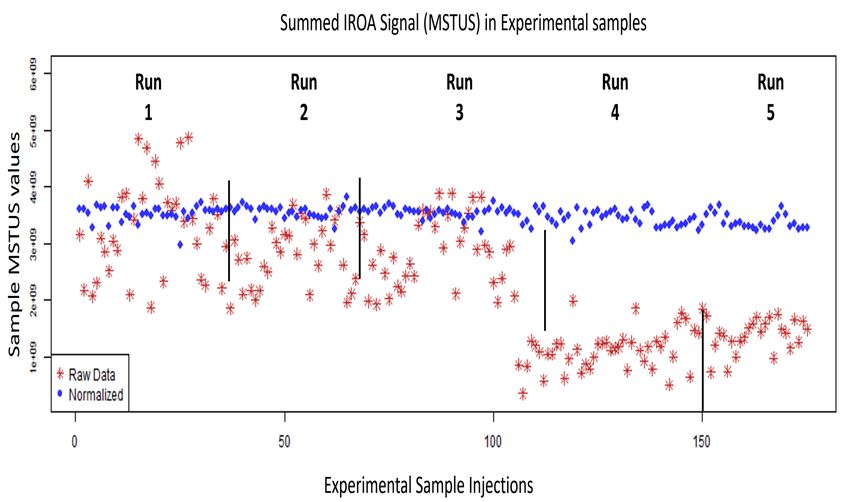

Figure 4. A dual-MSTUS normalization using all compounds. Normalization to reduce batch effect in the sample by using All compounds present in all individual samples.

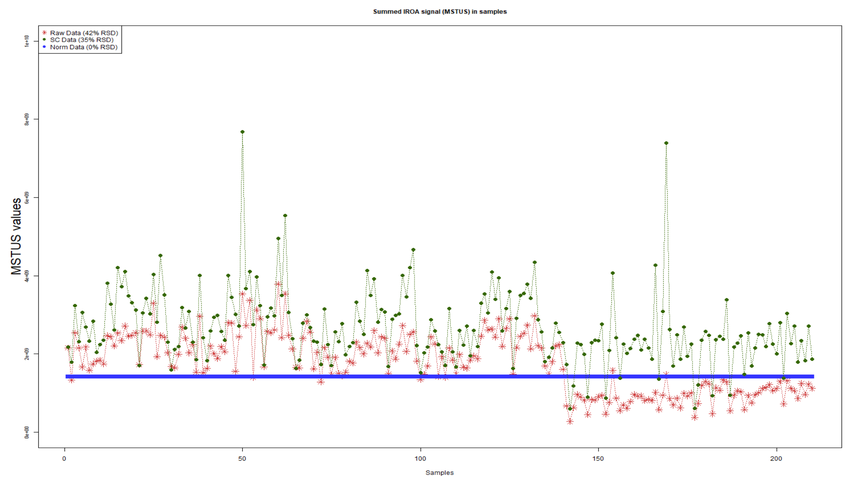

Figure 5. A dual-MSTUS normalization using Core compounds. Normalization to effectively remove batch effect in the sample by using all compounds common across all samples (Core compounds).

CONCLUSIONS

In the original MTUS protocol Warrack et al. removed all of the artifactual data in an effort to reduce "signal noise", but they selected the normalization value in an arbitrary fashion. In the case of the TruQuant protocol demonstrated here, the IROA workflow is dedicated to evaluate all the peaks used in the analysis matched to an IS of biological origin. As there are two isotopically differentiable population of peaks (Experimental and IS), and since the IS signal suffers the same variances as the Experimental signal, i.e. suppression, injection fragmentation etc., "Dual MSTUS" allows the removal of even more noise. Further, in this study we have observed that instead of considering all compounds present in each individual sample, compounds present across all samples yield better normalization. This normalization is not only more specific than the original but is less sensitive to error which reduces the %RSD (Relative Standard Deviation) values to zero percentage. Unlike most normalization procedures that employ statistical batch averaging approaches, this method corrects for the sources of error in both sample preparation and instrumentation. This approach could be a valuable method for large clinical studies involving multiple batches.

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.**
3. **Warrack et al. "Normalization Strategies for metabonomic analysis of Urines", J. Chrom B 2009 877(5) 547-552. doi: 10.1016/j.jchromb.2009.01.007**

Acknowledgement: IROA-labeled materials were produced for IROA Technologies under agreement by CIL.
Conflict of Interest Disclosure: Chris Beecher and Felice de Jong, founders of IROA Technologies supplied material and participated in the discussions.