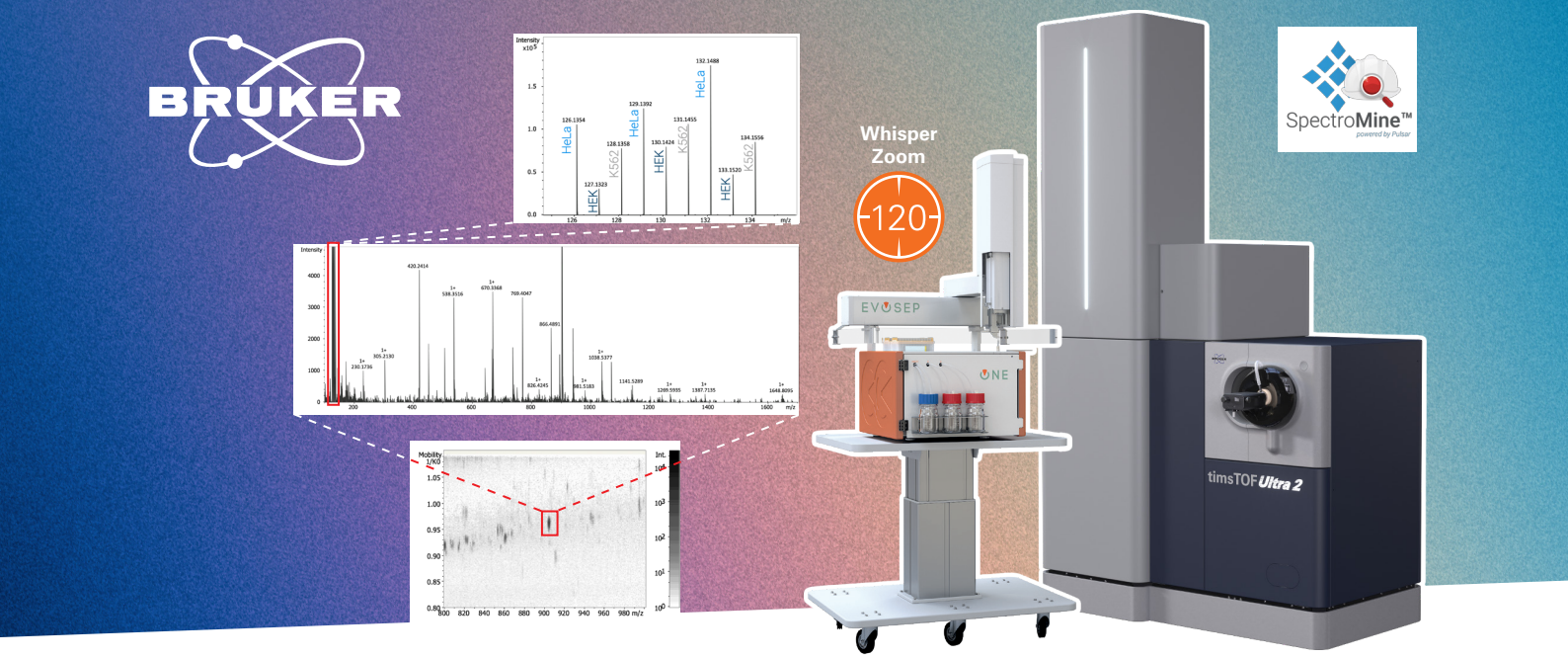

Exceeding 1000 cells per day – scalable single cell analysis using the Evosep Whisper Zoom method on the timsTOF Ultra 2

Scalability is one of the crucial factors for single cell proteomics to allow data collection to achieve reliable statistical power for elucidating heterogeneity in ultra-high sensitivity single cell and near single cell tissue input applications.

Abstract

Multiplexing as well as fast chromatography with low overhead time aids in upscaling proteome analyses. Here, we demonstrate applicability of the new Whisper Zoom 120 samples per day (SPD) method for exceeding >1,000 SPD in a multiplexing approach with data acquisition with a TMTpro-optimized MS/MS stepping dda-PASEF® method on the timsTOF Ultra 2 and data analysis using SpectroMine 4.5. Protein digests of HeLa, HEK293, and K562 cell lysates were labeled with a subset of the 16-plex TMTpro for 9-plexicity at single cell level equivalents analyzing 1080 samples per day at a depth of up to 1500 protein groups per cell within a 5 min active gradient.

Keywords:
TMTpro, Whisper Zoom 120 SPD, timsTOF Ultra 2, 1000 SPD, SpectroMine 4.5

Introduction

Ultra-high sensitivity mass spectrometry has vitalized single cell proteomics over the past few years. Especially label-free approaches in data independent acquisition mode have dramatically increased proteome depth. However, label-free applications are commonly limited in throughput and their scalability typically ends in the low hundreds of samples range. High sensitivity LC applications using the new Whisper Zoom methods in 20, 40, 80, and 120 samples per day (SPD) are a great addition to scaling up label-free proteomics with reduced sample overhead time of 2-3 min. However, the maximum throughput per day reachable in label-free is 120 SPD.

Christoph Krisp[1,2], Dorte Bekker-Jensen[3], Ole Bjeld Hørning[3], Nicolai Bache[3], Torsten Müller[2], Markus Lubeck[2], Daniel Hornburg[4], [1]MDC-Bruker Center of Excellence for Single Cell Omics, Max-Delbrück Center - Max Delbrück Center - Berlin Institute for Medical Systems Biology (MDC-BIMSB), Berlin, Germany; [2]Bruker Daltonics GmbH & Co. KG, Bremen, Germany; [3]Evosep, Odense, Denmark; [4]Bruker Scientific LLC, Billerica, MA; USA

Multiplexing is hence the next logical step. Applications like non-isobaric label-based multiplexing such as plexDIA (1, 2) or mDIA (3) is a suitable combination of upscaling by labeling coupled to sensitive and deep proteome coverage implementing DIA scan modes. However, non-isobaric multiplexing leads to richer but also more complex MS1 and MS2 level information without increasing precursor intensity with an increase of plexing and is currently still limited to low level multiplexing, leading to several hundred cells analyzable per day. Isobaric tag-based multiplexing like TMT (4), iTRAQ (5), or EASI-Tag (6) allow for higher multiplexing rates and higher scalability potential and the benefit of precursor intensity increase with the number of plexing without increasing MS1 complexity. However, due to MS2 reporter-based quantification, it is typically incompatible with DIA acquisition methods because of precursor co-isolation and is thus analyzed in data dependent acquisition mode. To increase proteome coverage, sample fractionation is commonly applied, leading to extended analysis time per sample and low throughput. Further, to increase detectability of precursors in MS2 level multiplexed single cell experiments, reference channels are added which aid in precursor detection but may also cause skewing of data depending on which channels are used for the reference sample. Further, MS2 based multiplexing suffers more frequently from batch effects caused by missing data not by random, effects hard to account for by data normalization and imputation. In particular for cell phenotyping in large sample sets such as tissue dissection experiments, cell release by tissue dissociation, or circulating cell isolations like peripheral blood mononuclear cells, higher throughput is essential and some reduction in ultimate proteome coverage may be acceptable in exchange for increased throughput.

Trapped ion mobility spectrometry (TIMS) adds an orthogonal separation technique to liquid chromatography, reducing complexity by separating out molecules with near identical molecular weight but different shape, mobility offset mass aligned (MOMA), drastically reducing precursor co-isolation and thus ratio compression effects on reporter level typically seen in non-CCS enables mass spectrometers. This is an ideal feature for TMT-based multiplexing approaches especially for low sample inputs like single cell, where each sample manipulation, e.g. transfer steps, or fractionation, leads to substantial losses and allows for fast chromatographic separation.

Here, we highlight the applicability of TMTpro in a 9-plex setting at single cell level equivalence analyzed with fast and sensitive chromatographic separation using Whisper Zoom 120 SPD on the Evosep with a TMTpro optimized MS/MS stepping dda-PASEF acquisition method on the timsTOF Ultra 2, realizing throughput of 1080 individual samples in one day.

Label	Condition 1	Condition 2	Condition 3	Condition 4
126	HeLa	K562	HEK293	K562
127 ^{15}N	HEK293	HeLa	K562	HEK293
128 ^{15}N	K562	HEK293	HeLa	HeLa
129 ^{15}N	HeLa	K562	HEK293	K562
130 ^{15}N	HEK293	HeLa	K562	HEK293
131 ^{15}N	K562	HEK293	HeLa	HeLa
132 ^{15}N	HeLa	K562	HEK293	K562
133 ^{15}N	HEK293	HeLa	K562	HEK293
134 ^{15}N	K562	HEK293	HeLa	HeLa

Table 1
TMT labeling conditions
Four different combinatorial TMTpro 9-plex labeling conditions to assess.

Methods

Tryptic digests from cell lysate protein extracts of HEK 293 (human embryonic kidney cells, in house digest), HeLa (Cervical cancer, Pierce) and K562 (Lymphoma, Promega) were pipetted into a 96 well plate (10 ng per well, 40 single cell equivalents). Samples were dried and reconstituted in 10 µL 0.015% n-Dodecyl β-D-maltoside and 0.1M triethylammonium bicarbonate (TEAB) followed by labeling with a subset of TMTpro 16-plex (Thermo Fisher Scientific) using the labels 126, 127^{15}N, 128^{15}N, 129^{15}N, 130^{15}N, 131^{15}N, 132^{15}N, 133^{15}N, and 134^{15}N to generate a 9-plex TMT experiment with 1 Da spacing. Label switches were introduced as shown in Table 1.

Labeling was quenched after 1h incubation with 0.5% hydroxylamine and pooled per condition, dried, and reconstituted in 400 µL

0.1% formic acid (FA). Evotips Pure were washed and equilibrated as recommended in the preparation protocol, for sample loading, 10 µL 0.1% FA and 10 µL of one of the four conditions (2.25 ng, representative of a 9-plex TMTpro multiplexed single cell experiment without carrier channel) were added. Evotips were placed on an Evosep One (Evosep Biosystems) and separated using the Whisper Zoom 120 sample per day (SPD) method using a 5 cm C18 IonOpticks Aurora Rapid 75 column (5 cm x 75 µm x 1.7 µm). Eluting peptides were transferred into a timsTOF Ultra 2 with precursor detection with a TMTpro optimized MS/MS stepping method in dda-PASEF in 300 Hz speed mode and high sensitivity mode enabled. Data analysis was performed in SpectroMine 4.5 using a human protein sequences database (Uniprot reviewed, no isoforms, 20,370 entries), setting carbamidomethylation of cysteine residues, and TMTpro on peptide N-terminal residues as well as Lysine residues as fixed modification and oxidation of methionine residues and protein N-terminus acetylation as variable modifications (Figure 1).

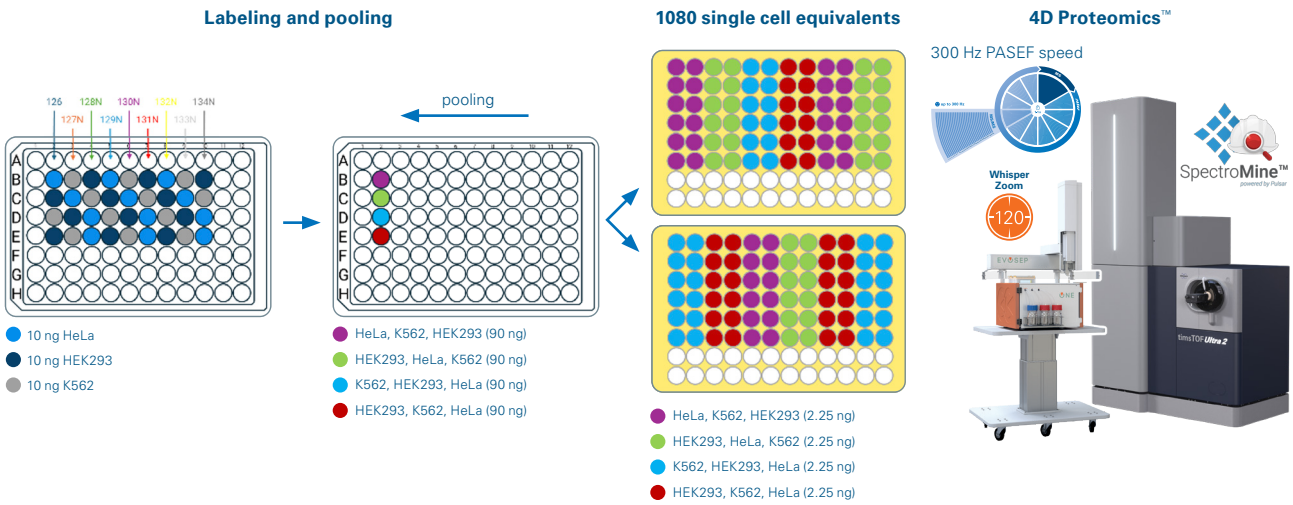

Figure 1

TMT Multiplexing schematic.

Schematic depicting TMTpro subset sample labeling workflow to achieve 9-plexicity for LC-MS/MS based sample analysis using the Evosep One in Whisper Zoom 120 SPD and precursor detection with a timsTOF Ultra 2 in MS/MS stepping dda-PASEF mode followed by data processing in SpectroMine 4.5.

Results and discussion

In total, 1080 individual samples (360 HEK 293, 360 HeLa, 360 K562) were analyzed in one day across 120 9-plex TMTpro samples using the Whisper Zoom 120 SPD method. The active gradient time was 5 min, 10 min total gradient length, and 2 min overhead time between sample injections. To simulate a 9-plex TMTpro multiplexed experiment without carrier channel, 2.25 ng (9 x 250 pg) from the HEK 293, HeLa, and K562 TMT labeled pool was loaded onto the Evotip. The timsTOF Ultra 2 was operated in dda-PASEF using a TMTpro with a fast gradient optimized MS/MS stepping method. Here, the precursor selection polygon was drawn closely around the densest multiply charged precursor area. Precursors selected for MS/MS characterization were in the first PASEF ramp for MS/MS fragmented with collision energies, transfer times, and pre-pulse storage times optimal for the peptide backbone of TMTpro labeled peptides, followed by a second PASEF ramp for MS/MS fragmented with collision energies, transfer times, and pre-pulse storage times optimal for TMTpro reporter ions in the low *m/z* range. Acquired MS/MS spectra per precursor are merged to obtain a precursor MS/MS spectrum containing both, the peptide backbone fragment information for identification as well as TMTpro reporter ion for quantification. The timsTOF Ultra 2 was operated in 300 Hz speed mode with 3 x two PASEF ramps per cycle leading to a cycle time of 0.74 s, perfectly matching the gradient speed (Figure 2).

Data processing in SpectroMine 4.5 resulted in the identification of 2963 protein groups identified in at least 1 TMT experiment, 2754 protein groups in at least 2 TMT batches, 1312 protein groups in 50% of all batches, and 684 protein groups were identified in all batches

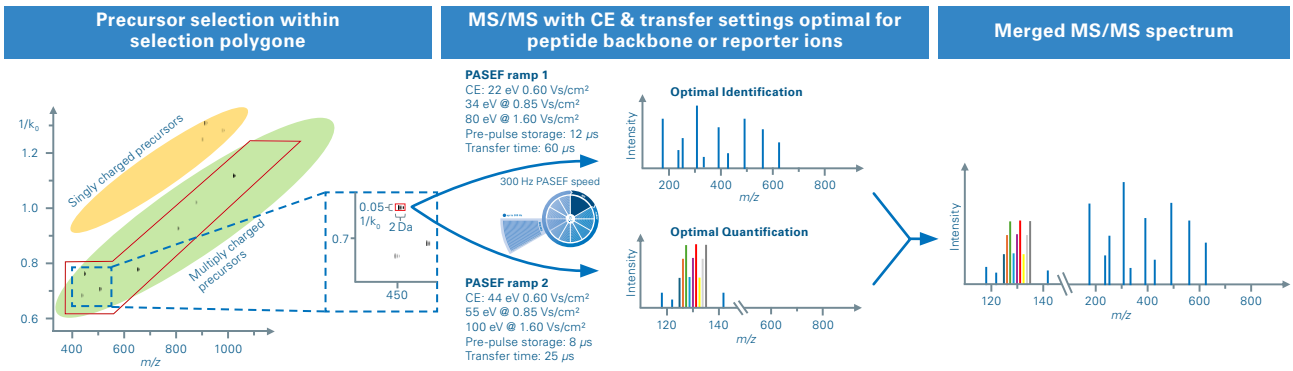

Figure 2

MS/MS stepping dda-PASEF based precursor selection and detection.

Schematic visualization of the precursor selection process for MS/MS stepping dda-PASEF acquisition. Precursor fragmentation with two different collision energy (CE), ion transfer and pre-pulse storage settings analyzed within two consecutive dda-PASEF ramps for optimal peptide fragment and TMTpro reporter detection.

(Figure 3A). The number of identified protein groups per cell type was consistent across all cell types and across the TMT batches with an average of 1450 protein groups per cell type (Figure 3B). Pearson correlation scores were obtained across all 120 TMT batches for proteins quantified in the same cell type (>0.93) and comparing across all three compared cell types (0.8-0.9) (Figure 3C), indicative for significant differences in the protein abundance distribution in these cells. The reproducibility within the TMT batches was excellent with on average 7% coefficient of variation (CV) per cell type and on average 9% CV per TMT batch (Figure 3C). Batch to batch variance was much higher, however, since the expected biological differences between the cell types are rather large, no further batch effect normalization was performed.

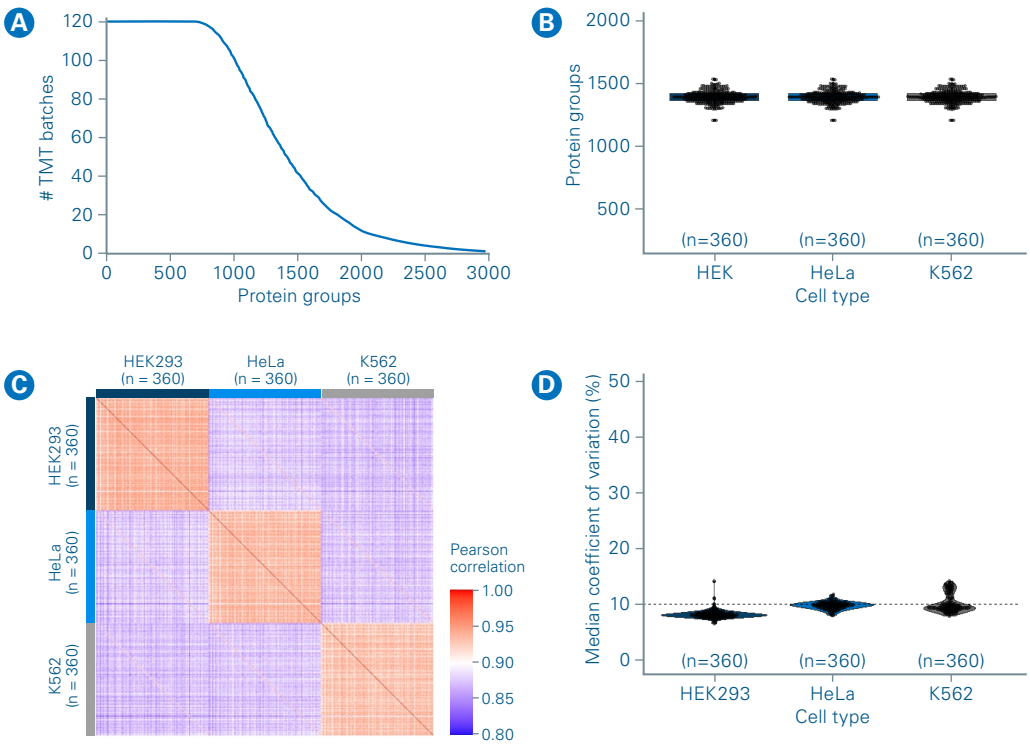

Figure 3

Database search results and reproducibility assessment.

(A) Data completeness assessment across the 120 TMT batches, (B) cell type-based protein group identification rates across the 1080 individual samples. (C) Heatmap visualization of the Pearson correlation scores for protein abundance patterns across the 1080 samples with highest correlation within the cell type group. (D) Violin plot showing the coefficient of variation distribution within the three preparation replicates per cell line and TMT batch with median values of less than 10% for all three cell types.

Principal component analysis of all protein groups identified in all samples across all TMT batches (676 protein groups) demonstrated excellent separation of the 3 cell types in the first two components. However, TMT label-based biases in the individual cell type clustering can be observed (Figure 4A). KEGG pathway analysis was performed on the 676 proteins identified in all 1080 samples using the STRING (7). The top 10 KEGG pathways represent processes viable

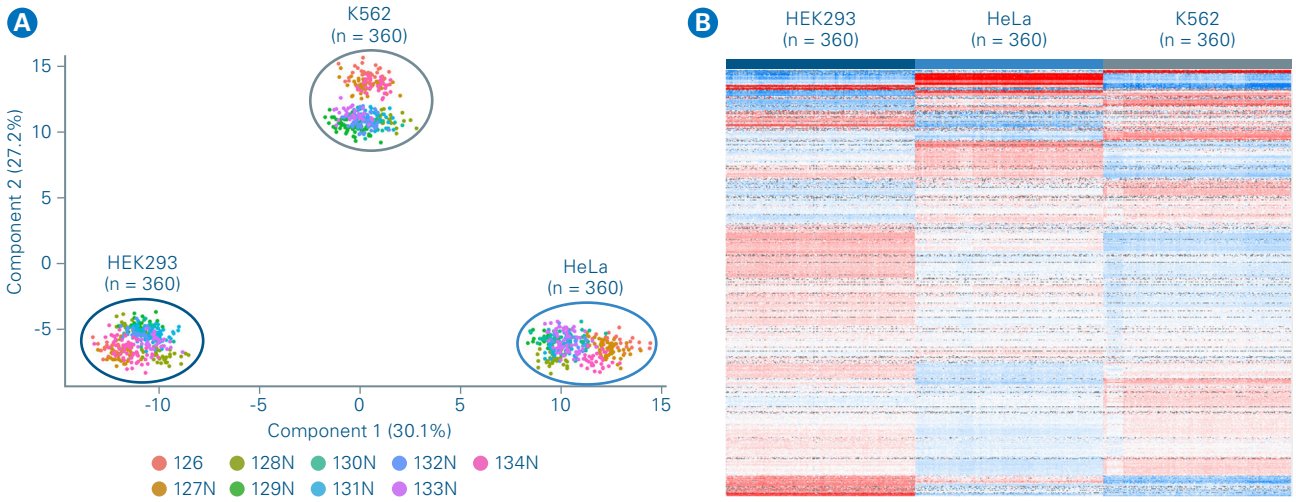

Figure 4

Cell type specific quantitative differences.

Ⓐ Sample projection in the first and second component of a principal component analysis of all proteins identified in all 1080 samples (676 protein groups) showing clear cell type separation Ⓑ Heatmap visualization of normalized protein abundance profiles in HeLa, HEK 293 and K562 digests including proteins found in at least 50% of all samples.

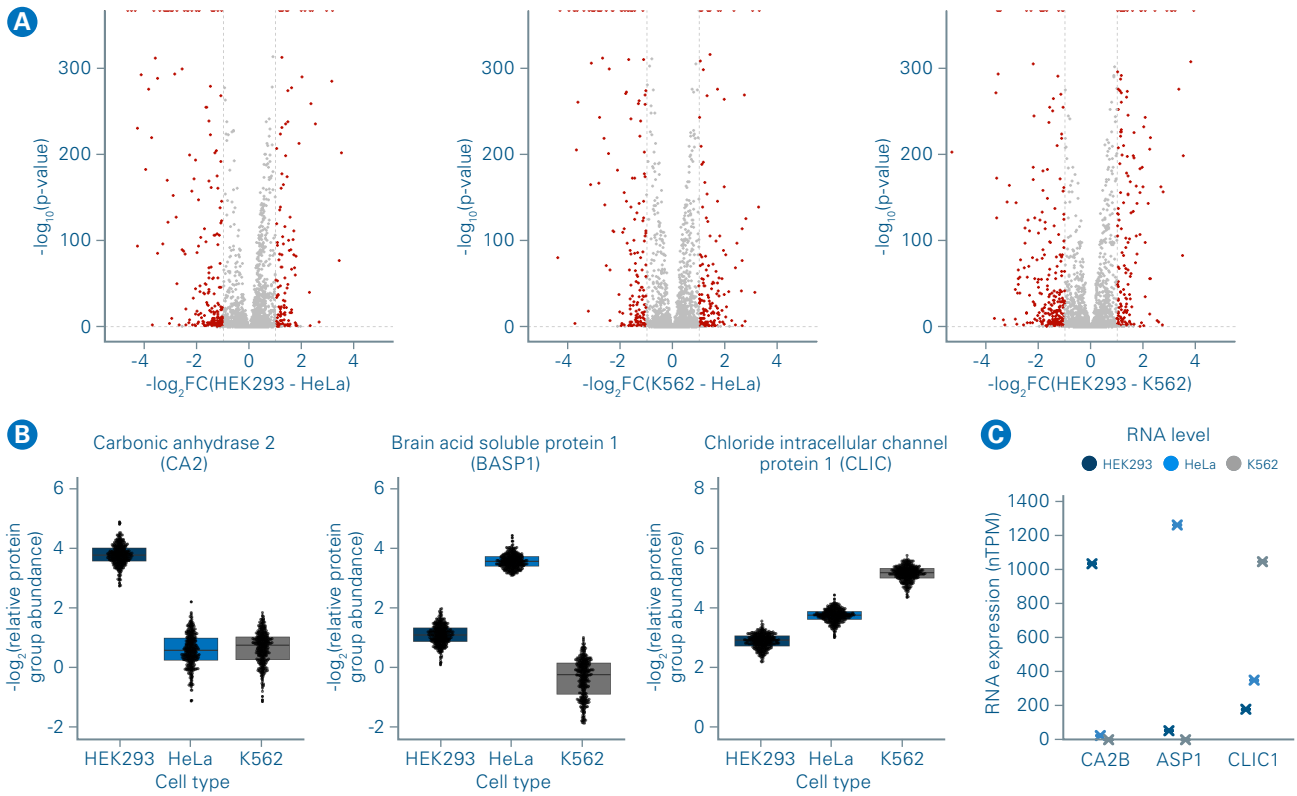

Figure 5

Cell type specific protein abundance profiles.

Ⓐ Volcano plot depicting p-value significant (-log(p-value) > 1.30103) and abundance difference of abs(log_2(fold chance)) > 1 proteins identified in HEK 293, HeLa, or K562. Ⓑ Representative cytosolic protein abundance profiles of proteins among the most significantly different when comparing the three cell lines to each other; carbonic anhydrase 2 (CA2) for HEK 293, brain acid soluble protein 1 (BASP1) for HeLa, and chloride intracellular channel protein 1 (CLIC1). Ⓒ Normalized RNA transcription level of CA2, BASP1 and CLIC1 as reported in the Human Protein Atlas (proteinatlas.org, 8).

cells highly depend on such as carbon metabolism, TCA cycle, DNA replication, proteasome, or spliceosome functions (Table 2). Heatmap representation of relative protein abundance depicts differences and similarities between the three cell types (Figure 4B).

Protein abundances in the three investigated cell lines were quite different as depicted in the Volcano plots in Figure 5A. To evaluate if differences seen between the three cell types are biologically driven and not digest preparation biased, a cytoplasmic protein with the lowest p-value and high protein abundance difference to the two other cell types was selected and the relative abundance profiles compared to RNA expression levels for this protein in the corresponding cell line reported in the Human Protein Atlas (proteinatlas.org, 8). Carbonic anhydrase 2 (CA2), representative for HEK 293, brain acid soluble protein 1 (BASP1), representative for HeLa cells, and chloride intracellular channel protein 1 (CLIC1), representative for K562 cells, showed normalized transcript per Million (nTPM) values around 1000 nTPM and low values in the other two cell types. The protein abundance across the 1080 individual samples analyzed here represent comparable profile patterns (Figure 5B and 5C) driving differentiator in the cell lines investigated here are based on biological differences.

KEGG Pathway	KEGG ID	# of associated proteins identified	Strength	FDR
Proteasome	hsa03050	24 of 43	1.21	2.73E-17
Ribosome	hsa03010	67 of 131	1.18	9.41E-47
Spliceosome	hsa03040	52 of 132	1.06	1.01E-31
Aminoacyl-tRNA biosynthesis	hsa00970	17 of 44	1.05	1.95E-10
Citrate cycle (TCA cycle)	hsa00020	10 of 28	1.02	5.53E-06
Non-homologous end-joining	hsa03450	4 of 12	0.99	0.014
Biosynthesis of amino acids	hsa01230	22 of 73	0.95	1.12E-11
Pyruvate metabolism	hsa00620	11 of 36	0.95	5.49E-06
Carbon metabolism	hsa01200	32 of 116	0.91	7.34E-16
DNA replication	hsa03030	10 of 36	0.91	3.08E-05

Table 2
Top ten KEGG pathways represented by the 676 protein groups identified across all 10180 samples.

Conclusion

- Scalability in Single-Cell Proteomics: Multiplexing and fast chromatography enhance scalability, crucial for robust data and statistical power in sensitive single-cell applications.
- High Throughput with Whisper Zoom: The Whisper Zoom method allows over 1000 samples per day using TMTpro multiplexing and timsTOF Ultra 2.
- Single-Cell Proteomics Efficiency: The method analyzes 1080 samples/day with up to 1500 protein groups per cell in a 5-minute gradient using 9-plex TMTpro labeling.
- Benefit of the timsTOF platform robustness as the true game-changer for achieving genuine high throughput. Being fast is pointless if you can't maintain that performance over time.

References

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

[2] AppNote LCMS-222, 1911577, 2024

[3] Thielert M (2023). Mol Syst Biol. **19**:e11503

[4] Li J, et al. (2020). Nat Methods. **17**:399–404

[5] Wiese S. et al. (2007). Proteomics. **7**:340-350

[6] Virreira Winter S, et al. (2018). Nat Methods. **15**:527–530

[7] Szklarczyk et al. (2023). Nucleic Acids Res. **51**(D1): D638-D646

[8] Pontén F, et al. (2008). J Pathol. 387-393

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 2
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/

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 2** 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® 19** module integrated into Bruker ProteoScape

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**