Dear Dr. Farley and the Editorial Board of Scientific Reports,

We read with great interest two recent papers published in *Scientific Reports* by Mesnage and colleagues (Mesnage, 2016, Mesnage 2017). Both papers utilize quantitative proteomics to identify changes in protein expression in response to 1) genetic modification of maize and 2) ingestion of glyphosate in rats, respectively. Over the past several years, mass spectrometry-based proteomics has developed into an incredibly powerful tool for the examination of complex systems such as those interrogated in these two studies. We compliment the authors on their use of tandem-mass tags (TMT) for quantitative analysis. However, this method is technically challenging and requires an adherence to standards developed over the past several years by the computational proteomics community in order to ensure accurate identification and quantitation of proteins, and, importantly, avoid erroneous results of data analysis. In our analysis of these papers, we found that the methods used to calculate relative quantification of proteins were fundamentally flawed and do not conform to the standards of the field, thereby undermining the interpretation and conclusions of both studies.

Proteins are improperly quantified

These papers are not technically sound. The conclusions may be accurate when the data is properly analyzed, but in its current state, they cannot be supported by the calculations used. Protein expression changes are used to argue that 1) glyphosate causes fatty liver disease in rats and 2) that transgenic maize is substantially different that unmodified maize. Even with the lack of methodological details in these papers (described below), it is clear that the authors do not quantify proteins.

For standard protein abundance calculation based on TMT labelling, averaging across all peptides of a protein between two or more experimental conditions gives abundance ratios of proteins. Individual peptides may show outlier behavior, and are often not an accurate indicator of the behavior of full proteins, which are composed of multiple peptides. In their seminal review of TMT quantification, Rauniyar and Yates clearly state that, “Peptide abundance ratios are calculated by combining data from multiple fractions across MS runs and then averaging across peptides to give an abundance ratio for each parent protein.” In TMT quantification, single peptides are highly variable, with robust quantification derived from the accumulated evidence from multiple peptides and conclusions about protein levels from single peptides are often discarded as unreliable (Rauniyar, 2014). More observations per protein lead to smaller standard error (Mahoney, 2011). Here, however, the authors rely on quantitation of individual mass spectra as a proxy for protein abundance measurements. If any individual peptide from a protein was found differently across samples, even just a single observation, the entire protein was counted as differentially expressed.

The following points describe quantification errors in more detail:

1. Though the publications clearly claim to calculate and interpret differences in protein expression, these papers instead calculate fold changes for individual peptides. Even peptides with the same sequence and different post-translational modifications are kept

separate from each other for quantification. We can provide a clear demonstration of this fact. As a standard of the field, protein quantification should be based on the combination of data from multiple peptides mapping to each protein (Rauniyar, 2014).

2. The representation of individual peptides as proxy for full proteins is clearly demonstrated by Supplementary Dataset 3 of Mesnage 2016, where each of the listed perturbed proteins is misannotated with a column 'Mass (Da)' that actually represents the unitless mass/charge ratio (m/z) of individual peptides, and not mass (Da) of a peptide or protein.
3. If any one single peptide from a protein was found differently across samples, even just a single observation, the entire protein was counted as differentially expressed.
4. In certain cases, different peptides from the same protein measure increased and decreased levels, resulting in the same protein listed multiple times, paradoxically both increased and decreased across conditions.
5. Common ambiguous and isobaric peptide assignments are not corrected. In certain cases, multiple entries are included in the results table based on the observation of a single shared or isobaric peptide (isoleucine/leucine ambiguity, deamidation, etc.). In other cases, non-unique peptides that match multiple proteins in the maize or rat proteome are listed with a single protein ID. While protein groups are commonly assigned, how the major protein was selected from multiple possible is unclear. As stated by Rauniyar and Yates, 2014, "Significant quantification errors arise if a quantified peptide is not unique to its corresponding protein".

As the conclusions of both these papers depend on flawed quantification of proteins, the results presented in both papers are likely incorrect.

Insufficient methodological details

Mass spectrometry papers require detailed methods so that others can reproduce and understand analyses. While we were able to deduce some of the methods used, we have been unable to fully replicate the calculations due to insufficient description of data analysis. These papers provide only very superficial methods that do not adhere to the standards of the field, and leave many outstanding questions:

1. In Mesnage 2016, the database of protein sequences used to identify peptides is described as the maize proteome, however, supplementary tables inexplicably identify fungal maize pathogens. There is no mention of a common contaminant database, or why non-maize proteins were included in the proteomic database.
2. There are several known sources of quantification error for TMT labelling quantification, including 1) efficiency of labeling, 2) isotopic contamination of isobaric mass tags, and 3) convolution of multiple co-isolated TMT signals within the isolation window (Sandberg 2014). These are well known issues with TMT-based quantitation. Approaches to correct for these errors are well established in the proteomics community, but were not addressed by the authors.
3. Important details regarding MS instrument parameters are missing. We are particularly interested in these settings due to the apparent low number of high confidence PSMs identified in these two paper. Counts of PSMs are an order of magnitude lower than would be expected for these experiments. In Mesnage 2016, there are only 3,000 PSMs for each maize replicate, while Mesnage 2017 has 6,000 for each rat replicate. A comparable study identified 35,000 and 45,000 PSMs corresponding to 3300 and 3900 proteins, respectively (Plubel, 2017).

4. The methods for the database search list an MS2 fragment mass tolerance of 0.8 ppm. For standard TMT analysis, the MS2 mass tolerance is normally set to 0.8 Daltons. While this parameter is dependent on the resolution of the MS2 spectra, 0.8 ppm is unusually strict at any resolution. If the tolerance described is accurate, it would explain the very low number of peptide-spectral matches (PSMs) identified at high confidence.
5. There is no description of error modeling to determine the false discovery rate of PSMs nor the criteria for high confidence PSMs.
6. Details regarding the calculation of peptide/protein abundances are completely missing. While Mesnage 2017 mentions an uncited SumScale normalization, Mesnage, 2016 does not discuss how quantitative TMT data was normalized or fold change ratios calculated.

No access to raw data

The Scientific Reports instructions to authors states that: "*An inherent principle of publication is that others should be able to replicate and build upon the authors' published claims. Therefore, a condition of publication in Scientific Reports is that authors are required to make materials, data and associated protocols promptly available to readers...*", and policy for all Nature journals mandates public deposition of raw data in a repository affiliated with the PRIDE (Perez-Riverol, 2015) proteomics data consortium. When asked for raw data for Mesnage, 2016, the authors were only able to provide a highly processed dataset. It is standard in the field of mass spectrometry proteomics for spectral files (i.e. .RAW, .mzXML, .mzML etc.), protein sequence database, and unprocessed search results to be deposited in an online repository. Access to this data is required to fully evaluate evidence for protein abundance. If the authors cannot provide access to the raw spectral data, it is our strong opinion that the two papers must be withdrawn.

Unprocessed data and methodological details are absolutely required to characterize the analytical step taken in these two papers. Lack of raw data notwithstanding, processed files show incorrect interpretation of mass spectrometry data, resulting in unsupported claims. We urge the editors of Scientific Reports to take appropriate action. At minimum, we request that the editors enforce their data availability mandate so that we can undertake a full reanalysis of Mesnage, 2016 and Mesnage, 2017. In summary, we would like the editor to require the authors to remedy the problems we have identified:

1) Errors in relative protein quantification across conditions
2) Inadequate/missing description of methods, preventing replication of data analysis
2) No access to raw data

Claire D. McWhite*
Daniel R. Boutz, Ph.D.
The University of Texas at Austin
*claire.mcwhite@utexas.edu

https://www.nature.com/authors/policies/availability.html

Mesnage R, Agapito-Tenfen SZ, Vilperte V, et al. An integrated multi-omics analysis of the NK603 Roundup-tolerant GM maize reveals metabolism disturbances caused by the transformation process. *Scientific Reports*. 2016;6:37855. doi:10.1038/srep37855.

Mesnage R, Renney G, Séralini G-E, Ward M, Antoniou MN. Multiomics reveal non-alcoholic fatty liver disease in rats following chronic exposure to an ultra-low dose of Roundup herbicide. *Scientific Reports*. 2017;7:39328. doi:10.1038/srep39328.

Perez-Riverol Y, Alpi E, Wang R, Hermjakob H, Vizcaíno JA. Making proteomics data accessible and reusable: Current state of proteomics databases and repositories. *Proteomics*. 2015;15(5-6):930-950. doi:10.1002/pmic.201400302.

Plubell DL, Wilmarth PA, Zhao Y, Fenton AM, Minnier J, Reddy AP, Klimek J, Yang X, David LL, Pamir N. Extended Multiplexing of Tandem Mass Tags (TMT) Labeling Reveals Age and High Fat Diet Specific Proteome Changes in Mouse Epididymal Adipose Tissue. *Molecular and Cellular Proteomics.* 2017;16(5):873-890. doi: 10.1074/mcp.M116.065524.

Rauniyar N, Yates JR. Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics. *Journal of Proteome Research*. 2014;13(12):5293-5309. doi:10.1021/pr500880b.

Russell C, Heslegrave A, Mitra V, Zetterberg H, Pocock J, Malcolm A. Ward W, Pike I. Combined tissue and fluid proteomics with Tandem Mass Tags to identify low-abundance protein biomarkers of disease in peripheral body fluid: An Alzheimer's Disease case study. *Rapid Communications in Mass Spectrometry.* 2017;31:153–159. doi:10.1002/rcm.7777.

Sandberg A, Branca R, Lehtiö J,Forshed J.Quantitative accuracy in mass spectrometry based proteomics of complex samples: The impact of labeling and precursor interference. *Journal of Proteomics.* 2014;96:133-144. doi: 10.1016/j.jprot.2013.10.035

Conflict of interest statement
Neither DRB nor CDM declare any conflicts of interest related to these papers.