High-throughput single-cell proteomics quantifies the emergence of macrophage heterogeneity

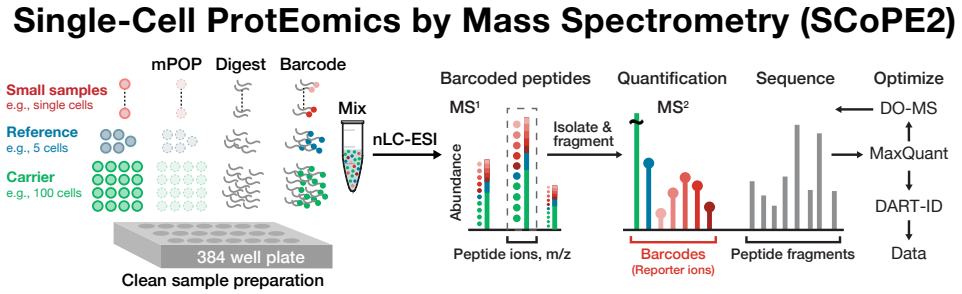
Harrison Specht,[1,2] Edward Emmott,[1,2] Toni Koller,[1,2] & Nikolai Slavov[1,2,3,✉]

[1]Department of Bioengineering, Northeastern University, Boston, MA 02115, USA
[2]Barnett Institute, Northeastern University, Boston, MA 02115, USA
[3]Department of Biology, Northeastern University, Boston, MA 02115, USA

✉ Correspondence: nslavov@alum.mit.edu
∈ Data, code & protocols: scope2.slavovlab.net

Abstract

The fate and physiology of individual cells are controlled by networks of proteins. Yet, our ability to quantitatively analyze protein networks in single cells has remained limited. To overcome this barrier, we developed SCoPE2. It integrates concepts from Single-Cell ProtEomics by Mass Spectrometry (SCoPE-MS) with automated and miniaturized sample preparation, substantially lowering cost and hands-on time. SCoPE2 uses data-driven analytics to optimize instrument parameters for sampling more ion copies per protein, thus supporting quantification with improved count statistics. These advances enabled us to analyze the emergence of cellular heterogeneity as homogeneous monocytes differentiated into macrophage-like cells in the absence of polarizing cytokines. We used SCoPE2 to quantify over 2,000 proteins in 356 single monocytes and macrophages in about 85 hours of instrument time, and the quantified proteins allowed us to discern single cells by cell type. Furthermore, the data uncovered a continuous gradient of proteome states for the macrophage-like cells, suggesting that macrophage heterogeneity may emerge even in the absence of polarizing cytokines. Our methodology lays the foundation for quantitative analysis of protein networks at single-cell resolution.



Introduction

Tissues and organs are composed of functionally specialized cells. This specialization of single cells often arises from the protein networks mediating physiological functions. Yet, our ability to comprehensively quantify the proteins comprising these networks in single-cells has remained relatively limited.[1] As a result, the protein levels in single cells are often inferred from indirect surrogates – sequence reads from their corresponding mRNAs.[2,3] Such inference is unreliable because the levels of many proteins are set primarily by regulating their degradation or synthesis.[4,5]

Single-cell RNA sequencing methods have empowered the classification of single cells, characterization of spatial heterogeneity, and discovery of new cell types.[2,6] These methods for single-cell detection of transcripts depend critically on amplifying nucleic acids. Since the amplification process starts by sampling only a few copies per transcript, the estimated mRNA abundances are affected by large sampling (counting) errors.[2,3] These errors are inherent to sampling a few copies per transcript and cannot be reduced by the amplification process. They can only be reduced by sampling more (ideally all) copies of a gene.[7]

Sampling many protein copies per gene may be feasible since most proteins are present at over 1,000-fold more copies per cell than their corresponding transcripts.[1,7,8] This high abundance also obviates the need for amplification. Since amplification may introduce noise, obviating amplification is a desirable aspect. Thus, the high copy number of proteins may allow their quantification without amplification.

However most technologies for quantifying proteins in single cells rely on antibodies, which afford only limited specificity.[1] Although mass-spectrometry (LC-ESI-MS/MS) has enabled accurate, high-specificity, and high-throughput quantification of proteins from bulk samples,[9–11] its application to single cells is in its infancy.[1,7] To apply these powerful technologies to the analysis of single cells, we developed Single-Cell ProtEomics by Mass Spectrometry (SCoPE-MS).[7,12] SCoPE-MS introduced the concept of using carrier proteins barcoded with tandem-mass-tags (TMT), which serves three important roles: (i) reducing sample loss, (ii) enhancing the detectability of ions during MS1 survey scans and (iii) providing fragment ions for peptide sequence identification. By combining this concept with MS-compatible cell lysis, we established the feasibility of applying multiplexed LC-ESI-MS/MS to quantify proteins from single cells.

However, the cost, throughput, and reliability of SCoPE-MS data fell short of our vision of single-cell proteomics.[7] Our vision requires quantifying thousands of proteins and proteoforms across thousands of single cells at an affordable cost and time. Such data could support clinical

applications, such as biomarker discovery. Moreover, these data could permit inferring direct casual mechanisms underlying the functions of protein networks.[7] The more cells and proteoforms are quantified, the fewer assumptions are needed for this analysis. Thus, our goal in developing SCoPE2 was to increase the number of cells and proteins analyzed at affordable cost while sampling a sufficient number of ion copies per protein to make quantitative measurements.

To achieve this goal, we followed previously outlined opportunities.[7] In particular, we overhauled multiple experimental steps, including cell isolation, lysis and sample preparation.[13] Furthermore, we developed data-driven computational methods for optimizing the acquisition of MS data (DO–MS; Data-driven Optimization of MS)[14] and for interpreting these data once acquired, e.g., for enhancing peptide identification (DART-ID; Data-driven Alignment of Retention Times for IDentification).[15] These advances combine synergistically into a next generation SCoPE-MS version, SCoPE2, that affords substantially improved quantification and throughput.

These advances enabled us to ask fundamental questions: Do homogeneous monocytes produce homogeneous macrophages in the absence of polarizing cytokines? Are macrophages inherently prone to be heterogeneous, or is their heterogeneity simply reflecting different progenitors and polarizations induced by different cytokines? These questions are cornerstones to our understanding of macrophage heterogeneity that plays important roles in human pathophysiology: Depending on their polarization, macrophages can play pro-inflammatory (usually ascribed to M1 polarization) or anti-inflammatory roles (usually ascribed to M2 polarization), and be involved in tissue development and maintenance.[16] Some studies suggest that rather than separating into discrete functional classes, the M1 and M2 states represent the extremes of a wider spectrum of end-states.[16,17] We found, surprisingly, that the individual macrophage-like cells were highly heterogeneous even though they originated from homogeneous monocytes exposed to identical environmental conditions.

Results

The overall work-flow of SCoPE2 is illustrated in Fig. 1a. Single cells are isolated in individual wells, lysed, and the proteins digested to peptides. The peptides from each single cell are covalently labeled (barcoded) with isobaric tandem-mass-tags (TMT), and therefore labeled peptides with the same sequence (and thus mass) appear as a single mass/charge cluster in the MS1 scans. The MS instrument isolates such clusters and fragments them. The fragmentation generates reporter ions

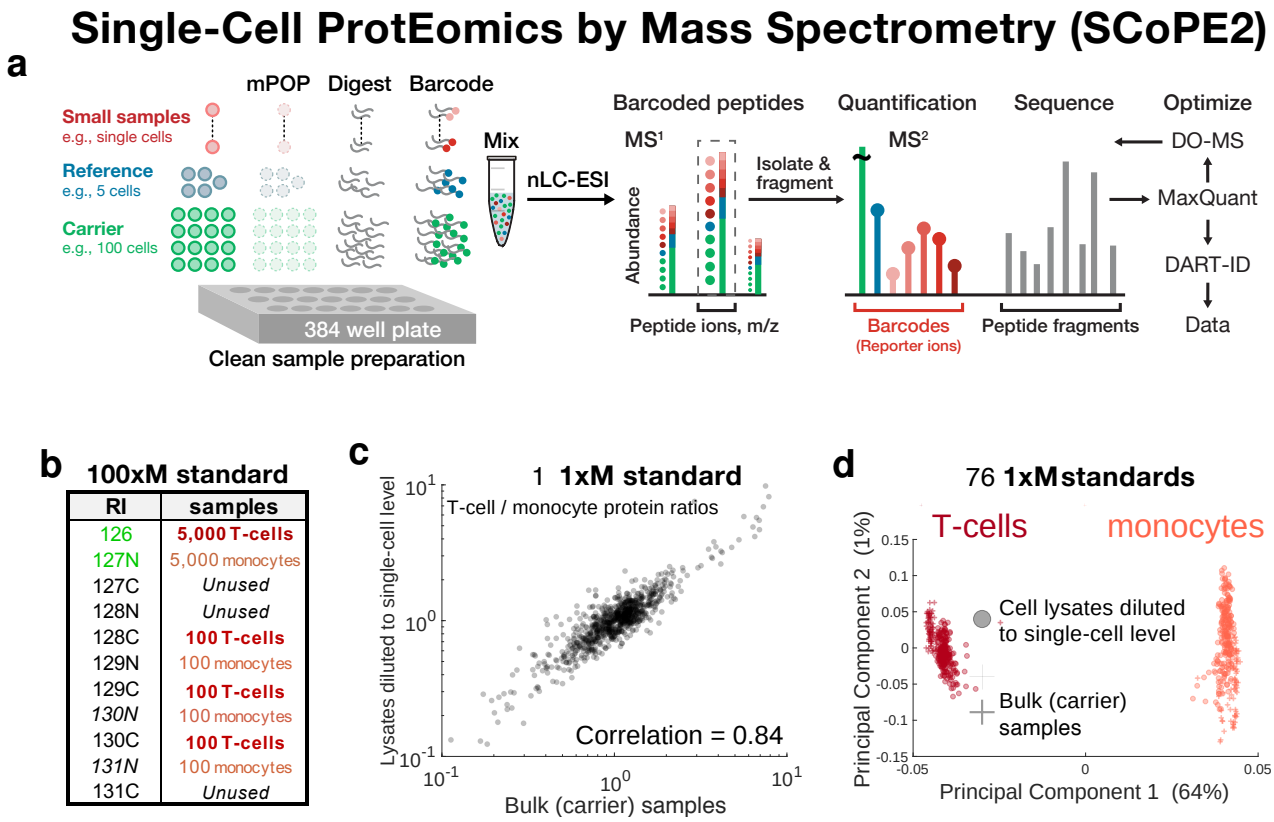


Figure 1 | Developing and benchmarking SCoPE2 with standards
(**a**) Conceptual diagram and work flow of SCoPE2. Cells are sorted into multiwell plates and lysed by mPOP. The proteins in the lysates are digested with trypsin, the resulting peptides labeled with TMT, combined, and analyzed by LC-MS/MS. SCoPE2 sets contain reference channels that allow merging single-cells from different SCoPE2 sets into a single dataset. The LC-MS/MS analysis is optimized by DO-MS,[14] and peptide identification enhanced by DART-ID.[15] (**b**) Schematic for the design of a 100xM standard sets. Monocytes (U937 cells) and T-cells (Jurkat cells) were FACS-sorted at the indicated numbers and labeled with tandem-mass-tags having the indicated reporter ions (RI). (**c**) Comparison of protein fold-change between T-cells and monocytes estimated from the small-samples and from the carrier samples of a 1xM standard, i.e., 1 % sample from the 100xM standard described in panel a. The relative protein levels measured from samples diluted to single-cell levels are very similar to the corresponding estimates from the carrier (bulk) samples. (**d**) Principal component analysis separates samples corresponding to T-cells (Jurkat cells) or to monocytes (U-937 cells). The small samples (which correspond to 100 cells diluted 100-fold to single cell-level) cluster with the carrier samples, indicating that relative protein quantification from all samples is consistent and based on cell type. All quantified proteins were used for this analysis and each protein was normalized separately for the carrier channels and the small sample channels.

(RI), whose abundances reflect protein abundances in the corresponding samples (single cells), Fig. 1a. Key advances of SCoPE2 over SCoPE-MS include:

- Instead of lysing cells by focused acoustic sonication, SCoPE2 lyses cells by Minimal ProteOmic sample Preparation (mPOP).[13] mPOP uses a freeze-heat cycle that extracts proteins efficiently in pure water, thus obviating cleanup before MS analysis. mPOP allows sample preparation in multiwell plates, which enables simultaneous processing of many samples in parallel with inexpensive PCR thermocyclers and liquid dispensers. This advance over SCoPE-MS allows SCoPE2 to reduce lysis volumes 10-fold, from $10\mu l$ to $1\mu l$, to reduce the cost of consumables and equipment over 100-fold, and to increase throughput of sample preparation over 100-fold by parallel processing.

- SCoPE2 also introduces a reference channel composed of a reference sample used in all sets. This reference is about 5-fold more abundant than the small-sample channels (i.e., a single cell proteome) so that the higher abundance results in improved ion counting-statistics while remaining comparable to that of single cells, and thus likely to be within the linear range of quantification.

- SCoPE2 introduces modified liquid chromatography, MS instrumentation, and parameters, as detailed in the Methods. Such changes include shorter gradients (for increased throughput) and a narrower isolation window (0.7 Th) for improved ion isolation.

- SCoPE2 employs a data-driven computational framework, DO-MS, to optimize instrument settings and thus data acquisition.[14] DO-MS uses MaxQuant[10,18] output to rationally determine instrument parameters that improve the sampling of peptide ions, e.g., by increasing the probability that the instrument samples a peptide peak coming from the nLC at its apex.[14]

- Once the MS data are acquired, SCoPE2 can use additional features of the data to enhance their interpretation. For example, SCoPE2 uses a principled Bayesian framework (DART-ID)[15] to incorporate retention time information for increasing the confidence of assigning peptide sequences.

These advances work synergistically to enhance the ability of SCoPE2 to quantify proteins in single cells. Below, we exemplify the improvements, starting with the application of DO-MS.

Optimizing SCoPE2 with standards

The quality of LC-MS/MS data strongly depends on numerous interdependent parameters (e.g., chromatographic packing, LC gradient steepness, and ion accumulation time). To optimize such parameters, we applied DO-MS on 1xM standards. These 1xM standards are 1 % samples of a bulk sample (100xM) with carriers of 5,000 cells and small-samples each comprised of 100 cells, as shown in Fig. 1b. Thus, 1xM standards approximated idealized SCoPE2 sets and enabled us to focus on optimizing LC-MS/MS parameters using identical samples, i.e., independent from the biological variations of SCoPE2 sets.

First, we optimized our analytical column configuration and LC gradient settings. Each 1xM injection was analyzed for only 60 minutes since our goal was to optimize the number of proteins quantified across many cells, rather than merely the number of proteins quantified per sample.[7] By varying chromatographic parameters and benchmarking their effects with DO-MS, we minimized elution peak widths. Sharper elution peaks increase the sampled copies of each peptide per unit time and reduce the probability that multiple peptides are simultaneously isolated for MS2 analysis, Fig. 1a. Concurrent with optimizing peptide elution profiles, we optimized the data-dependent acquisition MS settings, such as minimum MS2 intensity threshold, MS2 injection time, and number of ions sent for MS2 analysis per duty cycle (i.e., TopN), to increase the probability of sampling the apex of the elution peak of each peptide. This optimization increased the number of ions copies sampled from each peptide.[14]

Benchmarking SCoPE2 with standards

The 1xM standards also permitted estimating the instrument measurement noise – independent from biological and sample preparation noise – in the context of SCoPE2 sets. This noise estimate was motivated by our concern that factors unique to ultra-low abundance samples, such as counting noise,[1,3,7] may undermine measurement accuracy. Of particular importance is the ability of SCoPE2 to quantify changes of proteins across cell-types rather than the overall protein abundance[5]; the overall correlation (computed by averaging across proteins) tends to be very high across human tissues and hard to interpret since it confounds different sources of variance.[5] Thus, we compared the fold-changes of proteins between T-cells and monocytes (Jurkat / U-937 protein ratios) estimated from just two cell lysates diluted to single-cell levels against the corresponding ratios estimated from the bulk samples used as carriers, Fig. 1b. The high concordance of these

estimates (Spearman $\rho = 0.84$) strongly indicates that the instrument noise in quantifying proteins by SCoPE2 is small, consistent with our arguments that the abundance of proteins in single mammalian cells is high enough to minimize the sampling (counting) noise.[7]

To further evaluate relative quantification, beyond the results for two samples diluted to single-cell level (Fig. 1b), we consolidated the data from 76 1xM standards and computed all pairwise correlations. This 592-dimensional matrix was projected onto its first two principal components (PC). The largest PC accounts for 64% of the total variance in the data and perfectly separates all samples corresponding to T-cells or monocytes. Crucially, the cell lysates diluted to single-cell levels separate the same way as the carrier samples, indicating that the separation is driven by cell-type specific protein differences. Thus, the SCoPE2 design can reliably quantify protein abundances at single-cell levels.

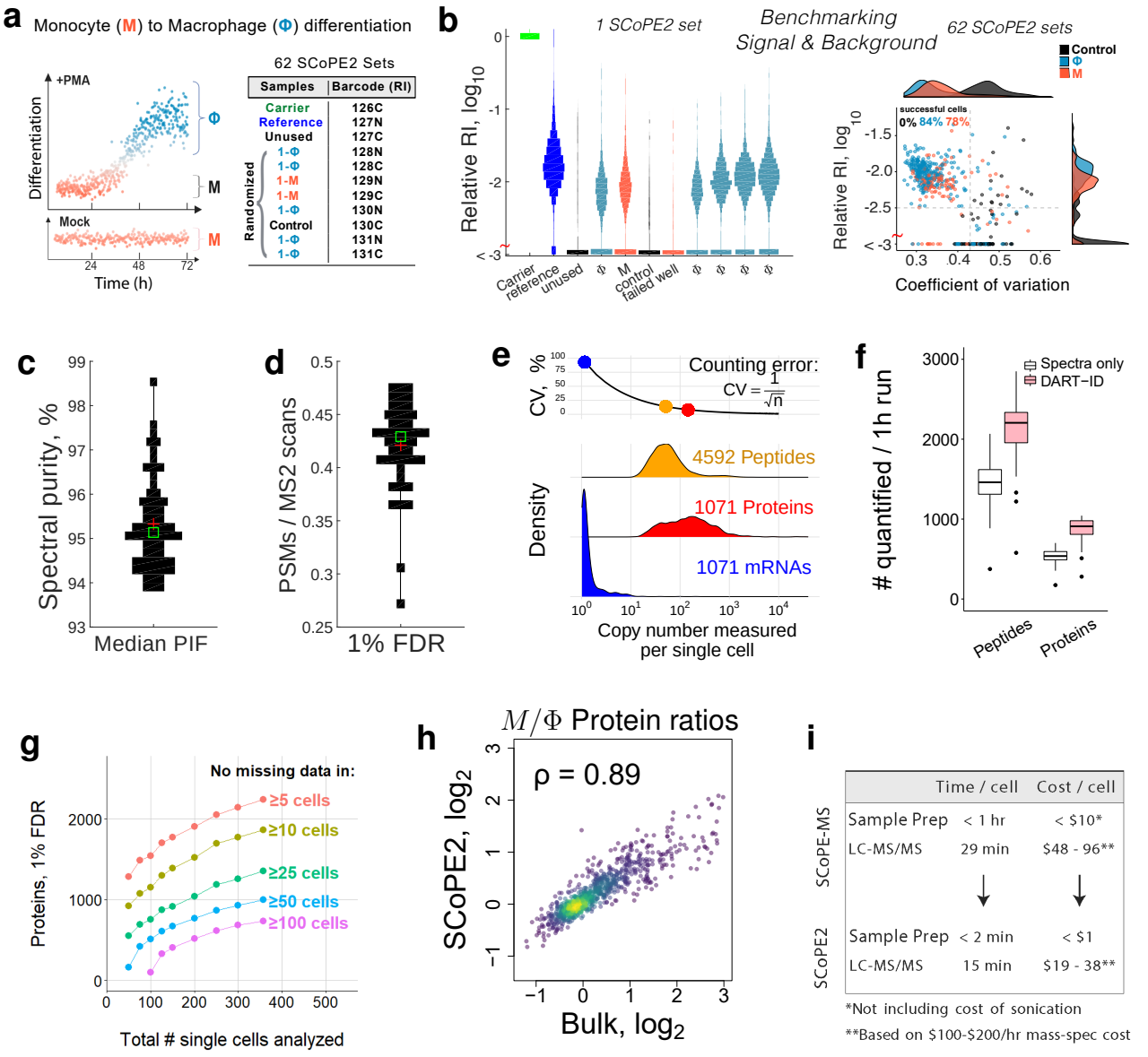


Figure 2 | **Proteins quantified in single cells** (**a**) Monocytes were differentiated into macrophages by PMA treatment, and FACS-sorted cells prepared into 62 SCoPE2 sets. (**b**) Distributions of reporter ion (RI) levels for a singe SCoPE2 set exemplify the detected signal from wells containing a cell – macrophages (blue) or monocytes (red) – and from a control well, which had no cell but was treated identically as the wells with cells. About 80% of the analyzed single cells have consistent quantification (lower CV) and higher RI signal than the control wells. (**c**) The Precursor-Ion-Fraction (PIF) for MS2 spectra is uniformly high. (**d**) Over 40% of MS2 spectra are assigned to peptide sequences at 1 % FDR. (**e**) Number of unique-barcode reads per mRNA[19] or ions per peptide/protein for a set of 1071 genes. The higher copy numbers measured for proteins support more reliable counting statistics compared to mRNAs. (**f**) Number of identified and quantified peptides and proteins in single cells from SCoPE2 sets analyzed on 60min nLC gradients. All identifications are shown with and without DART-ID[15], both at 1% FDR. (**g**) As the number of analyzed single cells increases, so does the number of proteins with complete quantification, i.e., without missing data. (**h**) The relative protein levels (monocyte / macrophage protein ratios) estimated from the single cells correlates strongly to the corresponding measurements from bulk samples. (**i**) Estimated costs (USD) and LC-MS time for SCoPE-MS and SCoPE2 workflows.

Quantifying proteins in single cells

Having demonstrated that proteins from 1xM standards can be quantified with low noise, we next applied SCoPE2 to the analysis of single cells. As a model system, we chose monocytes differentiating to macrophage-like cells in the presence of phorbol-12-myristate-13-acetate (PMA), Fig. 2a. We chose this system since it provides a clear benchmark – the ability to identify two closely related but distinct cell types. This system also presents an open research question: Are macrophage-like cells produced from this differentiation as homogeneous as the monocytes from which they originate or more heterogeneous? To answer this question independently from the heterogeneity inherent to primary monocytes, we used a homogeneous monocytic human cell line, U-937[20].

The SCoPE2 work-flow (Fig. 1a) can be used with manually picked cells, FACS-sorted cells or cells isolated by microfluidic technologies that minimize the volume of the droplets containing cells[7]. Here, we used a MoFlo Astrios EQ cell sorter to sort single cells into 384-well plates, one cell per well, Fig. 2a. The single cells were sorted and prepared into 62 SCoPE2 sets, as described in Fig. 1a and the Methods. The sorting followed a randomized layout to minimize biases, and sample preparation was automated as described in Methods.

First, we sought to identify single cells for which the peptide signal is (i) significantly above the background and (ii) internally consistent. To estimate the background signals, we incorporated control wells into SCoPE2 sets, Fig. 2a. These control wells did not receive cells but were treated identically to wells with single cells, i.e., received trypsin, TMT additions, and hydroxylamine. The distributions of RIs indicate that the majority of the peptides have over 10-fold lower RI in the control wells compared to the wells with single cells, Fig. 2b. Some single cells also have low RI signal, perhaps because of failures of sorting or sample preparation. As a measure for internal consistency, we computed the coefficient of variation (CV) for all peptides originating from the same protein, i.e., standard deviation / mean for the RI ratios relative to the reference channel. Then the median CV for all proteins from a SCoPE2 channel provides a measure for its consistency of relative protein quantification. Plotting each SCoPE2 channel – corresponding to a single cell or a control well – in the space of CVs and relative RIs reveals that about 80% of the single cells have higher RIs and lower CVs than the control wells, Fig. 2b. These 356 single cells were analyzed further.

Next, we explored the purity of the MS2 spectra shown in Fig. 2c. Thanks to the DO-MS optimization, all runs had spectra with purity significantly above 90% and over 40% of these spectra

could be assigned to a non-contaminant peptide sequence, Fig. 2d. This high rate of confident spectral assignment was aided by DART-ID.[15]

In developing and optimizing SCoPE2, we prioritized maximizing the number of ion copies used to quantify each protein in each single cell because a low copy-number of ions results in significant counting noise.[7] This noise arises because, similar to single-cell RNA-seq, SCoPE2 samples a subset of the molecules from a single cell. This sampling process contributes to a counting error: The standard deviation for sampling n copies is $\sqrt{n}$ (from the Poisson distribution), and thus the relative error, estimated as standard deviation over mean, is $\sqrt{n}/n = 1/\sqrt{n}$, Fig. 2e. Thus our optimization aimed to increase ion delivery not merely the number of identified peptides.[14]

To estimate our sampling error, we sought to convert the RI abundances (i.e., the barcodes from which SCoPE2 estimates peptide abundances, Fig. 1a) into ion copy numbers. To do so, we extracted the signal to noise ratios (S/N) for RIs and multiplied them by the number of ions that induces a unit change in S/N. For our Q-Exactive basic orbitrap operated at 70,000 resolving power, a S/N ratio of 1 corresponds to about 6 ions.[21,22] Thus for our system, a S/N ratio of 50 corresponds to 300 ions; see Methods for details. The results in Fig. 2e indicate that SCoPE2 samples 10-100 fold more copies per gene than single-cell RNA-seq,[19] which corresponds to much smaller sampling (counting) errors.

On average, SCoPE2 quantifies over 2,000 peptides corresponding to about 1,000 proteins per single cells per 1h gradient, Fig. 2f. While longer gradients can increase this coverage, they will reduce the number of cells analyzed per unit time. Since most single-cell analysis requires analyzing large number of single cells, we focused on short nLC gradients and on maximizing the number of proteins quantified across many cells. Indeed, we found that the number of peptides quantified across many cells, and thus suitable for biological analysis, increases with the number of analyzed cells, Fig. 2g.

To evaluate the accuracy of relative protein quantification in single cells, we computed the fold change between the average protein levels in monocytes and macrophages, averaging *in silico* across the single cells within a cell-type. These protein fold-changes from single cells were compared to the corresponding fold-changes measured from bulk samples, i.e., averaging across single cells by physically mixing their lysates, Fig. 2h. The strong correlation indicates that SCoPE2 can accurately measure protein fold-changes in single cells.

A key aim of SCoPE2 is to reduce cost and analysis time. This aim motivated many of our choices, including the use of commercial multiwell plates (as opposed to specialized tubes as

we did previously[12]), the use of multiplexing, and the reduction of nLC gradients to 1 hour. These efforts allowed SCoPE2 to reduce the cost and time for sample preparation by over 10-fold, Fig. 2i. It also reduced the LC-MS/MS time, and thus its cost, about 3 times. The estimated cost in Fig. 2i is based on MS facility fees and is lower for our in-house LC-MS/MS analysis.

Single-cell proteomes indicate increased heterogeneity of macrophages

Next we turn to the question of whether the homogeneous monocytes differentiated to similarly homogeneous macrophages, Fig. 2a. We sought to answer this question by applying unsupervised clustering methods to all quantified proteins, without filtering any proteins. Specifically, we wanted to characterize the cellular heterogeneity without assuming that cells fall into discrete clusters.

As a first and simplest approach, we performed principal component analysis (PCA) using all quantified proteins, Fig. 3a. The results indicated that the first PC accounts for over 60 % of the total variance and separates the cells into two mostly discrete clusters. Color-coding the cells by their labels indicates that the clusters correspond to the monocytes and the macrophages, Fig. 3a. The macrophage cluster appears more spread-out, suggesting that the differentiation increased the cellular heterogeneity. Indeed, the pairwise correlations between monocytes are significantly larger than those between macrophages, Fig. S1. To evaluate the abundance of monocyte and macrophage associated proteins within the single cells, we color-coded each cell by the median abundance of proteins identified to be differentially abundant from analyzing bulk samples of monocytes and macrophages, Fig. 3b.

Macrophages exibit a continuum of proteome states

As a second approach, we performed unsupervised spectral clustering of the cells. This network-based analysis allows to identify relationships between the cells based on analyzing the Laplacian matrix associated with the network of cells; see Methods. Since we did not want to assume discrete clusters of cells, we sorted the cells based on their corresponding elements in the Laplacian vector as defined by eq. 1 and shown in Fig. 3c. This analysis indicated again that the monocytes and macrophages form two mostly discrete clusters, with only a few cells having intermediate proteome states. Indeed, the elements of the Laplacian vector are bimodally distributed, Fig. 3d. This cell clustering is based on hundreds of proteins with differential abundance between monocytes and macrophage-like cells, shown in Fig. 3c: Proteins with higher abundance in monocytes are

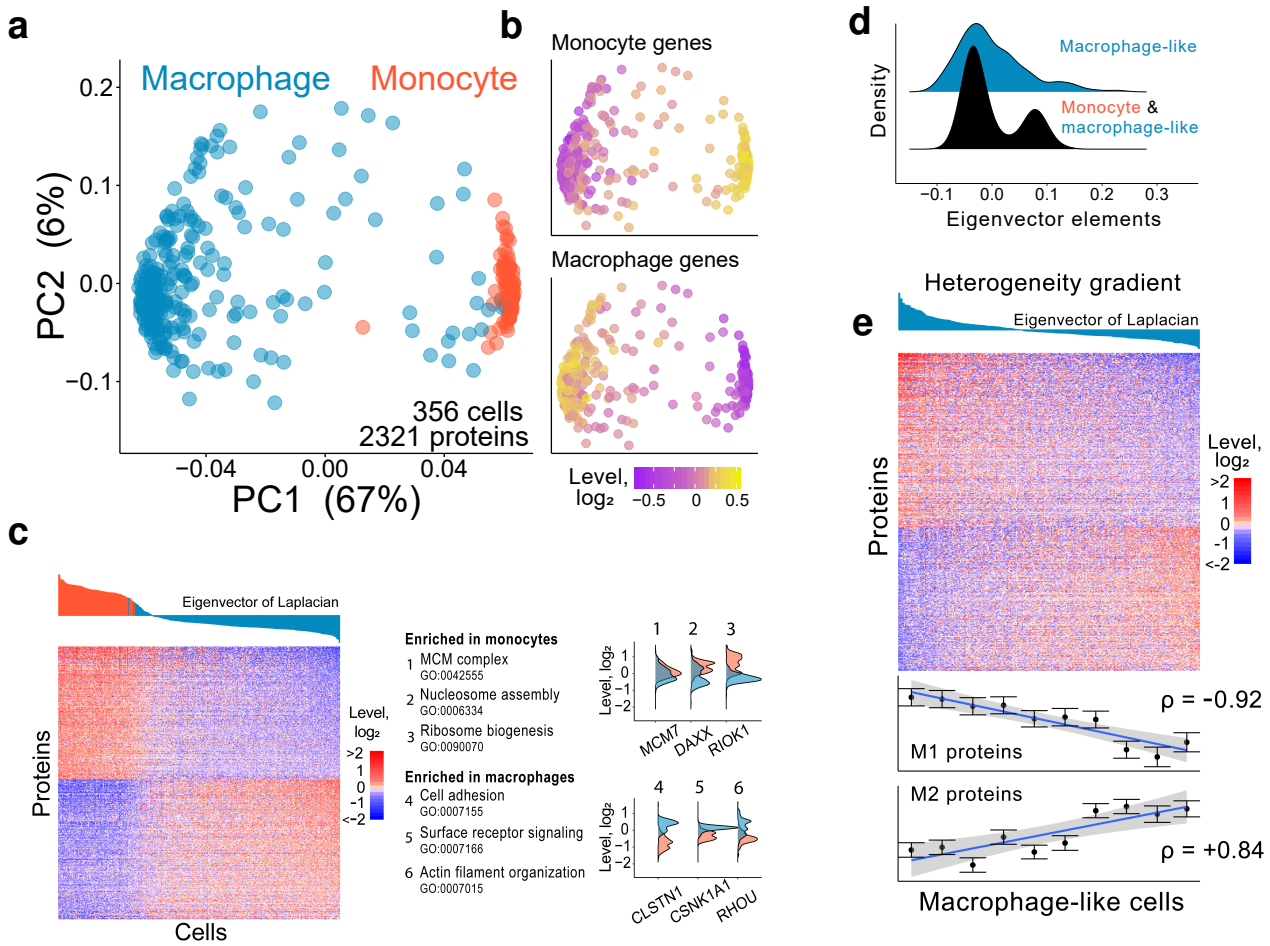


Figure 3 | Single-cell proteomes define a continuum of macrophage-like states (**a**) A weighted principal component analysis (PCA) of 356 single cells using all 2,321 proteins quantified across multiple single cells. Missing values were imputed using k-nearest neighbor (k=3). Cells are colored by cell type. (**b**) The cells from the PCA in panel a are colorcoded based on the abundance of monocyte and macrophage genes, defined as the 30 most differential proteins between bulk samples of monocytes and macrophages. (**c**) Heatmap of 464 proteins (the top 20 %) whose abundance varies the most between two clusters of cells identified by unsupervised spectral clustering of all quantified proteins and cells. The cells are permuted based on their rank in the corresponding Laplacian vector from the spectral clustering, see eq. 1. Gene set enrichment[23] identified overrepresented functions for the proteins enriched within each cell type. These functions are displayed alongside representative protein distributions from each gene set. (**d**) The distributions of elements of the Laplacian eigenvectors computed frrom eq. 1 for all cells (monocyte and macrophage-like cells) and for macrophage-like cells alone. (**e**) The unsupervised spectral analysis from panel c was applied only to the macrophage-like cells, revealing a gradient of macrophage heterogeneity. Cells were ordered based on the corresponding elements of the Laplacian eigenvector, eq. 1. The top 20% of proteins with the largest fold change between the first 40 cells and last 40 cells are displayed (464 proteins). Genes enriched in M1 or M2 polarized primary human macrophages[24] were identified in the single-cell data and their median value over bins of 26 cells (same order as in the above heatmap) displayed. Error bars denote standard error of data points in each bin.

enriched for proliferative functions, including the MCM complex involved in DNA replication and ribosome biogenesis[23]. Proteins with higher abundance in macrophages are enriched for immune

functions and cell adhesion proteins. Proteins from these functional groups – such as the transcription corepressor DAXX, the DNA replication licensing factor MCM7, and the Rho family of GTPases RHOU – have abundances strongly associated with either monocytes or macrophages, Fig. 3c. These enrichment results are consistent with the functional specialization of monocytes and macrophages and further validate the ability of SCoPE2 data to recapitulate known biology.

To explore the heterogeneity within the macrophages, we applied the same spectral analysis as in Fig. 3c, but this time only to the macrophage-like cells. Interestingly, the distribution of the elements of the associated Laplacian eigenvector (defined by eq. 1) is very broad and unimodal (Fig. 3d), suggesting that the cellular heterogeneity observed in this population is better described by a continuous spectrum rather than by discrete clusters. Indeed, the heatmap of protein levels for macrophage-like cells ordered based on the Laplacian eigenvector shows that most proteins change gradually across a continuous spectrum, Fig. 3e. Analyzing the proteins from this gradient, we observed a remarkable trend: Genes previously identified as differentially expressed between M1 and M2-polarized primary human macrophages[24] are also differentially expressed between single macrophage cells. For example, the cells at the left edge of Fig. 3e show high expression of genes upregulated in M1-polarized macrophages, decreasing monotonically from the left to right of Fig. 3e. Genes upregulated M2-polarized primary human macrophages appear to be expressed in a reciprocal fashion, with lower expression at the left edge of Fig. 3e, increasing monotonically across the figure.

Discussion

SCoPE2 enables scalable, robust and affordable quantification of about 1,000 proteins per single cell, and over 3,000 proteins across many cells. This coverage is achieved with 1 hour of analysis time per SCoPE2 set (about 15 min / cell), which allowed us to analyze hundreds of cells on a single instrument in just 85 hours. Most exciting for us, SCoPE2 succeeded in delivering and quantifying hundreds of ion copies from most detected proteins. This observation strongly supports the feasibility of single-cell LC-MS/MS protein quantification without amplification. Indeed, we benchmarked over 80% reliability[5] for measured protein fold-changes, Fig. 1c and Fig. 2h.

The reliability of data from SCoPE2 opens the potential not only to identify and classifying cell sub-populations, but to go beyond such descriptive analysis: We believe that accurate protein quantification across thousands of single cells may provide sufficient data for studying post-

transcriptional regulation in single cells and for inferring direct causal mechanisms in biological systems.[7]

To have such an impact, SCoPE2 analysis must be robust and accessible. A step in this direction is replacing the expensive and time-consuming lysis used by SCoPE-MS[12] with mPOP[13], Fig. 1a. Another step is DO-MS that makes it easier to implement and adapt SCoPE2 to different samples and LC-MS systems.[14] A further step is the analysis identifying successful cells shown in Fig. 2b. We believe that these steps bring us closer to the transformative opportunities of single-cell proteomics.[7]

We demonstrated that U-937-derived macrophages showed increased heterogeneity compared to the monocyte form, Fig. 3a. Having been exposed to identical environmental conditions, single macrophage cells exhibited coordinated protein level changes, Fig. 3e. In the absence of further treatment with polarizing cytokines or lipopolysaccharide to specifically induce macrophage polarisation[25], the differentiated macrophage population existed in a continuum, showing reciprocal loss or gain of proteins previously identified as enriched in M1 or M2 macrophages[24], Fig. 3d. This observation suggests that polarization might be a propensity inherent to macrophages.

Data Availability:

The raw MS data and the search results were deposited in MassIVE (ID: MSV000082841) and in ProteomeXchange (ID: PXD010856).

- **Facilitating LC-MS/MS evaluation:** To facilitate evaluation of our RAW LC-MS/MS data, we include detailed distribution plots generated by DO-MS.[14] These plots allow quick assessment of the nLC, ions detected at MS1 and MS2 level, apex offsets, identification rates and other important LC-MS/MS features.

- **Facilitating data reuse:** To facilitate reanalysis of or data, we also made them available in easily reusable form, including 3 files in comma separated values (csv) format as follows:

  1. `Peptides-raw.csv` – peptides $\times$ single cells at 1 % FDR and including peptides identified by DART-ID. The first 2 columns list the corresponding protein identifiers and peptide sequences and each subsequent column corresponds to a single cell.

2. `Proteins-processed.csv` – proteins × single cells at 1 % FDR, imputed and batch corrected.
3. `Cells.csv` – annotation × single cells. Each column corresponds to a single cell and the rows include relevant metadata, such as, cell type if known, measurements from the isolation of the cell, and derivative quantities, i.e., rRI, CVs, reliability.

Supplemental website can be found at: scope2.slavovlab.net

Acknowledgments: We thank R. G. Huffman, A.T. Chen, D. Perlman, A. Petelski, A. Marneros, M. Jovanovic, J. Alvarez, Z. Niziolek, Y. Katz, A. Andersen, and B. Karger, for assistance, discussions and constructive comments. This work was funded by startup funds from Northeastern University, a New Innovator Award from the NIGMS from the National Institutes of Health to N.S. under Award Number DP2GM123497, and an iAward from Sanofi to N.S.

Competing Interests: The authors declare that they have no competing financial interests.

Correspondence: Correspondence and materials requests should be addressed to nslavov@alum.mit.edu

Author Contributions

Experimental design: H.S., and N.S.
LC-MS/MS: H.S., and T.K.
Sample preparation: H.S., and E.E.
Raising funding & supervision: N.S.
Data analysis: H.S., and N.S.
Writing & editing: H.S., E.E, and N.S.

Methods

Cell culture Jurkat (T-cells) and U-937 cells (monocytes) were grown as suspension cultures in RPMI medium (HyClone 16777-145) supplemented with 10% fetal bovine serum (FBS) and 1% pen/strep. Cells were passaged when a density of 10^6 cells/ml was reached, approximately every two days. Monocytes were differentiated to macrophage-like cells by adding phorbol 12-myristate 13-acetate (5nM final concentration) to the culture medium for 24 hours, then washing the newly-adherent cells with fresh media and allowing to recover for a further 48 hours before harvest. Mock-treated U-937 cells were passaged with fresh media at 24 hours and harvested along with the treated cells at 72 hours.

Harvesting cells U-937 cells having undergone PMA-induced differentiation were washed twice with ice-cold phosphate buffered saline (PBS) and dissociated by scraping. Cell suspensions of Jurakt cells or undifferentiated U-937 cells were pelleted and washed quickly with cold PBS at 4 ^oC. The washed pellets were diluted in PBS at 4 ^oC. The cell density of each sample was estimated by counting at least 150 cells on a hemocytometer.

Sample randomization and sorting SCoPE2 sets were designed such that, on average, there would be 5 single macrophages, 2 single monocytes and 1 control well per set. These were randomized over a 384-well plate such that there would be 32 SCoPE2 sets produced per plate. Single U-937 monocyte and macrophage-like cells were isolated and distributed into 384-well PCR plates using a Beckman Coulter MoFlo Astrios EQ Cell Sorter into $1\mu l$ of pure water with MassPREP peptides in 384-well PCR plates (ThermoFisher AB1384). Carrier channels containing 100 cells of each type were sorted likewise onto the same plate. The reference channel was prepared separately by sorting 10,000 cells of each type into a $500\mu\, l$ eppendorf tube and used in all sets across multiple plates.

SCoPE2 sample preparation Single cells and carrier cells were lysed by freezing at -80 ^oC for at least 5 minutes and heating to 90 ^oC for 10 minutes. Then, samples were centrifuged briefly to collect liquid; trypsin (Promega Trypsin Gold) and buffer triethylammonium bicarbonate (TEAB) (pH 8.5) were added to $10\ ng/\mu l$) and 100mM, respectively. The samples were digested for 4 hours in a thermal cycler at 37 ^oC (BioRad T100). Samples were cooled to room temperature and labeled with $1\ \mu l$ of 22mM TMT label (TMT11 kit, ThermoFisher, Germany) for 1 hour. The unreacted TMT label in each sample was quenched with $0.5\ \mu l$ of 0.5% hydroxylamine for 45 minutes at room temperature. Samples were centrifuged briefly following all reagent additions to collect liquid. The samples corresponding to one TMT11 plex were then mixed in a single glass

HPLC vial and dried down to $1\ \mu l$ in a speed-vacuum (Eppendorf, Germany) at 35^oC.

1xM standard preparation Jurkat and U-937 cells were harvested and counted as described above. Five thousand three hundred cells from each type were digested (100mM TEAB pH 8.5, $10\ ng/\mu l$ trypsin at 37 ^oC for 4 hours), divided into 5000, 100, 100, and 100 cell equivalents, labeled with TMT11, and combined such that there are two carrier channels of 5000 cell equivalents (one of Jurkat, one of U-937) and six channels of 100 cell equivalents, three of Jurkat and three of U-937 (Supplementary information). This sample was diluted 100x and aliquoted into glass HPLC vials. Material equivalent to 50 cells in the two carrier channels and 1 cell in the six other channels was injected for analysis by LC-MS/MS.

SCoPE2 Mass spectrometry analysis SCoPE-MS samples were separated via online nLC on a Dionex UltiMate 3000 UHPLC; $1\mu\ l$ out of $1.2\mu\ l$ of sample was loaded onto a 25cm x 75um IonOpticks Aurora Series UHPLC column (AUR2-25075C18A). Buffer A was 0.1% formic acid in water and buffer B was 0.1% formic acid in 80% acetonitrile / 20% water. A constant flow rate of 200nl/min was used throughout sample loading and separation. Samples were loaded onto the column for 20 minutes at 1% B buffer, then ramped to 5% A buffer over two minutes. The active gradient then ramped from 5% B buffer to 25% B buffer over 53 minutes. The gradient then ramped to 95% B buffer over 2 minutes and stayed at that level for 3 minutes. The gradient them dropped to 1% B buffer over 0.1 minute and stayed at that level for 4.9 minutes. Loading and separating each sample took 95 minutes total. All samples were analyzed by a Thermo Scientific Q-Exactive mass spectrometer from minute 20 to 95 of the LC loading and separation process. Electrospray voltage was set to 2,200V, applied at the end of the analytical column. To reduce atmospheric background ions and enhance peptide signal to noise ratio, an Active Background Ion Reduction Device (ABIRD, by ESI Source Solutons, LLC, Woburn MA, USA) was used at the nanospray interface. The temperature of ion transfer tube was 250 degrees Celsius and the S-lens RF level set to 80. After a precursor scan from 450 to 1600 m/z at 70,000 resolving power, the top 5 most intense precursor ions with charges 2 to 4 and above the AGC min threshold of 20,000 were isolated for MS2 analysis via a 0.7 Th isolation window. These ions were accumulated for at most 300ms. Then they were fragmented via HCD at a normalized collision energy of 33 eV (normalized to m/z 500, z=1) and the fragments analyzed at 70,000 resolving power. Dynamic exclusion was used with a duration of 30 seconds with a mass tolerance of 10ppm.

Analysis of raw MS data Raw data were searched by MaxQuant[10,18] 1.6.2.3 against a protein sequence database including all entries from the human SwissProt database (downloaded July 30,

2018; 20,373 entries) and known contaminants such as human keratins and common lab contaminants (default MaxQuant contaminant list). MaxQuant searches were performed using the standard work flow[26]. We specified trypsin/P digestion and allowed for up to two missed cleavages for peptides having from 7 to 25 amino acids. Tandem mass tags (11plex TMT) were specified as fixed modifications. Methionine oxidation (+15.99492 Da), asparagine deamidation (+0.9840155848 Da), protein N-terminal acetylation (+42.01056 Da) were set as a variable modifications. As alkylation was not performed, carbamidomethylation was disabled as a fixed modification. Second peptide identification was disabled. Calculate peak properties was enabled. All peptide-spectrum-matches (PSMs) and peptides found by MaxQuant were exported in the evidence.txt files. False discovery rate (FDR) calculations were performed in the R programming language environment.

DART-ID search Seventy-six replicate injections of the 1xM standard and 62 SCoPE2 sets were analyzed by DART-ID[15]. A configuration file for the search is included in Supplementary information.

Data filtering All subsequent data analysis was performed in the R (v3.5.2) programming language environment. All code used is available at https://github.com/SlavovLab/SCoPE2. The MaxQuant evidence.txt (with identification confidence updated by DART-ID) was filtered for posterior error probabilities (PEP) < 0.02, protein FDR $< 1\%$, reverse hits, contaminant hits, and precursor intensity fraction (PIF) > 0.8. Peptides with $> 10\%$ reporter ion intensity of the carrier channel reporter ion intensity (if carrier was used in that set) were removed.

Single cell data filtering Single cells with a median reporter ion intensity $> 5\%$ of its set's carrier channel were removed. Relative reporter ion intensities (rRI) were calculated for each peptide in each single cell relative to its set's carrier channel. The internal consistency of protein quantification for each single cell was evaluated by calculating the mean coefficient of variation for proteins (Leading razor proteins) identified with > 5 peptides for that cell. Coefficient of variation is defined as the standard deviation divided by the mean. Here the standard deviation was calculated for each protein in each single cell as the standard deviation of rRI of the 5+ peptides belonging to the same protein divided by the mean of the 5+ rRIs. The mean of the CVs was taken per single cell. Control wells were used to determine a reasonable cutoff value of CV, below which we could have higher confidence that that channel truly contained cellular material and not just signal from noise or contamination. Relative reporter ion intensities from the control wells were used to determine a reasonable rRI cutoff value, above which we could have higher confidence that the channel truly contained cellular material, and was not empty due to miss-sorting or failures during sample

preparation.

Data transformations After filtering, the peptide-level reporter ion intensities for the remaining single cells were arranged into a matrix of peptides x single cells (rows x columns). All single cell reporter ion intensities were normalized to the reference channel intensities in their respective sets. The columns then the rows were normalized by their respective median values. Rows with greater than 99% missing data were removed, then columns with greater than 99% missing data were removed. The values in the matrix were log_2 transformed, then the protein level quantification was calculated by mapping each peptide to its respective (leading razor) protein and taking the median value if there was more than 1 peptide mapped to that protein. Now the matrix has dimensions proteins x single cells (rows x columns). The data was again normalized by subtracting the column then row medians.

Imputing missing values Missing values in the protein x single cell matrix were imputed by k-nearest neighbor imputation ($k = 3$) using euclidean distance as a similarity measure between the cells.

Weighted principal component analysis The protein x single cell matrix is called $\mathbf{M}$. The fraction missing data for the ith protein, w_i, was calculated as the fraction missing data in all peptides belonging to that protein across all filtered single cells. The covariance matrix was calculated as $\frac{1}{\sum_i w_i}\mathbf{M}^T\mathbf{W}\mathbf{M}$, where $\mathbf{W} = diag(w_i)$. Principal component analysis was performed on this weighted matrix.

Spectral clustering of cells Spectral clustering was performed by first computing a matrix of positive pairwise weights, $\mathbf{W}$ between all cells. We defined the weight between two cells to be their Pearson correlation plus 1 so that all weights were positive. Then the Laplacian matrix is $\mathbf{L} = \mathbf{D} - \mathbf{W}$, where $\mathbf{D}$ is a diagonal matrix whose diagonal elements contain the sum of elements in the corresponding rows of $\mathbf{W}$, i.e., $D_{i,i} = \sum_j W_{i,j}$. Then trivially, the smallest eigenvalue of $\mathbf{L}$ is 0, and its corresponding eigenvector is the constant vector, e.g., the vector of ones. The second smallest eigenvalue corresponds to the non-constant vector $\mathbf{v}$ that minimizes eq. 1.

$$\mathbf{v}^T\mathbf{L}\mathbf{v} = \frac{1}{2}\sum_{i,j} w_{ij}(v_i - v_j)^2, \text{ so that } \mathbf{v}^T\mathbf{v} = 1 \tag{1}$$

Thus, computing this eigenvector corresponds to global convex optimization that assigns similar $\mathbf{v}$ elements to cells connected by high weights. The Laplacian vector used for sorting cells in Fig. 3 was the eigenvector with the smallest non-zero eigenvalue.

Single cell RNA-seq data Unique molecular identifier counts for monocyte cells were extracted from Kang et al. following the Seurat immune alignment tutorial available at satijalab.org[19].

Converting signal-to-noise to ion counts Signal-to-noise (S/N) was extracted from the raw files by Proteome Discoverer v2.3. Ion counts were calculated by multiplying the S/N by an estimate for the number of ions that induces a unit change in S/N for our instrument. This factor was estimated as 3.5 for an orbitrap at a resolving power of $240,000$. This factor scales with the square root of the ratio of the resolving power, so for our instrument run at $70,000$ resolving power, the number of ions that induces a unit change in S/N is $3.5 \times \sqrt{\frac{240{,}000}{70{,}000}} = 6.5$[21,22].

Gene set enrichment analysis Gene set enrichment analysis was performed using the online tool GOrilla[23]. Genes (proteins) were ordered by fold change between their means in each cluster (for comparing monocytes and macrophage-like cells) or by fold change between the mean of the first 40 and last 40 cells as ordered by spectral clustering (for comparing macrophage-like cells).

Determining genes enriched in monocytes and macrophage-like cells from bulk proteomic data For each protein profiled by bulk proteomic methods a two-sided t-test was performed comparing the relative protein level between the two cell types (20 replicates per cell type). Fold change between the two cell types was calculated by taking the difference in means and the top 60 most differential proteins (30 "up-regulated" in monocytes, 30 "up-regulated" in macrophage-like cells) with a p-value less than 0.01 from the t-test were taken. This list of proteins constitute the "monocyte genes" and the "macrophage genes" displayed in Fig. 3b.

Supplemental Figures

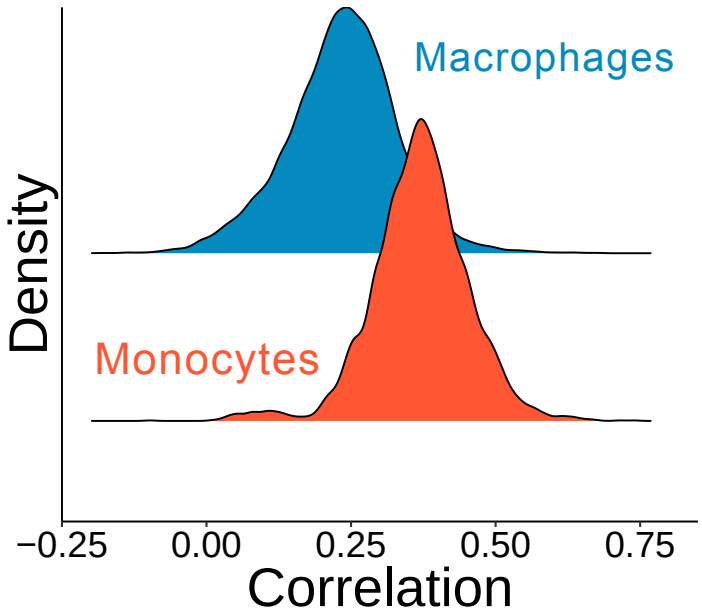


Figure S1 | Distributions of pairwise correlations within cell types.
To evaluate the homogeneity of monocytes and macrophages, we computed and plotted all pairwise Pearson correlations within each cell type, i.e., all pairwise correlations between monocytes and all pairwise correlations between macrophage-like cell. The correlations are computed from the fold-changes relative to the reference so that they do not conflate variability between different proteins and emphasize the differences between the cells.[5] The correlations between macrophage-like cell are significantly lower than between monocytes, consistent with the more dispersed macrophage cluster in Fig. 3a. These results indicate that the macrophage-like cells are more heterogeneous than the monocytes from which they differentiate.

References

1. Levy, E. & Slavov, N. Single cell protein analysis for systems biology. *Essays In Biochemistry* **62.** doi:10.1042/EBC20180014 (4 2018).
2. Shapiro, E., Biezuner, T. & Linnarsson, S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nature Reviews Genetics* **14,** 618 (2013).
3. Grün, D., Kester, L. & Van Oudenaarden, A. Validation of noise models for single-cell transcriptomics. *Nature methods* **11,** 637 (2014).
4. Liu, Y., Beyer, A. & Aebersold, R. On the dependency of cellular protein levels on mRNA abundance. *Cell* **165,** 535–550 (2016).

5. Franks, A., Airoldi, E. & Slavov, N. Post-transcriptional regulation across human tissues. *PLoS computational biology* **13,** e1005535 (2017).

6. Ben-Moshe, S. & Itzkovitz, S. Spatial heterogeneity in the mammalian liver. *Nature Reviews Gastroenterology & Hepatology,* 1 (2019).

7. Specht, H. & Slavov, N. Transformative opportunities for single-cell proteomics. *Journal of Proteome Research* **17,** 2563–2916 (8 June 2018).

8. Milo, R., Jorgensen, P., Moran, U., Weber, G. & Springer, M. BioNumbers-the database of key numbers in molecular and cell biology. *Nucleic acids research* **38,** D750–D753 (2010).

9. Cravatt, B. F., Simon, G. M. & Yates Iii, J. R. The biological impact of mass-spectrometry-based proteomics. *Nature* **450,** 991 (2007).

10. Cox, J. *et al.* Andromeda: a peptide search engine integrated into the MaxQuant environment. *Journal of proteome research* **10,** 1794–1805 (2011).

11. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. en. *Nature* **537,** 347–355. ISSN: 1476-4687 (Sept. 2016).

12. Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass-spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biology* **19,** 161 (2018).

13. Specht, H. *et al.* Minimal sample preparation for high-throughput proteomics. *bioRxiv*. doi:10.1101/399774 (2019).

14. Huffman, G., Chen, A. T., Specht, H. & Slavov, N. DO-MS: Data-Driven Optimization of Mass Spectrometry Methods. *J. of Proteome Res.* doi:10.1021/acs.jproteome.9b00039 (2019).

15. Chen, A., Franks, A. & Slavov, N. DART-ID increases single-cell proteome coverage. *PLoS Comput Biol.* doi:10.1371/journal.pcbi.1007082 (2019).

16. Martinez, F. O. & Gordon, S. The M1 and M2 paradigm of macrophage activation: time for reassessment. *F1000Prime Reports* **6,** 13 (2014).

17. Ginhoux, F., Schultze, J. L., Murray, P. J., Ochando, J. & Biswas, S. K. New insights into the multidimensional concept of macrophage ontogeny, activation and function. *Nature Immunology.* **17,** 34–40 (2016).

18. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* **26,** 1367–1372 (2008).

19. Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nature Biotechnology* **36,** 89–94 (2018).

20. Sundstrom, C. & Nilsson, K. Establishment and characterization of a human histiocytic lymphoma cell line (U-937). *Int. J. Cancer* **17,** 565–577 (1976).

21. Eiler, J. *et al.* Analysis of molecular isotopic structures at high precision and accuracy by Orbitrap mass spectrometry. *International Journal of Mass Spectrometry* **422,** 126–142 (2017).

22. Makarov, A. & Denisov, E. Dynamics of ions of intact proteins in the Orbitrap mass analyzer. *Journal of the American Society for Mass Spectrometry* **20,** 1486–1495 (2009).

23. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10,** 48 (2009).

24. Martinez, F. O., Gordon, S., Locati, M. & Mantovani, A. Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. eng. *Journal of Immunology (Baltimore, Md.: 1950)* **177,** 7303–7311. ISSN: 0022-1767 (Nov. 2006).

25. Abdulhadi, F. *Differentiation of U-937 Monocytes to Macrophage-Like Cells Polarized into M1 or M2 Phenotypes According to Their Specific Environment: A Study of Morphology, Cell Viability, and Cd Markers of an In Vitro Model of Human Macrophages.* MA thesis (Wright State University, 2014).

26. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature protocols* **11,** 2301 (2016).