NIH Public Access
Author Manuscript
Anal Chem. Author manuscript; available in PMC 2014 October 15.

Published in final edited form as:
Anal Chem. 2013 October 15; 85(20): . doi:10.1021/ac401559v.

Measurement of DNA concentration as a normalization strategy for metabolomic data from adherent cell lines

Leslie P. Silva[#†], **Philip L. Lorenzi**[#†], **Preeti Purwaha**[†], **Valeda Yong**[†], **David H. Hawke**[§], and **John N. Weinstein**[*,†]

[†]Department of Bioinformatics & Computational Biology, University of Texas MD Anderson Cancer Center, 1400 Pressler St., Houston, TX, 77030, United States

[§]Department of Pathology, University of Texas MD Anderson Cancer Center, 7435 Fannin St., Houston, TX, 77054, United States

[#] These authors contributed equally to this work.

Abstract

Metabolomics is a rapidly advancing field, and much of our understanding of the subject has come from research on cell lines. However, the results and interpretation of such studies depend on appropriate normalization of the data; ineffective or poorly chosen normalization methods can lead to frankly erroneous conclusions. That is a recurrent challenge because robust, reliable methods for normalization of data from cells have not been established. In this study, we have compared several methods for normalization of metabolomic data from cell extracts. Total protein concentration, cell count, and DNA concentration exhibited strong linear correlations with seeded cell number, but DNA concentration was found to be the most generally useful method for the following reasons: 1) DNA concentration showed the greatest consistency across a range of cell numbers; 2) DNA concentration was the closest to proportional with cell number; 3) DNA samples could be collected from the same dish as the metabolites; and 4) cell lines that grew in clumps were difficult to count accurately. We therefore conclude that DNA concentration is a widely applicable method for normalizing metabolomic data from adherent cell lines.

INTRODUCTION

Metabolites, defined here as endogenous small molecules with molecular mass below 1200 Da, are central to intermediary metabolism and are the building blocks of cell components, including carbohydrates, proteins, RNA, and DNA. The estimated number of biological metabolites is between 3,000 and 10,000, a large number of which are still unidentified.[1-2] They play active roles in all cell processes. The need (and ability) to monitor metabolites is rapidly growing, in part because metabolite levels are considered the most direct link to phenotype and function.[3-4]

Untargeted metabolomics aims to identify and measure all metabolites in a biological sample. In mammals, metabolomics technologies have been used to study diseases, define pathophysiological processes, and discover biomarkers.[5-11] In the context of cancer, for example, it is increasingly apparent that important signaling pathways and oncogenes act through changes in metabolites.

[*]Corresponding author: jweinste@mdanderson.org, Fax: 713-563-4242.

Supporting Information **Available.** Additional information as noted in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Metabolites encompass a wide variety of chemical properties, making simultaneous extraction, separation, and measurement difficult.[12] Nevertheless, untargeted metabolomics makes it possible to associate previously unrecognized metabolites with unique phenotypes and, therefore, to elucidate biomarkers and gain insights into disease pathogenesis.[12] Nuclear magnetic resonance (NMR) and mass spectrometry (MS), the two most accepted methods for measurement of metabolites, are highly complementary, but the combination of liquid chromatography and MS (LC-MS) is much more sensitive for metabolite detection; it permits thousands of metabolites to be identified at low concentrations.[1]

Because of unwanted variation that may be introduced during sample preparation or instrumental analysis, data normalization is an especially important aspect of the metabolomics workflow. It is often necessary to compare samples across time points, concentration levels, differences in cell type, or differences among patients. Differences in the amount of cell material included in the sample can yield inaccurate conclusions in the absence of an appropriate normalization method. Common choices for normalization of metabolomics data include cell number, total protein, total metabolite signal (e.g., the total ion chromatogram (TIC)),[13] median metabolite signal, or a housekeeping metabolite. Which is best? The answer depends on a number of experimental considerations and the biological question(s) being asked. For example, cell number is commonly considered an appropriate normalizer for a variety of assay types, but its accurate measurement for adherent cells requires either in situ imaging or trypsinization. The former can be cumbersome and is difficult for cells that grow in clumps; the latter introduces its own series of problems, most notably the requirement for samples that are separate from the ones used for metabolite measurement, because trypsinization has been reported to introduce metabolomic artifacts including the deamidation of asparagine to aspartic acid.[14] The requirement for additional samples increases processing time and the amounts of reagents and cells required for an experiment. Given those limitations of cell counting, total protein is often used for normalization of metabolomic data.[15-16] However, the buffers and organic solvents required to quench cellular metabolism and extract metabolites have been noted to result in inaccurate protein measurements,[17] probably due to precipitation of protein. Therefore, like trypsin-assisted cell counting, total protein as a basis for normalization requires protein quantitation from a separate, parallel sample. Subsequent metabolite extraction and protein precipitation under acidic conditions presents a viable alternative; however it has been reported that the overall number of metabolites extracted under acidic extraction conditions (e.g., using perchloric acid) is as much as 70% lower than that from organic extraction methods.[18] Another option for data normalization is the use of housekeeping metabolites (e.g., metabolites whose concentrations are proportional to cell number).[19] Housekeeping metabolites may be conceptually attractive, but their utility is limited by the need to validate them for each cell line and treatment being studied; a proper “housekeeper” must remain constant across all cell lines and treatments in the experiment.

The purpose of this study was to identify an efficient, accurate, and precise method for normalizing metabolomic data from cancer cell lines. With emphasis on normalizers that can be accurately measured using the same sample used for measurement of metabolite levels, we investigated several methods. We show that two commonly used methods for cell number normalization—protein concentration and median metabolite intensity—exhibit large error. We demonstrate that DNA concentration from a metabolomic sample cell pellet, which is normally discarded, satisfies the requirements for an efficient, accurate, and precise normalizer for metabolomic data.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

EXPERIMENTAL SECTION

Chemicals

LC-MS reagents and metabolite standards were purchased from Sigma-Aldrich (St. Louis, MO, USA) unless otherwise noted. HPLC-grade water (H_2O) and acetonitrile (ACN) were purchased from Thermo Fisher Scientific (Pittsburgh, PA, USA).

Sample Preparation (Protein Assay)

To determine the accuracy of total protein measurement, serial dilutions of 0.25, 0.5, 1, 2, and 4 million OVCAR-8 cells were seeded in 10 cm dishes (Corning, Tewksbury, MA) in triplicate. Cells were incubated at 37 °C and 5% CO_2 for four hours to allow attachment to the dish. Proteins were then extracted using two distinct methods. In the first set of triplicate samples, cells were washed twice with 4 mL of phosphate-buffered saline (PBS) and extracted by adding 200 μL of bicine/CHAPS buffer (Protein Simple, Santa Clara, CA) containing protease and phosphatase inhibitors (Millipore, Billerica, MA) to the dish. Cells were scraped and transferred into a microfuge tube, and incubated on ice for 20 min, followed by a brief (1 s) vortex on a Disruptor Genie (Scientific Industries Inc., Bohemia, NY) at 4 °C. Cell debris was pelleted by centrifugation for 10 min at 20,000 *g* and 4 °C. Supernatant was snap-frozen in 20-μL aliquots using liquid nitrogen and stored at −80 °C. A second triplicate set of cells was harvested by washing twice with dry ice-cooled 60% methanol containing 0.85% (w/v) ammonium bicarbonate (AMBIC) and extracting with 500 μL of dry ice-cooled methanol:chloroform:water (7:2:1) (v/v/v).[20] Cells were scraped and transferred into a microfuge tube on ice, followed by vortexing for 30 s at 4 °C. Cell debris was pelleted by centrifugation for 2 min at 20,000 *g* and 4 °C. Lysate was snap-frozen in 20-μL aliquots using liquid nitrogen and stored at −80 °C. The bicinchoninic acid (BCA) protein assay was performed on thawed aliquots according to the manufacturer's protocol (Thermo Fisher Scientific, Rockford, IL, USA).

Sample Preparation (Cell Pellet Mass Assay)

As with the protein extraction protocol, serial dilutions of 0.25, 0.5, 1, 2, and 4 million OVCAR-8 cells were seeded in 10 cm dishes in triplicate, and cells were allowed to attach for four hours at 37 °C and 5% CO_2. Microfuge tubes were labeled and pre-weighed using an analytical balance (Sartorius, Elk Grove, IL). Cells were harvested using the protocol described in the previous section for methanol/chloroform/water extraction. The extraction mixture was placed in the pre-weighed tube on ice. The tube was then vortexed for 30 s, and cell debris was centrifuged for 2 min at 20,000 *g* at 4 °C. The supernatant was discarded, and the cell pellet was dried at room temperature. The labeled tubes were then re-weighed on the analytical balance, and cell pellet mass was determined by subtracting the initial tube mass.

Sample Preparation (Cell Pellet Protein Assay)

Cells were seeded and harvested as described for the cell pellet mass assay. The pellet was dried at room temperature and resuspended in bicine/CHAPS buffer containing protease and phosphatase inhibitors. A BCA protein assay was performed to quantitate total protein concentration of the reconstituted cell pellet.

Sample Preparation (Cell Pellet DNA Assay)

Cells were seeded as described above, then harvested by washing twice with dry ice-cooled 60% methanol containing 0.85% (w/v) AMBIC and extracted with 500 μL of dry ice-cooled methanol:chloroform:water (7:2:1) (v/v/v). [20] The cells were then scraped and transferred into a microfuge tube on ice, followed by vortexing on a Disruptor Genie for 30 s. Cell

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

debris was pelleted by centrifugation for 2 min at 20,000 *g* and 4°C. The lysate was immediately used for metabolomic sample preparation (below) or snap-frozen and stored at −80 °C. The pellet was air dried at room temperature and subsequently used for DNA extraction. A genomic DNA mini-kit for blood/cultured cells (IBI Scientific, Peosta, IA) was used for all DNA extractions. Briefly, 150 μL of red blood cell (RBC) lysis buffer was added to the dried cell pellet and incubated at room temperature overnight to allow full solubilization. The remainder of the protocol was performed to manufacturer's specifications, and the DNA was eluted with 50 μL of elution buffer. DNA concentrations were analyzed using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific).

Sample Preparation (Metabolomic Analysis)

Cells were seeded and harvested as described in the DNA assay section. 200 μL of cell lysate was added to 600 μL of 100% methanol containing 10 μM xylitol as an internal standard on a Phenomenex Strata protein precipitation plate (Phenomenex, Torrance, CA, USA). The resulting filtrate was dried down using a SpeedVac Concentrator (Thermo Fisher Scientific). Dried preparations were stored at −80 °C until analysis.

Sample Preparation (Time Series Study)

OVCAR-8, U251, and MDA-MB-435 cell lines were seeded using 90,000 cells per 10-cm dish in triplicate and incubated at 37 °C with 5% CO_2 for 48 h. The OVCAR-4 cell line was seeded in a similar fashion using 260,000 cells per 10 cm dish. At 48 h, the media were replaced with fresh media, and cells were harvested for analysis of DNA concentration (as described above) after incubation for an additional 0, 8, 24, or 48 h.

LC-MS Analysis

Metabolomic preparations for all four cell lines were reconstituted in 50 μL of mobile phase A (A) (water + 0.1% formic acid (FA)) immediately prior to analysis. The LC-MS system consisted of a Dionex Ultimate 3000 UPLC coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). 5 μL of sample was drawn into a 20-μL sample loop (partial loop fill mode) and injected onto an Agilent Eclipse Plus C18 column (1.8 μm × 2.1 mm × 50 mm). Metabolites were separated using a 25-min linear gradient of 100% A to 100% ACN + 0.1%FA (B), with an initial 5 min hold at 100% A and a flow rate of 100 μL/min. Metabolites were detected in positive ion mode with the *m/z* range set at 50–1000 Da, an acquisition time of 10 ms, and resolution of 60,000.

Data Analysis

Data were acquired with Xcalibur software (Thermo Fisher Scientific) and processed through the XCMS Online server.[21] Metabolites were annotated using an in-house standard reference library of approximately 150 compound standards.

Four metabolites from three polar compound classes were selected (glutamic acid, uracil, N-acetylputrescine and creatine) and their intensities normalized by dividing by the respective cell number that was determined by each technique. When normalizing to protein concentration or DNA concentration, metabolite intensities from individual samples were normalized to corresponding protein concentration or DNA concentration prior to averaging the normalized values of the replicates. For the other methods, which used surrogate cell dishes, metabolite intensities from individual samples were divided by the respective averaged protein concentration or cell count. For normalization by median ion intensity, the median ion intensity was calculated from all features in the XCMS report, and the intensities of the four selected metabolites were divided by the respective median ion intensity obtained for each individual sample.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

RESULTS AND DISCUSSION

Metabolomic studies on cultured cells require accurate methods for data normalization. We focused here on identifying a method that is a strong surrogate for cell number under sample preparation conditions commonly used for metabolomic analyses.

DNA concentration is a robust surrogate for cell number

Protein concentration has been correlated with cell number and used for data normalization.[15-16] However, accurate quantitation of protein concentration from cells extracted with metabolomics-compatible solvents (specifically methanol and acetonitrile) is questionable, as demonstrated here. When we determined protein concentration using a bovine serum albumin (BSA) standard and plotted it as a function of seeded cell number, unfavorable results were obtained for both a protein-compatible extraction (bicine/CHAPS) and a metabolite-compatible extraction (methanol/acetonitrile) (Figures 1A and B). Figures 1A and B both illustrate a linear increase in protein concentration with increasing seeded cell number. However, when the slope of each curve was calculated using a log/log scale, which should result in a slope equal to 1 if a 2-fold increase in cell number corresponds to a 2-fold increase in signal, the slopes in Figures 1A and B were 0.75 and 0.44, respectively. The deviation from unity by 25% for the bicine/CHAPS buffer was somewhat surprising and suggested that bicine/CHAPS extracts cellular protein with only 75% efficiency under the conditions tested. When the 0.25 million cell data point was excluded, the log/log slope resulted in less than 20% error, suggesting poor protein recovery at cell numbers below 0.50 million. That method, however, still requires the seeding and processing of additional cell dishes, which is not ideal for metabolomic preparations. The greater than 50% deviation exhibited by the organic extraction was less surprising, since organic solvents cause protein precipitation. The right-hand panels of Figures 1A and B illustrate the problems in a different way; the y-axis was changed to protein concentration divided by seeded cell number, and the result was not a flat line as would be expected for an ideal protein extraction. Interestingly, the trends in both cases stabilized with larger cell numbers, again suggesting that a minimum of 0.50 million cells should be used for preparation of metabolomic samples by our protocol.

OVCAR-8 samples from 10 cm dishes seeded with 0.25 – 4 million cells were also subjected to trypsinization followed by a trypan blue assay, and the cell counts were plotted against seeded cell number (Figure 1C). Although that method produced a linear curve (R^2 of 0.99), we noted substantial cell loss during the processing steps. For example, in samples in which 4,000,000 cells were seeded, just over 3,000,000 cells were counted, suggesting a recovery problem. The log/log slope of 1.20 and the disproportionally low signals yielded by the lower cell numbers in the right panel of Figure 1C suggest that the inaccuracy of trypan blue counting increases with decreasing cell number. Visual inspection of the dishes before processing indicated that no cell death had occurred during the 4-hour attachment time and that cells were indeed attached to the culture dish. It is possible that more efficient detachment methods could yield improved accuracy. However, it is noteworthy that samples used for such cell counting must be seeded separately from the samples to be processed for metabolomic data analysis, because trypsinization introduces a large number of artifacts to the metabolomic preparations and requires additional processing time, which in turn leads to conversion and degradation of metabolites. We conclude that trypan blue counting of trypsin-detached adherent cells is not an optimal method for normalizing metabolomic data. An alternative method—counting attached cells by in situ imaging—may circumvent the problems noted here but was not tested.

OVCAR-8 cell samples were prepared for analysis of DNA concentration using commercial DNA mini-prep kits. We initially compared kits from IBI Scientific and Qiagen

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(Germantown, MD). However, the dynamic range and overall DNA recovery were greater with the IBI Scientific kit (data not shown), so it was used for all further DNA preparations. DNA concentration was measured using a NanoDrop spectrophotometer and plotted against seeded cell number (Figure 1D). Like the other measures described thus far, DNA concentration yielded a strong linear relationship with seeded cell number ($R^2 = 0.98$), and the log/log slope of 0.90 was the closest to unity among the methods tested. When DNA concentration values were divided by the number of cells seeded (right panel, Figure 1D), deviations from a flat line were smaller than with other methods—less than 10% over four orders of magnitude. To calculate the percent deviation from the experimental mean, an average of all measured values divided by the seeded cell number for each normalization strategy was calculated (AVG), and the following equation was used:

$$\% \quad Deviation \quad from \quad Mean = \left(\frac{IND - AVG}{AVG}\right) \times 100$$

where IND represents each individual measured value divided by the number of cells seeded. Deviation from the mean for all four normalization strategies is shown in Figure 2, which illustrates that normalization by DNA concentration was the only method that yielded less than 20% deviation from the mean across the full range of seeded cell numbers. We note that all of the methods performed relatively poorly at 0.25 million cells, so all deviations reported from here onward refer to the results obtained for 0.5 million to 4 million cells.

Cell pellet mass and protein concentration determined from the cell pellet are not good surrogates for cell number

Because we noted an association between cell number and visual size of the cell pellet in the tubes after extraction and pelleting (data not shown), we decided to assess cell pellet mass as a possible normalization strategy. Prior to metabolite extraction, tubes were pre-weighed using an analytical balance. Metabolites were then extracted as described previously, the cell pellet was dried, and the tubes were re-weighed. The mass of the cell pellet was then calculated and plotted against cell number (Supplementary Figure 1A). The Pearson correlation was only 0.45, possibly because the mass range of the analytical balance was not sufficient for accurate weighing the small cell pellet yielded by 0.25 to 4 million cells.

Similarly, we re-suspended the *dried* cell pellet in bicine/CHAPS buffer for determination of protein concentration. However, the BCA assay yielded large variation among replicates and poor linear regression ($R^2 = 0.51$; Supplementary Figure 1B), suggesting incomplete re-solubilization of the protein from larger cell numbers. We concluded that neither protein concentration from the cellular debris nor the mass of the cell debris itself would be effective for normalization of metabolomic data from adherent cell lines.

Metabolite intensity is proportional to cell number

Metabolite extracts from the samples used for DNA concentration analysis were also analyzed by LC-MS. Four annotated metabolites (glutamic acid, uracil, N-acetylputrescine, and creatine) representing different compound classes were chosen for subsequent analysis. Deconvoluted intensities for those four metabolites were strongly correlated with cell number, with R^2 values of 0.99 (Figure 3). The log/log slope for each metabolite was calculated as described above, and all slopes were within 20% of unity. The strong relationship between metabolite intensity and seeded cell number emphasizes the importance of accounting for, and normalizing to, cell number when there is potential for cell number variation among samples at the time of processing.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The four metabolites in the left-hand panels of Figure 3 were then normalized using protein concentration, cell count, or DNA concentration and plotted together in the right-hand panels. The ideal normalization method would yield a flat horizontal line in the plot of normalized metabolite intensity vs. seeded cell number. Normalization by DNA concentration was the only method that showed less than 20% deviation from the mean normalized value across all four metabolites. Overall, DNA concentration from the same dish or cell count from a parallel dish both appeared to perform comparably well over four orders of magnitude. Significant deviation from the mean was noted with the protein normalization strategies, with as high as 60% deviation from the mean across multiple seeded cell numbers.

When we normalized metabolite intensities to the median ion intensity observed within an individual sample, deviations from the mean were as high as 150% (Supplementary Figure S2). That observation could potentially be attributed to internal standards, which were present at 10 μM, which accounts for a large percentage of the overall signal observed at low cell number yet a low percentage of the overall signal observed at high cell number. Since internal standards are a necessary component of our metabolite extraction protocol, normalization by median ion intensity was not considered to be a feasible option.

DNA concentration is a robust surrogate of cell number in multiple adherent cell types

To ensure that the normalization by DNA concentration method was applicable to a variety of adherent cell lines, we also tested the correlation of DNA concentration with seeded cell number in another ovarian cancer line (OVCAR-4), a glioblastoma line (U251), and a melanoma line (MDA-MB-435). The log/log slope of DNA concentration vs. cell number deviated from unity by less than 20% in all four cell lines. The R^2 values for the four cell lines were 0.98 or higher (Figure 4). Together, these findings suggest that measurement of DNA concentration is a robust normalization strategy across a wide range of cell types.

DNA concentration increases linearly with cell incubation time

To mimic a typical time-series metabolomic experiment, we then investigated whether DNA concentration would correspond to cell number over a 48-hour growth period. Using the same four cancer cell lines mentioned above, cells were seeded and allowed to attach for 48 h, at which point the media were changed, and cells were harvested at 48, 56, 72, and 96 h. DNA concentrations for each time point are plotted in Figure 5. Within the set of replicates at each time point, metabolite intensities were normalized to the largest DNA concentration value in the set. OVCAR-4, U251, MDA-MB-435, and OVCAR-8 cells are reported to have doubling rates of, 41.1, 23.8, 25.8, and 26.1 h, respectively.[22] DNA concentrations were determined for the four cell lines and were found to correspond with those doubling times. These DNA values were used to calculate the experimental doubling times for each cell line by fitting to an exponential growth curve.[23] The doubling times for each cell line based on DNA concentration were 40.99 (green squares), 25.19 (red triangles), 30.20 (purple inverted triangles), and 14.57 (blue circles) for OVCAR-4, U251, MDA-MB-435, and OVCAR-8 cells, respectively. The experimental doubling times for OVCAR-4, U251, and MDA-MB-435 obtained from DNA concentration were close to their published values, but the experimental doubling time obtained for OVCAR-8 cells was slightly unexpected; cells doubled at almost twice the published rate. That phenomenon corresponds to previous, unpublished observations from our laboratory; OVCAR-8 cells tend to proliferate more rapidly with increased confluence. The growth rates of U251 and MDA-MB-435 cell lines, in contrast, tend not to be affected by confluence, as supported by the strong consistency between the reported cell doubling times and the DNA doubling times we measured.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

CONCLUSIONS

It is often necessary to compare metabolomic data across time series, concentration series, different cell types, or tissues from different patients. When processing metabolomic data derived from those and other types of experiments, the choice of normalization strategy can dramatically influence the accuracy of the biological conclusions. We report here the development and validation of a useful normalization method for such data. Notably, we observed large variation in measured protein concentration from metabolite-extracted samples, yet such values are commonly used for normalization of metabolomic data from adherent cell lines. Since organic solvents are required to extract the largest number of detectable features,[3, 24] the preferred normalization method is to use DNA concentration, because it provides quantitatively robust accuracy across a range of cell types and does not require a separate, parallel sample, thereby allowing the metabolomic workflow to be streamlined. The method was also accurate and robust across four different cell lines and across time series data.

Despite the demonstrated effectiveness of DNA as a parameter for use in normalization, instances in which the technique would fall short should be considered. For instance, if comparing across cell lines that differ in ploidy or number of nuclei, the difference must be accounted for in the data normalization strategy. That may or may not compromise the utility of the method. For instance, treatment with DNA damaging agents would potentially arrest cell cycle, possibly leading to DNA:cell number ratios that differ from the control (untreated) sample.[25-26] In particular cases such as those, our results suggest that normalization by cell counting (trypan blue assay) from parallel dishes might be the best alternative. Finally, it must be noted that all normalization methods fail at very low cell number due to insufficient sensitivity. Accordingly, our results suggest that a minimum of 0.5 million cells should be used for accurate normalization of metabolomic data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by U.S. National Cancer Institute (NCI) grant numbers CA143883 and CA083639, the Chapman Foundation, and the Michael and Susan Dell Foundation (honoring Lorraine Dell). We thank David A. Pirman for editorial and scientific advice and Alexander (Sasha) Raskind for fruitful discussions and suggestions.

REFERENCES

(1). Spratlin JL, Serkova NJ, Eckhardt SG. Clinical Cancer Research. 2009; 15(2):431–440. [PubMed: 19147747]

(2). Vuckovic D. Analytical and Bioanalytical Chemistry. 2012

(3). Yanes O, Tautenhahn R, Patti G, Siuzdak G. Analytical Chemistry. 2011; 83:2152–2213. [PubMed: 21329365]

(4). Patti G, Yanes O, Siuzdak G. Nature Reviews Molecular Cell Biology. 2012; 13:263–272.

(5). Plumb R, Rainville P, Potts W, Castro-Perez J, Johnson K, Wilson I. Rapid Communications in Mass Spectrometry. 2007; 21:4079–4164. [PubMed: 18022959]

(6). Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, Brown M, Knowles JD, Halsall A, Haselden JN, Nicholls AW, Wilson ID, Kell DB, Goodacre R, Consortium THSM. Nature Protocols. 2011; 6(7):1060–1083.

(7). Rainville PD, Stumpf CL, Schockcor JP, Plumb RS, Nicholson JK. Journal of Proteome Research. 2007; 6:552–558. [PubMed: 17269712]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

(8). Veselkov KA, Vingara LK, Masson P, Robinette SL, Want E, Li JV, Barton RH, Boursier-Neyret C, Walther B, Ebbels TM, Pelczer I, Holmes E, Lindon JC, Nicholson JK. Analytical Chemistry. 2011; 83:5864–5872. [PubMed: 21526840]

(9). Dietmair S, Timmins NE, Gray PP, Nielsen LK, Kromer JO. Analytical Biochemistry. 2010; 404:155–164. [PubMed: 20435011]

(10). Kotarsky H, Keller M, Davoudi M, Leveen P, Karikoski R, Enot DP, Fellman V. PLoS One. 2012; 7(7):1–10.

(11). Hirayama A, Nakashima E, Sugimoto M, Akiyama S-I, Tomita M, Yuzawa Y, Soga T. Analytical and Bioanalytical Chemistry. 2012; 404(10):3101–3109. [PubMed: 23052862]

(12). Patti G. Journal of Separation Science. 2011; 34(1):3460–3469. [PubMed: 21972197]

(13). Manna SK, Krausz KW, Bonzo JA, Idle JR, Gonzalez FJ. Journal of Proteome Research. 2013; 12:2269–2281. [PubMed: 23586774]

(14). Du Y, Wang F, May K, Xu W, Liu H. Analytical Chemistry. 2012; 84:6355–6360. [PubMed: 22881398]

(15). Evans AM, DeHaven CD, Barrett T, Mitchell M, Migram E. Analytical Chemistry. 2009; 81:6656–6667. [PubMed: 19624122]

(16). Fong MY, McDunn J, Kakar SS. PLoS One. 2011; 5:e19963. [PubMed: 21625518]

(17). Simpson R. Cold Spring Harbor Protocols. 2006

(18). Sellick CA, Knight D, Croxford AS, Maqsood AR, Stephens GM, Goodacre R, Dickson AJ. Metabolomics. 2010; 6:427–438.

(19). Cao B, Aa J, Wang G, Wu X, Liu L, Li M, Shi J, Wang X, Zhao C, Zheng T, Guo S, Duan J. Analytical and Bioanalytical Chemistry. 2011; 400(2983-2993)

(20). Beltran A, Suarez M, Rodriguez MA, Vinaixa M, Samino S, Arola L, Correig X, Yanes O. Analytical Chemistry. 2012; 84:5838–5844. [PubMed: 22697410]

(21). Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. Analytical Chemistry. 2006; 78(3): 779–787. [PubMed: 16448051]

(22). dtp.nci.nih.gov

(23). Roth, V. http://www.doubling-time.com/compute.php

(24). Sellick CA, Hansen R, Stephens GM, Goodacre R, Dickson AJ. Nature Protocols. 2011; 6(8): 1241–1249.

(25). Barakat K, Gajewski M, Tuszynski JA. Current Topics in Medicinal Chemistry. 2012; 12(12): 1376–1390. [PubMed: 22794314]

(26). Bennett MJ, Chan G, Rattner JB, Schriemer DC. Cell Cycle. 2012; 11(16):0–9.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

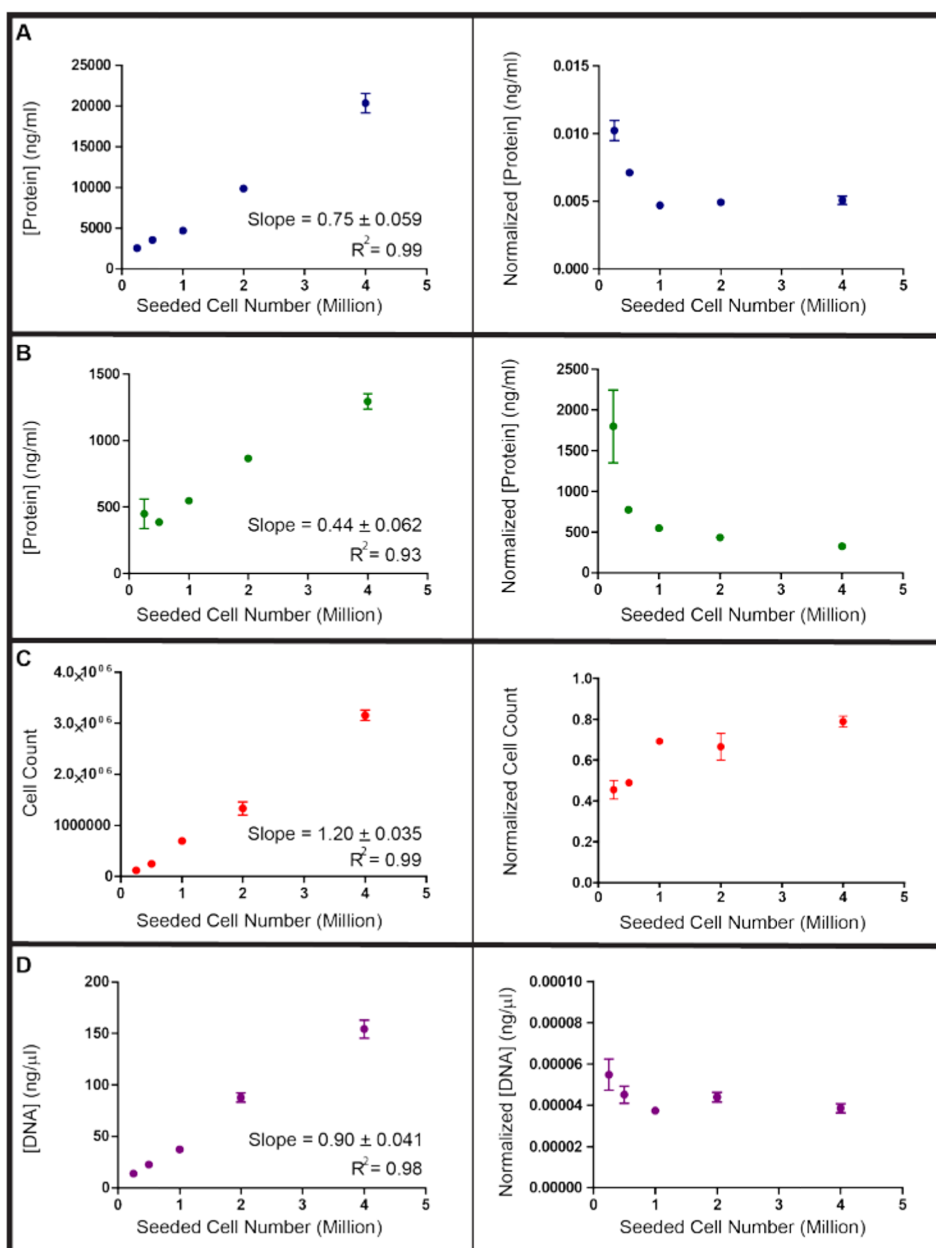

Figure 1.

Comparison of different normalization methods. Triplicate (n=3) 10-cm dishes of OVCAR-8 cells were seeded at 0.25, 0.5, 1, 2, or 4 million cells per dish, the cells were allowed to attach, and four different assays were performed. (A) total protein extracted from a surrogate dish with bicine/CHAPS buffer; (B) total protein extracted from the cell pellet after extraction with MeOH/ACN; (C) cells counted with trypan blue; and (D) total DNA extracted from the cell pellet after extraction with MeOH/ACN. The left-hand panels diplay the log/log slope and R^2 for each normalization type; the right-hand panels display assay values divided by the actual seeded cell number.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

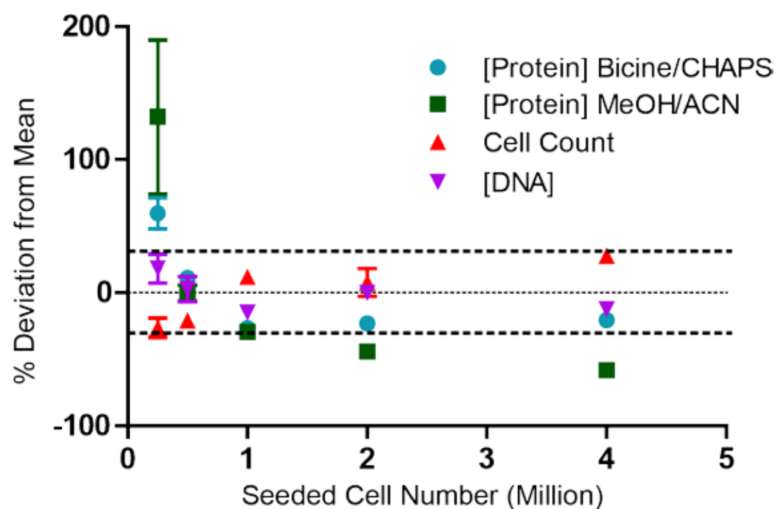

Figure 2.

Scatter plot displaying the percent deviation from the mean for the four tested normalization methods. The y-axis indicates averaged value of each normalization method divided by the seeded cell number. Thick dashed lines represent 20% deviation from the mean normalized value for each strategy.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

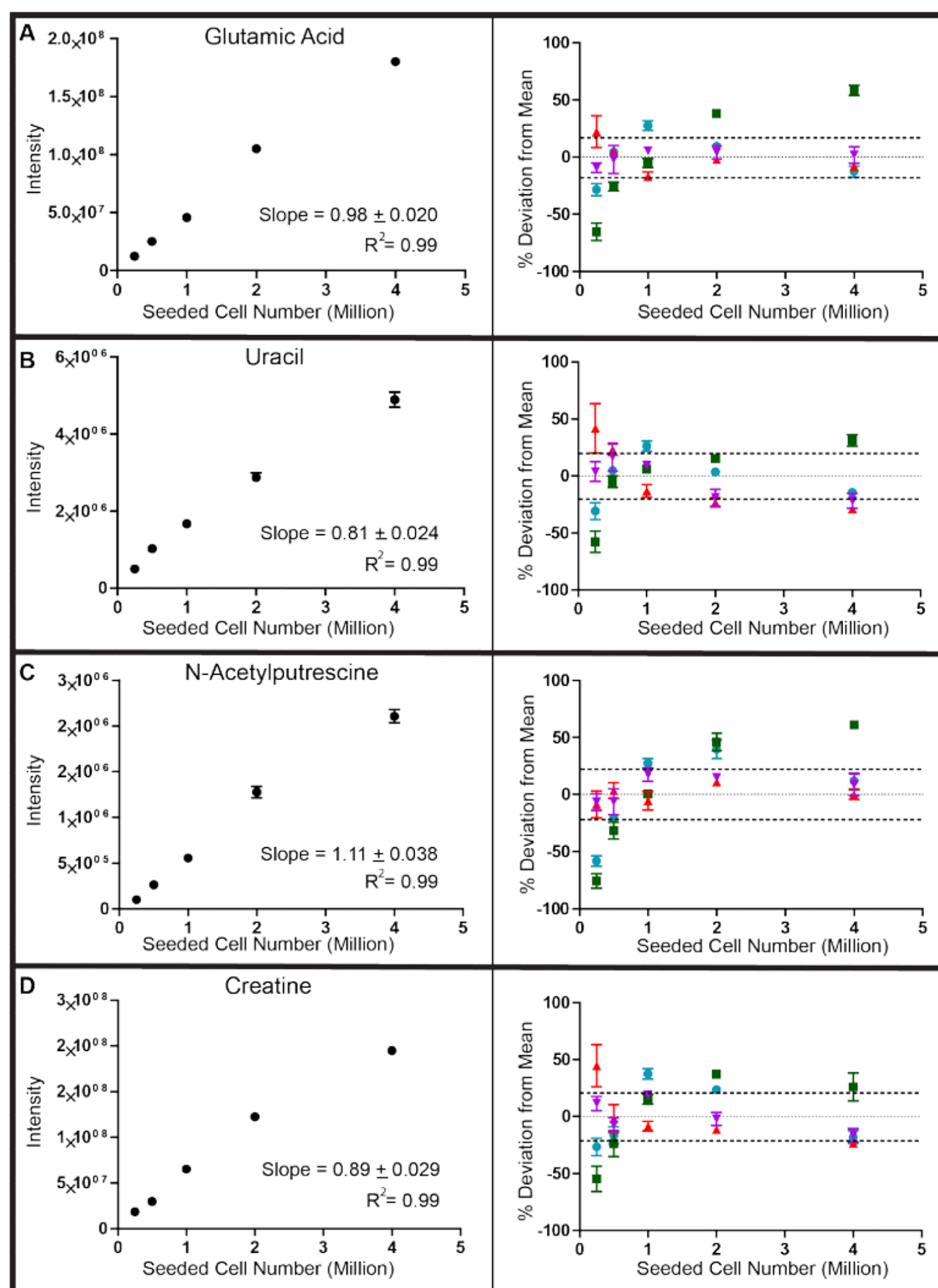

Figure 3.

Metabolite intensity as a function of cell number. In the left-hand panels, values for four metabolite intensities are plotted against seeded cell number. In the right-hand panels, scatter plots display the percent deviation from the mean for metabolite intensity normalized by each of the four normalization methods tested. Blue circles and green squares represent normalization by protein concentration in bicine/CHAPS buffer and MeOH/AMBIC, respectively. Red triangles represent normalization by cell counting, and purple inverted triangles represent normalization by DNA concentration. Dashed lines indicate 20% deviation from the mean.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

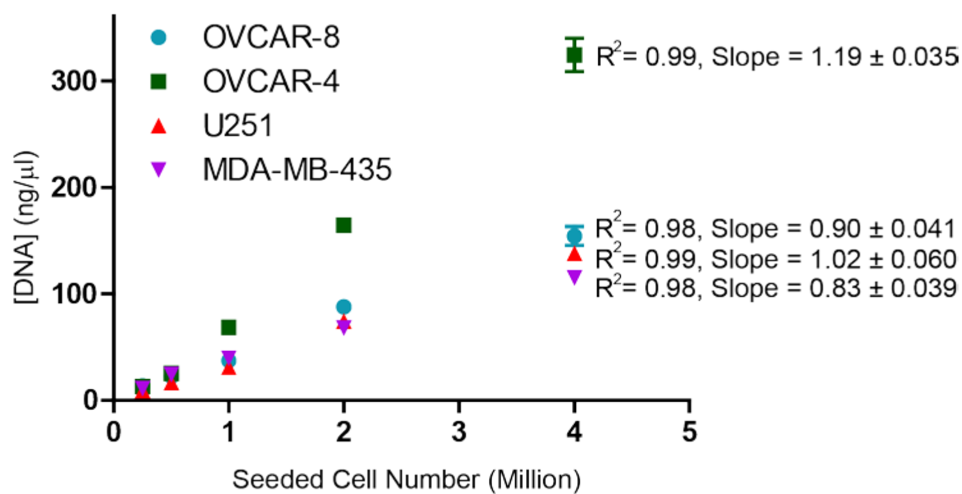

Figure 4.

DNA concentration as a function of cell number for four different cell lines. DNA was extracted from the cell pellet of metabolomic preparations from OVCAR-8,OVCAR-4, U251, and MDA-MB-435 cancer cells. Coefficient of determination (R^2) and slope values are listed beside the respective data points for each cell line.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

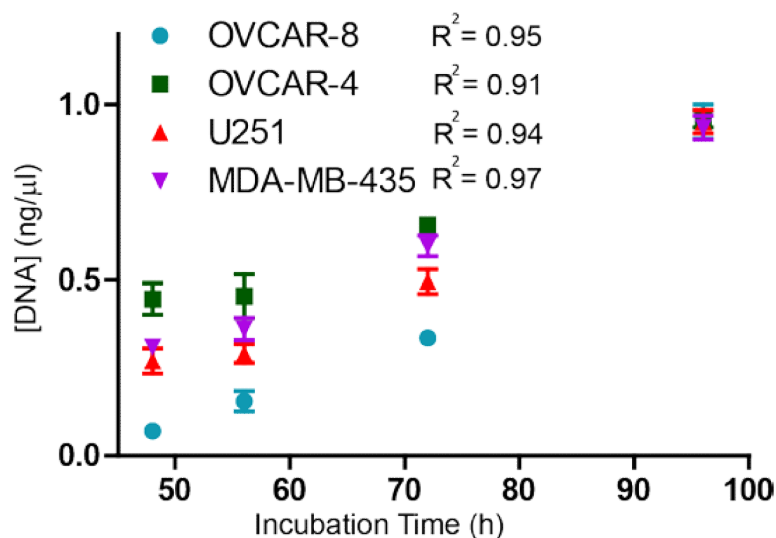

Figure 5.

DNA concentration as a function of time in four different cell lines. Cells were seeded and incubated for 48 h at 37 °C and 5% CO_2. At 48 h, media were replaced with fresh media, and cells were harvested at 48 (immediately), 56, 72, or 96 h. The DNA doubling times for OVCAR-4, U251, MDA-MB-435 and OVCAR-8 were calculated to be 40.99, 25.19, 30.20, and 14.57 h, respectively, calculated based on an exponential growth curve (corresponding to “log-phase” growth).[23]