Published in final edited form as:

Anal Chem. 2019 October 15; 91(20): 13119–13127. doi:10.1021/acs.analchem.9b03349.

High-Throughput Single Cell Proteomics Enabled by Multiplex Isobaric Labelling in a Nanodroplet Sample Preparation Platform

Maowei Dou^{1,†}, Jeremy Clair^{2,†}, Chia-Feng Tsai², Kerui Xu¹, William B. Chrisler², Ryan L. Sontag², Rui Zhao¹, Ronald J. Moore², Tao Liu², Ljiljana Pasa-Tolic¹, Richard D. Smith², Tujin Shi², Joshua N. Adkins², Wei-Jun Qian², Ryan T. Kelly^{1,3}, Charles Ansong^{2,*}, Ying Zhu^{1,*}

¹Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99354, United States

²Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99354, United States

³Department of Chemistry and Biochemistry, Brigham Young University, Provo, Utah 84604, United States

Abstract

Effective extension of mass spectrometry-based proteomics to single cells remains challenging. Herein we combined microfluidic nanodroplet technology with tandem mass tag (TMT) isobaric labeling to significantly improve analysis throughput and proteome coverage for single mammalian cells. Isobaric labeling facilitated multiplex analysis of single cell-sized protein quantities to a depth of ~1,600 proteins with median CV of 10.9% and correlation coefficient of 0.98. To demonstrate in-depth high throughput single cell analysis, the platform was applied to measure protein expression in 72 single cells from three murine cell populations (epithelial, immune, and endothelial cells) in <2 days instrument time with over 2,300 proteins identified. Principal component analysis grouped the single cells into three distinct populations based on protein expression with each population characterized by well-known cell-type specific markers. Our platform enables high throughput and unbiased characterization of single cell heterogeneity at the proteome level.

*Corresponding authors: Dr. Ying Zhu, ying.zhu@pnnl.gov, Dr. Charles Ansong, charles.ansong@pnnl.gov.

†Authors contributed equally to this work.

Publisher's Disclaimer: "Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.

Supporting Information

The Supporting Information is available free of charge on the [ACS Publications website](#).

Keywords

Mass spectrometry; microfluidics; tandem mass tag; proteomics; single cell

Introduction

Multicellular organisms contain diverse cell types and subtypes with distinct functions. Understanding the heterogeneity or cell subpopulations at single cell level is of great interest for biomedical research.^{1, 2} Single cell analysis provides potential insights for understanding unique features of cell populations, cellular lineage, function, differentiation, impacts of microenvironments, and rare cell types in a complex cellular system.^{3, 4} Single cell RNA-seq assays have been widely employed to characterize gene expression patterns.⁵ However, RNA measurements provide an incomplete picture of the molecular state of a cell as proteins mediate the bulk of cellular function. A compelling body of literature^{6–10} has shown the correlation between RNA and protein abundance is at best moderate. Profiling the proteome (entire complement of the proteins expressed within a cell at a given time) at single cell level is recognized as a substantial technical challenge. As there is no whole proteome amplification method available, measurements largely rely on the effectiveness of the workflow, as well as the overall sensitivity of instrumental platform. Currently, protein profiling in single cells mainly utilizes targeted, antibody-based approaches,^{11, 12} such as Mass Cytometry (CyTOF)¹³, Proseek Multiplex¹⁴, and single-cell western blots¹⁵. However, these approaches share common shortcomings in that only a limited number of proteins (less than 100 targeted proteins) can be analyzed simultaneously, and the quantification accuracy is largely determined by the specificity of available antibodies.

Mass spectrometry (MS)-based proteomic analyses are capable of quantifying thousands of proteins in an unbiased manner.¹⁶ However, due to insufficient sensitivity, large sample losses during preparation, and low throughput, the extension of MS-based proteomics to single cell studies remains challenging.^{11, 17} While MS sensitivity has significantly increased with the development of nanoflow separations, electrospray ionization, and more sensitive mass spectrometers,¹⁷ approaches to efficiently process single cells and deliver the sample to MS with minimal sample loss have been largely ineffective. Great effort has been put to improve sample recovery including the use of low-binding tubes, MS-compatible surfactants, immobilized enzyme reactor (IMER)-based approaches,¹⁸ and the development of single-pot workflows such as e.g. the Oil-Air-Droplet chip, and on-valve high temperature trypsin digestion.^{17, 19–21}

We recently developed a microfluidic nanodroplet sample preparation approach, termed nanoPOTS (Nanodroplet Processing in One-pot for Trace Samples), that enables single cell proteomic sample processing and analysis using ultrasensitive nanoLC-MS/MS.^{22–24} NanoPOTS reduces the sample processing volume to less than 200 nanoliters (nL) and the total exposure surfaces to $\sim 1 \text{ mm}^2$, thus significantly limiting surface losses during sample processing, which is critical for single cell proteomics.^{22, 25} Using the nanoPOTS platform, an average of 670 protein groups were confidently identified from single HeLa cells²³ and 160 proteins from a single spiked circulating tumor cells isolated from whole blood²⁶. While

the current nanoPOTS workflow yielded good single cell proteome coverage, the analytical throughput was relatively low at ~8 single cells per day. Higher throughput is needed to facilitate efficient analysis of large populations of single cells to provide insights into cellular heterogeneity.

Tandem mass tag (TMT) isobaric labeling is an approach developed for multiplexed identification and quantification of proteins from multiple different samples in a single LC-MS analysis/run.²⁷ Individual samples are first labeled with discrete TMT reagents (also referred to as TMT channels) and then combined into a multiplexed set prior to analysis by high-resolution LC-MS. The proteins in each sample can be identified and quantified based on the TMT reporter ions. Up to 11 samples can be simultaneously measured in single LC-MS runs using TMT 11plex, providing an attractive potential approach/strategy to significantly enhance the throughput of single cell proteomics platform. This concept was initially demonstrated in the proteomic study of single blastomeres isolated from *Xenopus laevis* embryos, where three blastomeres were quantified in single MS run.²⁸ Recently, several TMT-based approaches have been developed that facilitate quantification of low abundance proteins. Russell et al. developed the TMTcalibrator™ approach labeling diseased tissue in parallel with protein from body fluids.^{29–31} The high-level proteins in tissues facilitated the triggering of MS/MS fragmentation to enable quantifying of low-level proteins in body fluids. Budnik *et al.*³² developed a single cell analysis method (SCoPE-MS) by implementing isobaric labeling of single cells in concert with 200 cells as carrier. Tian *et al.*³³ developed a deep proteomic analysis approach to discover single amino acid variants in a small numbers of cancer cells. Focusing on post-translational modification analysis, Yi *et al.*³⁴ developed a boosting to amplify signal with isobaric labeling (BASIL) strategy and applied it to study phosphoproteomic dynamics in mass-limited clinical specimens. Common to all these studies, a “boosting” sample containing much larger amount (30–500 fold) of peptide mass was labeled with one TMT channel and the mass-limited study samples were labeled with the remaining TMT channels of the multiplex set. The more abundant boosting peptides provided rich fragment information for peptide identification, while reporter ions provided quantification information for study samples. While these TMT-based boosting methods establish the feasibility of TMT labeling for proteomics analyses of mass-limited samples down to single cells, they uniformly suffer from the drawbacks of minute samples being processed conventionally. Specifically, the attendant sample losses due to non-specific adsorption/adhesion on surfaces.

To address this challenge, we combined the nanoPOTS approach with TMT isobaric labeling method, which was previously used in SCoPE-MS workflow,³² to improve both proteomic sample processing efficiency and analysis throughput for single cells. Single cells were isolated by flow cytometry, processed in nanowells, labeled with TMT 10-plex, and finally combined for LC/MS analysis. We systematically investigated the labeling reproducibility and the effect of boosting ratios on single cell quantification. The optimized system was applied to measure protein expression in 72 single cells from three cultured murine cell populations (epithelial, immune, and endothelial cells). We demonstrated our single cell proteomics approach provides deep proteome coverage for single cell analysis, can discriminate mammalian cell types and identify cell-specific protein markers.

Experimental Section

Single cell isolation

We employed the BD Influx II cell sorter (BD Biosciences, San Jose, CA) to directly isolate single cells into nanowells.²³ A customized matrix was generated in the cell sorter control software to match the nanowell array design. Fluorescent beads were used to optimize sorting parameters and to confirm successful deposition in each well. To select only viable cells, murine cells were labeled with a membrane-permeable live-cell labeling dye (100 nM, Calcein AM, eBioscience, Thermo Fisher) for 30 min in the dark. After cell collection, the nanowell chips were stored at -80°C or directly submitted for proteomic sample preparation.

Proteomic sample preparation in nanodroplets

The nanoPOTS-based proteomic sample preparation procedures were described previously.²² A home-built robotic liquid handling system was employed to dispense nanoliter reagents into nanowells for proteomic processing. The nanowell chip was placed in a closed chamber that was maintained at 95% relative humidity during dispensing procedures to minimize liquid evaporation in nanowells.

The nanoPOTS-based sample preparation was carried out as the following steps. (1) Cell lysis, protein extraction, and disulfide reduction: 100 nL solution containing 0.2% DDM and 5 mM TCEP in 100 mM TEAB was added into each of the nanowells and the nanowell chip was incubated at 70°C for 30 min. (2) Alkylation of sulfhydryl groups: 50 nL 30 mM IAA in 100 mM TEAB was dispensed into each nanowell and the nanowell chip was incubated at room temperature in dark for 30 min. (3) Lys-C digestion: 50 nL Lys-C solution containing 0.5 ng Lys-C in 100 mM TEAB was added into each nanowell and the nanowell chip was incubated at 37°C for 4 h. (4) Tryptic digestion: 50 nL trypsin solution containing 0.5 ng trypsin in 100 mM TEAB was added into each nanowell and the nanowell chip was incubated at 37°C for 6 h. (5) TMT labeling: each of the 100 nL dissolved TMT reagents (0.8 mg in 41 μL acetonitrile) was added into the corresponding nanowells; for the boosting channel well where relatively large amount of peptides mixture was preloaded, 300 nL of the dissolved TMT-131 (0.8 mg in 41 μL acetonitrile) was added; the nanowell chip was incubated at room temperature for 1 h. (6) TMT quenching: 50 nL and 300 nL of 2% hydroxylamine solution was added into single cell channel and boosting channel wells respectively, and the nanowell chip was incubated at room temperature for 15 min. (7) Acidification: 50 nL 5% formic acid solution was added into nanowells and the nanowell chip was incubated at room temperature for 15 min. (8) Sample collection: the TMT labeled peptides from multiple nanowells were collected and combined into the booster well and mixed with boosting peptides. We also performed one round of nanowell wash to increase the recovery and the wash solution was also combined into the booster well. The combined samples were stored in a section of fused silica capillary (200 μm i.d., 360 μm o.d.) with both ends sealed with Parafilm. The capillary section was stored at -20°C prior to nanoLC-MS/MS analysis.

Nanoflow LC-MS/MS analysis for single cell samples

Both SPE precolumn (100 μm i.d., 360 μm o.d., 4 cm long) and LC column (30 μm i.d., 360 o.d., 50 cm long) were slurry-packed in house with 3- μm C18 packing material (300-Å pore size, Phenomenex, Terrence, CA).³⁵

The sample storage capillary was connected to the SPE precolumn using a PEEK union (Valco instruments, Houston, TX). Sample was loaded and cleaned up in the SPE precolumn using a nanoACQUITY UPLC pump (Waters, Milford, CT) by infusing Buffer A (0.1% formic acid in water) at a flow rate of 1000 nL/min for 10 min. The SPE precolumn was then reconnected to the LC column using a low-dead-volume PEEK union (Valco instruments, Houston, TX) for LC separation. The flow rate of the LC separation was 50 nL/min with the split from 300 nL/min using a nanoUPLC pump (Dionex UltiMate NCP-3200RS, Thermo Scientific, Waltham, MA). A linear 100-min gradient of 8–30% Buffer B (0.1% formic acid in acetonitrile) was used for the LC separation. Then the LC column was washed by ramping Buffer B to 45% in 15 min and 90% in 5 min, and finally re-equilibrated with 2% Buffer B for another 10 min.

An Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher) was employed for all data acquisition under automated data dependent acquisition mode. The separated peptides were ionized at a spray voltage of 2 kV and the ions were collected into an ion transfer capillary set at 150 °C. The RF lens was set at 30%. The MS1 scan was set at a mass range from 375 to 1575, a scan resolution of 120 k, an AGC target of 1E6, and a maximum injection time of 50 ms. Precursor ions with charges of +2 to +7 and intensities >20,000 were selected for MS/MS sequencing. Precursor ions were isolated with an m/z window of 0.7 Th and fragmented by high energy dissociation (HCD) set at 35%. Repeat sampling were reduced with an exclusion duration of 60 s and m/z tolerance of ± 10 ppm. MS/MS scan was carried out in the Orbitrap with an AGC target of 1E5. The MS/MS scan resolutions and maximum injection times were set as 60 k and 246 ms.

Data analysis of single cell samples

The spectral identifications were performed using MS-GF+³⁶ on the Thermo RAW files. Searches were performed in semi-tryptic mode with a tolerance set at 20 ppm, Methionine oxidation was set as variable modification (+15.9949 Da), TMT 10 plex (+229.1629 Da) and cysteine carbamidomethylation (+57.0215 Da) were set as fixed modifications. The maximum peptide length was set to be comprised at a minimum of 6 and a maximum of 50 residues. Only ions with precursor charge comprised between 2 and 5 were considered for the search. For the HeLa digest, the human UniProt KB database (fasta file downloaded 2017-04-12, 20,198 sequences) was used for the search; for all the other experiments, the murine UniProt KB database (fasta file downloaded 2017-04-12, 16,865 sequences) was used for the search, in both case a self-assembled list of human, bovine and porcine usual contaminants (16 sequences) was added to the database searched. MS-GF+ creates its own decoy database from the searched fasta files to enable to estimate the False Discovery Rates (FDR). The TMT10plex reporter intensities were extracted using MASIC³⁷ v3.0.7111 (<https://github.com/PNNL-Comp-Mass-Spec/MASIC/releases/tag/v3.0.7111>) with a reporter

tolerance of 0.003 Da. Internally, all the data was managed through the PRISM Data Management System.³⁸

The subsequent data processing was performed in R v3.5.1 using the package RomicsProcessor v0.1.0, which is available on Github (<https://github.com/PNNL-Comp-Mass-Spec/RomicsProcessor>) and Zenodo (<https://zenodo.org/record/3386527>). This package that comprise several modules enabling to perform data assembly, data visualization, data clustering, allows to perform data subsetting, various normalization methods and general statistics. This omics-oriented package enables to process data in a non-destructive fashion by saving the original data frame, its associated metadata, the transformative steps of the data processing and the processed data in the same multi-layered type of R object named “romics_object”. Additionally, a data processing folder was created for each data analysis performed. Each data processing folder contains the file generated by MS-GF+ and MASIC, the corresponding metadata, an annotated R markdown notebook, an html report containing the versions of all the packages utilized, all the figures generated in pdf format and the data at different points of the processing (i.e. after first assembly and at the end of the normalization steps). The general data normalization and interpretation pipeline included the following steps described below. The identified spectra (MS-GF+ table) were merged with their reporter intensities (MASIC table); a False Discovery Rate (FDR) filtering was applied in order to obtain an FDR below 1% at the spectra, PSM and protein levels (only the proteins identified with at least two unique PSMs were conserved); an isotope impurity correction was applied; the data was rolled up to the PSM and protein level (this type of normalization has the advantage of putting less weight in low abundant peptides); The resulting PSM and protein data frames were then log2 transformed; the method implemented in the package psmartR v0.9.0³⁹ was used to filter out technical or process-based and extreme biological outliers; after a median centering of the data, for the large single cell dataset, only the PSMs or Proteins with a maximum missingness of 40% within at least given single cell-type were conserved for quantification. Finally, a ComBat batch correction from the SVA package v3.30.0⁴⁰ was applied to remove the obvious TMT-set specific batch effects. The evaluation of the data was performed at multiple steps of the data analysis process using diverse metrics including the distribution of the Coefficient of Variation, the correlation values when working with similar samples, hierarchical clustering, and Principal Component Analysis (PCA). As PCA requires full data, the missing values were imputed using the method implemented in the package missMDA v1.14⁴¹ and the analysis was performed using the Package FactoMineR v1.41⁴²; the missingness of the data was restored prior to export the resulting and to perform pairwise two-tailed heteroscedastic student T.tests. Only the proteins with higher abundance in one condition compared to the two others (T.test $p < 0.05$) were considered enriched in the corresponding cell type.

Data availability

Data deposited and freely available at MassIVE data repository, MassIVE ID: MSV000084110.

Results and Discussion

Workflow for the nanoPOTS-based sample processing

To integrate the TMT labeling and boosting concept into the nanoPOTS technology, we re-designed the layout of the first generation nanoPOT chips previously described²². Figure S1 shows the layout of the second generation nanoPOTS chip. Here a series of smaller and larger diameter nanowells are patterned onto the nanoPOTS chip. The smaller diameter (1.2 mm) nanowells are utilized to hold single cells and the large diameter (1.8 mm) nanowells are utilized for boosting samples, as well as for sample combination (or multiplexing) within the same TMT set. The total surface area of the smaller diameter nanowells utilized to hold single cells or single cell-sized protein quantities is $\sim 1 \text{ mm}^2$. This is a much smaller effective total surface area compared to the conventional 0.5-mL tubes ($\sim 130 \text{ mm}^2$) utilized in other single cell and/or mass-limited analyses described,^{29, 32, 34} providing the benefit of significantly reduced opportunity for surface-adsorption-induced protein/peptide (i.e. sample) loss and thus presumably higher sensitivity.

Figure 1 shows the integrated workflow of nanoPOTS processing and TMT-labeling. Cells or protein lysate are deposited into nanowells and protein extraction (as needed), reduction, alkylation and digestion are performed as previously described.²³ 100 nL (for smaller nanowell) or 300 nL (for larger nanowell) of dissolved TMT reagents are then added to the appropriate nanowells. When using TMT10plex and implementing the boosting concept, up to 9 samples (including the boosting sample) can be combined (or multiplexed) and analyzed in a single LC-MS run as depicted in Figure 1. Here the boosting sample is labeled with the TMT 131 channel and the 130N channel is left empty due to isotopic contamination from the 131 channel. To maximize overall sensitivity of LC-MS analysis, the combined samples (i.e. multiplexed TMT set) are analyzed using a 30- μm -i.d. nanoLC column operated at 50 nL/min with an Orbitrap Fusion Lumos mass spectrometer. Given a typical 120-min LC gradient and 8 samples of interest (e.g. single cells) multiplexed in a single TMT10plex set, the nanoPOTS-TMT platform can analyze ~ 70 samples (e.g. single cells) per day, or over ~ 490 samples (e.g. single cells) per week. Such throughput level is approaching the first generation Fluidigm C1 single cell RNA-seq platform that can process up to 96 cells per day.⁴³

TMT-based quantitative proteomics for ultralow input biological materials

In traditional TMT-based quantitative proteomics, TMT-labeling is typically done with samples at ng– μg levels. However, single mammalian cells only contain sub-ng levels of protein (0.1 to 0.5 ng total protein), thus it is necessary to assess whether the TMT-labeling strategy will provide high quantification accuracy at this scale. We evaluated TMT-labeling of single cell-level protein material in our nanodroplet processing platform (nanoPOTS) using commercially available HeLa protein digest from Thermo Scientific. We reasoned that a commercially available and quality-controlled digest will more readily facilitate future benchmarking efforts by the scientific community, a current challenge. Single-cell-equivalent amounts (0.2 ng) of the HeLa digest were dispensed into seven of the small diameter nanowells and 10 ng of HeLa digest was dispensed into a larger nanowell. The single-cell-equivalent samples were then labelled with the first seven channels of the

TMT10plex reagent (i.e. from 126 to 129N) and the larger (10 ng) boosting sample was labelled with the 131 channel.

We first evaluated the TMT labeling efficiency by comparing the number of unlabeled peptides with that of labeled peptides. As shown in Table S1 and Supplementary Dataset 1, the labeling efficiency was ~99% for each of the duplicate LC-MS analyses, indicating that our nanoPOTS platform provides a highly efficient TMT labeling reaction at the nanoliter-scale and for the sub-ng protein amounts. The median values of log₂-transformed protein intensities from each of the 0.2-ng HeLa digests ranged from 14 to 14.4, and for the 130C and 130N empty channels the log₂-transformed median intensity was 9.9 and 10.8, respectively (Figure 2A and Figure S2A). The intensity values observed for the empty channels arise from a combination of isotope contamination and chemical noise, two well recognized phenomena.⁴⁴ Pair-wise analysis of the single-cell-equivalent TMT channels showed Pearson's correlation coefficients ranging from 0.97 to 0.99 (Figure 2B and Figure S2B). The median coefficient of variation (CV) of the seven single-cell-equivalent samples in each of the two multiplexed TMT sets analyzed by LC-MS was ~11% (Figure 2C). When the LC-MS data from the two multiplexed TMT sets (corresponding to 14 single-cell-equivalent samples) were combined, a higher CV with a median value of 24.1% was observed as anticipated with no batch correction. Principal component analysis (PCA) clearly showed most of the variance resulted from batch artifacts (Figure S2C). With the batch covariate identified as being the TMT set run, we employed an empirical Bayes method,⁴⁵ using the ComBat function from the R package sva⁴⁰ to perform batch correction. After batch correction, the two TMT sets were indistinguishable (Figure S2D) and the median CV of protein intensities observed in the two combined TMT datasets dropped to ~11% (Figure 2C). Across the two datasets 11,988 unique peptides and 1,604 proteins (requiring at least 2 unique peptides per protein) were identified for the single-cell-equivalent (0.2 ng) HeLa digests, with 79% of the peptides and 98% of the proteins common between the two datasets (Figure 2D). Taken together, these results indicate robust high-throughput quantification can be achieved for single cell proteomics by combining isobaric labeling and nanodroplet-based sample processing.

Effect of boosting ratios on the quantification of single cell proteomics

As demonstrated in previous studies,^{29, 32, 34} the incorporation of higher protein amounts in one or more channels of a multiplexed TMT-labelled set can help to trigger the fragmentation of otherwise low abundance peptides from mass-limited samples in the same multiplexed set. In this boosting strategy, the boosting ratio (i.e. ratio between the amount peptide in the boosting channel and in the remaining sample channels) reported in the literature ranges from 30 to 500-fold. However, the impact of boosting ratio (specifically the mass ratio of boosting peptides to single-cell-sized peptides) on quantification accuracy for single cell-level analyses has not been well-characterized and remains unclear. Therefore, we evaluated the effect of boosting ratios on quantification for single cell analyses.

We used TMT-labeled single cells from three cultured murine cell lines (C10, epithelial cells; RAW, macrophage cells; SVEC, endothelial cells). Three levels of boosting material were tested: 0 ng (no boosting), 5 ng and 50 ng. The protein digest used for boosting

contained equal amounts of proteins from each of the three cell types. Single mammalian cells are estimated to have an average protein content of ~0.2 ng,⁴⁶ thus the 5 ng and 50 ng protein digests correspond to boosting ratios of ~25 and 250, respectively. For each of the three tested boosting ratio scenarios, two independent TMT sets were prepared and analyzed (Table S2 and Supplementary Dataset 2). As expected, both peptide and protein identifications increased with increasing boosting ratios (Figure 3A–B). At 0-fold, 25-fold and 250-fold boosting ratios the number of identified proteins in single cells were 171, 890, and 1408, respectively. A label-free bulk analysis of the proteome of 5 replicates of the same three cell types (C10, RAW, SVEC) employed above revealed that their proteomes were sufficiently distinct to enable distinguishing of the three cell types based on protein expression alone (Figure S3A). We anticipated that the same would hold true at the single cell-level given reproducible and accurate quantification of single cell proteomes. For each boosting ratio scenario described above (i.e. 0-fold, 0 ng; 25-fold, 5 ng; 250-fold, 50 ng), an unsupervised PCA was performed on protein intensities extracted from TMT reporter ions. As shown in Figure 3C, while individual cells clustered well by type for the samples employing the 0-fold (0 ng) or 25-fold (5 ng) boosting scenario, this was not the case for the samples employing the 250-fold (50 ng) boosting scenario. Interestingly, although only 170 highly abundant proteins were quantified in single cell samples without boosting, the cell types grouped based on their proteome profile alone, indicating the nanoPOTS-TMT-based single cell proteomics platform has good potential to distinguish cell populations without the requirement of additional input. As noted above, while the 250-fold (50 ng) boosting scenario improved proteome coverage compared to the 25-fold (5 ng) boosting scenario, it however yielded reduced cell type specific clustering compared to the 25-fold (5 ng) boosting scenario.

We performed an analysis of the data to understand the factor(s) driving this observation. Examination of the data suggested the level of missing values (i.e. missingness) within a dataset was a main driver of the above observation. When the level of boosting was enhanced from the 25-fold to 250-fold, the level of missingness at the peptide level and protein level for single cell channels increased from 20% to 42% and 3% to 11% respectively (Figure S3B–C). Additionally, the level of unshared peptides and unshared proteins increased from 45% to 55% and from 6% to 11% respectively, with the greater boosting level (Figure S3D–E). We reasoned that the increased missing values caused the reduced single cell quantification performance at the higher 250-fold boosting scenario. Taken together, these results suggest there is an optimum boosting scenario that should ideally be determined for the cell types being studied prior to utilizing the boosting strategy.

TMT-labeling-based single cell proteomics profiling of murine cell populations

To further demonstrate an initial application of our platform for high throughput deep single cell proteome profiling, we applied it to quantify single cell proteomes from a larger population of cells. 72 single cells, corresponding to 24 C10 cells, 24 RAW cells, and 24 SVEC cells, were deposited onto nanoPOTS chips using FACS as we previously described²³ to perform the large scale single cell experiment. To minimize the incidence of dead cells, cellular debris and precipitates deposited into nanowells of the nanoPOTS chip and to ensure analysis of only viable cells, we utilized Calcein AM, a membrane permeant dye that

fluoresces only in metabolically active cells. Isolation and deposition of cells into nanowells of the nanoPOTS chip gated on the Calcein AM signal readily distinguished viable cells from dead cells and other contaminants. The single cells were deposited onto nanoPOTS chips and labeled with discrete TMT channels in the format shown in Table S3. A total of 4 nanoPOTS chips were utilized to accommodate the 72 single cells, with each nanoPOTS chip harboring 6 individual single cells distributed evenly across the three cell types.

From LC-MS/MS analyses of the 12 multiplexed TMT batches containing 72 single cells we identified 2,331 proteins at a <1% false discovery rate (FDR) and conservatively 1,597 proteins identified with at least 2 unique peptides at <1% FDR (Supplementary Dataset 3). For quantitative analysis of the single cell proteome data, we utilized a multiparametric method implemented in the pmarR package³⁹ to identify and exclude outliers from further downstream processing. A total of eleven cells were considered as technical, process-based or extreme biological outliers and were excluded from further analysis. From the remaining 61 single cells only proteins present in 60% of the datasets from a given cell type (C10, RAW, or SVEC cells), equivalent to proteins present in 44% of all datasets across all cell types, were considered for quantification resulting in 1,225 quantifiable proteins (Supplementary Dataset 3). The levels of all 1,225 proteins quantified in single cells were projected onto their principal components. Principal component analysis (PCA) showed the single cell proteomes grouped by cell type (Figure 4A).

To identify features driving the grouping of the three cell types we performed pairwise Student's tests ($p < 0.05$) utilizing non-imputed data. Pairwise Student's tests ($p < 0.05$) revealed 138 proteins were enriched in C10 cells compared to RAW and SVEC cells, 82 proteins were enriched in RAW cells compared to C10 and SVEC cells, and 56 proteins were enriched in SVEC cells compared to C10 and RAW cells (Figure 4B). Importantly, markers of each of the three cell types (C10 epithelial cells, RAW macrophage cells, SVEC endothelial cells) were present at higher abundance in the expected single cell populations as shown in Figure 4C and discussed below. Ezrin, a protein normally located in the apical membrane of lung epithelial cells⁴⁷ was significantly higher in C10 single cells compared to the other two cell types (T-test $p < 0.001$). This observation supported our earlier label-free nanoPOTS-based analysis of human sorted lung epithelial and mesenchymal cells where we also identified ezrin as significantly enriched in lung epithelial cells.²³ Moesin, a paralog of ezrin known to be enriched in the apex of some microvilliated epithelial cells⁴⁸ and Collagen I were also enriched in C10 single cells. CD68, CAPG, and LYZ1, three macrophage and/or immune cell markers were all enriched in RAW single cells. CD68, CAPG, and LYZ2, three macrophage and/or immune cell markers were all enriched in RAW single cells. CD68 is a macrophage and monocyte surface marker protein.⁴⁹ CAPG is a macrophage-capping protein.⁵⁰ LYZ2 is an antimicrobial enzyme that is exclusively expressed in monocytes, macrophages and granulocytes.⁵¹ SVEC4-10 is an endothelial cell line derived by SV40 (strain 4A) transformation of endothelial cells from axillary lymph node vessels with known antigens/markers of this cell line described by ATCC including H2 class I histocompatibility protein (H2-K1), Factor VIII related antigen and VCAM. While we did not detect the latter two proteins in our single cell proteomics dataset, H2-K1 was enriched in single SVEC4-10 cells. Prosaposin, a protein known to be high in lymphocytes⁵² and STAT1 described to be most abundant in lymphoid tissues on the Human proteome atlas, were also found enriched

in SVEC single cells. These results are largely in agreement with bulk-scale label-free abundance measurements (Figure S4).

Conclusion

We have combined nanoPOTS technology with TMT isobaric labeling to establish a platform that provides a robust and high-throughput proteomic workflow for effective multiplexed analysis of single mammalian cells. Reproducible quantitative proteome measurements across >60 single cells were achieved, with proteome coverage >1,200 proteins when conservatively requiring 2 unique peptides per protein quantified and presence in >40% of the cells analyzed. The enabling aspect of our platform is well demonstrated by the ability to group single cells by cell type based on their protein expression alone. We anticipate that our single cell proteomics platform should enable broad applications involving single cells or very limited starting quantities of material, e.g. understanding stem cell development, proteomic studies of isolated clinical specimens such as circulating tumor cells, or unbiased proteome imaging of tissue heterogeneity at cellular resolution.

While the current single cell proteomics platform provides significantly improved analytical throughput (~490 cells analyzed per week), we note that a new TMT16plex reagent from ThermoFisher coupled with improvements in platform automation and LC-MS gradient optimizations we are implementing is poised to increase current throughput by 3–4-fold. In addition to throughput, the single cell quantification accuracy can be further improved by incorporating gas phase fractionation technologies including high-field asymmetric waveform ion mobility spectrometry (FAIMS)⁵³ and real-time search MS³ (RTS-MS³)⁵⁴. We anticipate these ion fractionation strategies not only efficiently minimize ratio compression and improve proteome dynamic range, but also significantly reduce chemical noise and related false discovery. Finally, we also plan to reduce the sample processing volumes to low nanoliter or even picoliter scale, which could further improve protein recovery of single cells and reduce chemical contamination from reagents.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

Portions of this research were supported by a Laboratory Directed Research and Development award from Pacific Northwest National Laboratory (PNNL) to YZ, UC4 DK108101 (NIDDK) and U01 HL122703 (NHLBI) grants to CA, R21 CA223715 (NCI) grant to TS, R33 CA225248 (NCI) grant to RTK, DP3 DK110844 (NIDDK) grant to WJQ, U24CA210955 (NCI) grant to TL and RDS, and a P41 GM103493 (NIGMS) grant to RDS, the IARPA FunGCAT program (number:166156) to JNA. The experimental work was performed in the Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by the Department of Energy (DOE) and located at PNNL, which is operated by Battelle Memorial Institute for the DOE under Contract DE-AC05-76RL0 1830.

References

1. Zhang L; Vertes A, Single-cell mass spectrometry approaches to explore cellular heterogeneity. *Angew. Chem. Int. Ed* 2018, 57 (17), 4466–4477.

2. Heath JR; Ribas A; Mischel PS, Single-cell analysis tools for drug discovery and development. *Nat. Rev. Drug Discov* 2016, 15 (3), 204. [PubMed: 26669673]
3. Hou Y; Guo H; Cao C; Li X; Hu B; Zhu P; Wu X; Wen L; Tang F; Huang Y, Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* 2016, 26 (3), 304. [PubMed: 26902283]
4. Wilson NK; Kent DG; Buettner F; Shehata M; Macaulay IC; Calero-Nieto FJ; Castillo MS; Oedekoven CA; Diamanti E; Schulte R, Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. *Cell Stem Cell* 2015, 16 (6), 712–724. [PubMed: 26004780]
5. Buettner F; Natarajan KN; Casale FP; Proserpio V; Scialdone A; Theis FJ; Teichmann SA; Marioni JC; Stegle O, Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol* 2015, 33 (2), 155. [PubMed: 25599176]
6. Liu Y; Beyer A; Aebersold R, On the dependency of cellular protein levels on mRNA abundance. *Cell* 2016, 165 (3), 535–550. [PubMed: 27104977]
7. Schwanhäusser B; Busse D; Li N; Dittmar G; Schuchhardt J; Wolf J; Chen W; Selbach M, Global quantification of mammalian gene expression control. *Nature* 2011, 473 (7347), 337. [PubMed: 21593866]
8. Ghazalpour A; Bennett B; Petyuk VA; Orozco L; Hagopian R; Mungrue IN; Farber CR; Sinsheimer J; Kang HM; Furlotte N, Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genet.* 2011, 7 (6), e1001393. [PubMed: 21695224]
9. Moghieb A; Clair G; Mitchell HD; Kitzmiller J; Zink EM; Kim Y-M; Petyuk V; Shukla A; Moore RJ; Metz TO, Time-resolved proteome profiling of normal lung development. *Am. J. Physiol. Lung. Cell. Mol. Physiol* 2018, 315 (1), L11–L24. [PubMed: 29516783]
10. Du Y; Clair GC; Al Alam D; Danopoulos S; Schnell D; Kitzmiller JA; Misra RS; Bhattacharya S; Warburton D; Mariani TJ, Integration of transcriptomic and proteomic data identifies biological functions in cell populations from human infant lung. *Am. J. Physiol. Lung. Cell. Mol. Physiol* 2019.
11. Doerr A, Single-cell proteomics. *Nat. Methods* 2019, 16 (1), 20.
12. Macaulay IC; Ponting CP; Voet T, Single-cell multiomics: multiple measurements from single cells. *Trends. Genet* 2017, 33 (2), 155–168. [PubMed: 28089370]
13. Bendall SC; Simonds EF; Qiu P; El-ad DA; Krutzik PO; Finck R; Bruggner RV; Melamed R; Trejok A; Ornatsky OI, Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* 2011, 332 (6030), 687–696. [PubMed: 21551058]
14. Boylan KLM; Geschwind K; Koopmeiners JS; Geller MA; Starr TK; Skubitz APN, A multiplex platform for the identification of ovarian cancer biomarkers. *Clin. Proteom* 2017, 14 (1), 34.
15. Hughes AJ; Spelke DP; Xu Z; Kang C-C; Schaffer DV; Herr AE, Single-cell western blotting. *Nat. Methods* 2014, 11 (7), 749. [PubMed: 24880876]
16. Angel TE; Aryal UK; Hengel SM; Baker ES; Kelly RT; Robinson EW; Smith RD, Mass spectrometry-based proteomics: existing capabilities and future directions. *Chem. Soc. Rev* 2012, 41 (10), 3912–3928. [PubMed: 22498958]
17. Couvillion SP; Zhu Y; Nagy G; Adkins JN; Ansong C; Renslow RS; Piehowski PD; Ibrahim YM; Kelly RT; Metz TO, New mass spectrometry technologies contributing towards comprehensive and high throughput omics analyses of single cells. *Analyst* 2019, 144 (3), 794–807. [PubMed: 30507980]
18. Clair G; Piehowski PD; Nicola T; Kitzmiller JA; Huang EL; Zink EM; Sontag RL; Orton DJ; Moore RJ; Carson JP, Spatially-resolved proteomics: rapid quantitative analysis of laser capture microdissected alveolar tissue samples. *Sci. Rep* 2016, 6, 39223. [PubMed: 28004771]
19. Zhu Y; Piehowski PD; Kelly RT; Qian W-J, Nanoproteomics comes of age. *Expert. Rev. Proteomic* 2018, 15 (11), 865–871.
20. Li Z-Y; Huang M; Wang X-K; Zhu Y; Li J-S; Wong CCL; Fang Q, Nanoliter-scale oil-air-droplet chip-based single cell proteomic analysis. *Anal. Chem* 2018, 90 (8), 5430–5438. [PubMed: 29551058]

21. Shao X; Wang X; Guan S; Lin H; Yan G; Gao M; Deng C; Zhang X, Integrated proteome analysis device for fast single-cell protein profiling. *Anal. Chem* 2018, 90 (23), 14003–14010. [PubMed: 30375851]
22. Zhu Y; Piehowski PD; Zhao R; Chen J; Shen Y; Moore RJ; Shukla AK; Petyuk VA; Campbell-Thompson M; Mathews CE; Smith RD; Qian W-J; Kelly RT, Nanodroplet processing platform for deep and quantitative proteome profiling of 10–100 mammalian cells. *Nat. Commun* 2018, 9 (1), 882. [PubMed: 29491378]
23. Zhu Y; Clair G; Chrisler W; Shen Y; Zhao R; Shukla A; Moore R; Misra R; Pryhuber G; Smith R; Ansong C; Kelly R, Proteomic analysis of single mammalian cells enabled by microfluidic nanodroplet sample preparation and ultrasensitive nanoLC - MS. *Angew. Chem. Int. Ed* 2018, 130 (38), 12550–12554.
24. Zhu Y; Dou M; Piehowski PD; Liang Y; Wang F; Chu RK; Chrisler W; Smith JN; Schwarz KC; Shen Y; Shukla AK; Moore RJ; Smith RD; Qian W-J; Kelly RT, Spatially resolved proteome mapping of laser capture microdissected tissue with automated sample transfer to nanodroplets. *Mol. Cell. Proteomics* 2018, 17 (9), 1864–1874. [PubMed: 29941660]
25. Dou M; Zhu Y; Liyu A; Liang Y; Chen J; Piehowski P; Xu K; Zhao R; Moore RJ; Atkinson MA; Mathews CE; Qian W-J; Kelly RT, Nanowell-mediated two-dimensional liquid chromatography enables deep proteome profiling of <1000 mammalian cells. *Chem. Sci* 2018, 9 (34), 6944–6951. [PubMed: 30210768]
26. Zhu Y; Podolak J; Zhao R; Shukla AK; Moore RJ; Thomas GV; Kelly RT, Proteome profiling of 1 to 5 spiked circulating tumor cells isolated from whole blood using immunodensity enrichment, laser capture microdissection, nanodroplet sample processing, and ultrasensitive nanoLC–MS. *Anal. Chem* 2018, 90 (20), 11756–11759. [PubMed: 30269481]
27. Thompson A; Schäfer J; Kuhn K; Kienle S; Schwarz J; Schmidt G; Neumann T; Hamon C, Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem* 2003, 75 (8), 1895–1904. [PubMed: 12713048]
28. Lombard - Banek C; Moody SA; Nemes P, Single-cell mass spectrometry for discovery proteomics: quantifying translational cell heterogeneity in the 16 - cell frog (*Xenopus*) embryo. *Angew. Chem. Int. Ed* 2016, 55 (7), 2454–2458.
29. Russell CL; Heslegrave A; Mitra V; Zetterberg H; Pocock JM; Ward MA; Pike I, Combined tissue and fluid proteomics with Tandem Mass Tags to identify low - abundance protein biomarkers of disease in peripheral body fluid: An Alzheimer's Disease case study. *Rapid Commun. Mass Spectrom* 2017, 31 (2), 153–159. [PubMed: 27813239]
30. De Marco M; Venneri A, Volume and connectivity of the ventral tegmental area are linked to neurocognitive signatures of Alzheimer's disease in humans. *J. Alzheimers. Dis* 2018, 63 (1), 167–180. [PubMed: 29578486]
31. Dayon L. c.; Turck N; Kienle S; Schulz-Knappe P; Hochstrasser DF; Scherl A; Sanchez J-C, Isobaric tagging-based selection and quantitation of cerebrospinal fluid tryptic peptides with reporter calibration curves. *Anal. Chem* 2010, 82 (3), 848–858. [PubMed: 20058875]
32. Budnik B; Levy E; Harmange G; Slavov N, SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome. Biol* 2018, 19 (1), 161. [PubMed: 30343672]
33. Tan Z; Yi X; Carruthers NJ; Stemmer PM; Lubman DM, Single amino acid variant discovery in small numbers of cells. *J. Proteome. Res* 2018, 18 (1), 417–425. [PubMed: 30404448]
34. Yi L; Tsai C-F; Dirice E; Swensen AC; Chen J; Shi T; Gritsenko MA; Chu RK; Piehowski PD; Smith RD; Liu T; Qian W-J, A boosting to amplify signal with isobaric labeling (BASIL) strategy for comprehensive quantitative phosphoproteomic characterization of small populations of cells. *Anal. Chem* 2019, 91 (9), 5794–5801. [PubMed: 30843680]
35. Shen Y; Moore RJ; Zhao R; Blonder J; Auberry DL; Masselon C; Paša-Toli L; Hixson KK; Auberry KJ; Smith RD, High-efficiency on-line solid-phase extraction coupling to 15–150- μ m-id column liquid chromatography for proteomic analysis. *Anal. Chem* 2003, 75 (14), 3596–3605. [PubMed: 14570215]
36. Kim S; Pevzner PA, MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat. Commun* 2014, 5, 5277. [PubMed: 25358478]

37. Monroe ME; Shaw JL; Daly DS; Adkins JN; Smith RD, MASIC: A software program for fast quantitation and flexible visualization of chromatographic profiles from detected LC–MS (/MS) features. *Comput. Biol. Chem* 2008, 32 (3), 215–217. [PubMed: 18440872]
38. Kiebel GR; Auberry KJ; Jaitly N; Clark DA; Monroe ME; Peterson ES; Toli N; Anderson GA; Smith RD, PRISM: a data management system for high - throughput proteomics. *Proteomics* 2006, 6 (6), 1783–1790. [PubMed: 16470653]
39. Stratton KG; Webb-Robertson B-JM; McCue LA; Stanfill B; Claborne D; Godinez I; Johansen T; Thompson AM; Burnum-Johnson KE; Waters KM, psmartR: quality control and statistics for mass spectrometry-based biological data. *J. Proteome. Res* 2019, 18 (3), 1418–1425. [PubMed: 30638385]
40. Leek JT; Johnson WE; Parker HS; Jaffe AE; Storey JD, The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012, 28 (6), 882–883. [PubMed: 22257669]
41. Josse J; Huisson F, missMDA: a package for handling missing values in multivariate data analysis. *J. Stat. Softw* 2016, 70 (1), 1–31.
42. Lê S; Josse J; Huisson F, FactoMineR: an R package for multivariate analysis. *J. Stat. Softw* 2008, 25 (1), 1–18.
43. Gong H; Do D; Ramakrishnan R, Single-cell mRNA-seq using the Fluidigm C1 system and integrated fluidics circuits In *Gene Expression Analysis*, Springer: 2018; pp 193–207.
44. Rauniyar N; Yates Iii JR, Isobaric labeling-based relative quantification in shotgun proteomics. *J. Proteome. Res* 2014, 13 (12), 5293–5309. [PubMed: 25337643]
45. Johnson WE; Li C; Rabinovic A, Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007, 8 (1), 118–127. [PubMed: 16632515]
46. Milo R; Jorgensen P; Moran U; Weber G; Springer M, BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Res.* 2009, 38 (suppl_1), D750–D753. [PubMed: 19854939]
47. Li Q; Gao H; Xu H; Wang X; Pan Y; Hao F; Qiu X; Stoecker M; Wang E; Wang E, Expression of ezrin correlates with malignant phenotype of lung cancer, and in vitro knockdown of ezrin reverses the aggressive biological behavior of lung cancer cells. *Tumor Biol* 2012, 33 (5), 1493–1504.
48. Berryman M; Franck Z; Bretscher A, Ezrin is concentrated in the apical microvilli of a wide variety of epithelial cells whereas moesin is found primarily in endothelial cells. *J. Cell Sci* 1993, 105 (4), 1025–1043. [PubMed: 8227193]
49. Barros MHM; Hauck F; Dreyer JH; Kempkes B; Niedobitek G, Macrophage polarisation: an immunohistochemical approach for identifying M1 and M2 macrophages. *PloS one* 2013, 8 (11), e80908. [PubMed: 24260507]
50. Dabiri GA; Young CL; Rosenbloom J; Southwick FS, Molecular cloning of human macrophage capping protein cDNA. A unique member of the gelsolin/villin family expressed primarily in macrophages. *J. Biol. Chem* 1992, 267 (23), 16545–16552. [PubMed: 1322908]
51. Cross M; Mangelsdorf I; Wedel A; Renkawitz R, Mouse lysozyme M gene: isolation, characterization, and expression studies. *PNAS* 1988, 85 (17), 6232–6236. [PubMed: 3413093]
52. Darmon A; Maschmeyer P; Winau F, The immunological functions of saposins In *Advances in Immunology*, Elsevier: 2010; Vol. 105, pp 25–62. [PubMed: 20510729]
53. Schweppe DK; Prasad S; Belford MW; Navarrete-Perea J; Bailey DJ; Huguet R; Jedrychowski MP; Rad R; McAlister G; Abbatiello SE, Characterization and optimization of multiplexed quantitative analyses using high-field asymmetric-waveform ion mobility mass spectrometry. *Anal. Chem* 2019, 91 (6), 4010–4016. [PubMed: 30672687]
54. Erickson BK; Mintseris J; Schweppe DK; Navarrete-Perea J; Erickson AR; Nusinow DP; Paulo JA; Gygi SP, Active instrument engagement combined with a real-time database search for improved performance of sample multiplexing workflows. *J. Proteome. Res* 2019, 18 (3), 1299–1306. [PubMed: 30658528]

Figure 1.
Workflow of the nanoPOTS-TMT-based single cell proteomics platform.

Figure 2. Evaluation of isobaric-labelling-based quantitative proteomics using single-cell-sized (0.2 ng) HeLa digest. TMT channels from 126 to 129C were used to label 0.2-ng digest and channel 130N and 130C were left empty. (A) Distribution of protein intensities across all TMT channels in one TMT set. (B) Pairwise correlation of protein intensities across the 7 single-cell level protein digests. (C) The coefficients of variation of protein intensities within each TMT set, between two TMT sets, and two sets after batch correction. (D) Venn diagrams showing the overlap of peptide and protein identifications in 0.2-ng protein digests between TMT sets.

Figure 3. Effect of boosting ratios (0, 25 ×, and 250 ×) on the quantification performance of single cell proteomics. (A) Total number of identified peptides. (B) Total number of identified proteins; (C) Unsupervised PCA showing cell grouping based on protein expressions in single cultured murine cells (C10, Raw, and SVEC).

Figure 4.

(A) Unsupervised PCA showing the single cell grouping based on protein expression. (B) Heatmap of proteins enriched in each cell population (pairwise t.test<0.05). (C) comparison of cell type specific marker abundances across the three cell populations.