UC Santa Barbara

Title

Initial recommendations for performing, benchmarking and reporting single-cell proteomics experiments.

Permalink

Journal

Authors

Gatto, Laurent
Aebersold, Ruedi
Cox, Juergen
et al.

Publication Date

DOI

Peer reviewed

Initial recommendations for performing, benchmarking and reporting single-cell proteomics experiments

Laurent Gatto[1], **Ruedi Aebersold**[2], **Juergen Cox**[3], **Vadim Demichev**[4], **Jason Derks**[5], **Edward Emmott**[6], **Alexander M. Franks**[7], **Alexander R. Ivanov**[8], **Ryan T. Kelly**[9], **Luke Khoury**[5], **Andrew Leduc**[5], **Michael J. MacCoss**[10], **Peter Nemes**[11], **David H. Perlman**[12], **Aleksandra A. Petelski**[5,13], **Christopher M. Rose**[14], **Erwin M. Schoof**[15], **Jennifer Van Eyk**[16], **Christophe Vanderaa**[1], **John R. Yates III**[17], **Nikolai Slavov**[5,13]

[1]Computational Biology and Bioinformatics Unit, de Duve Institute, Université Catholique de Louvain, Brussels, Belgium.

[2]Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland.

[3]Max Planck Institute of Biochemistry, Martinsried, Germany.

[4]Charité – Universitätsmedizin, Berlin, Germany

[5]Departments of Bioengineering, Biology, Chemistry and Chemical Biology, Single-Cell Proteomics Center and Barnett Institute, Northeastern University, Boston, MA, USA.

[6]Centre for Proteome Research, Department of Biochemistry and Systems Biology, University of Liverpool, Liverpool, UK.

[7]Department of Statistics and Applied Probability, University of California Santa Barbara, Santa Barbara, CA, USA.

[8]Department of Chemistry and Chemical Biology, Barnett Institute of Chemical and Biological Analysis, Northeastern University, Boston, MA, USA.

[9]Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT, USA.

[10]University of Washington, Seattle, WA, USA.

[11]Department of Chemistry and Biochemistry, University of Maryland, College Park, MD, USA.

[12]Merck Exploratory Science Center, Merck Sharp & Dohme Corp., Cambridge, MA, USA.

[13]Parallel Squared Technology Institute, Watertown, MA, USA.

[14]Department of Microchemistry, Proteomics and Lipidomics, Genentech Inc., South San Francisco, CA, USA.

[15]Department of Biotechnology and Biomedicine, Technical University of Denmark, Lyngby, Denmark.

[16]Cedars Sinai Medical Center, Los Angeles, CA, USA.

[17]Departments of Molecular Medicine and Neurobiology, the Scripps Research Institute, La Jolla, CA, USA.

Abstract

Analyzing proteins from single cells by tandem mass spectrometry (MS) has recently become technically feasible. While such analysis has the potential to accurately quantify thousands of proteins across thousands of single cells, the accuracy and reproducibility of the results may be undermined by numerous factors affecting experimental design, sample preparation, data acquisition and data analysis. We expect that broadly accepted community guidelines and standardized metrics will enhance rigor, data quality and alignment between laboratories. Here we propose best practices, quality controls and data-reporting recommendations to assist in the broad adoption of reliable quantitative workflows for single-cell proteomics. Resources and discussion forums are available at https://single-cell.net/guidelines.

New approaches and technologies for experimental design, sample preparation, data acquisition and data analysis have enabled the measurement of several thousand proteins in small subpopulations of cells and even in single mammalian cells[1-11]. These developments open exciting new opportunities for biomedical research[12], as illustrated in Fig. 1. In some systems, subpopulations of molecularly and functionally similar cells can be isolated and analyzed in bulk, which allows for deeper proteome coverage. Other systems, however, do not allow for such isolation due to continuous (rather than discrete) phenotypic states or due to unknown cell states or markers[13,14]. Such systems require single-cell analysis; it is particularly needed for discovering new cell types[15] and for investigating continuous gradients of cell states, which has already benefited from single-cell MS proteomics[6,16-18]. Furthermore, when a large number of single cells are analyzed, the joint distributions of protein abundances enable new types of data-driven analysis (Fig. 1) that may support inferences with minimal assumptions[12,19].

Despite these promising prospects, single-cell MS is sensitive to experimental and computational artifacts that may lead to failures, misinterpretation or substantial biases that can compromise data quality and reproducibility, especially as the methodologies become widely deployed. To minimize biases and to maximize quantitative accuracy and reproducibility of single-cell proteomics, we propose initial guidelines for optimization, validation and reporting of single-cell proteomic workflows and results.

The tandem MS methods for single-cell bottom-up proteomics span a range of techniques[13], including multiplexed and label-free methods, both of which can be performed by data-dependent acquisition[1,20] and data-independent acquisition (DIA)[7,10]. The initial recommendations presented here are relevant to all these methods, and we will note

any exceptions. Our initial recommendations for experimental design, data evaluation and interpretation, and reporting are intended to stimulate further community-wide discussions that mature into robust, widely adopted practices. Imaging and top–down MS methods are also advancing and reaching single-cell resolution[21,22], although they differ substantially from MS-based bottom–up proteomic methods and are outside the scope of these recommendations.

Experimental design

Best practices for single-cell MS proteomics can effectively build on established practices for bulk analysis[23,24]. Common best practices include staggering biological treatments, sample processing and analytical batches so that sources of biological and technical variation can be distinguished and accounted for during result interpretation. Similarly, randomization of biological and technical replicates and batches of reagents during sample processing (for example, mass tags for barcoding) are recommended to minimize potential artifacts and to facilitate their diagnoses. We also recommend including appropriately diluted bulk samples as technical quality controls. The following specific issues are relevant for the design of single-cell proteomic measurements.

Single-cell isolation

A primary goal of sample preparation should be to preserve the biological state of cells with minimal perturbations. This can be challenging for tissues and for adherent cell cultures as cell isolation may require vigorous dissociation or detachment procedures. Extracting single cells from tissue samples in some cases may require enzymatic digestion of proteins, which may cleave the extracellular domains of surface proteins. Potential artifacts arising from these manipulations should be considered and may be minimized by using more gentle dissociation procedures, such as chelation of cations stabilizing extracellular protein interactions. Dissociated single cells should be thoroughly washed to minimize contamination of MS samples with reagents used for tissue dissociation.

While proteins are generally more stable than mRNA[25], most good practices used for isolating cells for single-cell RNA sequencing (scRNA-seq) and flow cytometry[26], such as quick sample processing at low temperature (4 °C), are appropriate for proteomics as well. Timing and other parameters of the cell-isolation procedure may be impactful and therefore should be recorded so that technical effects associated with sample isolation can be accounted for in downstream analysis. We recommend collecting as much phenotypic information as possible from cells prepared and isolated in the same manner, including cellular images and any relevant functional assays that can be performed. Such phenotypic data allow for orthogonal measures of cell state to be combined with MS data and thus to strengthen biological interpretations. While isolating single cells of interest, we recommend also collecting bulk samples from the same cell population (if possible). Having such bulk samples will allow for the inclusion of positive controls and for benchmarking; these two topics will be discussed more in sections below.

Many studies have used flow cytometry for isolating cells from a single-cell suspension[9,10,16,27]. Flow cytometry can perform very well, as indicated by the successful results of such studies. Yet, in the absence of high-performing sorters and expert operators, it may be one of the least robust steps of the workflow[5]. Thus, verifying the ability to robustly isolate individual cells by flow cytometry may save much time from troubleshooting downstream analysis steps. Studies have also isolated single cells by cellenONE[28,29], and it supports gentler and more robust isolation than flow cytometry, which is particularly helpful with primary cells[18].

When analyzing the proteomes of single cells from tissues, the spatial context should be characterized as best as possible, including both the location of each cell in the tissue and the extracellular matrix around it. Although a great area of interest, such single-cell MS proteomic analyses are in their infancy. Feasible approaches for spatial analysis include tissue sectioning by cryotome and laser-capture microdissection (LCM), which can be used to extract individual cells[30]. LCM has been used for spatially resolved extraction and subsequent MS analysis of tissue regions[31]. The application of plexDIA and isotopologous carriers[7,32] are showing promise to extend this analysis to single cells extracted by LCM[33]. We recommend avoiding the use of protocols that require cleanup from detergents for tissue disruption and instead prefer methods using only MS-compatible reagents.

Reducing contamination

Minimizing sources of contaminating ion species that disproportionately affect the analysis of small samples is critical for single-cell proteomic measurements. Contaminating ions can result from many sources, including reagents used during sample preparation, impure solvents, extractables and leachables from sample contact surfaces, and especially carryover peptides from previous single-cell or bulk runs that may persist within liquid handling, instrument components, capillaries and stationary phases, such as needle-washing solutions and column-retained analytes in liquid chromatography (LC) and reservoirs in capillary electrophoresis. Typically, only about 1% of peptides persist on C18 column resin following a run, and they may appear in subsequent runs as a carryover 'ghost' signal[34]. Fortunately, these carryover peptides generally make a quantitatively insignificant contribution to consecutive samples of comparable amounts.

However, when bulk samples are interspersed with single-cell runs, carryover peptides from these bulk samples may substantially contaminate or even dwarf the peptide content derived from the single cells. Thus, contaminants from bulk sample runs are often incompatible with quantitative single-cell analysis on the same LC–MS system. Before analyzing single-cell samples, analytical columns must be evaluated rigorously and deemed free of carryover, as previously described[5,27]. Other non-peptidic contaminants, such as leached plasticizers, phthalates and ions derived from airborne contaminants, often appear as singly charged ions and can be specifically suppressed by ion-mobility approaches[7,27,35] or, in the case of airborne contaminants, by simple air-filtration devices, for example, an active background ion reduction device (ABIRD)[5].

Because the ratio of sample-preparation volume to protein content is significantly increased, the amount of reagents to protein content is also significantly increased when preparing single cells individually. Thus, reducing sample-preparation volumes mitigates the effect of contaminant ions originating from reagents such as trypsin or mass tags[2,36]. Indeed, reducing sample-preparation volumes to 2–20 nl proportionally reduces reagent amounts per single cell compared to multiwell-based methods, which in turn reduces the ion current from singly charged contaminant ions[6].

Sample preparation

Ideally, sample preparation should consist of minimal steps designed to minimize sample handling, associated losses and the introduction of contaminants. For bottom–up proteomic analyses, workflows must include steps of cell lysis–protein extraction and proteolytic digestion. Given the picogram levels of protein present in a single cell, it is crucial to minimize contaminants and maximize sample recovery for downstream analysis. Fortunately, the composition and geometries of single cells isolated from patients and animals lend themselves to disruption under relatively gentle conditions, such as a freeze–heat cycle[5,37,38] or nonionic surfactants[39,40]. Such clean lysis methods are preferable over MS-incompatible chemical treatments (for example, sodium dodecyl sulfate or urea) that require loss-prone cleanup before MS analysis[41]. It can be beneficial to miniaturize processing volumes to the nanoliter scale to minimize exposure to potentially adsorptive surfaces[2,6], although such approaches may have limited accessibility. By contrast, sample preparations using low-microliter volumes offer broadly accessible options[16,37,42] and are described in detailed protocols[5,38].

Regardless of the selected preparation workflow, it is recommended that cells be prepared in batches that are as large as possible to minimize technical variability in sample handling. To this end, several liquid-handling tools have been successfully coupled with single-cell proteomic workflows to increase throughput and reduce technical variability. In particular, the Formulatrix MANTIS and the Opentrons have been adapted for 384-well-plate-based sample preparation[5,37,42]. The cellenONE system has also been employed for several automated protocols using microfabricated multiwell chips[2,28,43] or using droplets on glass slides[29]. We expect this landscape to continuously evolve toward increased consistency and throughput of sample handling.

Maximizing sample delivery to mass analyzers

For sample-limited analyses, it is especially important to maximize ionization efficiency (the fraction of gas-phase ions created from solution-phase molecules) and the transmission of those ions to the mass analyzer. Lower volumetric flow rates produce smaller, more readily desolvated charged droplets at the electrospray source, leading to increased ionization efficiency[44,45]. As such, reducing the flow rate of separations from hundreds to tens of nanoliters per minute can increase measurement sensitivity, but currently these gains must be achieved with custom-packed narrow-bore columns and may compromise robustness and measurement throughput[20]. Maximizing separation efficiency is also important, as narrower

peaks increase the concentration of eluting peptides and simplify the mixture entering the mass spectrometer at a given time.

A number of commercial nanoLC systems and columns provide a reasonable combination of sensitivity and efficiency for single-cell proteomics, and these are recommended for most practitioners. Alternative high-resolution separation techniques employing orthogonal separation mechanisms, for example, capillary electrophoresis and ion mobility, as well as multidimensional techniques may potentially be employed as front-end approaches in MS-based single-cell proteomics[11,46]. Increasing ion transmission in the mass spectrometer is generally the purview of instrument developers and companies, and future gains in this area are expected to further benefit single-cell proteomics.

Lastly, when injecting samples for analysis by LC–MS, because of the low protein amount, it is often desirable to inject the entire sample. If the samples are resuspended in too small of a volume, the autosampler may miss portions of the sample or may inject air into the lines, which adversely affects chromatography. Thus, we recommended striking the correct balance of suspension volume that prevents air injections and maximizes sample delivery. This balance depends partially on the autosamplers, sample vials and their shape and size. One implementation shown to perform robustly includes injecting one-microliter samples from 384-well plates[5,6,18].

Controls

Experimental designs should provide an estimate of quantitative accuracy, precision and background contamination. Precise measurements may arise from reproducing systematic biases, such as integration of the same background contaminants. Measurement precision can therefore be assessed by repeat measurements. By contrast, benchmarking measurement accuracy requires positive controls, that is, proteins with known abundances. One approach to benchmarking is incorporating into the experimental design samples with known quantitative values to assess quantitative accuracy. These controls may be derived from independent measurements based on fluorescent proteins or well-validated affinity reagents. Other positive controls include spike-in peptides[18], proteins or even proteomes in predefined ratios as performed for LFQbench experiments[47].

When cells from clusters consisting of different cell types can be isolated, the relative protein levels of the isolated cells may be quantified with validated bulk assays and used to benchmark in silico averaged single-cell estimates, an approach used by multiple studies[5,9,16,18,29]. A positive control for sample preparation may include bulk cell lysates diluted to the single-cell level. Estimating protein amounts corresponding to single cells is challenging, and thus we recommend starting with cell lysate from precisely known cell numbers (for example, estimated by counting cells with a hemocytometer) and performing serial dilution to the single-cell level[5]. Negative control samples, which do not contain single cells, should be processed identically to the single-cell samples. Such negative controls are useful for estimating cross-labeling, background noise and carryover contaminants.

When matching between runs (MBR) is used to propagate sequence identification, MBR controls should be included. Empty samples contain few ions, if any, that may be associated with incorrect sequences. Thus, using empty samples may lead to underestimating MBR false discoveries. MBR may be evaluated more rigorously by matching samples containing either mixed-species proteomes or samples containing single-species proteomes and then estimating the number of incorrectly propagated proteins. Such MBR controls (samples of mixed yeast and bacterial proteomes or only yeast proteomes) have been used to benchmark sequence propagation within a run[7], and similar standards should be used for benchmarking MBR. While MBR is best evaluated in each study with samples designed to reflect the analyzed proteomes, the field may benefit from preparing community reference samples that were analyzed in multiple laboratories and used for benchmarking MBR algorithms.

Batch effects

Systematic differences between groups of samples (biological) and analyses (technical) may lead to data biases, which may be mistaken for cell heterogeneity, and thus complicate result interpretation or sacrifice scientific rigor. To estimate and correct batch effects, treatments and analytical batches must be randomized whenever possible[48]. We recommend that treatment and batches are randomized so that batch effects can be corrected (estimate and remove batch effects from data) or modeled (for example, include batch effect as a covariate in models). When randomization is not performed, biological and technical factors may be fundamentally inseparable. For experiments in which randomization was not performed, downstream statistical analyses should include the batch information as covariates. These considerations are similar to those for bulk experiments, which have been previously described[49]. Furthermore, we recommend that all batches include the same reference sample, which can be derived from a bulk sample diluted close to a single-cell level.

Statistical power

Studies should be designed with sufficient statistical power, which depends on effect sizes, on measurement accuracy and precision, and on the number of single cells analyzed per condition. Simple experiments with large effect sizes, such as analyzing different cell lines, can achieve adequate statistical power with a few dozen single cells. Such experiments were common as proof-of-principle studies demonstrating analytical workflows. By contrast, experimental designs including primary cells, smaller effect sizes (for example, protein variability within a cell type[6]) or multiple treatment groups or patient cohorts require a much larger number of single cells and patients to achieve adequate statistical power[50-52].

Methods for MS data acquisition

Existing methods can be grouped into label free, which analyze one cell per sample, and multiplexed, which analyze multiple cells per sample. Label-free methods benefit from simpler sample preparation, while multiplexed methods benefit from analyzing more cells per unit time[5]. When multiplexing is performed by isobaric mass tags, quantification is adversely affected by the co-isolation and co-fragmentation of precursors. This co-isolation

can be mitigated by targeting the apexes of elution peaks and using narrow isolation windows[16,18]. The co-isolation artifacts on quantification can be overcome by performing quantification on peptide-specific and sample-specific ions, as in the case of plexDIA, which multiplexes cells with non-isobaric mass tags[7,53]. Isobaric mass tags have been used in combination with a carrier sample, which reduces sample losses and facilitates peptide sequence identification[54]. This approach has raised concerns as high carrier amounts may allow confident peptide identification without sampling sufficient peptide copies from the single cells to achieve precise quantification[55,56]. To address these concerns, multiple groups have converged on guidelines for balancing the precision and throughput of single-cell analysis using isobaric carriers[55,56].

Cross-validation using different MS methods

We recommend, when possible, cross-validating protein measurements with different methods that share minimal biases. Often, such cross-validation may be performed using the same MS instruments, and the results may be directly reported and compared in the same paper. Such cross-validation studies are particularly useful for supporting new and surprising biological results.

As an example, Leduc et al.[6] observed a gradient of phenotypic states and protein covariation within a cluster of melanoma calls not primed for drug resistance. The authors cross-validated these observations by analyzing biological replicates of the melanoma cells both by isobaric multiplexing with pSCoPE[18] and by non-isobaric multiplexing with plexDIA[7]. The results from the two methods were directly compared and reported in parallel so that the degree of biological and technical reproducibility can be evaluated[6]. Cross-validation analysis can also benefit from using different sample-preparation methods or enzymes for protein digestion. In such cross-validation analyses, quantitative trends supported by multiple methods and biological replicates are more likely to reflect biological signals rather than method-specific artifacts.

Method selection and optimization

The MS methods and their parameters should be selected depending on the priorities of the analysis. Maximizing the number of cells analyzed is best achieved with short separation times and multiplexed methods[57]. Maximizing the proteome depth is best achieved with longer separation methods, while maximizing the number of copies sampled per protein is best achieved with MS[1]-based methods and longer ion-accumulation times[7,36]. Multiple objectives, such as increased consistency, dynamic range and coverage, may best be simultaneously optimized with intelligent data-acquisition strategies[18,36,57,58]. The size of the isobaric carrier used can also help emphasize project priorities, such as depth of proteome coverage versus copy number sampled per peptide[55,56]. Choosing optimal method parameters can be time consuming, and software for systematic, data-driven optimization can speed up such optimizations[59].

Data evaluation and interpretation

Defining and evaluating reproducibility

Data reproducibility and evaluation can be performed at several levels of increasing difficulty, namely, repeating, reproducing and replicating[60]. Repeating a computational experiment or an analysis simply consists of using the exact same data, code, software and environment (typically the same computer), assuming that these are still available. Reproducing an experiment or analysis is an attempt by a different person that will mimic the original setup by downloading data and code, without necessarily having access to the same software environment. Replication represents a further challenge in which the results are to be obtained using new code, implementation and/or software; it is only possible with extensive and detailed description of the performed analyses. This description must include the versions of all software and databases used as well as all search parameters, ideally saved as structured documents, for example, xml.

Batch effects and cellular uniqueness

Two factors should be considered when reproducing single-cell protein measurements. First, no two cells are identical. Thus, we may reasonably hope to reproduce clusters of cells and trends (such as protein-abundance differences between cell types or cell states) but not the exact molecular levels for each analyzed cell. Second, batch effects may increase the apparent level of reproducibility (when biases are shared between replicates, such as peptide adhesion losses or co-isolation) or decrease it (when biases differ between replicates, such as protein-digestion biases). Thus, assessments and reports of reproducibility need to be specific about precisely what is being reproduced and how this may be impacted by batch effects originating from all steps, from cell isolation to data processing.

Evaluating quantitative accuracy

Quantitative accuracy is a measure of how closely the measurements correspond to known true values, as in the case of proteomes mixed in experimenter-determined ratios (Fig. 2a). When the true abundances are not known, evaluating accuracy is not possible and is sometimes confused with repeatability or precision. Yet, these quantities can be quite different as illustrated in Fig. 2a. Similarly, high correlation between replicates may be interpreted as evidence that the measurements are quantitatively accurate. This interpretation is wrong: many systematic errors may lead to erroneous measurements that are nonetheless very reproducible. Thus, reproducibility alone is insufficient to evaluate data quality.

Because single-cell proteomics pushes the limits of sensitivity for MS-based measurements, the quality of measurements depends on the number of ions measured from each single-cell population[55,56]. For example, if too few ions are sampled, the stochasticity of sampling results in counting noise, that is, low-precision estimates and technical variation in estimated protein abundances, which should be clearly distinguished from biological variability[36]. Mixing ratios of 1:1 can be used to evaluate ion sampling and precision but not accuracy because this ratio is not sensitive to systematic biases, such as co-isolation and interference. Accuracy can be evaluated relative to ground truth ratios, as created by mixing the proteomes of different species in known ratios[7,47].

As described in the cross-validation section, MS methods that share minimal biases (for example, quantifying precursors at the MS[1] level versus quantifying reporter ions at the MS[2] level) can also help reduce biases. This approach can include relative quantification from established bulk methods if the analyzed cell types can be isolated as bulk samples, as demonstrated with cell lines[7,9,16]. On a smaller scale, accuracy may be estimated for a limited number of proteins by spiking corresponding peptides at known ratios[18] or by using measurements that are as independent as possible; such independent measurements include fluorescent proteins, the abundance of which is measured fluorometrically[1], or immunoassays with high specificity, such as proximity ligation assays that enhance specificity by using multiple affinity reagents per protein[61].

Quantitative precision and accuracy are different metrics, the importance of which is highly dependent on the analysis. For example, cell clustering benefits from high-precision measurements and may tolerate low quantitative accuracy. By contrast, protein covariation analysis[6,19] and biophysical modeling[12] are more dependent on quantitative accuracy. Thus, benchmarks should clearly distinguish between accuracy and precision and focus on the metric that is more relevant to the biological goals of the analysis.

Comparisons between absolute protein intensities conflate variance due to protein-abundance variation across the compared samples (conditions) and across different proteins and may result in misleading impressions[62]. For example, the high correlation between the proteomes of T cells and monocytes in Fig. 2b may be interpreted as indicating that the two proteomes are very similar. Yet, many proteins differ in abundance reproducibly between T cells and monocytes (Fig. 2c). Thus, correlations between estimates of absolute protein abundance should not be used as benchmarks for relative protein quantification.

Evaluating quantitative consistency

Outside of carefully designed benchmarking experiments, the true protein abundances are unknown, and thus the accuracy of quantification cannot be directly benchmarked. However, it is often possible to evaluate the reliability of MS measurements based on comparing the quantitative agreement between (1) different peptide fragments from the same peptide (Fig. 2d) or (2) different peptides originating from the same protein.

For example, the internal consistency of relative quantification for a peptide may be assessed by comparing the relative quantification based on its precursors and fragments, as shown for single-cell plexDIA data in Fig. 2d. The degree of (dis)agreement may be quantified by the coefficient of variation (CV) for these estimates. Similarly, the CV estimated from the relative levels of different peptides originating from the same protein may provide a useful measure of reliability. This analysis is limited by the existence of proteoforms[63,64] but nonetheless may provide useful estimates of data quality. Note that this CV is very different from the CV computed using absolute peptide intensities or the CV computed between replicates. In the latter case, when comparing CVs across different analytical or experimental conditions, it is imperative to account for varying dataset sizes; that is, a rigorous comparison between experimental methods would rely on peptides and proteins identified and quantified across all samples, rather than also including peptides and proteins identified uniquely in individual experiments[59].

Accounting for biological and technical covariates

Single cells differ in size and thus protein content. Consequently, cell size is a major confounder for the differences in protein intensities between cells[6]. The basic normalization strategy here consists of subtracting from log-transformed protein quantities the respective medians across the proteins quantified[16]. However, this normalization can be undermined if the subset of quantified proteins varies substantially across single cells. Such variation may stem from differences in total protein amounts between cells or experimental variability, which may lead to differences in the numbers of missing values and proteins accurately quantified. In case of such variation, normalization should be based on a common subset of proteins or against a common reference, as described by Franks et al.[62]. Thus, processing of single-cell MS proteomic data is likely to be improved in the future with the development of more advanced normalization strategies, which may build upon those developed for scRNA-seq experiments[65] to mitigate similar challenges. To compensate for imperfect normalization, we suggest including a variable representative of the cell size, such as total protein content estimated from LC-MS data or forward scatter from flow cytometry, as a covariate in downstream analyses.

Managing missing data

One of the common challenges in analyzing single-cell data is handling the presence of missing values[48,66]. These tend to be more prevalent in single-cell proteomics than in typical bulk experiments as some proteins may be below the limit of detection (especially in smaller cells) or may not be sent for MS^2 analysis in every single cell. The latter problems can be fundamentally resolved by using DIA or prioritized data acquisition, and such methods substantially increase data completeness[7,18,32].

The missing data are a source of uncertainty that should be propagated through the analysis and ultimately reflected in the final conclusions. Many analyses may be conducted using only the observed data (without using imputed values), which assumes that the observed data are representative of the missing data. Yet, it is often desirable to impute missing values as this enables additional downstream analysis and may allow for explicit modeling of the missingness mechanisms. Indeed, imputation should take into account the nature of missing data (for example, missing at random or not at random[67]) in determining appropriate imputation methods. The type of missingness is determined by the mechanism leading to missing values, which depends on the algorithm for peptide sampling during mass spectrometric analysis. Shotgun methods using the topN heuristic introduce missing values that are more likely to occur at random, as they originate from the stochastic selection of precursors for MS^2 scans. By contrast, DIA and prioritized methods send precursors for MS^2 scans deterministically, and most missing values likely correspond to peptides below the limit of detection rather than those missing at random.

Comprehensive imputation methods for single-cell proteomics are yet to be developed and benchmarked, but recommendations developed for bulk proteomic methods may serve as useful guides[67-69]. While some recently developed methods for scRNA data may be adapted to proteomics, ultimately, the field needs methods that are specifically tailored to the mechanisms leading to missing peptides and proteins. Multiple imputation can be

used to quantify uncertainty in the results for a given missing data method. Although computationally demanding, it is also prudent to impute using different missing data models to further characterize the sensitivity of the results to unverifiable assumptions about the missingness mechanism. A simple example of this strategy would be to perform downstream data analysis, such as principal-component analysis (PCA), on the imputed data and compare the results to the analysis performed on the unimputed data[16,18]. Results that are insensitive to different types of imputation models are more reliable, while those that are contingent on the validity of a particular assumption about missingness should be viewed with more skepticism.

Dimensionality reduction

High-dimensional single-cell data are often projected onto low-dimensional manifolds to aid visualization and to denoise data. While such projections can be useful, the reduced data representations are incomplete approximations of the full data and often lose aspects of the data, as illustrated in Fig. 2e by projecting a three-dimensional dataset into different two-dimensional projections. As such, different low-dimensional projections may selectively highlight certain aspects of the data while obscuring others (Fig. 2e). At worst, they may severely distort the original data[70]. Thus, we recommend using dimensionality reduction as an initial data-analysis step that requires further scrutiny. Conclusions derived from reduced data representations, such as clustering of cells, should be validated against the high-dimensional data. The validation can be as simple as computing and comparing distances between cells in a higher-dimensional space, as demonstrated with macrophage clusters defined based on single-cell RNA and protein data[71].

While dimensionality-reduction representations can be useful for visualization, clustering of cell types in low-dimensional manifolds is inadequate for benchmarking quantification. Such representations indicate whether the cells cluster in a low-dimensional space, but they indicate little about the factors, whether biological or technical, that could be driving the clustering. More fundamentally, low-dimensional data reductions often account for only a fraction of the total variance in the data and thus may exclude relevant sources of biological variability (Fig. 2e). Some methods, such as PCA, better preserve global distances and are thus more amenable to interpretation, as opposed to their non-linear counterparts, such as t-distributed stochastic neighbor embedding (t-SNE)[72] or uniform manifold approximation and projection (UMAP)[73]; in these two latter methods, the separation between cell types is sensitive to various tuning parameters, which may introduce subjectivity. Furthermore, only the small distances within clusters are interpretable. Thus, when results, such as cluster assignment, are based on a low-dimensional manifold, we additionally recommend showing the corresponding distances in higher-dimensional space, for example, as distributions of pairwise distances between single cells within and across clusters[71].

When dimensionality reduction is used for clustering cells, we recommend including positive controls. These controls may be bulk samples composed of purified cell types (if such isolation is possible) from the same population as the single cells of interest. Such positive controls should be prepared in tandem with the single cells. Next, both positive controls and single cells can be projected simultaneously on the low-dimensional manifold.

This type of analysis provides useful evidence for evaluating clustering[16,18] patterns: the degree to which the positive controls and the single cells of the same type cluster together indicates the consistency of the measurements. To further determine whether sample preparation is driving any clustering, we also recommend evaluating whether principal components correlate with technical covariates (such as batches, missing value rate or mass tags) and correcting for these dependencies if needed.

Managing and propagating uncertainty

As discussed above, assumptions about missing data and the application of dimensionality-reduction methods can substantially influence the final conclusions. Thresholds, such as filters for excluding single cells due to failed sample preparation or for excluding peptides due to high levels of interference, can also influence the results[16,48]. Such choices should be based on objective grounds, such as true and false discovery rates derived from controls. For example, negative controls allow establishing objective filters for failed single cells as already implemented in multiple pipelines[7,16,48]. When thresholds are set based on subjective choices, this should be explicitly stated, and the choices should be treated as a source of uncertainty in the final results. The sensitivity of the results to all experimental and methodological choices should clearly be conveyed.

Interpreting features of single-cell proteomic data

Algorithms underlying peptide identification have evolved along with technological advances in data generation to use the increasing set of features from bulk proteomic data. Features measured at the single-cell level may differ substantially from those of corresponding bulk samples as lowly abundant fragments may not be detected and other fragments may have lower signal relative to background noise[74]. Mitigating these challenges may benefit from directed efforts dedicated to developing robust models trained on features that have the greatest discriminatory power at the single-cell-level input. These models may incorporate additional features with search engine results, as implemented by mokapot[75] and DART-ID[76]. To guard against false identifications, we recommend scrutinizing any peptides identified in single cells but not identified in larger bulk samples from the same biological systems. Such identifications are likely incorrect, especially for DIA experiments. Thus the spectra supporting them (for example, extracted ion current) should be examined and data-analysis methods should be reassessed.

To improve proteome coverage, new search engines may be designed and optimized to exploit regular patterns in the data, such as the precisely known and measured mass shifts in the precursors and fragments of plexDIA data[77,78]. Indeed, current single-cell proteomic MS methods are capable of measuring tens of thousands of peptide-like features; however, only a small fraction (between 1% and 10%) of these features are assigned sequences at 1% FDR[20,56,77]. Anticipated models that successfully address these unique challenges will enable identification rates to approach those of bulk experiments and extend the utility of single-cell proteomics in biomedical research[32,77].

Reporting standards

The goal of reporting is to enable other researchers to repeat, reproduce, assess and build upon published data and their interpretation[79]. While reproduction and replication do not guarantee accuracy, they build trust in the analysis process through verifiability, thus strengthening confidence in the reported data and results. Replication requires sufficient documentation of metadata, and a good starting place for reporting metadata are formats developed for bulk MS data[23,80], including those specifically for proteomic data[81] and those prepared by journals[82,83] and societies[84], as well as for scRNA-seq data[85]. Nonetheless, single-cell MS proteomic data have additional aspects that should be reported, which are the focus of our recommendations. Below, we document what we believe is essential information needed to provide value to single-cell proteomic data, metadata and analysis results.

Experimental design and method description

We recommend that the detailed design of the experiments should be reported, which includes treatment groups, number of single cells per group, sampling methods and analysis batches (Fig. 3). The experimental design may be reported as a table listing each analyzed single cell on its corresponding row and each descriptor in its corresponding column. Specifically, columns document biological and technical descriptors, that is, variables that describe the biology of the measured cells and technical factors that are likely to influence the measurements.

Biological descriptors should contain sample type (such as single cell, carrier, empty or control sample) and biological group, such as treatment condition or patient or donor identifier, cell line, organism and organ or part of origin (if cells from multiple organisms or multiple organs are assayed) and biological characteristics for multisample and/or multicondition studies. When available, additional biological descriptors may include the cell type and/or cell state (for example, their spatial and temporal information in tissues), physical markers (for example, pigmentation, measured by flow cytometry), cell size and aspect ratio. These descriptors apply only to single-cell samples and thus will remain empty for some samples, such as negative controls. Note that some of these descriptors might be known before data acquisition (such as cell types based on different cell cultures or following from flow cytometry sorting) or be the results of downstream analyses (such as cell types or cell states inferred from clustering or differential abundance analysis).

Technical descriptors should include the raw data file names (Box 1) and acquisition dates, as well as variables describing the underlying technical variability. These descriptors include all batch factors related to cell isolation, sample preparation, peptide and protein separation (chromatography or electrophoresis batches), operator(s) and instruments, and mass tags (in case of labeled quantitation). Such a sample metadata table allows for quality control, for example, by enabling verification that the number of rows in the table matches the number of cells reported in the paper and that the number and names of raw data files extracted from the table are compatible with the files in the data repositories (see Box 1). We encourage researchers to document additional descriptors when needed, such as variables defining subsets of cells pertaining to distinct analyses.

This sample metadata table should be complemented by a text file (often called README) that further describes each of these descriptors and the overall experiment. An example README file is included in Supplementary Note 1 to facilitate standardization and data reuse. The README file should contain a summary of the study design and the protocols. The measurement units of descriptors (such as micrometers for cell sizes) should also be documented in the README file, as opposed to encoding them as a suffix in the descriptor's name.

Data and code sharing

Ideally, raw and processed MS data should be shared using open formats, such as HUPO Proteomics Standards Initiative community-developed formats dedicated to MS data: mzML[86] for raw data, mzIdentML[87] for search results and mzTab[88] or text-based spreadsheets for quantitative data. When binary formats from proprietary software are provided, they should be converted into an open and accessible format as well when possible. Raw data files and search results should be made available through dedicated repositories, such as PRIDE[81] and MassIVE[89].

Code repositories, such as GitLab or GitHub[90], are ideal to store and share code, scripts, notebooks and, when size permits, quantitative data matrices. When these become too large to be stored directly with the scripts that generate them, they should be made available in institutional or general-purpose open repositories, such as Zenodo or Open Science Framework, or on publicly available cloud storage. The latter, however, requires a commitment by the data provider to keep the data public. The README file (Supplementary Note 1) containing the description of the experimental design and the different locations holding data should be provided in all these locations. The manuscript material and method section and/or the supplementary information should provide experiment identifiers and links to all the external data and metadata resources.

While these data-sharing recommendations apply broadly to proteomic experiments, some are specific to single-cell proteomics (such as single-cell isolation) and some are made more important because of the aim to analyze tens of thousands of single cells per experiment[57]. Such sample sizes are required to adequately power the analysis of dozens of cellular clusters and states across many treatment conditions and individuals. The large sample sizes, in turn, considerably increase the importance of reporting batches, including all variations in the course of sample preparation and data acquisition, as well as the known phenotypic descriptors for each single cell. These reporting recommendations expand the essential descriptors in the metadata. Large study sizes also heighten the importance of reporting datasets from intermediate processing steps, such as search results and peptide $\times$ cell matrices, to reduce the computational burden on reproducing individual steps from the analysis.

Sharing data is necessary but insufficient for replication data reuse. Any analysis of data is likely to require the associated metadata. Furthermore, the exact processing of data should be documented and shared as it can profoundly influence the final results that are used to infer biological interpretations. Data processing can hardly (and should not need to) be

retro-engineered from the result files. Therefore, annotated scripts or notebooks used to process, prepare and analyze the data should be provided with the data.

Using software for standardizing workflows across laboratories facilitates reporting. Examples of such workflows include the scp R–Bioconductor package[48,91], the sceptre Python package[9], the SCoPE2 pipeline[16,92] or the Scripts and Pipelines for Proteomics[93]. Packages that allow comparing structured and repeatable data processing, including evaluating different algorithms for a processing step, provide further advantages[48,91]. Software platforms that support exporting the commands and parameters used should be strongly preferred because audit log and/or parameter files can help tracking and later reproducing the different processing steps, including software and the versions used at each step. We strongly advise against using non-reproducible software given the difficulty in capturing their operation.

Result reporting

Given the rapid evolution of the field, specific description of the methods should be favored over simply referring to other publications using 'as previously analyzed in ref.'. When reporting results, it should be made clear which data the result refers to. This is, for example, crucial when reporting CVs when CVs on log-transformed data are lower than those on the linear scale. CVs can be used to quantify very different quantities, such as repeatability between MS runs or consistency of protein quantification based on different peptides, and thus the exact quantity must be explicitly specified. Similarly, researchers should systematically report major features of the data that influence the results and how these were observed and addressed throughout the data analysis. These typically include missing values and batch effects. Reproducibility requires going beyond the minimalist 'material and method' sections that often fail to describe the processing of samples and data to enable their replication.

Often, studies include several sets of raw, identification and quantitation files, addressing different research questions, such as different instruments or MS settings, different cell types or growth conditions, and different individuals. A single dump of all files makes data reuse challenging. In such situations, it is advisable to split the file in different folders, following a consistent structure. The high-level README file, already mentioned above, should describe what each of these folders correspond to, and each folder should contain its own README file describing its content in detail and the specific points that these sets of files aim to address.

As described above, data-acquisition strategies are inextricably linked to both the number of proteins quantified and the quality of quantitation in single-cell proteomic experiments. While the reporting of MS acquisition details is not necessarily required for data reanalysis, acquiring similar data could be impractical or impossible if key details are not reported. This is even more evident with the rise of intelligent data-acquisition strategies that often have more advanced, non-standard parameters or use third-party (non-vendor)-supplied software. Luckily, most raw data files report the parameters used for analysis and some vendors have enabled method generation from a raw data file. However, for instances in which third-party software makes real-time decisions that alter mass spectrometer operation, the

software should be made available to the broader research community. Ideally this software would be open source. If it needs to be delivered as a compiled executable, the underlying algorithms should be described in such a way that others could reproduce a similar method. Furthermore, the reporting of parameters relevant to the decisions made in real time as well as the output of real-time decisions would ideally be provided. These considerations would enable faster implementation in laboratories attempting to replicate published results on their own instrumentation.

Conclusions and perspectives

These reporting guidelines might give the impression that a lot of additional work is expected when reporting on studies according to our recommendations, many of which apply to all proteomic studies. Yet, the recommendations merely highlight good scientific practice to be implemented continuously, starting when the research is designed, when the data are acquired, processed and eventually interpreted. When so implemented, they become habits enabling robust research rather than a burden to be addressed at the end of the research project. Data, metadata and analysis documentation and reporting happen at different stages of the analysis process and rely on each other. The investment that we are suggesting here is simply work that is spread across the research project, rather than extra work done at the very end of it[94].

We believe that the adoption of guidelines for performing and reporting single-cell proteomic studies by the scientific community and their promotion by journals and data archives is essential for establishing solid foundations for this emerging field. The suggested reporting standards will facilitate all levels of replication and thus promote the dissemination, improvement and adoption of single-cell technologies and data analysis. Sound data evaluation and interpretation will further promote the reuse of single-cell proteomic data and results outside of the laboratories that currently drive the domain and increase secondary added value of our experiments and efforts.

We hope and expect that the initial guidelines offered here will evolve with the advancement of single-cell proteomic technologies[77], the increasing scale and sophistication of biological questions investigated by these technologies and the integration with other data modalities, such as single-cell transcriptomics, spatial transcriptomics, imaging, electrophysiology, prioritized MS approaches and post-translational-modification-level and proteoform-level (that is, top–down) single-cell proteomic methods. We invite the community to discuss these guidelines and contribute to their evolution. We hope to facilitate such broader contributions via an online portal at https://single-cell.net/guidelines.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Data availability

We did not generate new data for this article. Source data are provided with this paper.

Code availability

We did not generate new code for this article. The code used for simulations and plotting is available at https://github.com/SlavovLab/SCP_recommendations.

References

1. Budnik B, Levy E, Harmange G & Slavov N SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. Genome Biol. 19, 161 (2018). [PubMed: 30343672] A demonstration of quantifying hundreds of proteins per single human cell (T lymphocytes) and proteogenomic analysis of stem cell differentiation. It also introduced the isobaric carrier approach.

2. Zhu Y et al. Nanodroplet processing platform for deep and quantitative proteome profiling of 10–100 mammalian cells. Nat. Commun 9, 882 (2018). [PubMed: 29491378] Introduced a microfabricated chip (nanoPOTS) for sample preparation and used it to prepare small bulk samples in sample volumes of about 200 nl.

3. Singh A Towards resolving proteomes in single cells. Nat. Methods 18, 856 (2021). [PubMed: 34354289]

4. Kelly RT Single-cell proteomics: progress and prospects. Mol. Cell. Proteomics 19, 1739–1748 (2020). [PubMed: 32847821]

5. Petelski AA et al. Multiplexed single-cell proteomics using SCoPE2. Nat. Protoc 16, 5398–5425 (2021). [PubMed: 34716448]

6. Leduc A, Huffman RG, Cantlon J, Khan S & Slavov N Exploring functional protein covariation across single cells using nPOP. Genome Biol. 23, 261 (2022). [PubMed: 36527135] Introduced a method for simultaneous sample preparation of thousands of single cells in droplets of about 20 nl on the surface of glass slides. It also demonstrated cross-validation based on using different MS methods.

7. Derks J et al. Increasing the throughput of sensitive proteomics by plexDIA. Nat. Biotechnol 41, 50–59 (2022). [PubMed: 35835881] Introduced a multiplexed DIA method (plexDIA) that implements parallel analysis of both peptides and single cells, which enabled multiplicative increase in throughput.

8. Lombard-Banek C et al. In vivo subcellular mass spectrometry enables proteo-metabolomic single-cell systems biology in a chordate embryo developing to a normally behaving tadpole (*X. laevis*). Angew. Chem. Int. Ed. Engl 60, 12852–12858 (2021). [PubMed: 33682213] This study used in vivo and dual proteo-metabolomics single-cell MS for single-cell molecular systems biology in live vertebrate embryos.

9. Schoof EM et al. Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. Nat. Commun 12,3341 (2021). [PubMed: 34099695] Analyzed primary cells using an isobaric carrier and modified SCoPE2 approach.

10. Brunner A-D et al. Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. Mol. Syst. Biol 18, e10798 (2022). [PubMed: 35226415]

11. Choi SB, Polter AM & Nemes P Patch-clamp proteomics of single neurons in tissue using electrophysiology and subcellular capillary electrophoresis mass spectrometry. Anal. Chem 94, 1637–1644 (2022). [PubMed: 34964611]

12. Slavov N Unpicking the proteome in single cells. Science 367, 512–513 (2020). [PubMed: 32001644]

13. Slavov N Single-cell protein analysis by mass spectrometry. Curr. Opin. Chem. Biol 60, 1–9 (2021). [PubMed: 32599342]

14. Beltra J-C et al. Developmental relationships of four exhausted $CD8^+$ T cell subsets reveals underlying transcriptional and epigenetic landscape control mechanisms. Immunity 52,825–841 (2020). [PubMed: 32396847]

15. Grün D et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature 525, 251–255 (2015). [PubMed: 26287467]

16. Specht H et al. Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. Genome Biol. 22, 50 (2021). [PubMed: 33504367] This study analyzed thousands of proteins in over a thousand single cells. It performed parallel RNA and protein measurements in single cells and identified the emergence of polarization in the absence of polarizing cytokines.

17. Zhu Y et al. Single-cell proteomics reveals changes in expression during hair-cell development. eLife 8, e50777 (2019). [PubMed: 31682227]

18. Huffman RG et al. Prioritized single-cell proteomics reveals molecular and functional polarization across primary macrophages. Preprint at *bioRxiv* 10.1101/2022.03.16.484655 (2022). Demonstrated that prioritized MS analysis increases the consistency, sensitivity and depth of protein quantification in single cells. It also enabled quantifying post-translational modifications and polarization in primary macrophages.

19. Slavov N Learning from natural variation across the proteomes of single cells. PLoS Biol. 20, e3001512 (2021).

20. Cong Y et al. Improved single-cell proteome coverage using narrow-bore packed nanoLC columns and ultrasensitive mass spectrometry. Anal. Chem 92, 2665–2671 (2020). [PubMed: 31913019] Demonstrated increased sensitivity by using narrow-bore analytical columns.

21. Soltwisch J et al. Mass spectrometry imaging with laser-induced postionization. Science 348, 211–215 (2015). [PubMed: 25745064]

22. Johnson KR, Gao Y, Greguš M & Ivanov AR On-capillary cell lysis enables top-down proteomic analysis of single mammalian cells by CE-MS/MS. Anal. Chem 94, 14358–14367 (2022). [PubMed: 36194750]

23. Taylor CF et al. The minimum information about a proteomics experiment (MIAPE). Nat. Biotechnol 25, 887–893 (2007). [PubMed: 17687369]

24. Donnelly DP et al. Best practices and benchmarks for intact protein analysis for top–down mass spectrometry. Nat. Methods 16, 587–594 (2019). [PubMed: 31249407]

25. Shao W et al. Comparative analysis of mRNA and protein degradation in prostate tissues indicates high stability of proteins. Nat. Commun 10, 2524 (2019). [PubMed: 31175306]

26. Reichard A & Asosingh K Best practices for preparing a single cell suspension from solid tissues for flow cytometry. Cytometry A 95, 219–226 (2019). [PubMed: 30523671]

27. Cong Y et al. Ultrasensitive single-cell proteomics workflow identifies >1000 protein groups per mammalian cell. Chem. Sci 12, 1001–1006 (2021).

28. Woo J et al. High-throughput and high-efficiency sample preparation for single-cell proteomics using a nested nanowell chip. Nat. Commun 12, 6246 (2021). [PubMed: 34716329]

29. Leduc A, Huffman RG, Cantlon J, Kahn S & Slavov N Exploring functional protein covariation across single cells using nPOP. Genome Biol. 10.1186/s13059-022-02817-5 (2022).

30. Brasko C et al. Intelligent image-based in situ single-cell isolation. Nat. Commun 9, 226 (2018). [PubMed: 29335532]

31. Mund A et al. Deep Visual Proteomics defines single-cell identity and heterogeneity. Nat. Biotechnol 40, 1231–1240 (2022). [PubMed: 35590073]

32. Derks J & Slavov N Strategies for increasing the depth and throughput of protein analysis by plexDIA. J. Proteome Res 10.1021/acs.jproteome.2c00721 (2023).

33. Rosenberger FA et al. Spatial single-cell mass spectrometry defines zonation of the hepatocyte proteome. Preprint at *bioRxiv* 10.1101/2022.12.03.518957 (2022).

34. Dolman S, Eeltink S, Vaast A & Pelzing M Investigation of carryover of peptides in nano-liquid chromatography/mass spectrometry using packed and monolithic capillary columns. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci 912, 56–63 (2013).

35. Fernandez-Lima F, Kaplan DA, Suetering J & Park MA Gas-phase separation using a trapped ion mobility spectrometer. Int. J. Ion Mobil. Spectrom 14, 10.1007/s12127-011-0067-8 (2011).

36. Specht H & Slavov N Transformative opportunities for single-cell proteomics. J. Proteome Res 17, 2565–2571 (2018). [PubMed: 29945450]

37. Specht H, Harmange G, Perlman DH & Emmott E Automated sample preparation for high-throughput single-cell proteomics. Preprint at bioRxiv 10.1101/399774 (2018). An automated method for simultaneously preparing hundreds of single cells for MS analysis.

38. Petelski AA, Slavov N & Specht H Single-cell proteomics preparation for mass spectrometry analysis using freeze–heat lysis and an isobaric carrier. J. Vis. Exp 10.3791/63802 (2022).

39. Zhu Y et al. Proteomic analysis of single mammalian cells enabled by microfluidic nanodroplet sample preparation and ultrasensitive nanoLC–MS. Angew. Chem. Int. Ed. Engl 57, 12370–12374 (2018). [PubMed: 29797682] A label-free MS analysis of hundreds of proteins in single HeLa cells.

40. Laganowsky A, Reading E, Hopper JTS & Robinson CV Mass spectrometry of intact membrane protein complexes. Nat. Protoc 8, 639–651 (2013). [PubMed: 23471109]

41. Marx V A dream of single-cell proteomics. Nat. Methods 16, 809–812 (2019). [PubMed: 31406385]

42. Liang Y et al. Fully automated sample processing and analysis workflow for low-input proteome profiling. Anal. Chem 93, 1658–1666 (2021). [PubMed: 33352054]

43. Ctortecka C et al. An automated workflow for multiplexed single-cell proteomics sample preparation at unprecedented sensitivity. Preprint at *bioRxiv* 10.1101/2021.04.14.439828 (2022).

44. Cole RB Some tenets pertaining to electrospray ionization mass spectrometry. J. Mass Spectrom 35, 763–772 (2000). [PubMed: 10934430]

45. Li S et al. An integrated platform for isolation, processing, and mass spectrometry-based proteomic profiling of rare cells in whole blood. Mol. Cell. Proteomics 14, 1672–1683 (2015). [PubMed: 25755294]

46. DeLaney K et al. Microanalysis of angiotensin peptides in the brain using ultrasensitive capillary electrophoresis trapped ion mobility mass spectrometry. Anal. Chem 94, 9018–9025 (2022). [PubMed: 35696295]

47. Navarro P et al. A multicenter study benchmarks software tools for label-free proteome quantification. Nat. Biotechnol 34, 1130–1136 (2016). [PubMed: 27701404]

48. Vanderaa C & Gatto L Replication of single-cell proteomics data reveals important computational challenges. Expert Rev. Proteomics 18, 835–843 (2021). [PubMed: 34602016] A replication study that bolstered the confidence in single-cell MS proteomics and outlined the need for developing standardized and optimized data-analysis pipelines.

49. uklina J et al. Diagnostics and correction of batch effects in large-scale proteomic studies: a tutorial. Mol. Syst. Biol 17, e10240 (2021). [PubMed: 34432947]

50. Pino LK et al. Calibration using a single-point external reference material harmonizes quantitative mass spectrometry proteomics data between platforms and laboratories. Anal. Chem 90, 13112–13117 (2018). [PubMed: 30350613]

51. Slavov N Increasing proteomics throughput. Nat. Biotechnol 39, 809–810 (2021). [PubMed: 33767394]

52. Lazic SE, Clarke-Williams CJ & Munafò MR What exactly is 'N' in cell culture and animal experiments? PLoS Biol. 16, e2005282 (2018). [PubMed: 29617358]

53. Singh A Sensitive protein analysis with plexDIA. Nat. Methods 19, 1032 (2022). [PubMed: 36068318]

54. Budnik B, Levy E, Harmange G & Slavov N SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. Genome Biol 10.1186/s13059-018-1547-5 (2018).

55. Cheung TK et al. Defining the carrier proteome limit for single-cell proteomics. Nat. Methods 18, 76–83 (2021). [PubMed: 33288958]

56. Specht H & Slavov N Optimizing accuracy and depth of protein quantification in experiments using isobaric carriers. J. Proteome Res 20, 880–887 (2021). [PubMed: 33190502]

57. Slavov N Scaling up single-cell proteomics. Mol. Cell. Proteomics 21, 100179 (2022). [PubMed: 34808355]

58. Furtwängler B et al. Real-time search assisted acquisition on a tribrid mass spectrometer improves coverage in multiplexed single-cell proteomics. Mol. Cell. Proteomics 21, 100219 (2022). [PubMed: 35219906]

59. Huffman RG, Chen A, Specht H & Slavov N DO-MS: data-driven optimization of mass spectrometry methods. J. Proteome Res 18, 2493–2500 (2019). [PubMed: 31081635]

60. National Academies of Sciences, Engineering, and Medicine et al. Understanding reproducibility and replicability. In Reproducibility and Replicability in Science (National Academies Press (US), 2019).

61. Levy E & Slavov N Single cell protein analysis for systems biology. Essays Biochem. 62, 595–605 (2018). [PubMed: 30072488]

62. Franks A, Airoldi E & Slavov N Post-transcriptional regulation across human tissues. PLoS Comput. Biol 13, e1005535 (2017). [PubMed: 28481885]

63. Plubell DL et al. Putting Humpty Dumpty back together again: what does protein quantification mean in bottom–up proteomics. J. Proteome Res 21, 891–898 (2022). [PubMed: 35220718]

64. Malioutov D et al. Quantifying homologous proteins and proteoforms. Mol. Cell. Proteomics 18, 162–168 (2019). [PubMed: 30282776]

65. Lytal N, Ran D & An L Normalization methods on single-cell RNA-seq data: an empirical survey. Front. Genet 11, 41 (2020). [PubMed: 32117453]

66. Hicks SC, Townes FW, Teng M & Irizarry RA Missing data and technical variability in single-cell RNA-sequencing experiments. Biostatistics 19, 562–578 (2018). [PubMed: 29121214]

67. Lazar C, Gatto L, Ferro M, Bruley C & Burger T Accounting for the multiple natures of missing values in label-free quantitative proteomics data sets to compare imputation strategies. J. Proteome Res 15, 1116–1125 (2016). [PubMed: 26906401]

68. Dabke K, Kreimer S, Jones MR & Parker SJ A simple optimization workflow to enable precise and accurate imputation of missing values in proteomic data sets. J. Proteome Res 20, 3214–3229 (2021). [PubMed: 33939434]

69. Bramer LM, Irvahn J, Piehowski PD, Rodland KD & Webb-Robertson B-JM A review of imputation strategies for isobaric labeling-based shotgun proteomics. J. Proteome Res 20, 1–13 (2021). [PubMed: 32929967]

70. Chari T, Banerjee J & Pachter L The specious art of single-cell genomics. Preprint at *bioRxiv* 10.1101/2021.08.25.457696 (2021).

71. Specht H et al. Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. Genome Biol. 10.1186/s13059-021-02267-5 (2021).

72. van der Maaten L & Hinton G Visualizing data using *t*-SNE. J. Mach. Learn. Res 9, 2579–2605 (2008).

73. McInnes L, Healy J & Melville J UMAP: uniform manifold approximation and projection for dimension reduction. Preprint at *arXiv* 10.48550/arXiv.1802.03426 (2018).

74. Boekweg H et al. Features of peptide fragmentation spectra in single-cell proteomics. J. Proteome Res 21, 182–188 (2022). [PubMed: 34920664]

75. Fondrie WE & Noble WS mokapot: fast and flexible semisupervised learning for peptide detection. J. Proteome Res 20, 1966–1971 (2021). [PubMed: 33596079]

76. Chen AT, Franks A & Slavov N DART-ID increases single-cell proteome coverage. PLoS Comput. Biol 15, e1007082 (2019). [PubMed: 31260443]

77. Slavov N Driving single cell proteomics forward with innovation. J. Proteome Res 20, 4915–4918 (2021). [PubMed: 34597050]

78. Framework for multiplicative scaling of single-cell proteomics. Nat. Biotechnol 41, 23–24 (2023). [PubMed: 35851377]

79. Wilkinson MD et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3, 160018 (2016). [PubMed: 26978244]

80. Dai C et al. A proteomics sample metadata representation for multiomics integration and big data analysis. Nat. Commun 12, 5854 (2021). [PubMed: 34615866]

81. Perez-Riverol Y et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. 50, D543–D552 (2022). [PubMed: 34723319]

82. Carr S et al. The need for guidelines in publication of peptide and protein identification data: Working Group on Publication Guidelines for Peptide and Protein Identification Data. Mol. Cell. Proteomics 3, 531–533 (2004). [PubMed: 15075378]

83. Chalkley RJ, MacCoss MJ, Jaffe JD & Röst HL Initial guidelines for manuscripts employing data-independent acquisition mass spectrometry for proteomic analysis. Mol. Cell. Proteomics 18, 1–2 (2019). [PubMed: 30602589]

84. Omenn GS Reflections on the HUPO Human Proteome Project, the flagship project of the Human Proteome Organization, at 10 years. Mol. Cell. Proteomics 20, 100062 (2021). [PubMed: 33640492]

85. Füllgrabe A et al. Guidelines for reporting single-cell RNA-seq experiments. Nat. Biotechnol 38, 1384–1386 (2020). [PubMed: 33188371]

86. Martens L et al. mzML—a community standard for mass spectrometry data. Mol. Cell. Proteomics 10, R110.000133 (2011).

87. Vizcaíno JA et al. The mzIdentML data standard version 1.2, supporting advances in proteome informatics. Mol. Cell. Proteomics 16, 1275–1285 (2017). [PubMed: 28515314]

88. Griss J et al. The mzTab data exchange format: communicating mass-spectrometry-based proteomics and metabolomics experimental results to a wider audience. Mol. Cell. Proteomics 13, 2765–2775 (2014). [PubMed: 24980485]

89. Wang M et al. Assembling the community-scale discoverable human proteome. Cell Syst. 7, 412–421 (2018). [PubMed: 30172843]

90. Perez-Riverol Y et al. Ten simple rules for taking advantage of Git and GitHub. PLoS Comput. Biol 12, e1004947 (2016). [PubMed: 27415786]

91. Vanderaa C & Gatto L scp: mass spectrometry-based single-cell proteomics data analysis. Bioconductor https://bioconductor.org/packages/release/bioc/html/scp.html (2020).

92. Slavov N & hspekt. SlavovLab/SCoPE2: zenodo release 20201218 (v1.0). Zenodo 10.5281/zenodo.4339954 (2020).

93. Specht H, Huffman RG, Derks J, Leduc A & Slavov N Scripts and Pipelines for Proteomics (SPP) (GitHub, 2020).

94. Quintana D Five Things About Open and Reproducible Science that Every Early Career Researcher Should Know 10.17605/OSF.IO/DZTVQ (2020).

95. Slavov N Measuring protein shapes in living cells. J. Proteome Res 20, 3017 (2021). [PubMed: 33988997]

96. President and Fellows of Harvard College. File Naming Conventions https://datamanagement.hms.harvard.edu/collect/file-naming-conventions (Harvard Medical School, 2023).

BOX 1

Despite its apparent simplicity, file naming deserves thoughtful consideration

Files names should be unique (unlikely to be used in other studies) and linked to the measurements in the file; additional good practices are summarized in ref. [96]. We suggest thinking about file naming and file-naming conventions to easily identify groups of files pertaining to specific metadata elements or experiments. A systematic file-naming convention allows files to be both machine and human readable and searchable. File names should avoid using any special characters and use the same character (such as a dash or an underscore, rather than spaces) to separate the different elements of the file names. If using dates to list files chronologically, the YYYYMMDD format should be used. Finally, these naming conventions and any abbreviations used as part of the file names need to be documented in the main README file; see an example provided as Supplementary Note 1.
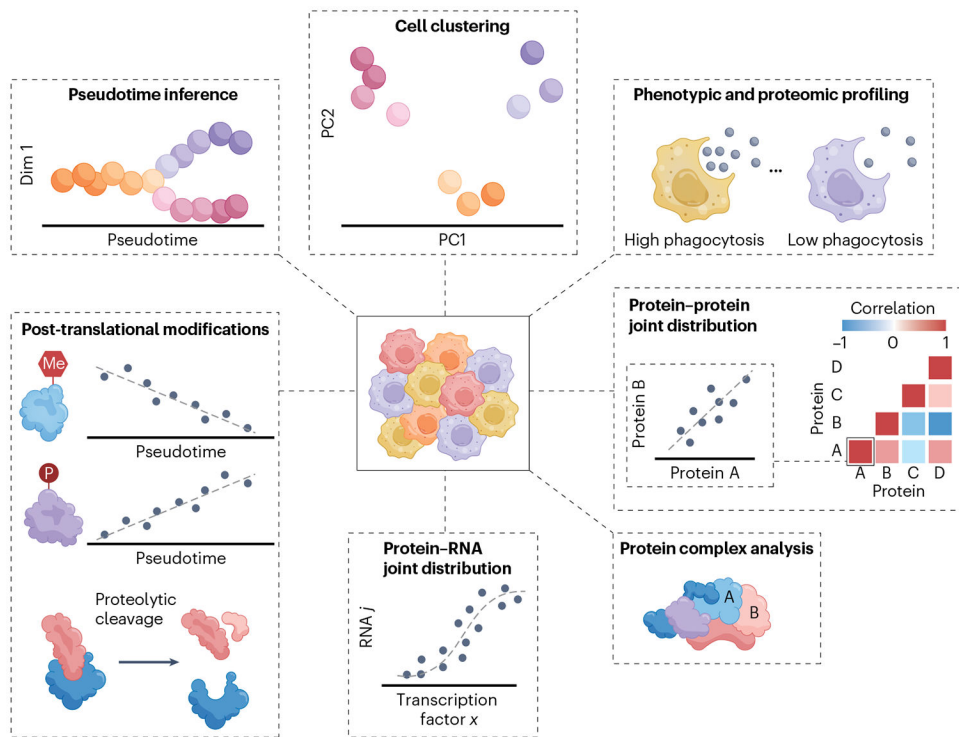
Fig. 1 |. Emerging applications of single-cell proteomics by MS.
Single-cell proteomic measurements can define cell type and cell state clusters[9], support pseudotime inference, link protein levels to functional phenotypes, such as phagocytic activity[18], quantify protein covariation and apply it to study protein complexes[1,6,19], analyze protein conformations[95] and quantify protein modifications, such as phosphorylation and proteolysis[5,6,18]. Furthermore, integrating protein and RNA measurements from the same biological systems (as in refs.[1,16]) allows inferring transcriptional and post-translational regulation[1,16] and investigating the covariation of transcription factors and downstream target transcripts[16]. Dim, dimension; PC, principal component.
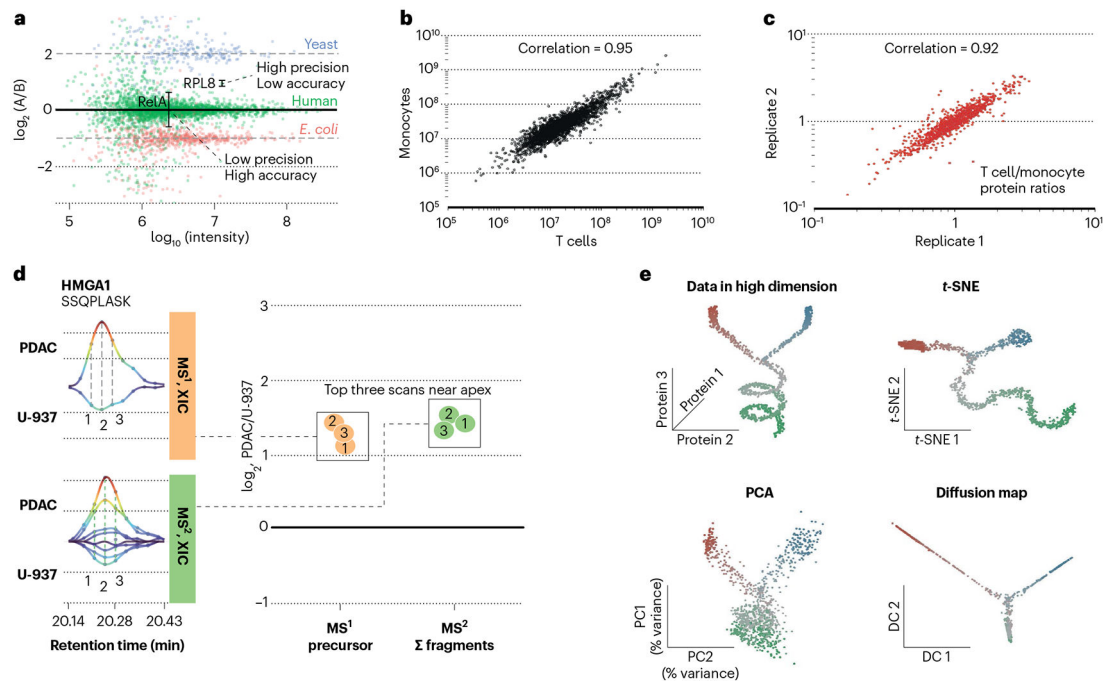
Fig. 2 |. Evaluating and interpreting single-cell proteomic data.

a, Quantitative accuracy of protein ratios between samples A and B measured by label-free DIA analysis relative to the corresponding mixing ratios denoted by dotted lines[7]. Some proteins are quantified with high precision but low accuracy (for example, ribosomal protein L8 (RPL8)), while others are quantified with high accuracy and low precision (for example, RelA). *E. coli, Escherichia coli*. The proteomes of T cells and monocytes correlate strongly (**b**) despite the fact that many proteins are differentially abundant between the two cell types (**c**). Data for **b,c** are from Specht et al.[37]. **d**, Extracted ion chromatograms (XIC) from single-cell MS measurements by plexDIA for a peptide from the high mobility group protein A1 (HMGA1). Such data allow quantifying peptides at both MS^1 and MS^2 levels, which can be used to evaluate the consistency and reliability of the quantification. This example data from Derks et al.[7] show that relative levels estimated from precursors (peach color) agree with the relative levels estimated from the corresponding summed-up fragments (green color). At both MS^1 and MS^2 levels, three estimates are obtained based on the three scans closest to the elution peak apex. The fold changes are between pancreatic ductal adenocarcinoma (PDAC) and monocyte (U-937) cells. **e**, Different dimensionality-reduction methods approximate the data in different ways. We simulated three-dimensional data for three cell states, where one cell state (green) progressively diverges to two distinct cell states (blue and red, top left). Projecting the data to two dimensions loses information. Specifically, PCA loses the non-linear cycling effect and mixes early (green) and intermediate (gray) cells, *t*-SNE does not correctly capture the distances between the three populations, and diffusion maps do not capture the noise in the data and compress the early state cells. DC1 and DC2 correspond to diffusion components 1 and 2. The code for this simulation is available at github.com/SlavovLab/SCP_recommendations.
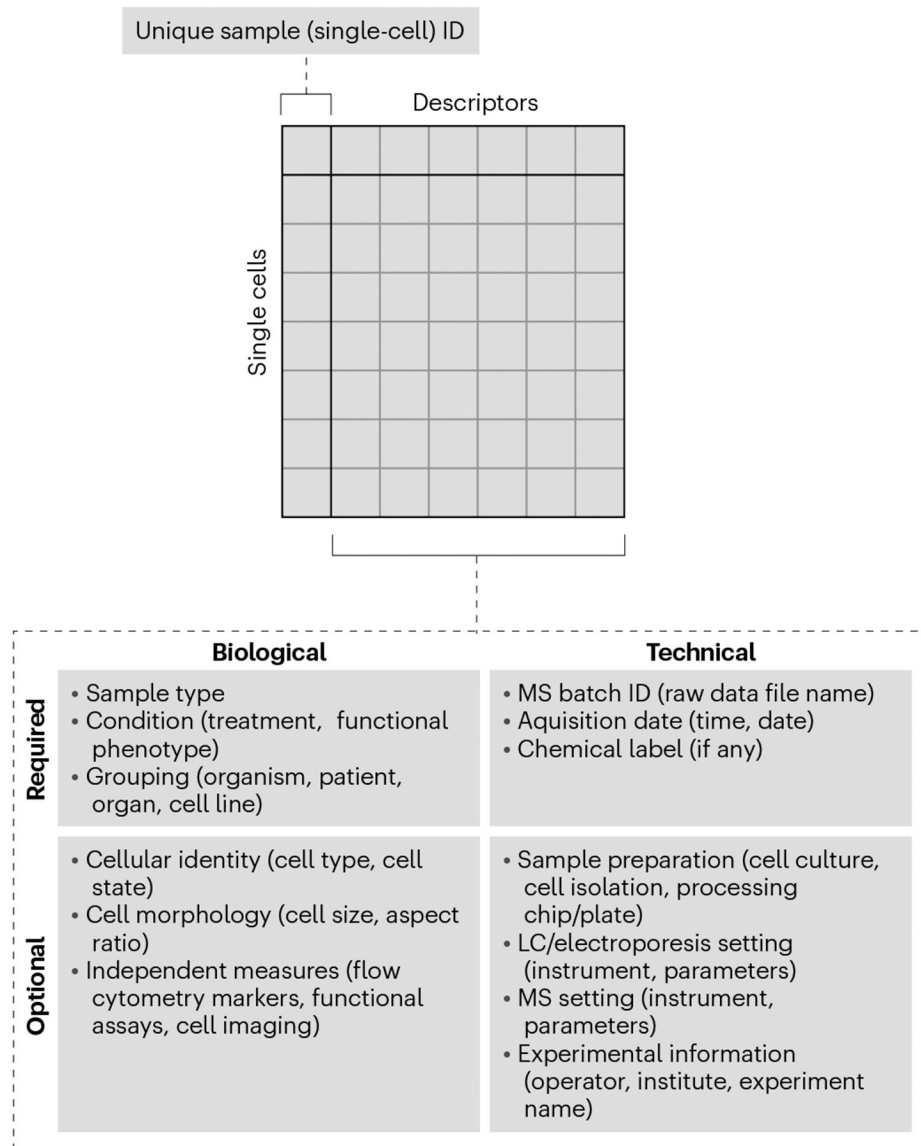
Fig. 3 |. Suggested descriptors of single-cell proteomic samples.

Metadata should include the experimental design table with rows corresponding to single cells and columns corresponding to the required and optional features listed here (an example is provided as source data). Attributes provided in parentheses are given as examples or for clarification. The green shading highlights required descriptors, while gray shading includes a non-exhaustive list of optional descriptors, which may also include spatial (for example, position in tissues) and temporal information for the cells when available. The descriptors (and their units, when relevant) should be documented in the experiment's dedicated README file.