plexDIA for scalable single cell analysis – improving quantification and sensitivity by TIMS separation

Single cell proteomics in a biological research context needs to deliver high sensitivity for deep proteome coverage while maintaining quantitative accuracy at scale.

Abstract

Labeling approaches outperform label-free analyses, however, multiplexing on the fragment level typically results in a reduced proteome depth due to its dependence on data dependent acquisition. Precursor based non-isobaric labeling, as the here shown plexDIA approach, uses non-isobaric labeling and is hence compatible with data independent acquisition, overcoming stochastic precursor sampling and reduces missing not at random type data incompleteness.

Here, we applied plexDIA in a fully automated 3-label approach performed on CFPAC-1 pancreatic cancer, WM-989 melanoma, and THP-1 leukemia cell lines (126 single cells in total) using the timsTOF Ultra in dia-PASEF mode. This approach allowed to analyze almost 100 samples per day with a proteome depth of more than 3000 proteins per cell and quantification of > 80,000 precursors per individual plexDIA run. Comparison to bulk material demonstrated preserved fold changes between the three cell types.

Keywords:
plexDIA, dia-PASEF,
timsTOF Ultra, nanoElute2

Introduction

Over the last few years single cell proteomics by mass spectrometry has significantly increased the number of peptides and proteins that can be detected and confidently identified from single mammalian cells. Using these rapid advances for biomedical research also requires simultaneously sustaining an increasing quantitative accuracy and throughput. For example, studies [1] have demonstrated the feasibility of calculating and interpreting protein correlations across single cells if sufficient quantitative accuracy, precision, and throughput can be achieved and maintained.

Achieving high quantitative accuracy, precision, and throughput is the focus of the workflow outlined here. Specifically, it demonstrates a workflow that allows for flexible and automated sample preparation that can be adapted to any desired set of mass tags for multiplexing single cells. The multiplexing enables increased throughput in linear proportion to the number

Christoph Krisp[1], Andrew Leduc[2], Luke Khoury[2], Nikolai Slavov[2], Torsten Müller[1], Daniel Hornburg[3], Markus Lubeck[1], Gary Kruppa[4]; [1]Bruker Daltonics GmbH & Co. KG, Bremen, Germany; [2]Departments of Bioengineering, Biology, Chemistry and Chemical Biology, Single Cell Proteomics Center, and Barnett Institute, Northeastern University, Boston, MA, USA; [3]Bruker Scientific LLC, San Jose, CA, USA; [4]Bruker s.r.o., Brno, Czech Republik.

of labels used while preserving the protein coverage and quantitative accuracy as previously demonstrated with plexDIA [2]. This workflow and the quantitative accuracy also benefit significantly from the efficient ion utilization enabled by trapped ion mobility spectrometry, which enables the isolation, fragmentation, and analysis of a large fraction of the ions delivered to the mass spectrometer. Furthermore, this technology operates with short duty cycles, increasing the frequency of sampling of the eluting precursor ions and fragments and therefore increasing the robustness and precision of quantification. We further increase this frequency by using multiple MS1 frames per duty cycle. The additional MS1 scans included in the duty cycles result in MS1 full range scans spaced by 300 ms, which results in about 10 points across a 3 second elution peak. This sampling frequency is significantly higher than the 3 points across a 3 second peak expected for a standard dia-PASEF method with about 1 s cycle time. The benefits of this frequent sampling have been demonstrated previously [3].

While multiplexing can substantially increase the throughput of mass spectrometry-based proteomics, especially for single cell applications, flexible and easy to implement, as well as scalable and sensitive sample preparation strategies are key for highest data quality. Yet, many of the sample preparation approaches for single-cell proteomics do not implement multiplexing or implement it assuming a fixed number of labels used.

To relax this limitation and increase throughput, this application note demonstrates an nPOP workflow that can accommodate different numbers of labels and different reagents simply by programming different configurations of droplets deposited on the surface of glass slides. In this workflow, each single cell is lysed, its proteins are digested and peptides are labeled in droplets on the surface of unpatterned glass slides. Since the droplets have volumes of 8 to 39 nL, the amounts of reagents used per single cell and the contact to surfaces can be minimized accordingly, crucial to increasing the sensitivity and sample purity. This workflow can simultaneously process thousands of single cells. The labeled peptides are then automatically pooled and transferred into multi-well plates that can be plugged into auto samplers (e.g., nanoElute 2), thus making the workflow compatible with full automation. Therefore, this workflow provides an automated option for scaling up single-cell proteomics to the automated analysis of thousands of single cells.

Methods

The CFPAC-1 pancreatic cancer, WM-989 melanoma, and THP-1 leukemia cell lines, for simplicity in the following termed PDAC, Melanoma, and Monocytes, were purchased from ATCC. Cells were thawed from liquid nitrogen and suspended directly in 1X PBS at a concentration of 300 cells per μL for cell sorting.

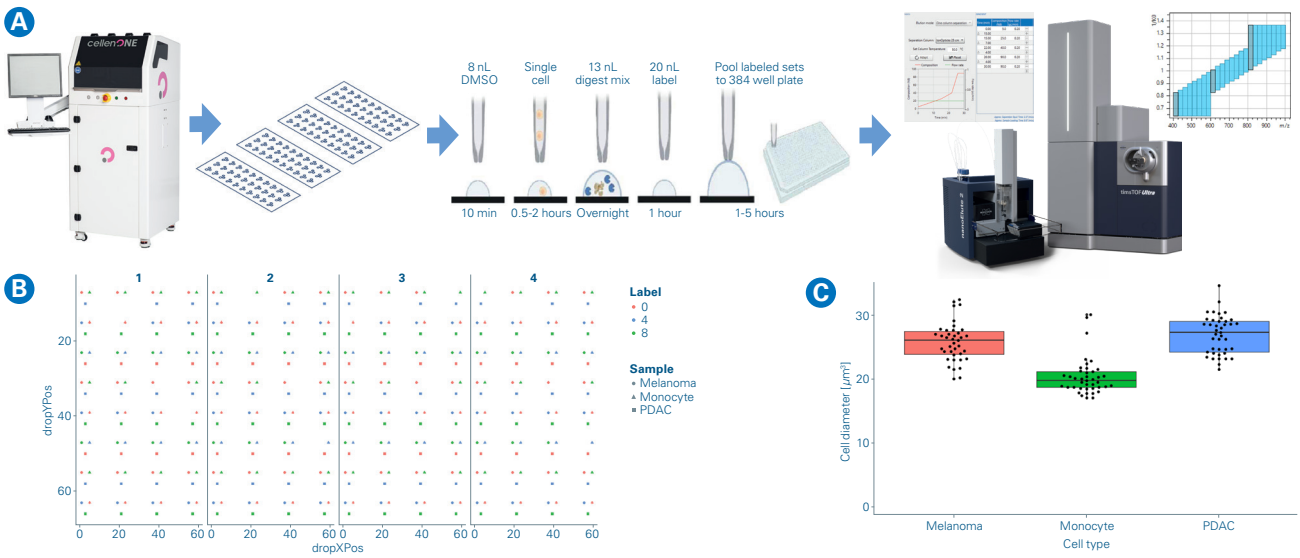

Figure 1

Workflow and meta information

(A) Schematic depicting sample preparation workflow and LC-MS/MS analysis strategy. (B) Schematic layout of cell positioning on the fluorocarbon-coated glass slides used in this experiment with label randomization information. (C) Cell diameter distributions within the cell type groups taken from the cellenONE report.

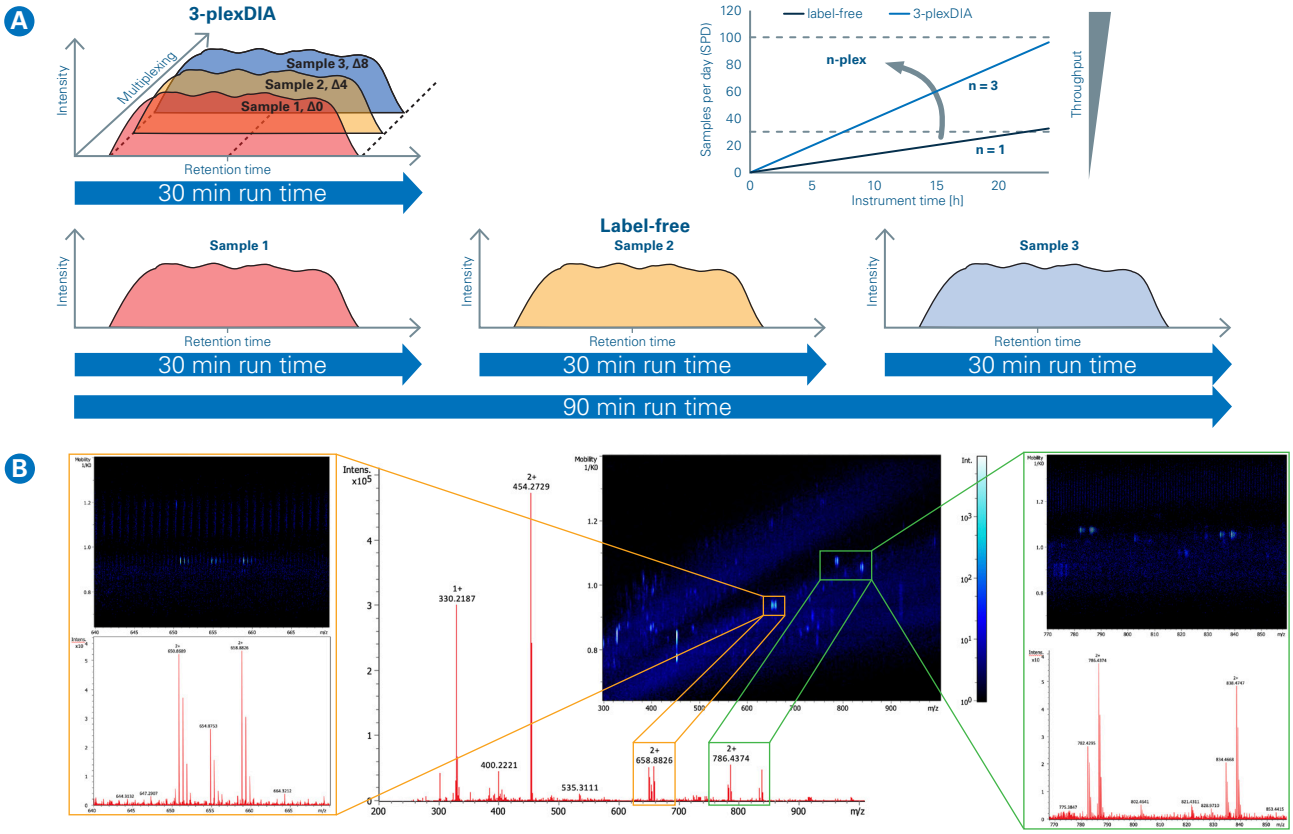

Figure 2

3-plexDIA analysis

Ⓐ Schematic highlighting the advantages of plexDIA over label-free analysis in sample throughput. Ⓑ Illustration of precursors seen in a *m/z* versus ion mobility ($1/k_0$) heatmap at a given retention time with associated MS spectrum of a precursor seen in all three cell types (left) and precursors only present in 2 of 3 cell types (right).

The nPOP sample preparation procedure was used to prepare single cells for multiplexed analysis by plexDIA. Briefly, the CellenONE cell sorter and liquid handler robotic system deposited single cells in 300 pL of 1x PBS into 8 nL droplets of 100% DMSO on the surface of a fluorocarbon coated glass slide for cell lysis. Then, 13 nL of master mix containing 100 ng/μL Promega trypsin gold, 5 mM HEPES pH 8.5 and 0.025% weight DDM was added to each single cell droplet on the slide. Droplet evaporation was prevented by setting the relative humidity to 75% and plate temperature to the dew point for overnight digest. The next day, cells were labeled by dissolving mTRAQ multiplexed labeling reagents in DMSO at a concentration of 1/40th unit per μL and dispensing 20 nL of either d0, d4 or d8 mass tags to each single-cell containing droplet. To facilitate the labeling reaction, 20 nL of 100 mM TEAB pH 8.5 was added and labeling proceeded for 1 hour. Single cell samples were then pooled by the CellenONE with 50% water 50% Acetonitrile and deposited to a 384 well plate. Samples were finally dried down and stored at -20°C until resuspension in 1 μL of water for LC/MS analysis. The samples were injected onto a 25 cm Aurora Ultimate column (75 μm, 1.7 μm, IonOpticks) using a nanoElute® 2. Peptides were eluted within a 25 min active gradient (30 min total acquisition time) and detected on a timsTOF Ultra in dia-PASEF mode using a 25 Da fixed window methods with 3 mass range focus switches per TIMS ramp with 8 TIMS ramps per cycle. Acquired data were searched against canonical human protein sequences including splice variants (Uniprot reviewed canonical sequences + isoforms) using DIA-NN version 1.8.1 [4] setting additional commands, as described by Derks *et al.* [2]: {fixed-mod mTRAQ, 140.0949630177, nK}, {channels mTRAQ,0,nK,0:0; mTRAQ, 4, nK, 4.0070994:4.0070994; mTRAQ, 8, nK, 8.0141988132:8.0141988132}, {peak-translation}, {original-mods}, {report-lib-info}, {ms1-isotope-quant}. This search used the spectral library that was previously generated from higher cell number plexDIA runs of Melanoma, PDAC, and Monocytes. Data were acquired and processed in the Single Cell Proteomics Center of Professor Nikolai Slavov at Northeastern University in Boston, MA, USA. Data analysis was performed using QuantQC for R (https://github.com/SlavovLab/QuantQC).

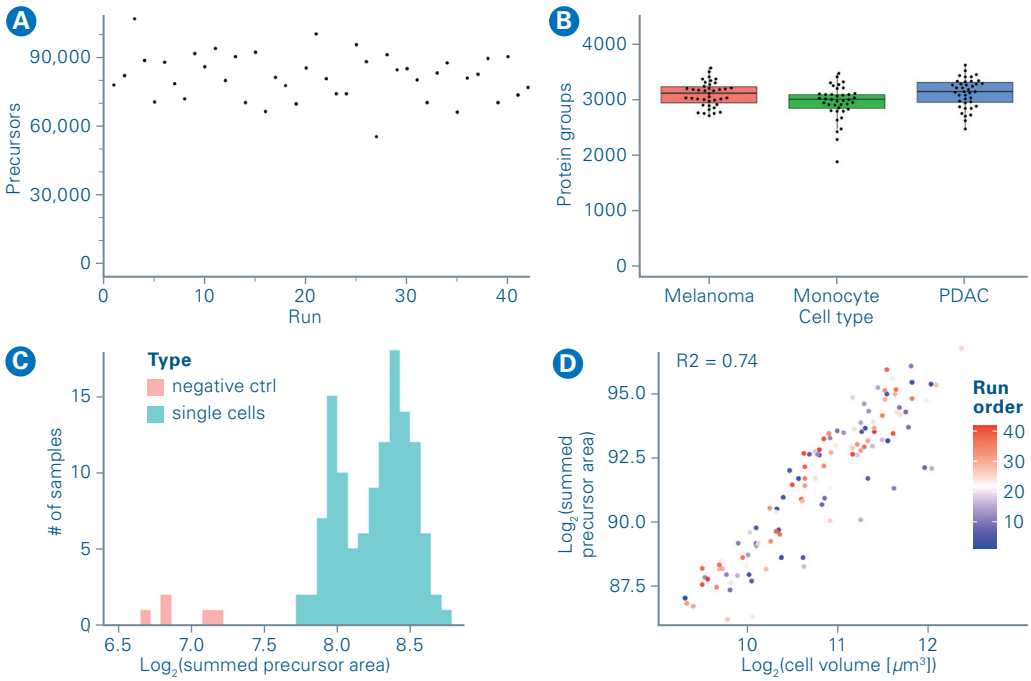

Figure 3

Results of data processing

Ⓐ Distribution of precursor identification rates across the 42 3-plexDIA experiments acquired at 96 SPD, Ⓑ Boxplot of protein group identifications per cell type acquired at 96 SPD, Ⓒ histogram of DIA-NN 1.8.1 $\log_{10}$(summed precursor area) of each analyzed single cell and the negative controls, Ⓓ Correlation of the estimated cell volume (μm^3) versus the $\log_2$(summed precursor area) of each analyzed single cell.

Results and discussion

Samples prepared for plexDIA analysis were analyzed on a timsTOF Ultra. Multiplexing increases the sample throughput per day and scales with the number of labels. Including overhead time and a gradient of 30 min, label free analysis result in a throughput of 32 runs per day (SPD), while a 3-plex approach allows for the analysis of 96 SPD and reducing the overhead time per sample by two-thirds (Figure 2A). Intrinsically, these samples are much more complex multiplying the number of detectable precursors with the number of plexes. plexDIA is therefore benefitting from TIMS separation, leading to reduced presence of interfering molecules of similar mass in precursor and fragment level. This enables accurate quantification, increased precursor detection and better association of non-isobaric mass-tagged peptides, as they appear at the same collisional cross section (CCS) value with mass-tag specific *m/z* offset. Figure 2B shows examples of ion mobility resolved 3-plex labeled precursors at different $1/k_0$ and *m/z* positions in a mobility vs. mass-to-charge based heatmap.

In total, 42 plexDIA runs including 126 samples in a 3-plex manor with 41 PDACs, 41 monocytes, 39 melanoma cells and 5 negative controls (no cell) were analyzed at 96 SPD using the nanoElute 2 timsTOF Ultra setup. Data processing was performed with DIA-NN 1.8.1 [4] with additional settings for the used labels and to enable MS1 translation for MS1 based quantification (see methods or Derks *et al.* [2]). Processing of the 42 runs resulted on average in the identification and quantification of 80,000 precursors (sum of precursors across channels; Figure 3A) in 25 min of active gradients. This translates to on average 3000 protein groups for Monocytes, 3100 protein groups for Melanoma cells, 3150 protein groups for PDAC cells, and in total 4486 protein groups identified in this experiment (Figure 3B). Intercalated MS1 scans in the dia-PASEF method increased the number of data points per peak from mean 4.3 to 14.6.

The distribution of $\log_{10}$ summed precursor intensity per cell (Figure 3C) demonstrated two maxima, the one at a value of 8 is manly derived from the comparably small monocytes (mean $\log_{10}$(summed precursor area) of 8.05) and the other at a value of 8.5 is represented

by the larger Melanoma (mean $\log_{10}$(summed precursor area) of 8.35) and PDAC cells (mean $\log_{10}$(summed precursor area) of 8.48). The negative controls on the other hand show $\log_{10}$ summed precursor intensity at least 10-fold lower than the single cells, demonstrating a clean background and an accurate data extraction of individual channel information out of the 3-plexDIA samples.

The estimated volume per cell (μm^3) correlated well with the sum of precursor intensities calculated for each cell (Figure 3D) with a correlation score of 0.73. This indicates that with an increase in cell volume the protein content increases proportionately.

Principal component analysis (Figure 4A) after k-nearest-neighbor (KNN) based data imputation clearly distinguished the three cell types. The Monocytes formed one cluster, whereas the Melanoma cells formed two, one representing most Melanoma cells and one small subcluster of 7 cells. The DPAC cells may also form a subcluster.

Protein abundance ratios calculated between the three cell types based on single cell data were compared to the ratios of bulk experiments performed in label free mode on these cells and showed good agreement in protein abundance ratios for Monocytes/PDAC, Melanoma/PDAC and Monocytes/Melanoma (Figure 4B).

Protein abundance profiles across the three cell types were compared, demonstrating the expected large differences in protein abundances in the three cell lines (Figure 5A-C). Among the proteins showing the largest abundance increase in melanoma compared to PDAC cell line and the Monocyte cell line were the cell surface proteins neural cell adhesion molecule L1 (L1CAM) and the CD44 antigen. L1CAM is discussed as a potential target for malignant melanoma therapy [5] and CD44 is known to be at high abundance in metastatic melanoma [6]. The monocyte cell line showed highest abundance difference to the melanoma and PDAC cell lines for carbonic anhydrase 2 (CA2) and adenylyl cyclase-associated protein 1, both are involved in immune response regulation in monocyte [7].

Interleukin 18 and aldehyde dehydrogenase 1A1 were elevated in abundance in the PDAC cell line compared to the Melanoma and Monocyte cell lines. IL18 has been shown to be elevated in pancreatic diseases like pancreatitis but also pancreatic cancer [8]. Aldehyde dehydrogenases are commonly elevated in solid tumors including pancreatic cancer.

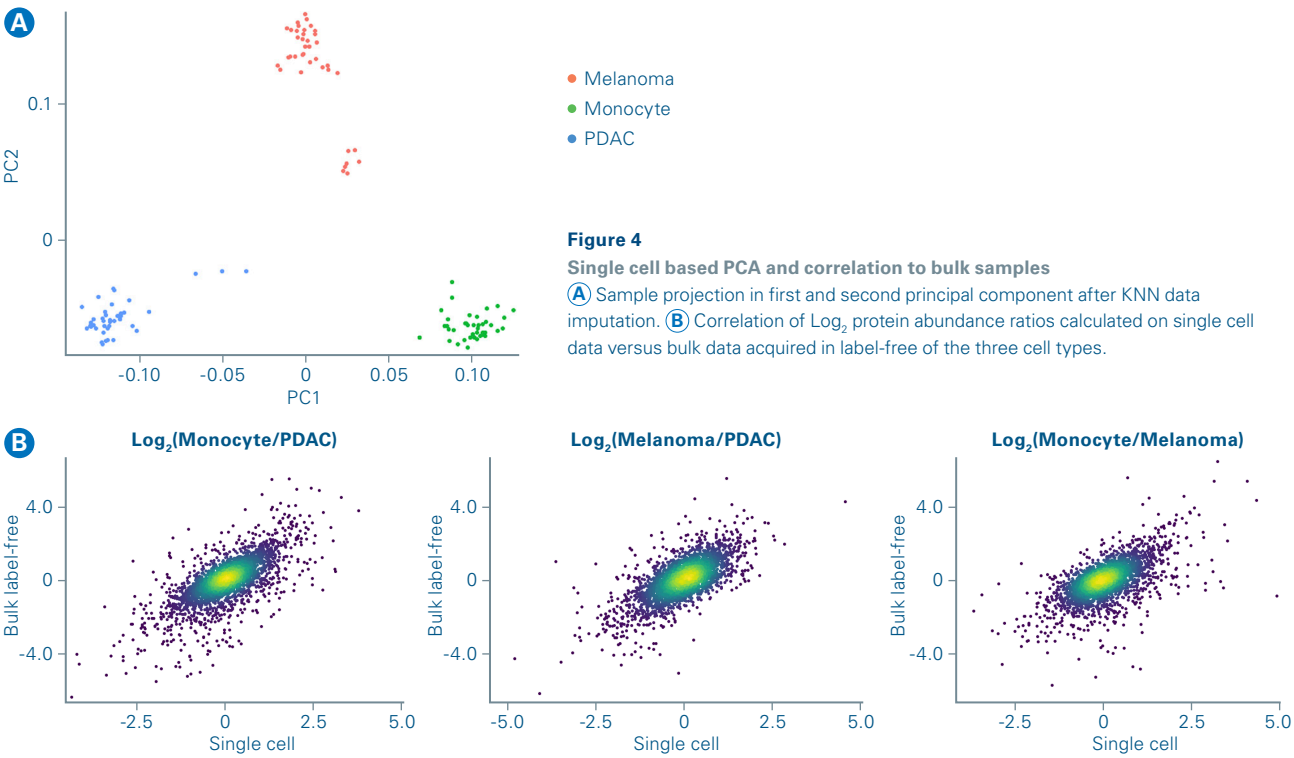

Figure 4

Single cell based PCA and correlation to bulk samples

(A) Sample projection in first and second principal component after KNN data imputation. (B) Correlation of Log_2 protein abundance ratios calculated on single cell data versus bulk data acquired in label-free of the three cell types.

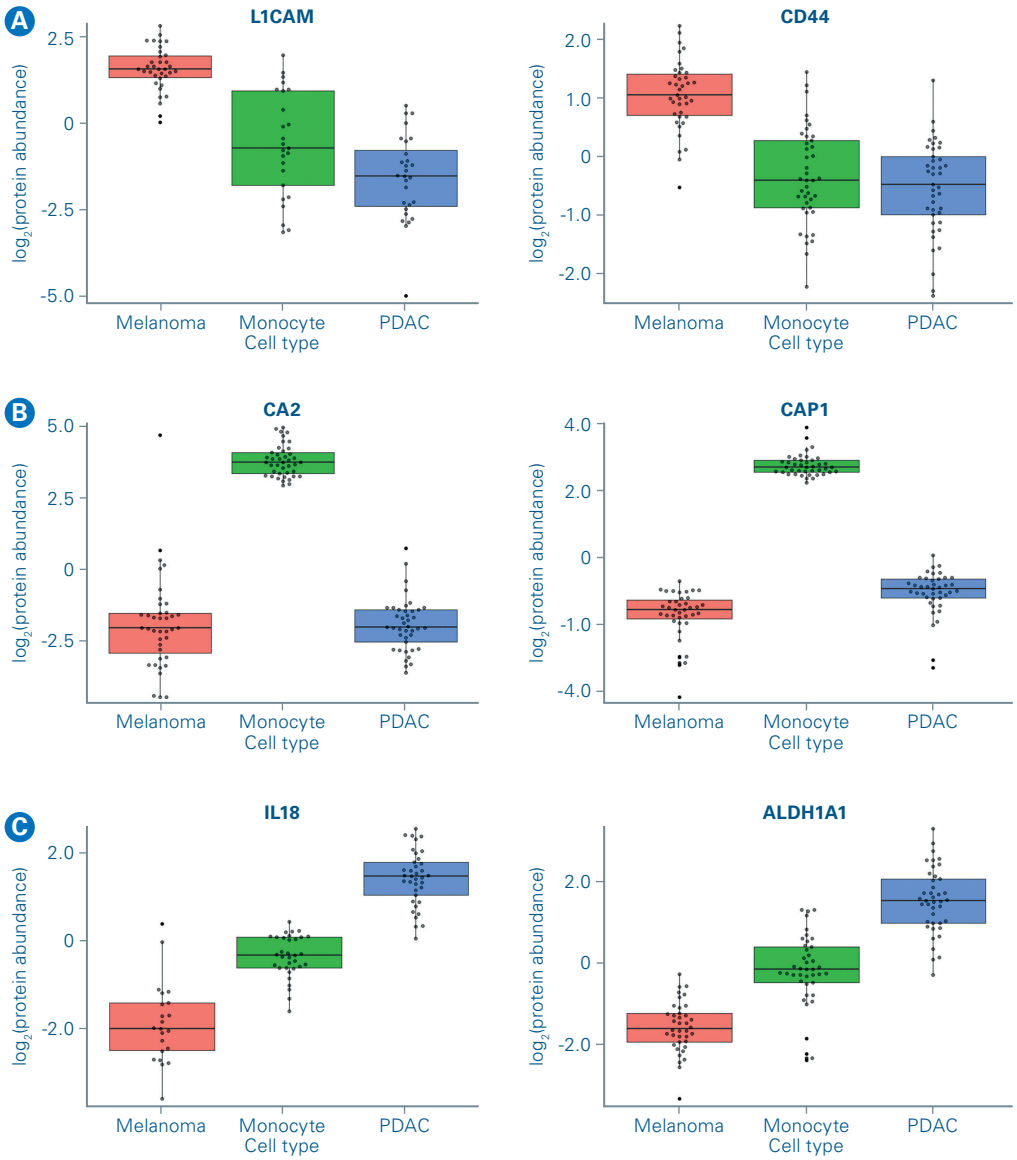

Figure 5

Cell type specific protein abundance profiles

Selection of proteins showing cell type specific protein abundance profiles. (A) Neural cell adhesion molecule L1 (L1CAM) and the CD44 antigen for the Melanoma cell line, (B) Carbonic anhydrase 2 (CA2) and adenylyl cyclase-associated protein 1 (CAP1) for the Monocyte cell line, and (C) interleukin 18 (IL18) and aldehyde dehydrogenase 1 A1 (ALDH1A1) for the PDAC cell line.

Conclusion

- plexDIA for scalable single cell analysis at 96 samples per day speed.
- TIMS separation of MOMA events for interference reduced MS1 and MS2 spectra.
- Identification of > 80,000 precursors from a plexDIA run with mean 14.5 data points per peak enabled by 4x MS1 scan intercalated dia-PASEF method.
- More than 3000 protein groups identified per cell type and about 4400 protein groups in total identified.
- Sample clustering according to cell type with tissue specific protein abundance increases.

References

[1] Leduc A, *et al.* Genome Biol, 2022, **23**:261

[2] Derks J, *et al.* Nat Biotechnol 2023, **41**:50–59

[3] Wallmann G, *et al.* J Proteome Res. 2023, **22**(10):3149-3158

[4] Demichev V, *et al.* Nat Commun 2022, **13**:3944

[5] Ernst AK, *et al.* PLOS ONE, 2018, **13**(2):e0192525

[6] Dietrich A, *et al.* Eur J Cancer. 1997, **33**(6):926-30

[7] Lee S, *et al.* Cell Metab. 2014, 4, **19**(3):484-97

[8] Li Z, *et al.* Cytokine Growth Factor Rev. 2019, **50**:1-12

Further reading

Deep Proteomic Insights from bulk to single cells
www.bruker.com/en/applications/academia-life-science/proteomics/single-cell-proteomics.html

Setting and Maintaining the Single Cell Proteomics Benchmark with the timsTOF Ultra in action
www.bruker.com/de/news-and-events/webinars/2023/setting-and-maintaining-the-single-cell-proteomics-benchmark.html

Speed meets Sensitivity with the timsTOF Ultra
www.bruker.com/en/products-and-solutions/mass-spectrometry/timstof/timstof-ultra.html

Everything you need. Maximum Sensitivity, Speed, and Robustness with PepSep Ultra and Ultra XL
www.bruker.com/en/products-and-solutions/mass-spectrometry/lc-ms/pepsep-columns.html

Immediately dive into results of your experiment
www.bruker.com/en/products-and-solutions/mass-spectrometry/ms-software/proteoscape.html

The gold standard for DIA proteomics analysis: Spectronaut®.
biognosys.com/software/spectronaut/

Expanding capabilities, see the timsTOF SCP in action
www.youtube.com/watch?v=3zKvc73HmQ4&t=57s

PASEF: discover the advantage of SPEED with Confidence
www.bruker.com/de/products-and-solutions/mass-spectrometry/timstof/pasef.html

Bruker 4D-Proteomics™ Ecosystem

Benefit	Feature
Universal range of applications	**NanoElute 2** to deliver robustness and performance without compromise.
Ultra separation	PepSep Ultra and PepSep Ultra XL deliver ultra sensitivity and reproducibility at PASEF® speeds, ensuring high-quality results.
Single Cell Handling	**cellenONE®** system to provide hands-free end-to-end cell sorting and sample preparation for single cells
Ultimate performance	CSI ion sources: **CaptiveSpray Ultra** moves boundaries and delivers robust performance and simplicity for outstanding sensitivity.
Higher throughput, deeper profiles	**TIMS** spreads sample complexity across the mobility dimension for fast turnaround, without compromise.
Enhanced confidence with TIMS	**Mobility offset, mass aligned** (MOMA) for confident identification and quantification of isobaric or nearly isobaric peptides.
Powerful insights	**Bruker ProteoScape**™: Real-time data processing for large experiments.
Limitless processing	**Spectronaut®** simplifies complex DIA workflows with unmatched sensitivity and accelerated data analysis.

4D-Proteomics™ Ecosystem by Bruker, empowering unmatched speed, selectivity, sensitivity, and robustness for the analysis of large cohorts of single cell cohorts.

For Research Use Only. Not for use in clinical diagnostic procedures.

Bruker Switzerland AG

Fällanden · Switzerland
Phone +41 44 825 91 11

Bruker Scientific LLC

Billerica, MA · USA
Phone +1 (978) 663-3660

ms.sales.bdal@bruker.com – www.bruker.com

Learn more at **www.bruker.com/timstof-ultra**