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

Laurent Gatto, Ruedi Aebersold, Juergen Cox, Vadim Demichev, Jason Derks, Edward Emmott, Alexander M. Franks, Alexander R. Ivanov, Ryan T. Kelly, Luke Khoury, Andrew Leduc, Michael J. MacCoss, Peter Nemes, David H. Perlman, Aleksandra A. Petelski, Christopher M. Rose, Erwin M. Schoof, Jennifer Van Eyk, Christophe Vanderaa, John R. Yates III, and Nikolai Slavov

Supporting website: https://single-cell.net/guidelines
Correspondence: nslavov@northeastern.edu

Abstract

Analyzing proteins from single cells by tandem mass spectrometry (MS) has 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. Establishing community guidelines and standardized metrics will enhance rigor, data quality, and alignment between laboratories. Here we propose best practices, quality controls, and data reporting guidelines to assist in the broad adoption of reliable quantitative workflows for single-cell proteomics.

New approaches and technologies for experimental design, sample preparation, data acquisition, and data analysis have enabled the measurement of several thousand proteins across many single mammalian cells[1–11]. These developments open up exciting new opportunities for biomedical research[12], as illustrated in Fig. 1. Yet, 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 artifacts and to maximize data reproducibility in single-cell proteomics, we offer guidelines for optimization, validation, and reporting of single-cell proteomic workflows and results.

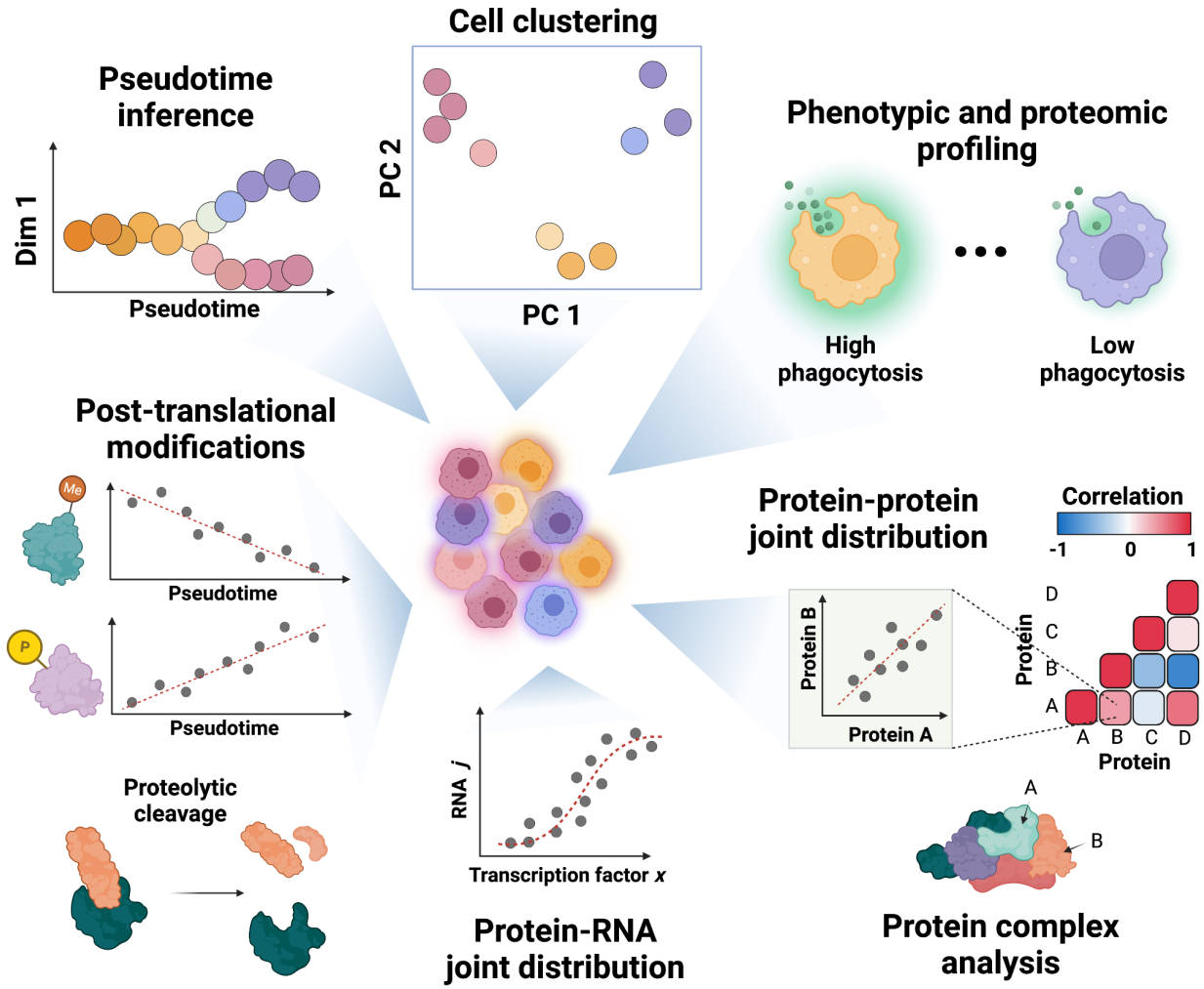

Figure 1 | Emerging applications of single-cell proteomics by mass spectrometry. Single-cell proteomic measurements can define cell-type and cell-state clusters[9], support pseudotime inference, link protein levels to functional phenotypes, such as phagocytic activity[13], quantify protein covariation[1,6], infer transcriptional and posttranslational regulation[14], and analyze protein modifications, such as phosphorylation and proteolytic activities[13].

The tandem MS methods for single-cell bottom-up proteomics span a range of techniques[15], including multiplexed and label-free methods, both of which can be performed by data-dependent acquisition (DDA)[1,16] and data-independent acquisition (DIA)[7,10]. The initial recommendations presented here are relevant to all of these methods, and we will note any exceptions. Our initial recommendations intend to stimulate further community-wide discussions that mature into robust, widely adopted practices. Imaging MS methods are also advancing and reaching single-cell resolution[17], although they differ significantly from tandem MS methods and are outside the scope of this white paper. Our recommendations are topically grouped into "experimental design", "data evaluation and interpretation", and "reporting".

Experimental design

Best practices for single-cell MS proteomics can effectively build on established practices for bulk analysis[18,19]. 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 results interpretation. Similarly, randomization of biological and technical replicates and batches of reagents during sample processing (e.g., 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 the cells with minimal perturbations. This can be challenging if cells are adherent, require vigorous dissociation or isolation procedures, and especially when they must be isolated from native tissues. While generally proteins are more stable than mRNAs, most good practices utilized for single-cell RNA-seq isolation, such as quick sample processing at low temperature (4°C), are appropriate for proteomics as well. The timing and other parameters of the cell isolation procedure may be impactful and so should be recorded in order 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 mass spectrometry data and thus to strengthen biological interpretations.

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 carry-over peptides from previous single-cell or bulk runs that may persist within liquid-handling instrumental components, capillaries, and stationary phases, such

as needle washing solutions and column-retained analytes in liquid chromatography (LC) and reservoirs in capillary electrophoresis (CE). Typically, only about 1% of peptides persist on C18 column resin following a run, and they may appear in subsequent runs as carry-over “ghost” signal[20]. Fortunately, these carry-over peptides generally make a quantitatively insignificant contribution to consecutive samples of comparable amounts. However, when bulk samples are interspersed with single-cell runs, carry-over peptides from these bulk samples may significantly 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 carry-over, as previously described[5,21]. 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[7,21–23] or, in the case of airborne contaminants, by simple air-filtration devices (e.g., ABIRD)[5,24].

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 exposure to freeze-heat cycles[5,25] or nonionic surfactants[26,27]. Such clean lysis methods are preferable over MS-incompatible chemical treatments (e.g., sodium dodecyl sulfate or urea) that require loss-prone cleanup before MS analysis. 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. As such, low-loss sample preparations using microliter volumes offer broadly accessible options[14,25,28,29]. 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.

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[30,31]. As such, reducing the flow rate of separations from hundreds to tens of nanoliters per minute can increase measurement sensitivity, but these gains must be currently achieved with custom-packed narrow-bore columns and may compromise robustness and measurement throughput[16]. 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[32]. 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, e.g., capillary electrophoresis and ion mobility as well as multidimensional techniques may

potentially be employed as front-end approaches in MS-based single-cell proteomics[11,33,34]. 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.

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. Thus, benchmarking quantification accuracy requires positive controls, i.e., 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, proteins, or even proteomes in predefined ratios as performed for LFQbench experiments[35]. 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 of the cell types present and used to benchmark the in-silico averaged single-cell estimates. 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 and performing serial dilution to 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 background noise, especially in the case of highly variable cell-types within an experiment.

When matching between runs (BMR) 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 applying it to samples containing mixed species proteomes or single species proteomes and then estimating the number of incorrectly propagated species. Such MBR controls (samples of mixed yeast and bacteria 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.

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 results interpretation or sacrifice scientific rigor. To estimate and correct batch effects, treatments and analytical batches must be randomized whenever possible[36]. We recommend that treatment and batches are randomized, so that batch effects can be corrected (estimate and remove batch effects from data) or modeled (e.g. include batch effect as a covariate in models). When randomization is not performed, biological and technical factors may be fundamentally inseparable. For experiments wherein 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[37]. 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 the effect sizes, on the 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. In contrast, experimental designs including primary cells, smaller effect sizes (e.g., protein variability within a cell type[6]), multiple treatment groups or patient cohorts, require a much larger number of single cells to achieve adequate statistical power[38–40]. Thus, increasing the throughput of single-cell proteomics and incorporating power estimates in the experimental designs will become crucial as single-cell proteomics matures[41,42].

Method optimization. Depending on the priorities of the analysis, a method may seek to maximize the number of cells analyzed, the proteome depth achieved, or the number of copies sampled per protein. These objectives may best be achieved with different analytical workflows, multiplexed versus label-free analysis, different isobaric carrier levels, or by different parameter regimes, such as length of peptide separation, ion accumulation times, and intelligent data acquisition strategies[4,42–45]. Choosing optimal method parameters can be time consuming, and software for systematic, data driven optimization can speed up such optimizations[46].

1. **Summary: Experimental design recommendations**
    - 1.1. Include positive and negative controls that allow rigorous benchmarking
    - 1.2. Use strategies to maximize sample delivery while minimizing contamination
    - 1.3. Design experiments that minimize the influence of batch effects
        - 1.3.1. Decouple experimental batches from biological covariates, e.g., by randomization
        - 1.3.2. Minimize batch numbers and maximize batch sizes

Data evaluation and interpretation

Defining and evaluating reproducibility. We begin discussing data reproducibility and evaluation by briefly defining several levels of increasing difficulty, namely *repeating*, *reproducing*, and *replicating*[47]. Repeating a computational experiment or an analysis simply consists in 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 and environment. Replication represents a further challenge where the results are to be obtained using new code/implementation/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, e.g., xml.

The advent of containerised workflows now facilitates analysis replication without the need to go through the complicated process of setting up the exact same computational pipeline manually[48]. Still, many data analysis solutions, such as the R computational environment (R core team, 2021), are typically used in a non-containerised fashion, making the end-results potentially version-dependent and difficult to reproduce, especially in the long term. Thus, we recommend introducing workflow managers to facilitate the reproduction of single-cell proteomics data analysis. Of note, while the different reproducibility concepts are often described in the context of computational experiments and data analysis, they can also be extrapolated to experimental workflows. For example, detailed, dynamic, and version controlled protocols, such as those on the protocols.io platform, can facilitate experimental reproduction. Other containered options, such as Docker images, are a fruitful strategy for standardizing computational analysis of complex datasets, with cross-platform compatibility and robust version control.

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 single-cell proteomics, the quality of measurements depends on the number of ions measured from each single-cell population[43,44]. For example, if too few ions are sampled, the stochasticity of sampling results in counting noise, i.e., low precision estimates and technical variation in estimated protein abundances, which should be clearly distinguished from biological variability. Such counting noise also affects single-cell RNA-seq methods (that sample even fewer RNA copies per cell), and some of the models developed for RNA-seq data may help handle counting noise in MS data as well[49]. Mixing ratios of 1:1 can be used to evaluate ion sampling and precision but not accuracy since this ratio is not sensitive to systematic biases, such as co-isolation and interference. One option is to evaluate accuracy against ground truth ratios, as created by mixing the proteomes of different species in known ratios[7,35,50]. As another approach, measurements of relative protein abundance by established bulk methods can provide useful benchmarks for evaluating corresponding single-cell-level measurements[7,9,14]. On a smaller scale, accuracy may be estimated for a limited number of proteins by spiking corresponding peptides at known ratios or using measurements that are as independent as possible; such independent measurements include fluorescent proteins whose abundance is measured fluorometrically[1] or immunoassays with high specificity, such as proximity ligation assays that enhance specificity by using multiple affinity reagents per protein[51].

Comparisons between absolute protein intensities conflate the variance due to protein abundance variation across the compared samples (conditions) and across different proteins and may result in misleading impressions[52]. 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 (i) different peptide fragments from the same peptide (Fig. 2d) or (ii) 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 alternate proteoforms[53,54] 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; i.e. a fair comparison between experimental methods would rely on commonly identified peptides and proteins and the accuracy of the measurement thereof, rather than also including peptides and proteins identified uniquely in individual experiments.

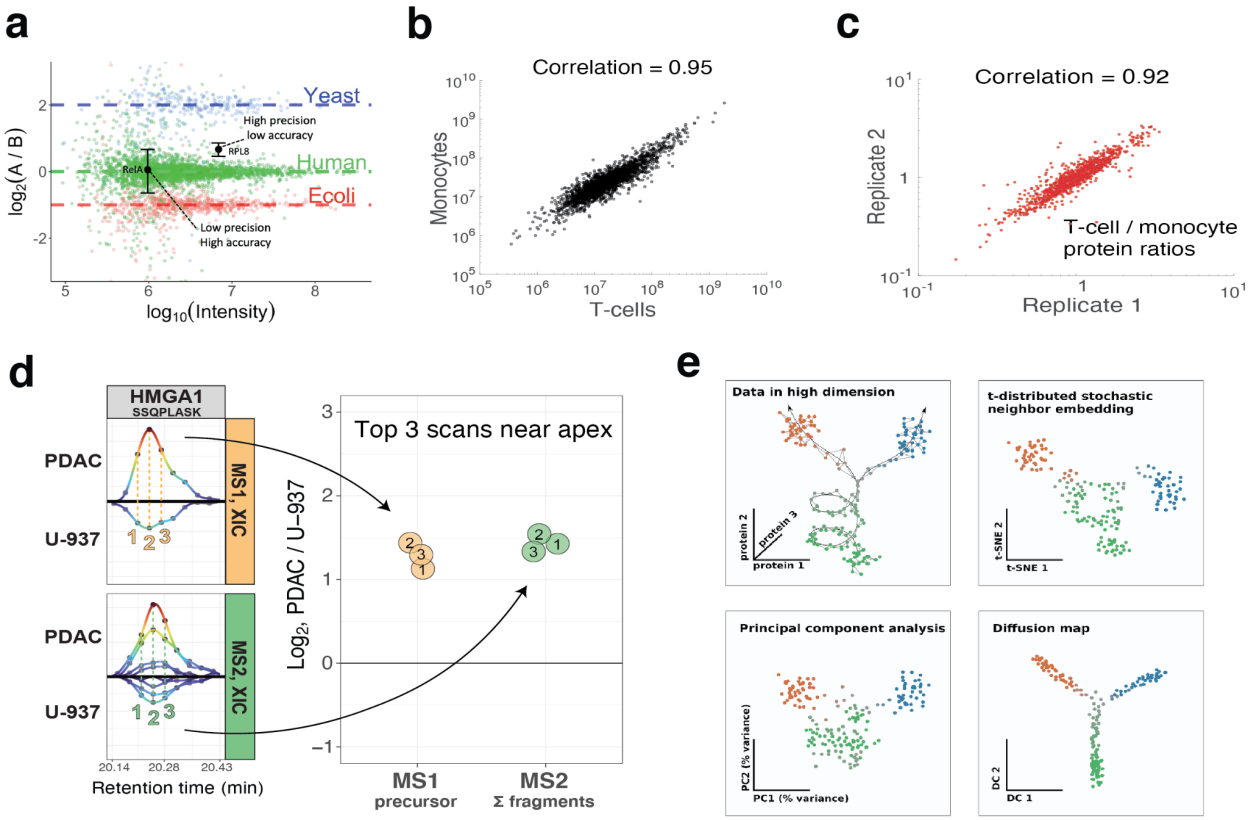

Figure 2 | Evaluating and interpreting single-cell proteomics data. a, Quantitative accuracy of protein ratios between samples A and B measured by label-free DIA analysis relative to the corresponding mixing ratios denoted by dotted lines[7]. Some proteins are quantified with high precision but low accuracy (e.g., RPL9), while others are quantified with high accuracy and low precision (e.g., RelA). The proteome 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 panels b and c are from Specht et al.[25] **d**, Extracted ion current from single-cell MS measurements by plexDIA. Such data allow quantifying peptides at both MS1 and MS2 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 MS1 and MS2 levels, 3 estimates are obtained based on the 3 scans closest to the elution peak apex. **e**, Different dimensionality reduction methods approximate the data in different ways. Cell states (represented by colors) are projected differently by different methods. PCA loses the non-linear cycling effect, tSNE does not correctly capture the distances between the 3 populations and diffusion maps do not capture the noise in the data.

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[14]. However, differences in total protein amounts between cells lead to differences in the numbers of missing values and proteins accurately quantified. Further bias can be introduced by the measurement noise affecting protein quantities at the lower abundance range. Thus, the processing of single-cell MS proteomic data is likely to be improved in the future with the development of more advanced normalization strategies, such as those developed for scRNA-seq experiments[55], which face a similar challenge. 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 FACS, 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[36,56]. These tend to be more prevalent in single-cell proteomics, compared to typical bulk experiments, due to differences in size and hence protein content between single cells. Another source of missing values is the fact that single-cell proteomics operates at the boundary of sensitivity of LC-MS instrumentation, and therefore a peptide quantified in some cells is below the detection limit in other cells. The missing data is 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, a common strategy for dealing with missing data is to impute missing values prior to any downstream analysis. Imputation must take into account the nature of missing data (e.g., missing at random or not at random[57]) in determining appropriate imputation methods. Comprehensive imputation methods for single-cell proteomics are yet to be developed and benchmarked, but recommendations developed for bulk proteomics methods may serve as useful guides[57–59]. While some recently developed methods for scRNA data may be adapted to proteomics, ultimately, the field needs methods which are specifically tailored to the mechanisms leading to missing peptides and proteins. Multiple imputation can be used to quantify the 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 PCA, on the imputed data and compare the results to the analysis performed on the unimputed data[13,14]. 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 the data, Fig. 2e. Nonetheless, the reduced data representations are at best incomplete approximations of the full data. 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[60]. Thus, we

recommend using dimensionality reduction only when it is essential. Conclusions derived from data reductions should be validated against the high-dimensional data.

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 driving the clustering. Some methods, such as PCA, are more amenable to interpretation, as opposed to their non-linear counterparts, such as t-distributed stochastic neighbor embedding (tSNE)[61] or uniform manifold approximation and projection (UMAP)[62]. With all methods, the projection of positive controls, such as bulk samples of purified cell types, on the low-dimensional manifold alongside single cells provides useful evidence for evaluating the clustering[13,14]. Thus, we recommend including such positive controls when possible. 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.

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. In addition, the separation between cells in low-dimensional manifolds is sensitive to many factors; even for the simplest dimensionality reduction methods, such as PCA, the separation depends on the number of proteins used in the analysis and assumptions about missing data. For nonlinear dimensionality reduction methods like tSNE and UMAP, the separation between cell types is sensitive to various tuning parameters. Furthermore, only the small distances in tSNE and UMAP manifolds are interpretable; the large distances are arbitrarily rendered for visualization purposes. All of these factors caution against basing conclusions solely on low-dimensional tSNE and UMAP manifolds. Therefore, we recommend performing data analysis in the high-dimensional space of all quantified proteins. When results are based on a low-dimensional manifold, such as cluster identification, the corresponding distances in the high-dimensional space should also be shown, e.g., as distributions of pairwise distances between single cells within and across clusters[63].

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[14,36]. 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[7,14,36]. When thresholds are set based on subjective choices, this should be explicitly stated and the choices treated as a source of uncertainty in the final results. When possible, the sensitivity of the results to all experimental and methodological choices should clearly be conveyed.

2. **Summary: Data evaluation and reporting recommendations**
    2.1. Explicitly distinguish between repeatability, reproducibility, and accuracy
    2.2. Use dimensionality reduction sparingly and validate its results
    2.3. The results must reflect the uncertainty in the data, methods, and parameters.

Reporting standards

The goal of reporting is to enable other researchers to repeat, reproduce, assess, and build upon published data and their interpretation[64]. 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 the metadata, and a good starting place for reporting metadata are formats developed for bulk MS data[18,65], including specifically for proteomics data[66], those prepared by journals[67–70] and societies[71], as well as for single-cell RNA-seq data[72]. Nonetheless, single-cell MS proteomics data have additional aspects that need to be reported, which are the focus of our recommendations. Below, we document the essential information needed to provide value to single-cell proteomics data, meta-data, and analysis results.

Experimental design. It is essential to provide readers with the detailed design of the experiment, which includes treatment groups, number of single cells per group, sampling, and analysis batches, Fig. 3. The experimental design should 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, i.e. variables that describe the biology of the measured cells and technical factors that are likely to influence the measurements. Mandatory biological descriptors contain sample type (such as single cell, carrier, empty, or control sample) and biological group, such as treatment condition or patient/donor identifier, cell line, organism and organ/part of origin (if cells from multiple organisms or multiple organs are assayed) and biological characteristics for multi-sample and/or multi-condition studies. Additional biological descriptors may include the cell type and/or cell state (e.g., their spatial and temporal information in tissues), physical markers (e.g., pigmentation, measured by FACS if available), 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 the acquisition of the data (such as cell types based on different cell cultures or following from FACS sorting) or be the results of downstream analyses (such as cell types or cell states inferred from clustering or differential abundance analysis). Mandatory technical descriptors include the raw data filenames and acquisition dates, as well as variables that describe the underlying technical variability, whether it is expected to be significant compared to the biological variability or not.

These descriptors include all batch factors related to cell isolation, sample preparation, peptide and protein separation (chromatography or electrophoresis batches), operator(s) and instruments (when multiple mass spectrometers were used), as well as chemical mass tags/labels (in case of labeled quantitation, e.g., TMTpro). Such a sample metadata table is also a simple and valuable quality control tool since it allows for 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 is compatible with the files in the data repositories (see below). We encourage researchers to document additional descriptors where 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. We include a standard README file and recommend using it to facilitate standardization and data reuse. The README file should contain a summary of the study design and the protocols. The measurement units of some of the 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.

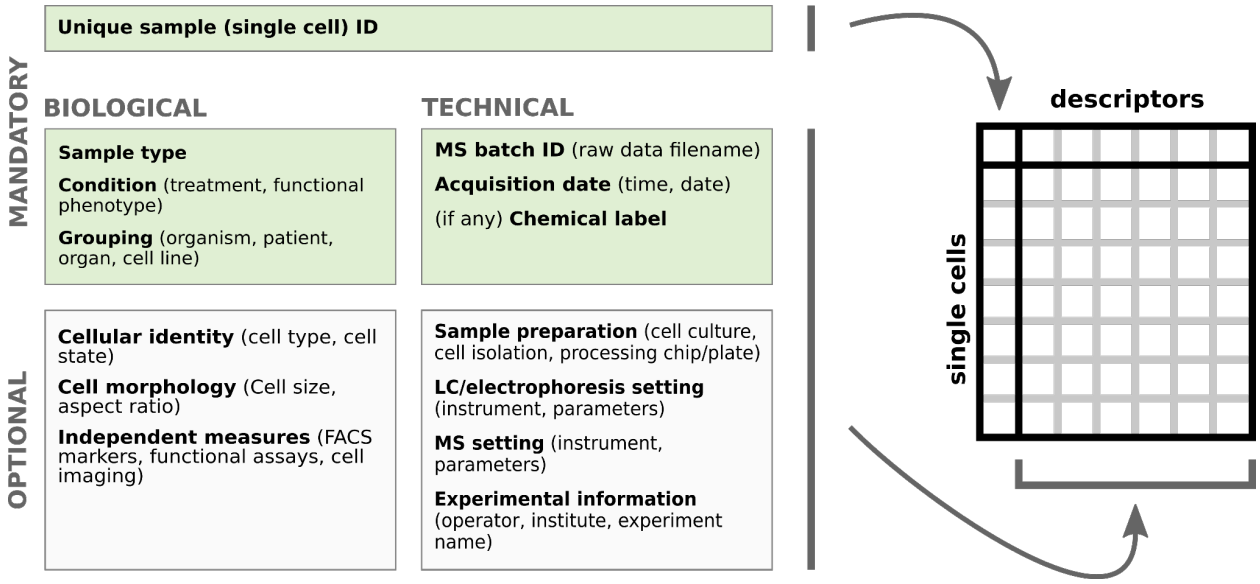

Figure 3 | Suggested descriptors of single-cell proteomics samples. The metadata should include the experimental design table with rows corresponding to single cells and columns to the mandatory and optional features listed here (an example is provided as Supporting File 2). Attributes provided in parenthesis are given as examples or for clarification. The green frames highlight mandatory descriptors, while the gray ones include a non-exhaustive list of optional descriptors, which may also include spatial (e.g., 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.

Ideally, the raw and processed MS data should be shared using open formats, such as HUPO Proteomics Standards Initiative community-developed formats dedicated to mass spectrometry data: mzML[73] for raw data, mzIdentML[74] for search results, and mzTab[75] or text-based spreadsheet for quantitative data. Solely providing binary formats from proprietary software should be avoided at all costs. Even when such software is used, un-processed and final result

files should be exported and shared with the community. Raw data files and search results should be made available through dedicated repositories, such as PRIDE[66] and MassIVE[76]. Code repositories, such as GitLab or GitHub[77], 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 (an example is provided as Supporting File 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 the experiment identifiers and links to all the external data and metadata resources. Editors and reviewers should systematically require the deposition of all data, metadata, and analysis details as a condition for paper acceptance and publication.

While these data sharing requirements apply broadly to proteomics 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[42]. 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 requirements 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 peptides x cells matrices, to reduce the computational burden on reproducing individual steps from the analysis.

BOX: Despite its apparent simplicity, file naming deserves thoughtful consideration. Files names should be unique (unlikely to be used in many other studies) and linked to the measurements in the file; additional good practices are summarized in ref.[78]. We suggest thinking about file naming and file naming conventions to easily identify groups of files pertaining to specific meta-data 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 filenames need to be documented in the main README file.

While sharing the data is necessary, it is by no means sufficient. It is well known that the processing of data has an important effect on 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 need to be provided with the data. Using software for single-cell proteomics, such as the *scp* R/Bioconductor package[36,79], the *sceptre* python package[9],or the SCoPE2 pipeline[80],

can help standardize workflows across laboratories. Packages that allow comparing structured and repeatable data processing, including evaluating different algorithms for a processing step, provide further advantages[36,79]. When using GUI-based software platforms that do not support exporting the commands that were executed as scripts, 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. Given the rapid evolution of the field, simply referring to other publications using 'as previously analyzed in [ref]' should also be avoided. When reporting results, it should be made clear which data the result refers to. This is, for example, crucial when reporting coefficients of variation (CVs), where CVs on log-transformed data are lower than 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 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. It is crucial to go beyond the minimalist 'Material and method' sections that systematically 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 subjects. A single dump or all files can't be considered an acceptable data sharing policy. In such situations, it is advisable to split these 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 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 proteomics experiments. While the reporting of MS acquisition details is not necessarily required for reanalysis of the data, 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 where 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, but 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 labs trying to replicate published results on their own instrumentation.

These reporting guidelines might give the impression that a lot of additional work is expected when reporting on studies according to our recommendation. They should not be regarded as a burden to be addressed at the end of the research project. Indeed, our 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. Data, meta-data, and analyses documentation and reporting happen at different stages of the analysis process, and rely on each other. The investment that we are arguing for here is simply work that is spread across the research project, rather than extra work done at the very end of it[81].

3. **Summary: Reporting standards**
    - 3.1. Use and document descriptors as part of the experimental design table to summarize how and why the experiments were conducted.
    - 3.2. Share all data, metadata, and processed data in recommended repositories to accurately describe what was generated as part of an experiment.
    - 3.3. Share data processing and analysis notebooks or scripts that are sufficient for repeating the complete analyses presented in a manuscript.
    - 3.4. Include a README file that describes all the data, metadata, and scripts and cross-reference the remote resources that store all the experiment's outputs.

Conclusions and perspectives

The adoption of these guidelines by the scientific community and their promotion by journals and data archives is essential for establishing solid foundations for the emerging field of single-cell MS proteomics and to uphold scientific rigor. 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 proteomics data and results outside of the labs 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 proteomics technologies[82], 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, PTM-level and proteoform-level (i.e., top-down) single-cell proteomics 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: single-cell.net/guidelines

Acknowledgements

We thank the numerous contributors to these initial recommendations and the community as a whole for the body of work that supports our recommendations. We thank R. Gray Huffman for feedback and detailed edits. The guidelines in this article were formulated in large part during

the workshops and through the discussions of the annual Single-Cell Proteomics Conference (https://single-cell.net).

Competing interests

The authors declare no competing interests.

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).
2. Zhu, Y. *et al.* Nanodroplet processing platform for deep and quantitative proteome profiling of 10–100 mammalian cells. *Nat. Commun.* **9**, 1–10 (2018).
3. Singh, A. Towards resolving proteomes in single cells. *Nature methods* vol. 18 856 (2021).
4. Kelly, R. T. Single-cell Proteomics: Progress and Prospects. *Mol. Cell. Proteomics* **19**, 1739–1748 (2020).
5. Petelski, A. A. *et al.* Multiplexed single-cell proteomics using SCoPE2. *Nat. Protoc.* **16**, 5398–5425 (2021).
6. Leduc, A., Gray Huffman, R., Cantlon, J., Khan, S. & Slavov, N. Exploring functional protein covariation across single cells using nPOP. *bioRxiv* 2021.04.24.441211 (2022) doi:10.1101/2021.04.24.441211.
7. Derks, J. *et al.* Increasing the throughput of sensitive proteomics by plexDIA. *Nat. Biotechnol.* (2022) doi:10.1038/s41587-022-01389-w.
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). *Angewandte Chemie* vol. 133 12962–12968 (2021).

9. Schoof, E. M. *et al.* Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nature Communications* vol. 12 (2021).
10. Brunner, A.-D. *et al.* Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *Mol. Syst. Biol.* **18**, e10798 (2022).
11. Choi, S. B., Polter, A. M. & Nemes, P. Patch-Clamp Proteomics of Single Neurons in Tissue Using Electrophysiology and Subcellular Capillary Electrophoresis Mass Spectrometry. *Anal. Chem.* **94**, 1637–1644 (2022).
12. Slavov, N. Unpicking the proteome in single cells. *Science* **367**, 512–513 (2020).
13. Gray Huffman, R. *et al.* Prioritized single-cell proteomics reveals molecular and functional polarization across primary macrophages. *bioRxiv* 2022.03.16.484655 (2022) doi:10.1101/2022.03.16.484655.
14. Specht, H. *et al.* Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome Biol.* **22**, 50 (2021).
15. Slavov, N. Single-cell protein analysis by mass spectrometry. *Curr. Opin. Chem. Biol.* **60**, 1–9 (2021).
16. 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).
17. Soltwisch, J. *et al.* Mass spectrometry imaging with laser-induced postionization. *Science* **348**, 211–215 (2015).
18. Taylor, C. F. *et al.* The minimum information about a proteomics experiment (MIAPE).

Nat. Biotechnol. **25**, 887–893 (2007).

19. Donnelly, D. P. *et al.* Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nat. Methods* **16**, 587–594 (2019).
20. 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).
21. Cong, Y. *et al.* Ultrasensitive single-cell proteomics workflow identifies> 1000 protein groups per mammalian cell. *Chem. Sci.* **12**, 1001–1006 (2021).
22. Fernandez-Lima, F., Kaplan, D. A., Suetering, J. & Park, M. A. Gas-phase separation using a trapped ion mobility spectrometer. *Int. J. Ion Mobil. Spectrom.* **14**, (2011).
23. Orsburn, B. C., Yuan, Y. & Bumpus, N. N. Single Cell Proteomics Using a Trapped Ion Mobility Time-of-Flight Mass Spectrometer Provides Insight into the Post-translational Modification Landscape of Individual Human Cells. *bioRxiv* 2022.02.12.480144 (2022) doi:10.1101/2022.02.12.480144.
24. Marie, A.-L. *et al.* High-Sensitivity Glycan Profiling of Blood-Derived Immunoglobulin G, Plasma, and Extracellular Vesicle Isolates with Capillary Zone Electrophoresis-Mass Spectrometry. *Anal. Chem.* **93**, 1991–2002 (2021).
25. Specht, H., Harmange, G., Perlman, D. H. & Emmott, E. Automated sample preparation for high-throughput single-cell proteomics. *BioRxiv* (2018).
26. 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).
27. Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of

intact membrane protein complexes. *Nat. Protoc.* **8**, 639–651 (2013).

28. Liang, Y. *et al.* Fully Automated Sample Processing and Analysis Workflow for Low-Input Proteome Profiling. *Anal. Chem.* **93**, 1658–1666 (2021).
29. Brunner, A. D., Thielert, M., Vasilopoulou, C. G. & Ammar, C. Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *BioRxiv* (2021).
30. Cole, R. B. Some tenets pertaining to electrospray ionization mass spectrometry. *J. Mass Spectrom.* **35**, 763–772 (2000).
31. 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).
32. Köcher, T., Swart, R. & Mechtler, K. Ultra-high-pressure RPLC hyphenated to an LTQ-Orbitrap Velos reveals a linear relation between peak capacity and number of identified peptides. *Anal. Chem.* **83**, 2699–2704 (2011).
33. 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).
34. K. DeLaney, D. Jia, L. Iyer, Z. Yu, S. B. Choi, P. J. Marvar, and P. Nemes. Microanalysis of angiotensin peptides in the brain using ultrasensitive capillary electrophoresis trapped ion mobility mass spectrometry. *Anal. Chem.*
35. Navarro, P. *et al.* A multicenter study benchmarks software tools for label-free proteome quantification. *Nat. Biotechnol.* **34**, 1130–1136 (2016).
36. Vanderaa, C. & Gatto, L. Replication of single-cell proteomics data reveals important

computational challenges. *Expert Rev. Proteomics* 1–9 (2021).

37. Čuklina, J. *et al.* Diagnostics and correction of batch effects in large-scale proteomic studies: a tutorial. *Mol. Syst. Biol.* **17**, e10240 (2021).
38. Pino, L. K. *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).
39. Davis, A., Gao, R. & Navin, N. E. SCOPIT: sample size calculations for single-cell sequencing experiments. *BMC Bioinformatics* **20**, 566 (2019).
40. Slavov, N. Increasing proteomics throughput. *Nature biotechnology* vol. 39 809–810 (2021).
41. Boekweg, H., Guise, A. J., Plowey, E. D., Kelly, R. T. & Payne, S. H. Calculating Sample Size Requirements for Temporal Dynamics in Single-Cell Proteomics. *Mol. Cell. Proteomics* **20**, 100085 (2021).
42. Slavov, N. Scaling Up Single-Cell Proteomics. *Mol. Cell. Proteomics* **21**, 100179 (2022).
43. Specht, H. & Slavov, N. Optimizing Accuracy and Depth of Protein Quantification in Experiments Using Isobaric Carriers. *J. Proteome Res.* **20**, 880–887 (2021).
44. Cheung, T. K. *et al.* Defining the carrier proteome limit for single-cell proteomics. *Nat. Methods* **18**, 76–83 (2021).
45. Furtwängler, B. *et al.* Real-Time Search Assisted Acquisition on a Tribrid Mass Spectrometer Improves Coverage in Multiplexed Single-Cell Proteomics. *Mol. Cell. Proteomics* **0**, (2022).
46. Huffman, R. G., Chen, A., Specht, H. & Slavov, N. DO-MS: Data-Driven Optimization of Mass Spectrometry Methods. *J. Proteome Res.* **18**, 2493–2500 (2019).

47. National Academies of Sciences, Engineering *et al. Understanding Reproducibility and Replicability*. (National Academies Press (US), 2019).

48. Wratten, L., Wilm, A. & Göke, J. Reproducible, scalable, and shareable analysis pipelines with bioinformatics workflow managers. *Nat. Methods* **18**, 1161–1168 (2021).

49. Brennecke, P. *et al.* Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* **10**, 1093–1095 (2013).

50. Muntel, J. *et al.* Comparison of Protein Quantification in a Complex Background by DIA and TMT Workflows with Fixed Instrument Time. *J. Proteome Res.* **18**, 1340–1351 (2019).

51. Darmanis, S. *et al.* Simultaneous Multiplexed Measurement of RNA and Proteins in Single Cells. *Cell Rep.* **14**, 380–389 (2016).

52. Franks, A., Airoldi, E. & Slavov, N. Post-transcriptional regulation across human tissues. *PLoS Comput. Biol.* **13**, e1005535 (2017).

53. Plubell, D. L. *et al.* Putting Humpty Dumpty Back Together Again: What Does Protein Quantification Mean in Bottom-Up Proteomics? *J. Proteome Res.* (2022) doi:10.1021/acs.jproteome.1c00894.

54. Malioutov, D. *et al.* Quantifying Homologous Proteins and Proteoforms. *Mol. Cell. Proteomics* **18**, 162–168 (2019).

55. Lytal, N., Ran, D. & An, L. Normalization Methods on Single-Cell RNA-seq Data: An Empirical Survey. *Front. Genet.* **11**, 41 (2020).

56. Hicks, S. C., Townes, F. W., Teng, M. & Irizarry, R. A. Missing data and technical variability in single-cell RNA-sequencing experiments. *Biostatistics* **19**, 562–578 (2018).

57. 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).

58. Dabke, K., Kreimer, S., Jones, M. R. & Parker, S. J. A Simple Optimization Workflow to Enable Precise and Accurate Imputation of Missing Values in Proteomic Data Sets. *J. Proteome Res.* **20**, 3214–3229 (2021).
59. Bramer, L. M., Irvahn, J., Piehowski, P. D., Rodland, K. D. & Webb-Robertson, B.-J. M. A Review of Imputation Strategies for Isobaric Labeling-Based Shotgun Proteomics. *J. Proteome Res.* **20**, 1–13 (2021).
60. Chari, T., Banerjee, J. & Pachter, L. The Specious Art of Single-Cell Genomics. *bioRxiv* 2021.08.25.457696 (2021) doi:10.1101/2021.08.25.457696.
61. van der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
62. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv [stat.ML]* (2018).
63. Specht, H., Emmott, E., Koller, T. & Slavov, N. High-throughput single-cell proteomics quantifies the emergence of macrophage heterogeneity. *bioRxiv 10.1101/665307* (2019) doi:10.1101/665307.
64. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016).
65. Dai, C. *et al.* A proteomics sample metadata representation for multiomics integration and big data analysis. *Nat. Commun.* **12**, 5854 (2021).
66. 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).
67. 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).

68. Abbatiello, S. *et al.* New Guidelines for Publication of Manuscripts Describing Development and Application of Targeted Mass Spectrometry Measurements of Peptides and Proteins. *Mol. Cell. Proteomics* **16**, 327–328 (2017).
69. Chalkley, R. J., MacCoss, M. J., Jaffe, J. D. & Röst, H. L. Initial Guidelines for Manuscripts Employing Data-independent Acquisition Mass Spectrometry for Proteomic Analysis. *Mol. Cell. Proteomics* **18**, 1–2 (2019).
70. Required Manuscript Content and Publication Guidelines: Molecular & Cellular Proteomics. https://www.mcponline.org/mass-spec-guidelines.
71. Omenn, G. S. Reflections on the HUPO Human Proteome Project, the Flagship Project of the Human Proteome Organization, at 10 Years. *Mol. Cell. Proteomics* **20**, 100062 (2021).
72. Füllgrabe, A. *et al.* Guidelines for reporting single-cell RNA-seq experiments. *Nat. Biotechnol.* **38**, 1384–1386 (2020).
73. Martens, L. *et al.* mzML--a community standard for mass spectrometry data. *Mol. Cell. Proteomics* **10**, R110.000133 (2011).
74. Vizcaíno, J. A. *et al.* The mzIdentML Data Standard Version 1.2, Supporting Advances in Proteome Informatics. *Mol. Cell. Proteomics* **16**, 1275–1285 (2017).
75. 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).
76. Wang, M. *et al.* Assembling the Community-Scale Discoverable Human Proteome. *Cell Syst* **7**, 412–421.e5 (2018).

77. Perez-Riverol, Y. *et al.* Ten Simple Rules for Taking Advantage of Git and GitHub. *PLoS Comput. Biol.* **12**, e1004947 (2016).

78. File naming conventions. https://datamanagement.hms.harvard.edu/collect/file-naming-conventions.

79. Vanderaa, C. & Gatto, L. scp: Mass Spectrometry-Based Single-Cell Proteomics Data Analysis. *Bioconductor* (2020) doi:10.18129/B9.bioc.scp.

80. Specht, H. *et al.* SCoPE2 processing pipeline. *Zenodo* (2020) doi:10.5281/zenodo.4339954.

81. Quintana, D. Five things about open and reproducible science that every early career researcher should know. (2020) doi:10.17605/OSF.IO/DZTVQ.

82. Slavov, N. Driving Single Cell Proteomics Forward with Innovation. *J. Proteome Res.* **20**, 4915–4918 (2021).