The spatial subcellular proteome is a dynamic environment; one that can be perturbed by molecular cues and regulated by post-translational modifications. Compartmentalization of this environment and management of these biomolecular dynamics allows for an array of ancillary protein functions. Profiling spatial proteomics has proved to be a powerful technique in identifying the primary subcellular localization of proteins. The approach has also been refashioned to study multi-localization and localization dynamics. Here, the analytical approaches that have been applied to spatial proteomics thus far are critiqued, and challenges particularly associated with multi-localization and dynamic relocalization is identified. To meet some of the current limitations in analytical processing, it is suggested that Bayesian modeling has clear benefits over the methods applied to date and should be favored whenever possible. Careful consideration of the limitations and challenges, and development of robust statistical frameworks, will ensure that profiling spatial proteomics remains a valuable technique as its utility is expanded.

1. Protein Subcellular Localization is a Key Component of Function

Correct subcellular localization affects protein function. This includes the availability of binding partners, cofactors, and substrates, as well as the presence of regulatory factors such as kinases. The interrogation of localization is therefore a necessary step to elucidate function. The localization of the proteome is highly dynamic, for example, trafficking of secretory pathway components through membrane-bound compartments and shuttling of proteins, such as transcription factors, between compartments based on their phosphorylation status. In response to activation of signaling pathways, there is mounting evidence that many proteins “moonlight” and perform multiple discrete functions depending on the cellular state and protein localization.

Altogether, this suggests that to understand protein function, we need to interrogate the proportional subcellular distribution of proteins and the role of post-translational modifications (PTMs) in regulating protein localization dynamics.

There is a multitude of experimental techniques to study proteome localization, including interactome mapping using proximity tagging, high throughput microscopy, and quantitative mass spectrometry. Here, we focus on profiling spatial proteomics, a high-throughput mass spectrometry-based technique to establish protein subcellular localization in cells by quantifying protein abundance within subcellular fractions created by biochemical fractionation such as centrifugation or detergent solubility. The principle of this approach is that proteins from the same subcellular niche will share a distinct abundance profile across the fractions. Protein localization can then be determined using semi-supervised analyses and prior information regarding sets of marker proteins from a limited selection of subcellular niches. This has proved to be a very powerful and flexible technique but, to date, it has largely been limited to analyzing the primary protein localization under a single condition. Profiling spatial proteomics is increasingly being used to map multiple localizations of proteins, dynamic localization upon perturbation, and the role of post-translational modifications. Now is an opportune moment to reflect on the current paradigm of profiling spatial proteomics and the inherent technical challenges and limitations of these powerful techniques. Here, we review the challenges associated with profiling spatial proteomics and recommend formal testing frameworks and modeling with explicit consideration of limitations of the method as a means to maximize its utility.

2. Identifying the Main Localization

A single protein copy can only be present in one localization at any given time, although it may transit or traffic within the cell during its lifecycle from its point of synthesis to the location(s) where it functions, onto where it is finally degraded. Since profiling spatial proteomics assays multiple cells, each containing multiple protein copies, it captures protein copies
at different stages of their lifecycle, as well as cells in different cell states. With profiling spatial proteomics, a pool of cells is fractionated into subcellular fractions which contain cellular material from multiple localizations in differing proportions. Thus, quantified abundance profiles represent an aggregation of protein localization across different cell cycle stages and cellular states.

Most proteins have been thought to adopt a single primary localization in which they are resident for most of their life cycle. Profiling spatial proteomics methods have largely focused on identifying this primary localization. By establishing localization-exclusive “marker” protein profiling, profiles of proteins with unknown localization can be assigned to their respective primary localizations. This approach was initially used to separate centrosomal proteins from non-specific proteins[27] and rapidly adopted for cell-wide localization assignment.[14,15] In these pioneer studies, the proteome coverage remained relatively low and assignment was carried out either by correlating profiles[15] or using partial least squares-discriminant analysis.[14]

Improvements in mass spectrometry have enabled deeper proteome coverage and this has facilitated the use of machine learning classifiers which learn the marker profiles for each localization and then assign proteins with unknown localization. The primary or main localization is then defined as the marker class which best reflects the profile of a given protein, with some filtering process to remove uncertain assignments.[28,29] This is a task to which many algorithms are suited, though, support vector machines (SVM) or neural networks are typically used.[18,19,30] An exploratory process of marker selection and detection of unannotated niches is usually performed prior to classification to ensure definition of an optimal set of marker classes. Existing generic algorithms can be used for this process, including K-means,[31] Mclust, hierarchical clustering,[31] DBSCAN,[32] and by visual inspection through hexbin plots or t-SNE.[33] Alternatively, a profiling spatial proteomics-specific method has been developed based on building an outlier statistic from iterative mixture modeling called phenoDisco.[34] To facilitate all these analyses in a reproducible environment, an extensive R package, pRoloc,[36] has been developed to visualize spatial proteomics profiles and reliably infer protein localization using machine learning,[35] and similar functionality has been added to Perseus.[36]

Machine learning-based classification with profiling spatial proteomics has been hugely successful at expanding knowledge of protein subcellular localization. Recent example publications have used SVM to assign 2855 protein groups to 14 subcellular niches in mouse E14TG2a embryonic stem cells,[18] 2423 proteins into nine membrane organelles in HeLa cells,[20] and 9286 proteins to four major localizations across five human cancer cell lines.[10]

3. Technical Considerations for the Main Localization Question

Although here we focus on considerations for studies determining multiple and dynamic localization of the proteome, workflows designed to return the main location of a protein are extremely worthwhile. The analysis of resulting data requires careful attention with perennial challenges associated with the classification approach taken.

Oliver M. Crook is a Wellcome funded Ph.D. student in the Cambridge Centre for Proteomics with Kathryn Lilley and is co-supervised by Paul Kirk in the MRC Biostatistics Unit. He is keen to work with all kinds of tricky data generated in system biology applications. He is developing Bayesian approaches to analyzing spatial proteomics data, so we can answer new questions with the data. Previously, he obtained a Master’s degree in mathematics at the University of Warwick focusing on dynamical systems and infectious disease modeling. When he is not working, he enjoys cooking and playing the flute.

Tom Smith is fascinated by the complexities of intracellular macromolecular processes, with a particular interest in the function of alternative splicing and the determinants of RNA and protein localisation. His doctoral studies focused on transgenerational stress memories in A. thaliana. He made the move into computational biology via a fellowship with CGAT (Oxford, UK), where he developed and applied bioinformatic tools to analyse diverse high-throughput data. In the group of Prof. Kathryn Lilley (Cambridge, UK), he brings together novel laboratory and data analysis approaches to interrogate the role of protein and RNA interactions in driving their respective subcellular localisation.

Mohamed Elzek is a physician scientist with a passion in everything omics-related, especially proteomics. He is now working on his Ph.D. at Kathryn Lilley’s group at the University of Cambridge, where his focus is on giving each protein and RNA molecule a postcode inside the cell through the process of DNA damage repair. He received his medical training at Alexandria University in his home country, Egypt. He then went on to gain experience in cancer proteomics research with Dr. Karin Rodland at the Pacific Northwest National laboratory in the United States and at Prof. Reudi Aebersold’s lab in Switzerland.
In the first instance, the selection of appropriate marker proteins is a vital step in the classification workflow and the categorization of markers is typically dataset dependent to ensure they properly represent a single localization. However, problems can arise when annotating compartments with only a few well annotated marker proteins including, amongst others: endosomes, poorly characterized localizations such as cytosolic granules, and cells with highly specialized organelles and/or with poor annotation.

Independently of the precise manner in which markers are selected, they inevitably represent a biased sample with a skew toward well documented proteins and sub-cellular locations. Therefore, the accuracy of marker classification does not indicate classification accuracy for non-marker proteins, as is sometimes suggested, as non-markers can be expected to be classified with lower accuracy. Application of SVM for classification purposes has been a very popular approach to date, however, the interpretation of SVM scores requires particular care. The SVM is a discriminative model, rather than a generative one, and so SVM scores do not represent probabilities. Indeed, the top SVM score for a multi-class classification may not be the most probable localization; that is, the SVM probabilities need not be consistent. If probabilities are desired, these can be approximated using, for example, quadratic optimization, as others have done in profiling spatial proteomics studies, but such approximations may be arbitrarily inaccurate. Alternatively, they can be estimated using additional hold-out data not used to train the classifier, but this is rarely available given the number of marker proteins per class.

3.1. The Extended Questions

While classification algorithms are well established and valuable to identify the main localization, they are inadequate to address the extended questions of mixed protein localization, differential localization upon cell state perturbation, and the interplay between post-translational modifications and protein localization. Below, we set out what we consider to be clear definitions for these terms as the intuitive interpretations from a biological standpoint may not match up to what we can assay with profiling spatial proteomics.

3.1.1. Multi-Localized Proteins

These include secretory pathway components which cycle through membrane organelles and RNA binding proteins such as HuR that shuttle between the cytosol and nucleus. Multi-localization is a poorly defined term in profiling spatial proteomics, with the definition arising from the schema used to assign the primary localization. Here, we define a protein to have multi-localization if it resides in more than one cellular compartment across the assayed cells. By this definition, the identification of primary localization for a protein is not, necessarily, mutually exclusive with an additional assertion that it has multi-localization. This definition is agnostic to whether the multi-localization is within or between cells, since profiling spatial proteomics cannot distinguish between these two possibilities - contrary to single cell methods, including (amongst others) microscopy-based approaches. Furthermore, we cannot distinguish the root of multi-localization; the process of pools of proteins dynamically interchanging is fundamentally different from discrete pools of translated proteins where the components never interchange, yet both will simply be observed as a multi-localization.

3.1.2. Protein Relocalization

This may occur in response to internal and external cues. We define these changes in either comparative or time series profiling spatial proteomics experiments as instances of differential localization. Alterations in protein synthesis or protein degradation may also be observed as changes in localization and hence we refer to this eventuality as differential localization to avoid any inference of active protein movement. Where a consistent change in localization occurs for the majority of protein copies assayed, one may observe a discrete differential localization from one localization to another. However, in many cases, the observation will be one of proportional changes in localization.

3.1.3. Post-Translational Modifications

These may regulate protein localization. A commonplace example is phosphorylation. In spatial proteomics studies, researchers typically aim to identify differences in protein localization governed by PTM status. A crucial consideration is that we cannot infer whether the modification is specifically regulating localization as it is also possible that the localization is regulated by other factors and the protein differentially modified according to its localization. Therefore, it is appropriate to avoid asserting these events as PTM-dependent localization; rather, these are best described as concurrent changes in PTM and protein localization.

All of the above questions, in theory, are answerable with current profiling spatial proteomics techniques. However, addressing these questions requires the development of new computational approaches. With these extended questions in hand, we turn to the important technical considerations for profiling spatial proteomics as we proceed beyond primary localization classification.

4. Technical Considerations for the Extended Questions

Before we address the potential solutions to the extended questions, we consider the additional technical considerations that need to be addressed when collecting and analyzing spatial proteomics data with these questions in mind (see Table 1).

First, as a matter of course, the proportion of features (peptide-spectrum matches, peptides or proteins) with missing values will necessarily increase as the overall number of samples increases for the more extensive experimental designs required. Here it is
Box 1. Proportional and Relative Quantification

It is not possible to estimate the proportion of protein resident in each of its subcellular niches using relative protein quantification. To see this, consider the following argument. During relative MS quantification there is a sampling rate $c$ which represents the proportion of protein analyzed, relative to the absolute amount of protein in that fraction. This sampling rate $c$ is corrupted by the loss of material as a result of protein extraction and the proportion of the resultant sample analyzed by MS. The former is unknown, whilst the latter is measurable - consequently, $c$ is unknowable.

First, let $f_a$ be the proportions of a protein resident in organelle $a$ across the fractions, likewise for $f_b$. Thus, $\sum f_a = 1$ and $\sum f_b = 1$, where $i$ denotes the $i$th fraction. Under MS quantification, the proportions are transformed according to $c$, such that we observe $c f_a$ and $c f_b$. These are normalized to obtained $\hat{c} f_a$ and $\hat{c} f_b$, to place them on the same scale (for example $\hat{c} = c / \sum c f_j$ for $j = a, b$).

We have defined the measured profile of a protein with respect to its true proportion and now consider a protein $f_i$ which has mixed localization between niches $a$ and $b$, with proportions $\pi$ and $1 - \pi$, respectively. Thus, $f_i = \pi f_a + (1 - \pi) f_b$ and so when measured by MS quantification we obtain

$$\hat{c}_a (\pi f_a + (1 - \pi) f_b)$$

However, mixing the relative proportions of observed profiles $\hat{c}_a f_a$ and $\hat{c}_b f_b$ results in

$$\pi \hat{c}_a f_a + (1 - \pi) \hat{c}_b f_b$$

which are not equal.

Important to consider that missing values may be “missing not at random,” for example, dependent on treatment, or “missing at random,” for example, due to stochastic processes inherent to data-dependent acquisition MS. The latter is especially problematic for label free quantitation (LFQ) approaches and despite efforts to identify the optimal imputation approach for LFQ, the resultant abundances are still relative to derived from the exact same peptide-spectrum match (PSM), which allows the quantification values for each fraction to be deemed an abundance profile across the fractions. Isobaric abundance profiles are typically scaled to generate an abundance profile across the fractions. Isobaric tagging allows the quantification values for each fraction to be derived from the exact same peptide-spectrum match (PSM), which reduces the variance of quantification significantly compared to LFQ. However, the resultant abundances are still relative to the amount of protein which was labeled. In either case, there are different total quantities of protein present in each experimental fraction and thus, we cannot estimate the mixture proportions from the observed profiles (see Box 1). To demonstrate this, we simulated multi-localization using previously published data. Relative quantification values were adjusted to approximate proportions of protein in each fraction. Multi-localization between the cytosol and mitochondria was then simulated by combining proportional quantification profiles for respective marker proteins and converted to relative abundances to observe how they would appear in a profiling spatial proteomics experiment. The comparison with directly combining relative profiles in the same ratios indicates that relative profiles do not capture multi-localizations accurately (Figure 1). Resolution of this problem requires conversion of relative protein abundances within fractions to proportional abundance across fractions but this is difficult to achieve by simply quantifying the total proportion of protein in each fraction, since extraction of the protein results in a loss of material. An alternative and promising approach for SILAC-compatible systems is to spike-in a consistent heavy labeled reference sample to achieve proportional quantification.

Third, when studying differential localization, it is important to consider that cell lysis, organelle morphology and/or cellular sub-structure may be considerably altered across conditions. For example, in order to determine the content of specific vesicles captured by specific golgins, Shin et al. relocated the golgins to the mitochondria by replacing their Golgi targeting domains with a mitochondrial transmembrane domains. This relocation leads to increased mitochondrial “zippering” and thus the mitochondria and interacting peroxisomes sediment at a lower centrifugation speed (Figure 2). Furthermore, we have previously observed that the truncated G1 and S phases in mouse embryonic stem cells have a significant impact on the resolution of Golgi profiles. Given that the morphology of many organelles is altered during mitosis, one would expect that conditions which alter the cell cycle stage distribution may significantly affect organelle morphology.

Finally, the lack of a suitable number of ground truths or strong prior expectations for all the extended questions severely hampers the development, implementation and comparison of tools to address them. Consider PTMs as an example; the role of phosphorylation in signaling pathways is well appreciated but the number of phosphorylation sites with known impact on localization which have been experimentally validated is limited when compared to our knowledge about the main subcellular localization of proteins. As such, computational methods to examine the...
role of phosphorylation in protein localization cannot take advantage of ground truths, or even strong prior expectations, and the assessment of the validity of the results from any method is not straightforward. With these considerations in mind, we now examine the attempts that have been made so far to address the extended questions with profiling spatial proteomics.

5. First Attempts to Address the Extended Questions

The classification approach has been highly effective at identifying the main localization of proteins; however, this framework has led the field to adopt sub-optimal methods to answer the extended problems (see Table 2).

As a first example, differing estimates have been reported for the proportion of proteins which are multi-localized using profiling spatial proteomics. We have previously suggested that approximately half the proteome is multi-localized, since it cannot be classified to a discrete localization, although we noted there are many explanations for this, not all of which relate to multi-localization. Taking a similar approach, Orre et al. observed consistent classification between biological replicates, with further analysis indicating that inconsistent classifications were likely due to inaccurate quantification. From this, they suggest that less than 10% of the proteome is multi-localized. In both these studies, a classification schema designed to determine the main localization was repurposed into an estimation of the proportion of multi-localized proteins and the disagreement is likely the result of poorly considered definitions for multi-localization. Indeed, as previously noted, successful identification of the main localization for a given protein is not mutually exclusive with that protein being multi-localized, and absence of primary classification has many explanations.

Similarly, dynamic localization has also been studied within the classification framework by treating the problem as two separate classification tasks and comparing classifications differentially between control and treatment. Whilst this is a valid approach, it can only identify clear cases of discrete differential localization and misses smaller changes in proportional localization.

To address the extended question of multi-localization, differential localization and PTMs in a classification-independent manner, informal methods have been introduced. These approaches have not clearly stated their assumptions or adequately justified their methodology and the community has not adopted a standardized rigorous approach. To elaborate, let us consider the approaches of Krahmer et al., where they attempt to address all three of these questions. First, they extend the protein correlation profiling approach with the goal of determining dual localizations (see Box 2). Second, to identify differences in the profiles between conditions, a test based on correlating intra- and inter-condition profiles was proposed (see Box 3). Finally, to analyze the spatial phosphoproteome, they apply several filtering steps along with their proposed method for correlating intra- and inter-condition profiles (see Box 4). In all three cases, assumptions are not explicitly stated or evaluated, the testing framework involves...
Table 1. Technical challenges in profiling spatial proteomics with particular relevance to addressing the extended problems. Missing values are problematic for the simple task of determining primary localization but even more so when addressing the extended problems. Similarly, ground truths are required for the simple task but these are usually readily available. Relative quantification and inconsistent organelle morphology are only challenging for the extended problems.

| Technical challenge                                      | Explanation                                                                 | Questions affected                          |
|----------------------------------------------------------|-----------------------------------------------------------------------------|---------------------------------------------|
| Missing values                                           | More samples result in more missing values, especially when comparing diverse conditions. | • Differential localization                |
| Relative quantification                                  | Protein quantification relative to the fraction cannot be readily converted to protein proportion in each fraction. | • PTMs                                      |
| Inconsistent organelle morphology and/or cell composition | Treatments can alter organelle Morphology/cell composition, invalidating implicit assumptions of the testing framework. | • Multi-localization                          |
| Insufficient ground truths or strong prior expectations   | We have strong expectations for protein main localization but far fewer expectations for the extended questions. Thus, proposed solutions cannot use prior knowledge and assessment of the quality of results obtained is difficult. | • Differential localization                 |

Table 2. Studies addressing the extended problems of multi-localization, differential localization, and the role of post-translational modifications. The computational approach used and study findings are briefly summarized.

| Study                              | Computational approach                                             | Findings                                                                 |
|------------------------------------|---------------------------------------------------------------------|-------------------------------------------------------------------------|
| HeLa cells (Itzhak et al.)         | Informal Mahalanobis distance-based test (MR method)               | 2824/5265 proteins multi-localized                                      |
|                                    |                                                                     | Four protein differentially localized in response to EGF stimulation    |
| HCMV infection in human lung fibroblasts (Beltran et al.) | Comparative SVM Classification                                    | 270/2775 proteins multi-localized                                       |
|                                    |                                                                     | 83 host protein differentially localized during infection              |
| Mouse neurons with EGF stimulation (Itzhak et al.) | SVM Classification                                                  | 66 protein differentially localized with an estimated FDR <10%           |
| AP-4 Knockout in HeLa cells (Davies et al.) | MR method                                                          | Three proteins differentially localized (1% FDR)                       |
| AP-5 Knockout in HeLa cells (Hirst et al.) | MR method and correlation based filtering                          | 26 proteins differentially localized (23% FDR)                        |
| Mouse hepatic steatosis (Krahmer et al.) | Correlation-based informal testing                                  | 910 proteins differentially localized (20% FDR)                        |
|                                    |                                                                     | 2542 phosphopeptides differentially localized (10% FDR)               |
| U-2 OS cell line (Geladaki et al.)  | SVM Classification                                                  | 58–65% of proteins are multi-localized (no primary localization assigned at estimated 5% FDR) |
| Five human cell lines and response to EGFR inhibition in HCC327 (Orre et al.) | Comparative SVM classification with heuristic filtering          | <10% protein multi-localized (inconsistent SVM classification between replicates) |
|                                    |                                                                     | 295 candidate differentially localized proteins (different SVM classification) |
|                                    |                                                                     | 13 differentially localized candidates after heuristic filtering     |
| Relocalization of golgin tethered transport vesicles in HEK 293T cells (Shin et al.) | Bayesian non-parametric two sample test and filtering by Mahalanobis distance to mitochondria | 139 proteins with profile shifted toward mitochondria                 |

heuristic filtering step(s) and the test itself is frequently not defined clearly.

As an alternative method to identify differential localization, Itzhak et al. proposed an informal testing approach denoted as the movement-reproducibility (MR) method.[20] This method uses the Mahalanobis distance between inter-condition profiles, as well as the correlation between these distances. Though this approach could be formalized, a null hypothesis is never stated. Given that the distances between proteins of the same organelle and between proteins of different organelles can vary considerably within an experiment it is also unclear, in general, how appropriate a distance approach is. To be more precise, what is meant by a big or interesting distance, as a formal test statistic, needs the context of which organelles the relocalization is between. Thus, it is hard to completely justify an approach agnostic to its spatial context.

To estimate the FDR of the MR method, the suggestion is to perform a “mock” experiment, comparing control versus control (requiring three additional replicates) so that the number of false positives can be estimated for a given cut-off value. This is based on the implicit but unstated assumption that organelle profiles remain similar across conditions. Assuming that most proteins
Box 2.
Krahmer et al. determining dual localizations

To determine dual localization Krahmer et al. take the following approach. For each protein

- The most likely primary organelle is determined using the SVM.
- The median profile of the primary organelle is combined with the median organelle profiles of the markers.
- The correlation value between the protein/peptide profile and these in silico profiles are determined.
- Proteins/peptides with correlation $>0.4$ are considered unreliable assignments.
- The alpha value is reported as a quantitative measure of second organelle localization.

Critique
The mixing of observed marker profiles can provide an expectation for multi-localization. However, as we have already demonstrated in Box 1, one has to be careful when making conclusion from mixed profiles. The first unstated part of the method is how the marker profiles are combined, presumably they were mixed in different possible proportions; however, they might have simply been averaged. Furthermore, it is unclear which method of correlation was applied. It appears that the authors incorrectly posit correlation $>0.4$ was unreliable rather than $<0.4$. Finally, the authors report an alpha value, but never state how this was computed nor what it actually represents. The distribution of this measure is never reported so that we can determine how the alpha value changes from proteins with known single or dual localizations.

Box 3.
Krahmer et al. determining changes in localization

To determine changes in localization Krahmer et al. take the following approach.

- Within each condition compute the correlation between the quantitative profiles and retain them only if their maximum correlation is greater than 0.5. Precisely, the maximum correlation for the $i$th protein is

$$\rho_{\text{max}} = \max(\rho(f_{ij1}, f_{ij2}), \rho(f_{ij2}, f_{ij3}), \rho(f_{ij1}, f_{ij3})),$$

where $f_{ijc}$ denotes the quantitative profiles for protein $i$ in replicate $j$ and in condition $c$. Then if $\rho_{\text{max}} > 0.5$ for protein $i$ then the profiles are retained for further analysis.
- Repeat this process for all the conditions.
- Take the top two most correlated profiles from each condition, and compute the average of the within conditions correlations and the average of the between condition correlations for the same replicates. Making this explicit, we write $f_{ijc}$ for protein $i$ in replicate $j$ and in condition $c$. Then compute the average of the within replicate correlations

$$\bar{\rho}_w = \frac{\rho(f_{ij1}^{(1)}, f_{ij1}^{(1)}) + \rho(f_{ij2}^{(2)}, f_{ij2}^{(2)})}{2},$$

and the average of the between condition correlations

$$\bar{\rho}_b = \frac{\rho(f_{ij1}^{(1)}, f_{ij2}^{(2)}) + \rho(f_{ij1}^{(1)}, f_{ij2}^{(2)})}{2}.$$

- Then, for each protein, compute the difference in these quantities.

$$\delta \rho = \bar{\rho}_b - \bar{\rho}_w$$

- The list is then ranked from largest to smallest.
- The approach is repeated for Spearman and Pearson correlations and the results combined.
- An FDR threshold of 0.2 is set.

Critique
The initial filtering steps are arbitrary and there is not a justification of why they are performed or what was the motivation for the threshold. Both Spearman and Pearson correlations are used which have different assumptions, so it is unclear what is the meaning of combined correlation results. If Pearson correlations are used then there is an implicit assumption of bivariate normality that is never checked. Furthermore, correlation cannot be averaged because they are not additive. Rather they obey the law of cosines and should be treated as such. It is also unclear how the different correlations were combine and what an appropriate null hypothesis is in this scenario. This makes it hard to understand how a $p$-value is computed and whether there was correction for multiple testing. Without clearly stating the assumptions it is challenging to assess how appropriate these methods are and how they apply to other datasets or even if the methodology presented is valid. Unfortunately, because of the lack of clarity and the use of propriety software it is impossible to reproduce the analysis.
(at least 90%) do not relocalize, these additional replicates are unnecessary as proper calibration of FDR can be achieved by computing an empirical null using large-scale data analysis methods such as permuting sample labels. More importantly however, the assumption of consistent organelle profiles was not tested in any studies where MR method was applied and, as shown here (Figure 2), organelle profiles can change significantly across conditions, and thus, this approach is not generalizable.

6. Modeling to Address the Extended Questions

The success of the MR method should not be undersold, having been used successfully to find differential localizations in several studies. However, the question is whether some protein relocalization has been overlooked and if the same results could be obtained with fewer experiments.

Formal testing procedures do have their place when performed appropriately. Consider the study of Shin et al. where a Bayesian non-parametric two sample test was used to determine whether protein profiles were perturbed between two conditions. Among those proteins with perturbed profiles those that demonstrated relocalization toward the Mitochondria were of particular interest within this study. However, the mitochondrial profile was condition-dependent; hence, a marker agnostic approach was not appropriate and instead the squared Mahalanobis distance to the profile of mitochondrial markers in each condition was used to identify the movements toward the mitochondria.

Multi-localization, differential localization and the role of PTMs in localization are challenging questions for profiling spatial proteomics. To develop a methodology to answer these questions they first need precise definitions. From here unified methods can be developed with clear assumption so that the extent of their applicability can be assessed. We believe that to address the extended questions of profiling spatial proteomics we must model the data. This brings with it new challenges. To date, classification approaches have usually made few assumptions and for data processing, for example, peptide spectral match (PSM) aggregation and profile normalization have received little attention. In particular, a variety of methods are used to normalize profiles (max signal, sum normalization, relative to heavy spike-in). For support vector machines and neural networks, so long as the proteins’ profiles still cluster, the precise normalization approach is unlikely to have a significant impact. However, when modeling the data, we often have to make assumptions about the underlying distributions and the normalization method may well invalidate assumptions about the distributions of the profiles upon which our models are based. As such, it will be necessary to devote more attention to the proper processing of spatial proteomics data. Furthermore, modeling replicates allows a direct assessment of classification confidence and reliability, in contrast to the common approach of concatenating replicates before SVM classification, which precludes such inferences.

7. Quantifying Uncertainty with Bayesian Modeling

Further benefit can be drawn by moving to a Bayesian inference framework, which enables the quantification of uncertainty. For the classification task, we have already presented solutions to this problem through the T-Augmented Gaussian Mixture model (TAGM; and its non-parametric counterpart) which attempts to directly model the data. The modeling framework allows treatment of markers as strong priors and learning of the truly representative distribution of the organelle from the data. This allows us to distinguish proteins that have confident localizations from those that are uncertain between two or more localizations. Modeling approaches are not without their limitations and often resulting in increased computational burden. For example, TAGM is currently limited to modeling the data at the protein level and discards potentially valuable information at the peptide and PSM level.

More elaborate models can enable simultaneous assignment of proteins to organelles and novelty detection by allowing proteins to either be assigned to an annotated organelle or one that has not been manually annotated. For example, Crook et al. propose a semi-supervised Bayesian approach and uncover a novel group of proteins trafficking from the ER to the early Golgi apparatus. Additionally, differential localization could also be elucidated using joint models across conditions, with uncertainty quantification to assist in ranking candidates for future experimental investigation. Similarly, the relationship between post-translational modification and differential localization could be examined by comparing profiles for

---

Box 4. Krahmer et al. determining changes in the phosphoproteome

To determine phosphoproteome changes, (Krahmer et al.) take the following approach.

1. The subcellular localization of phosphopeptides is identified using PCP.
2. A filter to retain those sites where phosphorylation changes across condition is applied.
3. For the remaining phosphosites filter to those within proteins whose localization changes according to determining changes in localization.

Critique

All of the assumptions of determining changes in localization are applied without evaluation that these assumptions hold true. Furthermore, there is no apparent requirement for the intersecting change in phosphorylation, change in protein localization and phosphopeptide localization to be concordant.

---

2020, 20, 1900392 1900392 (8 of 10) © 2020 The Authors. Proteomics published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
the modified and unmodified protein forms. Through careful consideration of the question in hand, well-designed models have the capacity to uncover many valuable insights which may be missed by ad hoc approaches.

8. Appropriate Models Depend on Experimental Design

The biological system of interest, experimental design and statistical or mathematical model in question are not independent or modular entities. They are better thought of as parts of a whole and can vitally inform each other. If quantification of protein proportions in different compartments is desired, then the model, design and system should reflect this, perhaps at the cost of other important quantities such as the resolution of the subcellular niches interrogated. Furthermore, subcellular fractionation may not need to be exactly replicable if the desire is to achieve the maximum possible separation of organelles, but the applied model should be aware of this choice. Finally, if the translocation of proteins of interest is between two organelles with similar biochemical properties, then the experiment can be designed to ensure maximal separation of these subcellular niches and prior information about these properties can be embedded into a statistical model. Design, modeling, and experiment are an iterative process that allows each to build upon the former to gain the most information possible.

9. Discussion and Outlook

The modern inception of profiling protein abundance across organelle-enriched fractions has been hugely successful in furthering our understanding of protein function.

As we re-purpose this method to examine multi-localization, differential localization, and the interplay of post-translational modification and localization, it is an opportune moment to consider how best to marry novel experimental and computational strategies. One framing of all these questions could be achieved by quantifying the proportion of the protein in each localization assayed. From these proportions, one could then assign the primary localization of a protein, whether it is multi-localized within a single condition, and whether the proportions are affected by condition or post-translational modification. However, our experimental constraints preclude straightforward determination of localization proportions. Thus, in the near term, we expect that the translocation of proteins of interest is between two organelles with similar biochemical properties, then the experiment can be designed to ensure maximal separation of these subcellular niches and prior information about these properties can be embedded into a statistical model. Design, modeling, and experiment are an iterative process that allows each to build upon the former to gain the most information possible.

Acknowledgements

O.M.C. is a Wellcome Trust Mathematical Genomics and Medicine student and acknowledges generous funding from the School of Clinical Medicine, Cambridge. T.S. and M.E. are supported by the Medical Research Council, Grant/Award number: 5TR00; Wellcome Trust, Grant/Award numbers: 110170/Z/15/Z, 110071/Z/15/Z.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

O.M.C. and T.S. contributed equally to this work. Author contributions are described according to CRediT standards. O.M.C. and T.S. contributed to the conceptualization, writing the original draft, writing the review and editing, visualization, and formal analysis. M.E. contributed to the conceptualization, writing the original draft, writing the review and editing. K.S.L. contributed to the conceptualization, writing the original draft, supervision, project administration, and funding acquisition.

Keywords

organelle, protein localization

Received: March 20, 2020
Revised: May 18, 2020
Published online: July 12, 2020

[1] T. J. Gibson, Trends Biochem. Sci. 2009, 34, 471.
[2] C. K. Barlowe, E. A. Miller, Genetics 2013, 193, 383.
[3] R. Puertollano, S. M. Ferguson, J. Brugarolas, A. Ballabio, EMBO J. 2018, 37.
[4] A. Y. Lee, W. Chen, S. Stippec, J. Self, F. Yang, X. Ding, S. Chen, Y. C. Juang, M. H. Cobb, Proc. Natl. Acad. Sci. USA 2012, 109, 16841.
[5] E. A. Balta, M. T. Wittmann, M. Jung, E. Sock, B. M. Haeberle, B. Heim, F. von Zweydlor, J. Heppt, J. von Wittgenstein, C. J. Gloeckner, D. C. Lie, Front. Mol. Neurosci. 2018, 11, 211.
[6] F. Christian, E. L. Smith, R. J. Carmody, Cells 2016, 5, 12.
[7] H. Kiu, S. E. Nicholson, Groundwater 2012, 50, 58.
[8] M. Mani, C. Chen, V. Ambree, H. Liu, T. Mathur, G. Zwieck, S. Zabad, B. Patel, J. Thakkar, C. J. Jeffery, Nucleic Acids Res. 2015, 43, D277.
[9] V. Hung, S. S. Lam, N. D. Udesi, T. Sivinika, C. Guzman, V. K. Mootha, S. A. Carr, A. Y. Ting, eLife 2017, 6, e24463.
[10] J. Y. Youn, W. H. Dunham, S. J. Hong, J. D. R. Knight, M. Bashkurov, G. I. Chen, H. Bagci, B. Rathod, G. MacLeod, S. W. M. Eng, S. Angers, Q. Morris, M. Fabian, J. F. Côté, A. C. Gingras, Mol. Cell 2018, 69, 517.
[11] P. J. Thul, L. Akesson, M. Wiking, D. Mahdessian, A. Geladaki, H. Ait Blai, T. Alm, A. Asplund, L. Björk, L. M. Breckels, A. Bäckström, F.
