Real-Time Search Assisted Acquisition on a Tribrid Mass Spectrometer Improves Coverage in Multiplexed Single-Cell Proteomics

Benjamin Furtwängler[*†‡§], *Nil Üresin*[†‡], *Khatereh Motamedchaboki*[¶], *Romain Huguet*[¶], *Daniel Lopez-Ferrer*[¶], *Vlad Zabrouskov*[¶], *Bo T. Porse*[†‡§], *Erwin M. Schoof*[*‖]

[†]The Finsen Laboratory, Rigshospitalet, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark

[‡]Biotech Research and Innovation Centre (BRIC), University of Copenhagen, Copenhagen, Denmark

[§]Novo Nordisk Foundation Center for Stem Cell Biology, DanStem, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark

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

[¶]ThermoFisher Scientific, San Jose, CA, USA

*Corresponding Authors:

Benjamin Furtwängler, email: benjamin.furtwangler@bric.ku.dk

Erwin M. Schoof, email: erws@dtu.dk

KEYWORDS

single-cell proteomics, isobaric tag quantification, multiplexing, SPS-MS3, TMT, real-time-search

ABSTRACT

In the young field of single-cell proteomics (scMS), there is a great need for improved global proteome characterization, both in terms of proteins quantified per cell and quantitative performance thereof. The recently introduced real-time search (RTS) on the Orbitrap Eclipse Tribrid mass spectrometer in combination with SPS-MS3 acquisition has been shown to be beneficial for the measurement of samples that are multiplexed using isobaric tags. Multiplexed single-cell proteomics requires high ion injection times and high-resolution spectra to quantify the single-cell signal, however the booster channel facilitates peptide identification and thus offers the opportunity for fast on the fly peptide identification before committing to the time intensive quantification scan. Here, we compared classical MS2 acquisition against RTS-SPS-MS3, both using the Orbitrap Eclipse Tribrid MS with the FAIMS Pro ion mobility interface and we present a new acquisition strategy that makes use of fast real-time searched linear ion trap scans to preselect MS1 peptide precursors for quantitative MS2 Orbitrap acquisition. Here we show that this strategy, termed RETICLE (**R**TS **E**nhanced Quan**t** of S**i**ngle **C**ell Sp**e**ctra) outperforms MS2 and RTS-SPS-MS3 in multiplexed scMS applications and enables the quantification of over 1000 proteins per cell at a MS2 injection time of 750ms using a 2h gradient, resulting in improved quality of scMS datasets.

INTRODUCTION

Recent developments in liquid chromatography coupled mass spectrometry (LC-MS) based proteomics revealed the high potential for its application on single cells[1–9]. A major breakthrough was the introduction of single-cell proteomics using mass spectrometry (scMS) via the SCoPE[1,2] method, where isobaric labeling is used to multiplex single-cells to be then measured in a single LC-MS run. Importantly, the addition of a booster channel (200-cell equivalent) provides significantly more peptides copies in addition to the single-cell channels and thus facilitates precursor ion identification in the fragmentation scan (MS2). Hence, the increased throughput permitted through isobaric multiplexing, which currently supports up to 18 channels with TMTpro[10,11], and the signal boosting effect of the booster channel, provide considerable advantages over label-free scMS. Nonetheless, multiplexed scMS suffers from its own limitations. For one, the booster channel decreases the quantitative performance of the adjacent channels due to signal spillover from isotopic impurities of the isobaric labels and furthermore increases the overall noise level relative to the very low abundant single-cell channels[5,9,12,13]. Moreover, long injection times are required to collect enough ions from the single-cell channels for robust ion statistics and accurate reporter ion quantification, resulting in slow scanning speed and consequently lower proteome coverage. Moreover, isobaric reporter ion quantification suffers from the well-studied problem of ratio-compression due to interference of co-isolated precursor ions[14,15]. This interference can be mitigated by applying gas-phase fractionation via the FAIMS interface[16,17], which reduces the complexity of the precursor ion stream entering the MS and removes +1 ion species (i.e. non-peptide contaminants). Even greater reduction of interference is achieved with the gas-phase purification of the MS2 fragment ions and subsequent secondary fragmentation to release the isobaric tag for quantification in a MS3 scan[15]. This feature is unique

to Tribrid MS instruments as synchronous precursor selection (SPS) is performed in the linear ion trap (LIT). Since the high resolution scan in the Orbitrap (OT) is only needed to measure accurate precursor masses in the MS1 scan and to resolve the reporter ions for quantification in the MS3 scan, peptide identification via the MS2 scan can be performed in the more sensitive LIT[18]. This results in fast and sensitive MS2 acquisition, as the Tribrid design enables parallelization of OT and LIT scans. Additionally, with the introduction of real-time search (RTS)[19,20] on the Orbitrap Eclipse Tribrid MS[21], only MS1 precursors that were identified as peptides of interest are subjected to the time consuming MS3 quantification, resulting in higher proteome coverage. Furthermore, since the MS1 precursor is identified, only the peptide fragments belonging to that peptide are subjected to SPS-MS3, reducing the co-isolation to a minimum and thus maximizing accuracy. Nonetheless, SPS-MS3 is less sensitive than MS2 quantification, because of the associated loss of ions during SPS and the second fragmentation step. This tradeoff between (i) increased accuracy through reduced interference and (ii) reduced accuracy through lower sensitivity favors RTS-SPS-MS3 in high-load applications[21]; however, it has yet to be investigated for the low abundant samples in multiplexed scMS where much higher ion injection times (ITs) are needed.

We recently presented our scMS workflow[7] and benchmarked the quantitative performance using the OCI-AML8227 cell-culture model[22]. This model maintains the hierarchical nature of Acute Myeloid Leukemia (AML) where a small population of self-renewing leukemic stem cells (LSC) differentiate to progenitors (PROG) and finally to terminally differentiated blasts (BLAST). Thus, this model system provides three distinct cell differentiation stages with differences on proteome level detectable by scMS. Furthermore, we investigated the protein fold-changes between these differentiation stages via bulk proteomics to a depth of nearly 7000 proteins, providing us with a reference set to benchmark scMS data in terms of quantitative accuracy.

In this work, we compared the performance of multiplexed scMS using MS2 acquisition against RTS-SPS-MS3, both using the Orbitrap Eclipse Tribrid MS with the FAIMS Pro ion mobility interface. Furthermore, we present a new acquisition strategy that makes use of fast real-time searched LIT scans to preselect MS1 peptide precursors for quantitative MS2 OT acquisition. Here we show that this strategy, termed RETICLE (**R**TS **E**nhanced Quan**t** of S**i**ngle **C**ell Sp**e**ctra) outperforms MS2 and RTS-SPS-MS3 in multiplexed scMS applications and enables the quantification of over 1000 proteins per cell at a MS2 injection time of 750ms using a 2h gradient, resulting in improved quality of scMS datasets.

EXPERIMENTAL PROCEDURES

Cell Culture and FACS Sorting

OCI-AML8227 cells were grown in StemSpan SFEM II media, supplemented with growth factors (Miltenyi Biotec, IL-3, IL-6 and G-CSF (10 ng/mL), h-SCF and FLt3-L (50 ng/mL), and TPO (25 ng/mL) to support the hierarchical nature of the leukemia hierarchy captured within the cell culture system. On day 6, cells were harvested (8e6 cells total), washed, counted, and resuspended in fresh StemSpan SFEM II media on ice at a cell density of 5e6 cells/ml. Staining was done for 30 mins on ice, using a CD34 antibody (CD34-APC-Cy7, Biolegend, clone 581) at 1:100 (vol/vol), a CD38 antibody (CD38-PE, BD, clone HB7) at 1:50 (vol/vol). Cells were washed with extra StemSpan SFEM II media, and subsequently underwent three washes with ice cold PBS to remove any remaining growth factors or other contaminants from the growth media. Cells were resuspended for FACS sorting in fresh, ice cold PBS at 2e6 cells/ml and stained with 7-AAD viability dye (1 ug/mL, Invitrogen). Cell sorting was done on a FACSAria III instrument, controlled by the DIVA software package (v.8.0.2) and operating with a 100 µm nozzle. Cells

from 3 different gates (CD34+CD38-, CD34+CD38+, CD34-) (**Supplementary Fig. 1**) were sorted at single-cell resolution into a 384-well Eppendorf LoBind PCR plate (Eppendorf AG) containing 1 µl of lysis buffer (50 mM Triethylammonium bicarbonate (TEAB) pH 8.5, 20% 2,2,2-Trifluoroethanol (TFE)). Directly after sorting, plates were briefly spun, snap-frozen on dry ice, and then boiled at 95 °C in a PCR machine (Applied Biosystems Veriti 384-well) for 5 mins. Plates were again snap-frozen on dry ice and stored at -80 °C until further sample preparation. The same procedure was followed for the booster plate, but instead of sorting single cells, 500 cells were sorted in four-way purity mode into each well without immunofluorescent preselection.

Sample Preparation of Diluted Standard

Peptide concentration of the TMTpro labeled OCI-AML8227 sample that was used to measure the MS3 reference library as previously described[7] was determined via Nanodrop and the sample was subsequently diluted to contain 250 pg of peptide in each of the 9 channels per injection. A bulk-sorted OCI-AML8227 booster was added as a 200-cell equivalent per injection. For the comparison of the close-out options, samples with a 100-cell equivalent booster were used.

Sample Preparation of Single-Cell Samples

After thawing, protein lysates from the single cells were digested with 2 ng of Trypsin (Sigma cat. nr. T6567), dissolved in 1 µl of 100 mM TEAB pH 8.5 containing Benzonase (Sigma cat. nr. E1014) diluted 1:5000 (vol/vol) to digest any DNA that would interfere with downstream processing. For the booster plates, the amount of trypsin was increased to 10 ng in order to digest the protein content of each well containing 500 cells. Plates were kept at 37 °C overnight to complete the protein digestion. All dispensing steps in this protocol were done using the Dispendix

I-DOT One instrument. After digestion, peptides were labeled with TMTPro reagents. 6 mM in 1 µl acetonitrile (ACN) of each label was added to the single-cell wells, while the 500-cell booster plate wells were labeled with 13 mM of TMTPro-126 reagent in each well. Subsequently, plates were kept at RT for 1 h. The labeling reaction was quenched with 1 µl of 1.25% or 2.5% Hydroxylamine for the single-cell and booster plate respectively for 15 min. Subsequently, the booster plate was pooled and desalted using a SOLAµ HRP 96-well plate. Eluted and desalted peptides were concentrated to dryness in an Eppendorf Speedvac, after which they were resuspended in A* (2% ACN, 0.1% TFA. The single-cell samples were then mixed from 14 single cells plus the equivalent of 200 booster channel cells. To ensure a balanced design, each sample contained 4 cells of CD34- and 5 cells of CD34+CD38- and CD34+CD38+ respectively. This pooling was performed using the Opentrons OT-2 liquid handler. The resulting peptide mix was concentrated in an Eppendorf Speedvac, and re-constituted in A* for individual Mass Spectrometry analysis.

Mass Spectrometry Data Acquisition of Diluted Standard

Peptides were loaded onto a µPAC trapping column (PharmaFluidics) at 5 ul/min, connected in a forward-flush configuration to a 50 cm µPAC analytical column (PharmaFluidics), with 100% Buffer A (0.1% Formic acid in water) using the UltiMate 3000 RSLCnano System (ThermoFisher), and the column oven operating at 35 °C. Peptides were eluted over a 90 min gradient, using 80% Acetonitrile, 0.1% Formic acid (Buffer B) going from 8% to 18% over 8 min at a flowrate of 500 nl/min, to 19% over 1 min with a simultaneous reduction of flowrate down to 150 nl/min, which was kept at this rate from there on. Then to 28% over 31 min, to 38% over 20 min, to 60% over 6 min, to 95% over 1 min, and kept at 95% for 23 min, after which all peptides

were eluted. Spectra were acquired with an Orbitrap Eclipse Tribrid Mass Spectrometer with FAIMS Pro Interface (ThermoFisher Scientific) running Tune 3.4 and Xcalibur 4.3. For all acquisition methods, FAIMS switched between CVs of -50 V and -70 V with cycle times of 3s and 1.5s, respectively. MS1 spectra were acquired at 120,000 resolution with a scan range from 375 to 1500 m/z, normalized AGC target of 300% and maximum injection time of 50ms. Precursors were filtered using monoisotopic peak determination set to peptide, charge state 2-6, dynamic exclusion of 120s with ±10 ppm tolerance excluding isotopes and different charge states, and a precursor fit of 70% in a windows of 0.7 m/z for MS2 and RETICLE and 60% in a 0.8 m/z window for RTS-MS3. Additionally, RTS-MS3 employed an intensity threshold of 5e3. Precursors selected for MS2 analysis were isolated in the quadrupole with a 0.7 m/z window for MS2 and RETICLE and 0.8 m/z for RTS-MS3. In the MS2 method, ions were collected for a maximum injection time of 500ms or 750ms and normalized AGC target of 500%, fragmented with 32 normalized HCD collision energy and MS2 spectra were acquired in the Orbitrap at 120,000 resolution with first mass set to 120 m/z. In the RTS-MS3 and RETICLE methods, ions were collected for a maximum injection time set to auto for RTS-MS3 and 23ms for RETICLE and normalized AGC target of 300% for both methods. Fragmentation was performed with 30 normalized CID collision energy and MS2 spectra were acquired in the LIT at scan rate rapid for RTS-MS3 and turbo for RETICLE. For the RETICLE method, LIT MS2 spectra were subjected to RTS using the homo sapiens database obtained from Uniprot (Swiss-Prot with isoforms) and Trypsin set as enzyme. Static modifications were TMTpro16plex on Lysine (K) and N-Terminus, and Carbamidomethyl on Cysteine (C). Oxidation on Methionine (M) was set as variable modification. Maximum missed cleavages was set to 0 and maximum variable mods to 1. FDR filtering was enabled, maximum search time was set to 40ms, and the scoring threshold was set to

1.4 XCorr, 0.1 dCn and 10 ppm precursor tolerance. Use as trigger only was checked and close-out was enabled with maximum number of peptides per protein set to 4. For the comparison of the close-out options, this number was either set to 4, 10 or close-out was disabled. Precursors identified via RTS were isolated in the quadrupole with a 0.7 m/z window, ions were collected for a maximum injection time of 500ms or 750ms and normalized AGC target of 500%, fragmented with 32 normalized HCD collision energy and MS2 spectra were acquired in the Orbitrap at 120,000 resolution with first mass set to 120 m/z. For the RTS-MS3 method, LIT MS2 spectra were subjected to RTS using the same settings as in RETICLE with differences being that maximum missed cleavages was set to 1, maximum variable mods to 2, maximum search time to 35ms and use as trigger only was deactivated, but TMT SPS MS3 Mode was activated. Subsequently, spectra were filtered with a precursor selection range filter of 400-1600 m/z, isobaric tag loss exclusion of TMTpro and precursor mass exclusion set to 5 m/z low and 3 m/z high. Precursors identified via RTS were isolated for a MS3 scan using the quadrupole with a 2 m/z window and ions were collected for a maximum injection time of 500ms or 750ms and normalized AGC target of 500%. SPS was activated and number of SPS precursors was set to 10. Isolated fragments were fragmented again with 45 normalized HCD collision energy and MS3 spectra were acquired in the Orbitrap at 120,000 resolution with a scan range of 100-500 m/z. Method trees are shown in **Supplementary Fig. 2**.

Mass Spectrometry Data Acquisition of Real scMS Samples

Real scMS samples were measured as described for the diluted standard with the following changes. Peptides were eluted over a 116 min gradient, using 80% Acetonitrile, 0.1% Formic acid (Buffer B) going from 8% to 15% over 8 min at a flowrate of 500 nl/min, to 15.6% over 1 min

with a simultaneous reduction of flowrate down to 150 nl/min, which was kept at this rate from there on. Then to 19% over 12 min, to 27% over 45 min, to 38% over 26 min, to 95% over 1 min, and kept at 95% for 23 min, after which all peptides were eluted. Cycle times for FAIMS switching were set to of 2.5s and 1.2s for -50 and -70 CV, respectively. For RTS, maximum number of peptides was set to 5 and all methods used maximum number of missed cleavages of 1 and maximum variable mods of 2. Carbamidomethyl on Cysteine (C) was not set, as these samples were not treated with TCEP/CAA.

Mass Spectrometry Raw Data Analysis

Raw files were analyzed with Proteome Discoverer 2.4 (ThermoFisher Scientific) with the built-in TMTPro Reporter ion quantification workflows using the standard settings if not further specified. Spectra were searched using the Sequest search engine using the homo sapiens database obtained from Uniprot (Swiss-Prot with isoforms). Static modifications were TMTpro16plex on Lysine (K) and N-Terminus, and for the diluted standard Carbamidomethyl on Cysteine (C) was set. Dynamic modifications were set as Oxidation (M), and Acetyl and Met-loss on protein N-termini. Fixed modifications were set to TMTPro on both peptide N-termini and K residues. Results were validated with Percolator and filtered to 1% FDR. For the MS2 and RETICLE methods, reporter ion quantification was performed on FTMS MS2 spectra and the same spectra were also sent to Sequest for identification, where they were searched with fragment mass tolerance of 0.02 Da. For the RTS-MS3 method, reporter ion quantification was performed on FTMS MS3 spectra and LIT MS2 spectra were searched with fragment mass tolerance of 0.6 Da. For the reporter ion quantification in all methods, normalization mode and scaling mode were set to None and average reporter s/n threshold was set to 0. Isotopic error correction was applied.

Data Analysis of Diluted Standard

Proteome Discoverer 2.4 result tables were loaded into python 3.7.9. Potential contaminant proteins were removed. The first three "single-cell" channels (127N, 128N, 129N) were excluded from the analysis as they might be contaminated from the isotopic impurities of the booster channel (126). Protein S/N values below 1.1 were set to missing values. Technical triplicates of each of the 6 methods were first normalized by equalizing the median s/n of the "single-cell" channels for each protein across replicates (one correction factor per protein). To calculate the silhouette coefficient, the normalized data were further processed by a median shift across cells, log2 transformation and removal of all proteins containing any missing value across all methods. Subsequently, protein data was transformed into the 17 first principal components (PCs). The silhouette coefficients of BLASTs and LSCs were calculated per method using the "silhouette_samples" function from Scikit-learn[23], providing the differentiation stage as labels and the first 17 PCs as matrix. All protein log2 fold-changes and coefficients of variation (CV) were calculated using the normalized raw data. Log2-fold-changes in the SPS-MS3 reference library were calculated by Proteome Discoverer.

Data Analysis of Real scMS Samples

scMS data was analysed as previously described[7]. Briefly, FACS.fcs files were processed in FlowJo 10.7.1 with the IndexSort 2.7 plugin to apply bi-exponential transform. FACS data and sort- and label layouts were used to create the metadata for each cell. Subsequent analysis was performed with python 3.7.9 and SCeptre version 1.0, which included normalization, cell filtering, imputation and embedding in PCA and UMAP space.

RESULTS

Design of Acquisition Methods

We created three different acquisition methods (MS2, RTS-MS3, RETICLE) for the Orbitrap Eclipse Tribrid Mass Spectrometer via the Xcalibur Intrument Setup (**Supplementary Fig. 2**). In the MS2 method, precursors for the quantitative MS2 scan are directly selected from the MS1 scan. In the RTS-MS3 method, precursors detected in the MS1 scan are subjected to an MS2 scan in the LIT, which is subsequently searched against a user-defined database. If the precursor is identified as a peptide of interest, it is subjected to a quantitative SPS-MS3 scan. The same procedure is applied in RETICLE, with the difference that precursors identified as a peptide of interest are subjected to a quantitative MS2 scan. Thus, MS2 and RETICLE acquisition yields high-resolution OT MS2 scans that are used for peptide identification and quantification in the subsequent data analysis, whereas RTS-MS3 yields low-resolution LIT MS2 scans for peptide identification and linked high-resolution OT MS3 scans that can only be used for quantification. To investigate how the ion injection time (IT) of the quantitative scan influences proteome coverage and quantitative performance, we tested each acquisition method with IT of 500ms and 750ms resulting in six methods (MS2 500ms, MS2 750ms, RTS-MS3 500ms, RTS-MS3 750ms, RETICLE 500ms, RETICLE 750ms).

Comparison of Acquisition Methods Using a Diluted Standard

To enable a reproducible comparison between MS acquisition methods we created a multiplexed sample that can be injected as technical replicate but also preserves the biologically relevant protein fold-changes between LSC, PROG, and BLAST. For this, the three differentiation stages were FACS sorted and equal numbers of cells were labeled with three TMTpro channels per

differentiation stage. Subsequently, the sample was diluted to contain 250 picogram of peptides per "single-cell" channel per injection and a booster channel was added as a 200-cell equivalent (**Fig 1a**). This sample was measured in triplicates for each of the six methods. The analysis of the raw files showed that with RTS-MS3 and RETICLE, less quantification scans were recorded (**Fig 1b**). Nonetheless, these RTS-triggered scans had considerably higher identification rates, resulting in more quantified proteins in the "single-cell" channels when comparing the same ITs. The higher identification rate in RETICLE compared to RTS-MS3 showed the advantage of using the high-resolution long-injected OT MS2 scans for peptide identification and simultaneous quantification, as this resulted in the highest protein coverage. Nonetheless, RTS-MS3, with its reliance on LIT spectra for peptide identification, performed similar to the MS2 method in terms of protein coverage. Another benefit of RTS is the possibility to limit the number of peptides per protein on the fly using the close-out filter, which improved the distribution of quantified peptides across proteins by reducing the fraction of proteins with only one peptide or with disproportionally many peptides (e.g. over 7 peptides) (**Fig. 1c**).

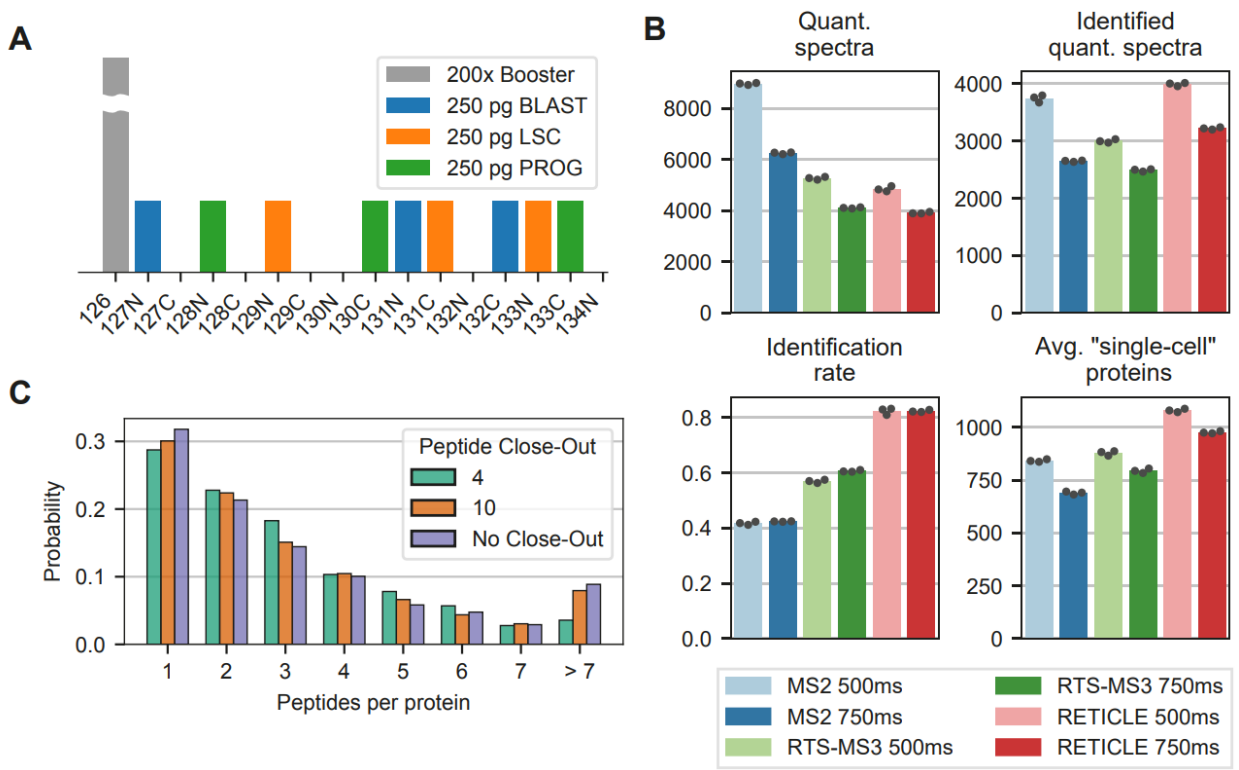


Figure 1. Comparison of scMS performance between the acquisition methods. A) Design of the multiplexed diluted standard. B) Number of acquired quantification spectra differs between the methods, as does the identification rate thereof, influencing the average number of proteins quantified in the "single-cell" channels. C) Close-out option implemented in RTS influences the distribution of peptides across proteins.

To evaluate the quantitative performance of the different methods, we measured the separation of BLAST and LSC based on differences in abundances of the same proteins using the silhouette coefficients (**Fig 2a**). We found that increased ITs overall improved the separation, exemplifying the importance of sampling enough ions from the very low abundant single-cell channels. Interestingly, RETICLE 750ms outperformed MS2 750ms, likely because of its advantageous distribution of peptides across these 345 proteins (**Supplementary Fig. 3**); this difference was less pronounced in MS2 500ms vs. RETICLE 500ms. RTS-MS3 outperformed RETICLE, indicating improved quantification of these proteins. To investigate whether the differences in separation

power can be attributed to quantitative accuracy or precision, we calculated the absolute log2FC differences between BLAST and LSC in scMS when compared to the bulk-measured reference dataset and the coefficient of variation (CV) of the protein abundances in LSC (i.e. technical variance) (**Fig. 2b**). The results showed that there are differences in accuracy between the three methods, with MS2 performing worse than the two RTS methods and RTS-MS3 being the most accurate, but there were no large differences between ITs. However, when comparing the CVs, it became apparent that higher ITs increased the precision of the measurement. These results indicate that the observed improvement in separation power with increased ITs can be attributed more to improved precision rather than a gain in quantitative accuracy, and that the reduction of interferences via RTS-MS3 acquisition results in more accurate quantification compared to MS2 methods. To validate if these conclusions would remain valid when more proteins are included in the comparison, we performed a pairwise comparison of each method with RETICLE 750ms (**Fig. 2c**), which lead to similar results. Taken together, the results show that RTS can increase scMS proteome coverage and quantitative performance when applied via the RETICLE method and that RTS-MS3 can improve quantitative accuracy in scMS while yielding comparable proteome coverage to the MS2 method.

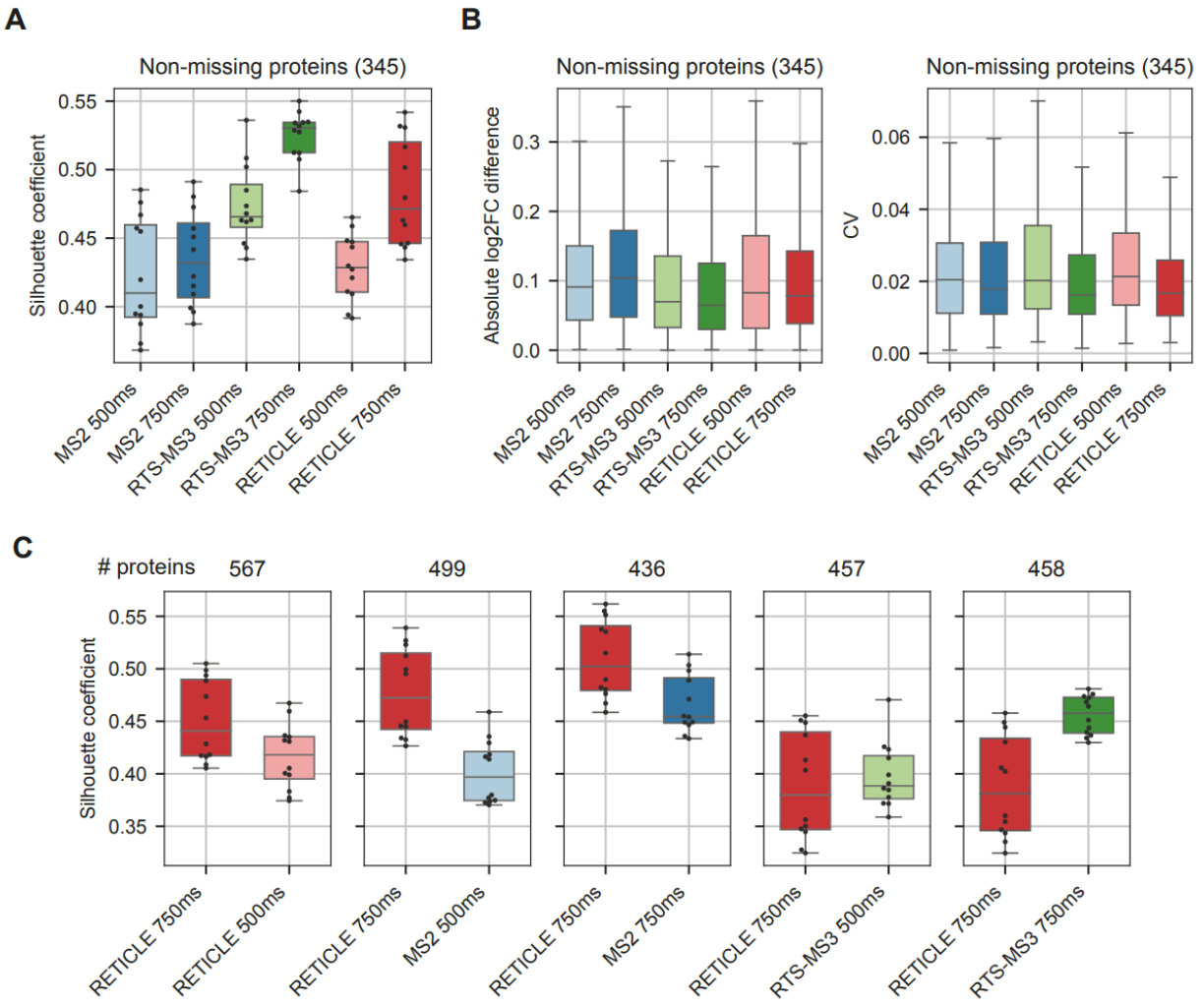


Figure 2. Comparison of quantitative performance between the acquisition methods. A) Separation of BLAST and LSC using scMS data measured as the silhouette coefficient of each "single-cell" in PCA space. Only proteins that were completely quantified across all methods were used. B) Left: Absolute difference of the log2FC between BLAST and LSC as measured in scMS and the bulk-measured reference dataset to assess quantitative accuracy of the acquisition methods. Right: CV of protein abundances in LSC to assess the quantitative precision (i.e. technical variance) of the acquisition methods. C) Pairwise comparison of separation of BLAST and LSC between RETICLE 750ms and all other methods. Only proteins that were completely quantified across the two compared methods were used.

Comparison of Acquisition Methods Using Real scMS Samples

To investigate the performance of the different acquisition methods with real scMS samples, we measured over 110 single cells for each of the six methods. For this, we FACS sorted single cells from three different populations (CD34+CD38-, CD34+CD38+, CD34-) (**Fig. 3a**) from the OCI-AML8227 cell-culture system into 384-well plates and prepared the cells as previously described[7]. After data acquisition, each dataset was processed with our scMS data analysis pipeline SCeptre[7], which includes data normalization, cell filtering, imputation and embedding. The results showed that with RETICLE, considerably more proteins were identified across the whole dataset, whereas RTS-MS3 identified similar numbers of proteins as MS2 (**Table 1**). Furthermore, RETICLE 750ms measured the highest number of high-coverage proteins, even surpassing MS2 500ms. When comparing the number of proteins measured in each cell (**Fig. 3b**), it became apparent that RETICLE outperformed MS2 by 36% (MS2 750ms: 718, RETICLE 750ms: 976; median proteins per cell) and RTS-MS3 performed similarly to MS2. Interestingly, in contrast to the diluted standard, the number of proteins per cell in the RTS methods did not change much between ITs, indicating that even higher ITs could be feasible. We subsequently investigated the ability of each dataset to separate the three different populations. For this, the datasets were processed further by imputing missing values and embedding into UMAP space (**Fig. 3c**). For this, low-coverage proteins were removed prior to imputation to achieve the best possible silhouette score (mean of all silhouette coefficients) for each dataset. The silhouette coefficients of the cells showed that the RETICLE methods achieved the highest separation of the three populations in UMAP space and that RTS-MS3 performed similarly to MS2 (**Fig. 3d**). A similar observation was made in PCA space (**Supplementary Fig. 4**), although differences were smaller. To avoid bias via differences in the number of cells or proteins across datasets, we repeated the analysis with datasets where cell

numbers were equalized and only overlapping proteins were used for the embedding (**Supplementary Fig. 5**). Surprisingly, differences were comparably small in UMAP and PCA space, indicating that the improvements in separation are also a result of the increased proteome depth achieved by the RETICLE methods. Taken together, these results suggest that the higher accuracy of RTS-MS3 might only provide minor improvements to data quality in our current scMS workflow, whereas RETICLE provides considerable advantages over classical MS2 acquisition by greatly increasing proteome coverage while simultaneously enabling higher MS2 ITs, which conversely improves separation of populations in scMS due to more and better biological signal.

Table 1. Comparison of real scMS samples

	MS2 500ms	**MS2 750ms**	**RTS-MS3 500ms**	**RTS-MS3 750ms**	**RETICLE 500ms**	**RETICLE 750ms**
# Quantification spectra	94,077	64,839	51,644	47,798	43,578	39,855
Identification rate quant. spectra	0.34	0.36	0.51	0.53	0.77	0.79
# Cells	120	118	118	127	118	128
# Proteins	1,514	1,205	1,537	1,457	2,123	1,991
# Proteins with >70% coverage	649	557	558	560	667	688

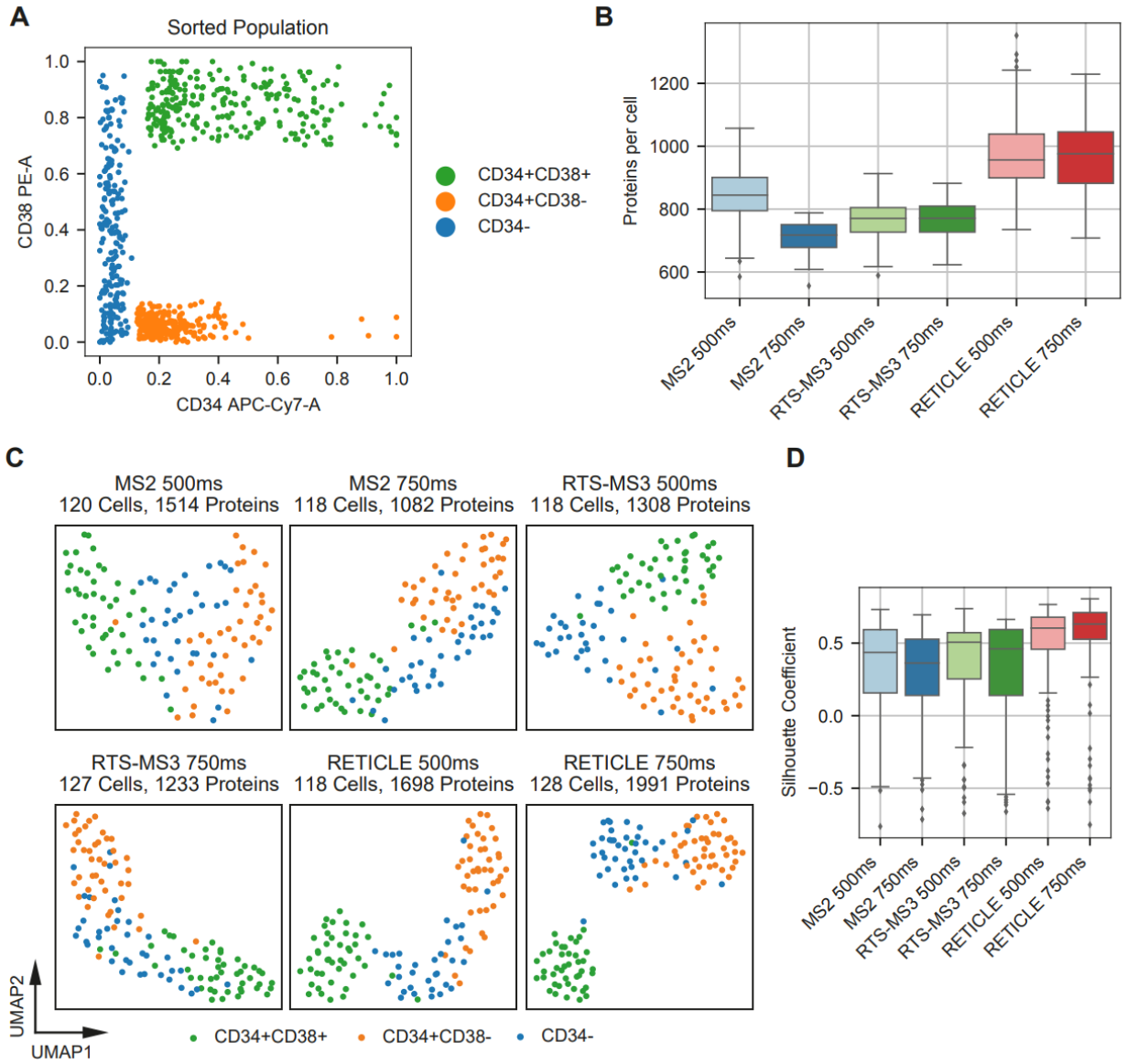


Figure 3. Comparison of acquisition methods using real scMS samples. A) FACS strategy to sample single cells from three different populations of the OCI-AML8227 cell-culture system. B) Boxplot of the number of proteins quantified in each cell across the methods. C) UMAP of scMS data after processing with SCeptre (normalization, cell filtering and imputation). D) Separation of the different populations measured by the silhouette coefficients of all cells in UMAP space.

DISCUSSION

Here, we evaluated the application of RTS-MS3 on the Eclipse Tribrid Mass Spectrometer to multiplexed single-cell proteomics (scMS) and presented a novel acquisition strategy, RETICLE, which makes use of RTS but performs more sensitive MS2 level quantification and identification compared to RTS-MS3. We demonstrated that RTS on Tribrid instrument enables the identification of peptide precursors before performing the time intensive quantification scan. This advantage played out especially well in multiplexed scMS, as the booster channel facilitates peptide identification and thus enables short LIT ion injection times, whereas the quantification of the single-cell channels requires much higher ion injection times and high-resolution OT spectra. We further demonstrated that the additional time spent on LIT scans, compared to a classical MS2 method, is beneficial, as the quantification spectra have much higher identification rates and only peptides of interest are quantified. Thus, RTS offers the potential to target specific proteins of interest or to exclude proteins that are not relevant for the specific research question. In terms of quantitative performance, we further exemplified the need for very high OT ion injection times to increase the precision and accuracy of the measurements, which can have a strong impact on the ability to separate different cell populations. Furthermore, we showed that RTS-MS3 yields the highest quantitative accuracy, however the positive impact on real scMS data was rather small. This could be due to the facts that interference is already greatly reduced by the FAIMS interface and the “Precursor Fit” purity filter, and that mostly highly abundant peptides are measured. Another reason could be that the gain in accuracy is offset by a loss in precision due to lower sensitivity of MS3 scans. Thus, for the overall scMS data quality it seems to be more beneficial to increase proteome depth and OT ion injection times, as achieved with RETICLE. Nonetheless, there could be applications where RTS-MS3 would be beneficial, e.g. when RTS is used to target

specific lower abundant proteins with the aim to perform highly accurate measurements. To implement RTS-MS3 or RETICLE for scMS, we note that depending on the sample preparation, chromatographic setup and the level of boosting, ion injection times for the LIT and OT should be adapted to reach optimal performance. Furthermore, cycle times between FAIMS switching should be adapted to the chromatographic peak widths. We also note that in some cases, RETICLE can also benefit from an intensity threshold filter to trigger LIT scans.

In conclusion, not only does RETICLE identify more proteins than MS2 and RTS-MS3, it also enables much higher ion injection times and as a result, is able to provide better resolving power for separating biological phenotypes. We hope our results inspire further pursuit of intelligent data acquisition strategies, with a lot of potential to be gained in the realm of single-cell proteome analysis.

Data Availability

The mass spectrometry data will be deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository. All data will also be available from the corresponding authors.

Code availability

The complete analysis will be available on Github (github.com/bfurtwa/RETICLE).

Author Contributions

B.F. and E.M.S conceived the RETICLE method, designed and carried out experiments, and wrote the manuscript. B.F. conducted data analysis. N.Ü. performed FACS sorting and sample preparation. K.M, R.H, D.L.F and V.Z helped develop the instrument methods, and contributed to

the overall functioning of the method. B.T.P. oversaw the study. All authors proofread and contributed to the manuscript.

Competing interests

K.M, R.H, D.L.F and V.Z are or were employees at ThermoFisher Scientific. All other authors declare no competing interests.

Acknowledgments

This work was supported by the Independent Research Fund Denmark, the Svend Andersen Foundation, the Candys Foundation, the Danish Cancer Society and through a center grant from the Novo Nordisk Foundation (Novo Nordisk Foundation Centre for Stem Cell Biology, DanStem; Grant Number NNF17CC0027852). B. Furtwängler is the recipient of a fellowship from the Novo Nordisk Foundation as part of the Copenhagen Bioscience Ph.D. Programme, supported through grant NNF19SA0035442. We thank the BRIC FACS facility for support with the flow cytometry work. We thank the DTU Proteomics Facility for support with the MS analyses.

REFERENCES

1. Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* **19**, (2018).
2. Specht, H. *et al.* Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome Biol.* **22**, 50 (2021).
3. Zhu, Y. *et al.* Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. *Nat. Commun.* **9**, (2018).
4. Zhu, Y. *et al.* Single-cell proteomics reveals changes in expression during hair-cell development. *Elife* **8**, e50777 (2019).
5. Tsai, C. F. *et al.* An Improved Boosting to Amplify Signal with Isobaric Labeling (iBASIL) Strategy for Precise Quantitative Single-cell Proteomics. *Mol. Cell. Proteomics* **19**, 828–838 (2020).
6. Cong, Y. *et al.* Ultrasensitive single-cell proteomics workflow identifies >1000 protein groups per mammalian cell. *Chem. Sci.* **12**, 1001–1006 (2021).
7. Schoof, E. M. *et al.* Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nat. Commun.* **12**, 3341 (2021).
8. Brunner, A.-D. *et al.* Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *bioRxiv* 2020.12.22.423933 (2020). doi:10.1101/2020.12.22.423933
9. Hartlmayr, D. *et al.* An automated workflow for label-free and multiplexed single cell proteomics sample preparation at unprecedented sensitivity. *bioRxiv* 2021.04.14.439828 (2021). doi:10.1101/2021.04.14.439828

10. Li, J. *et al.* TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide measurements across 16 samples. *Nat. Methods* **2**, 1–6 (2020).

11. Li, J. *et al.* TMTpro-18plex: The Expanded and Complete Set of TMTpro Reagents for Sample Multiplexing. *J. Proteome Res.* **20**, 2964–2972 (2021).

12. Cheung, T. K. *et al.* Defining the carrier proteome limit for single-cell proteomics. *Nat. Methods* **18**, 76–83 (2021).

13. Specht, H. & Slavov, N. Optimizing Accuracy and Depth of Protein Quantification in Experiments Using Isobaric Carriers. *J. Proteome Res.* **20**, 880–887 (2021).

14. Ting, L., Rad, R., Gygi, S. P. & Haas, W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat. Methods* **8**, 937–940 (2011).

15. McAlister, G. C. *et al.* MultiNotch MS3 Enables Accurate, Sensitive, and Multiplexed Detection of Differential Expression across Cancer Cell Line Proteomes. *Anal. Chem.* **86**, 7150–7158 (2014).

16. Pfammatter, S., Bonneil, E. & Thibault, P. Improvement of Quantitative Measurements in Multiplex Proteomics Using High-Field Asymmetric Waveform Spectrometry. *J. Proteome Res.* **15**, 4653–4665 (2016).

17. Schweppe, D. K. *et al.* Characterization and Optimization of Multiplexed Quantitative Analyses Using High-Field Asymmetric-Waveform Ion Mobility Mass Spectrometry. *Anal. Chem.* **91**, 4010–4016 (2019).

18. Heil, L. R., Remes, P. M. & MacCoss, M. J. Comparison of Unit Resolution Versus High-Resolution Accurate Mass for Parallel Reaction Monitoring. *J. Proteome Res.* acs.jproteome.1c00377 (2021). doi:10.1021/acs.jproteome.1c00377

19. Erickson, B. K. *et al.* Active Instrument Engagement Combined with a Real-Time Database

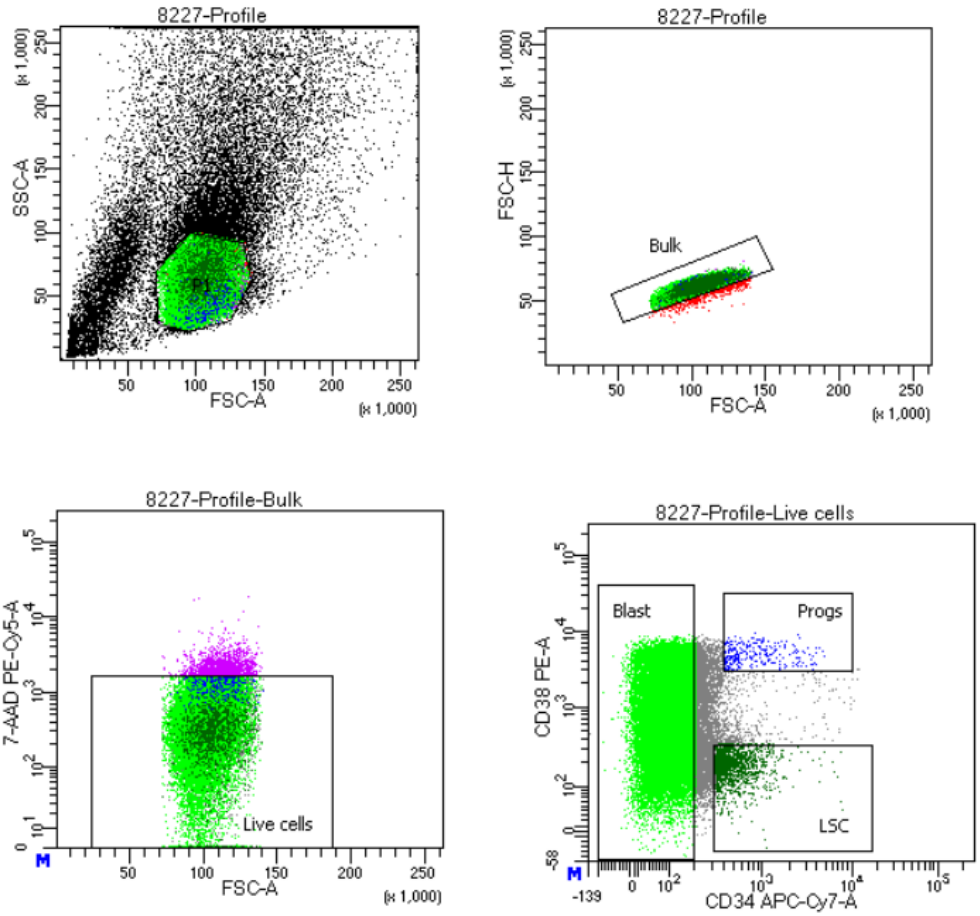
Search for Improved Performance of Sample Multiplexing Workflows. *J. Proteome Res.* **18**, 1299–1306 (2019).

20. Schweppe, D. K. *et al.* Full-Featured, Real-Time Database Searching Platform Enables Fast and Accurate Multiplexed Quantitative Proteomics. *J. Proteome Res.* **19**, 2026–2034 (2020).

21. Yu, Q. *et al.* Benchmarking the Orbitrap Tribrid Eclipse for Next Generation Multiplexed Proteomics. *Anal. Chem.* **92**, 6478–6485 (2020).

22. Lechman, E. R. *et al.* MiR-126 Regulates Distinct Self-Renewal Outcomes in Normal and Malignant Hematopoietic Stem Cells. *Cancer Cell* **29**, 214–228 (2016).

23. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).



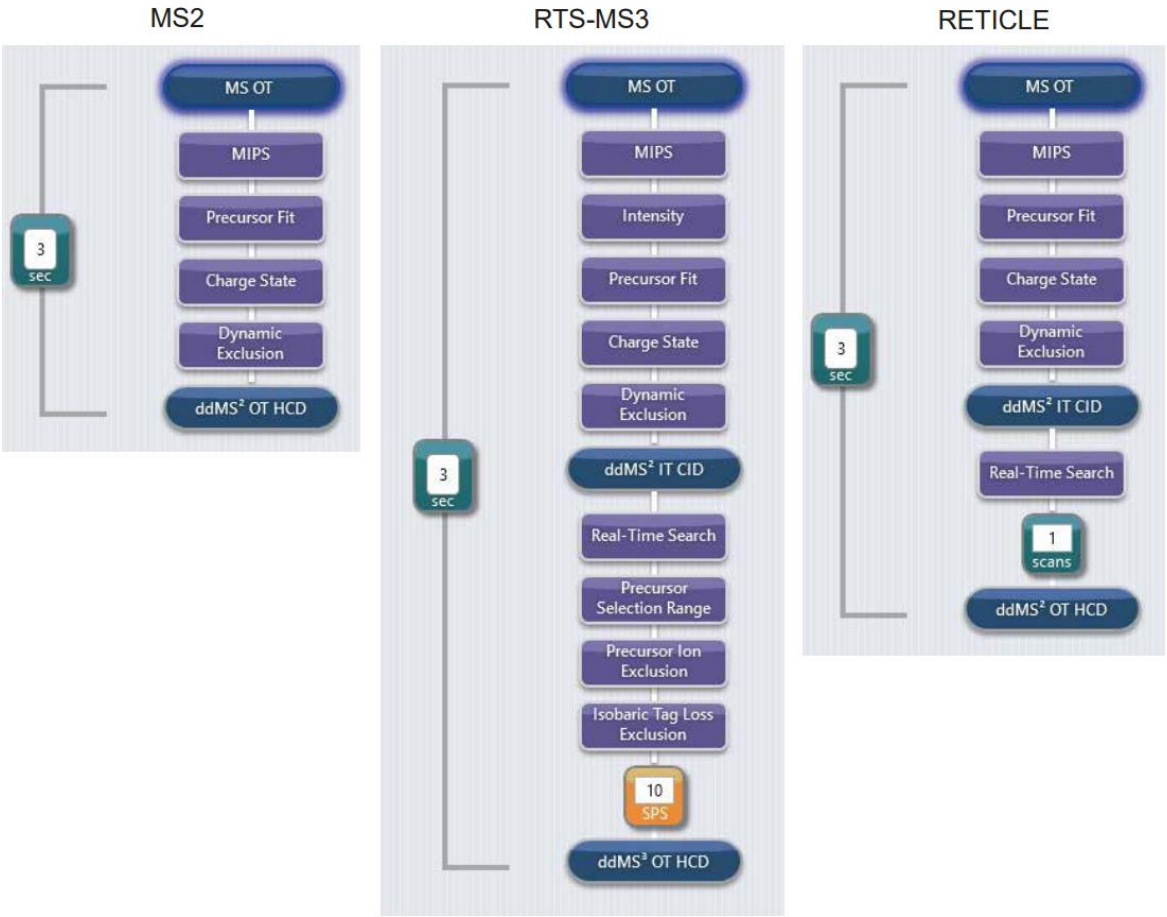
Tube: Profile

Population	#Events	%Parent	%Total
All Events	60,000	####	100.0
P1	41,198	68.7	68.7
Bulk	40,378	98.0	67.3
Live cells	38,858	96.2	64.8
Blast	32,555	83.8	54.3
Progs	284	0.7	0.5
LSC	1,337	3.4	2.2

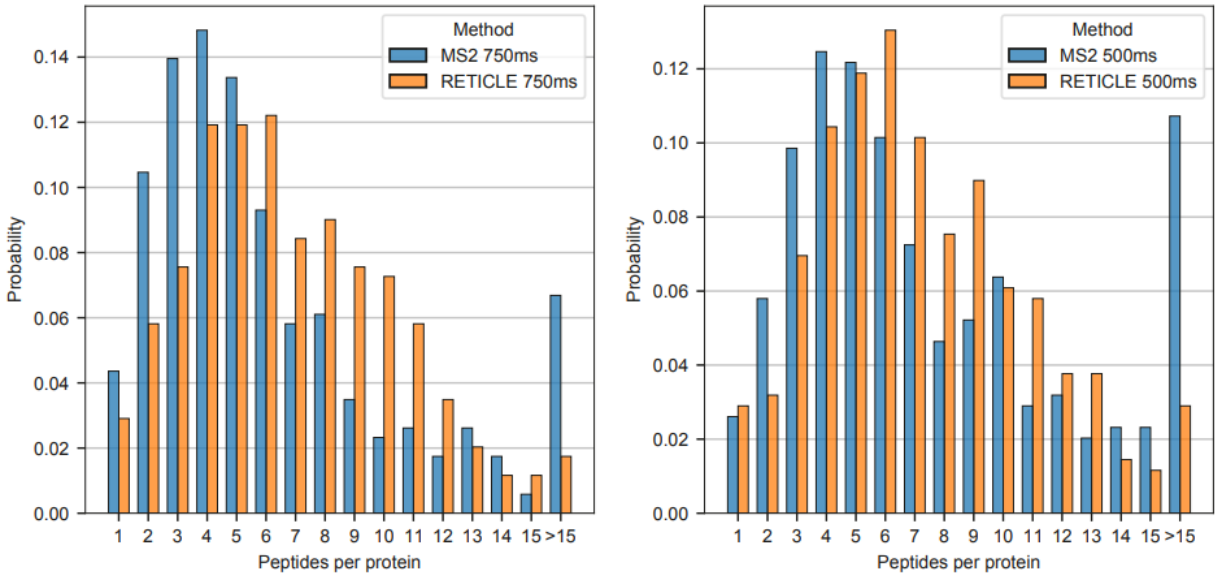
Experiment Name:	8227 single cell sort_07Jun...
Specimen Name:	8227
Tube Name:	Profile
Record Date:	Jun 7, 2021 11:37:30 AM
$OP:	Administrator
GUID:	2877c1ee-60e8-4cc3-96ac...

Population	#Events	%Parent
Live cells	38,858	96.2
Blast	32,555	83.8
Progs	284	0.7
LSC	1,337	3.4

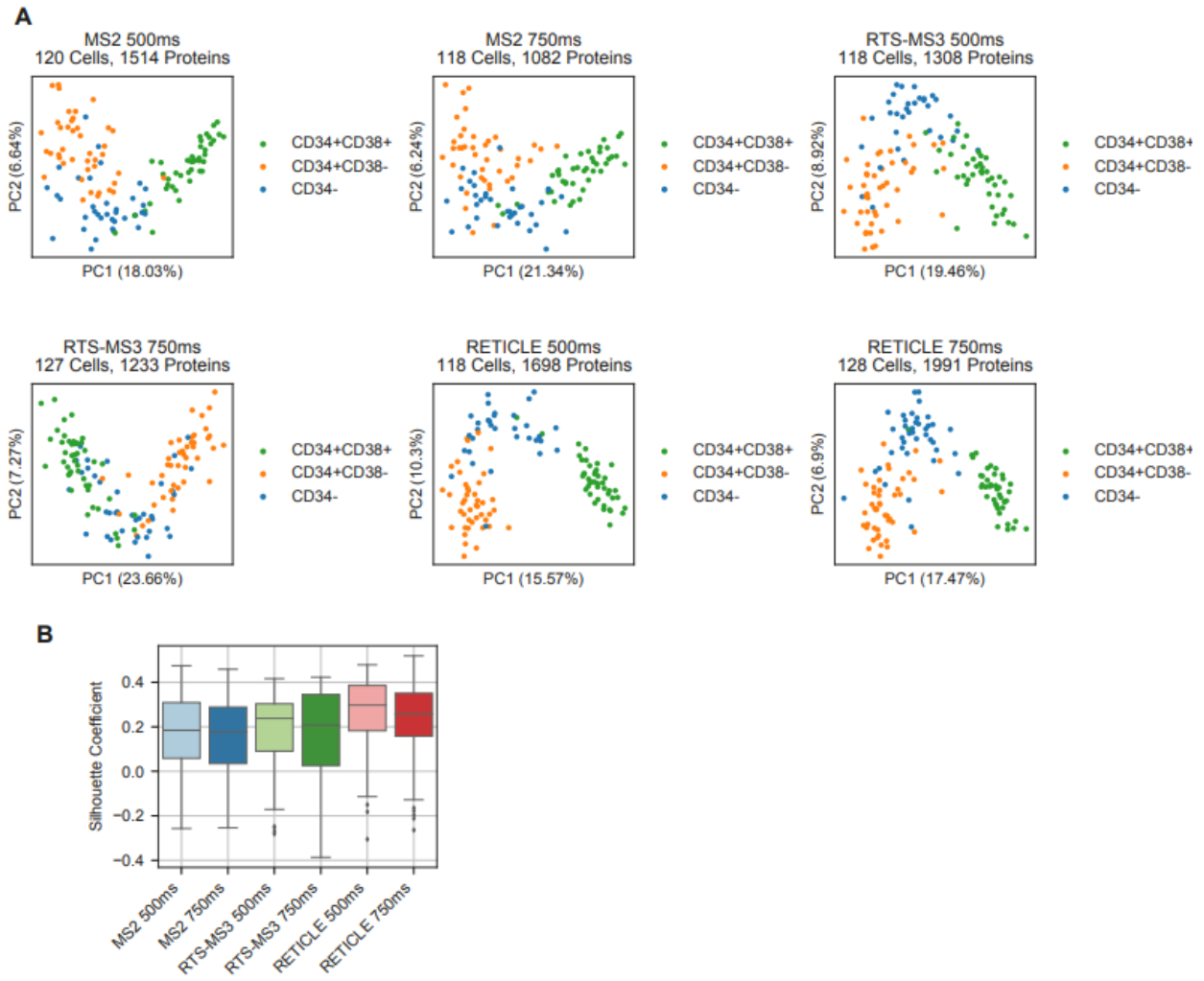
Supplementary Figure 1. FACS strategy for single-cell sort of OCI-AML8227 cell-culture system.



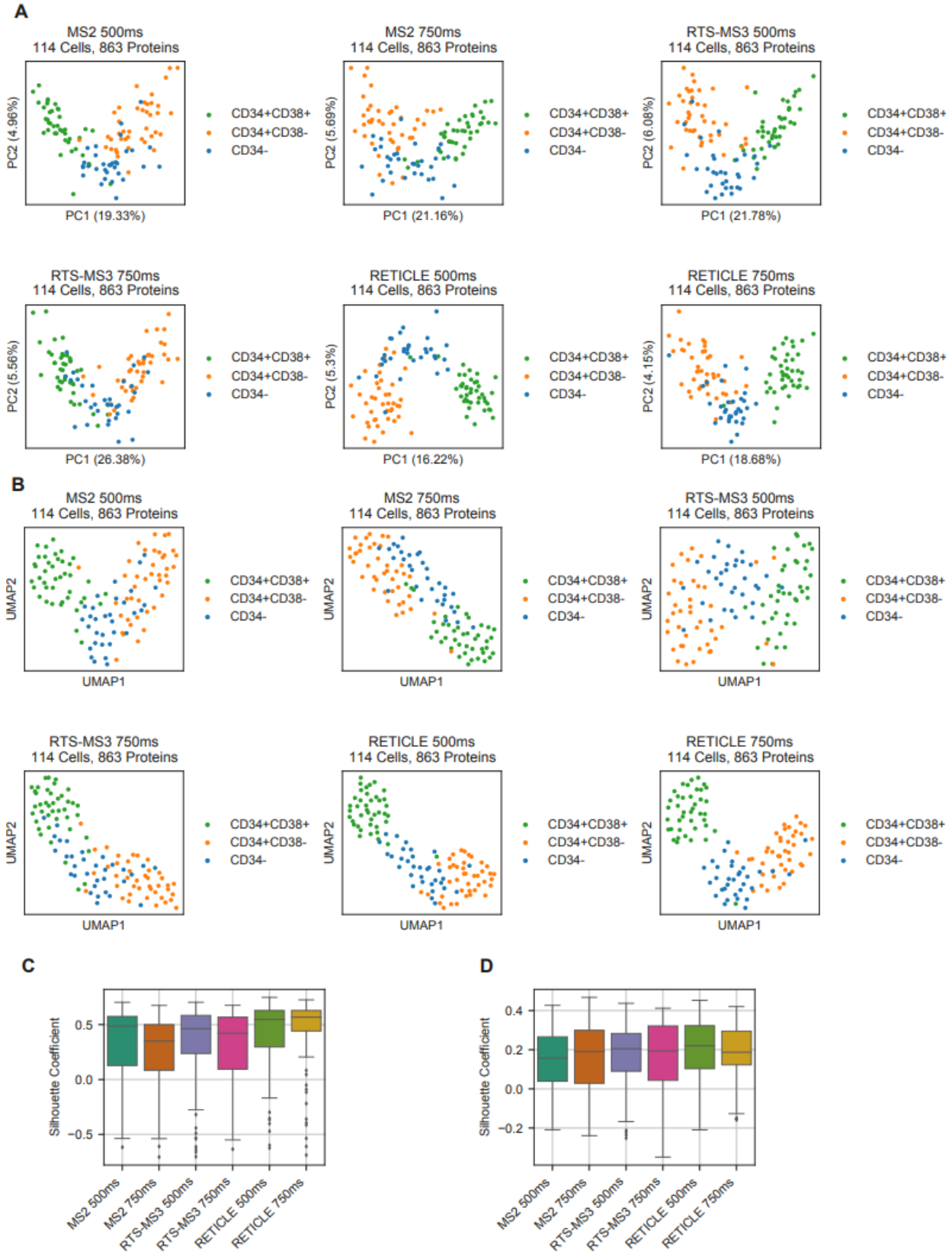
Supplementary Figure 2. Method trees of the three acquisition methods.



Supplementary Figure 3. Comparison of the distribution of the number of peptides per protein.



Supplementary Figure 4. A) PCA of scMS data after processing with SCeptre (normalization, cell filtering and imputation). B) Separation of the different populations measured by the silhouette coefficients of all cells in PCA space (first 20 PCs).


A
MS2 500ms
114 Cells, 863 Proteins
PC2 (4.96%)
PC1 (19.33%)
MS2 750ms
114 Cells, 863 Proteins
PC2 (5.69%)
PC1 (21.16%)
RTS-MS3 500ms
114 Cells, 863 Proteins
PC2 (6.08%)
PC1 (21.78%)
RTS-MS3 750ms
114 Cells, 863 Proteins
PC2 (5.56%)
PC1 (26.38%)
RETICLE 500ms
114 Cells, 863 Proteins
PC2 (5.3%)
PC1 (16.22%)
RETICLE 750ms
114 Cells, 863 Proteins
PC2 (4.15%)
PC1 (18.68%)
CD34+CD38+
CD34+CD38-
CD34-
B
UMAP2
UMAP1
C
Silhouette Coefficient
0.5
0.0
−0.5
MS2 500ms
MS2 750ms
RTS-MS3 500ms
RTS-MS3 750ms
RETICLE 500ms
RETICLE 750ms
D
Silhouette Coefficient
0.4
0.2
0.0
−0.2

Supplementary Figure 5. A) PCA and UMAP (B) of scMS data after processing with SCeptre (normalization, cell filtering and imputation) of only proteins that overlap between all methods (quantified in at least 3 cell per method). C) Separation of the different populations measured by the silhouette coefficients of all cells in UMAP space (first 2 dimensions) and PCA space (first 20 PCs) (D).