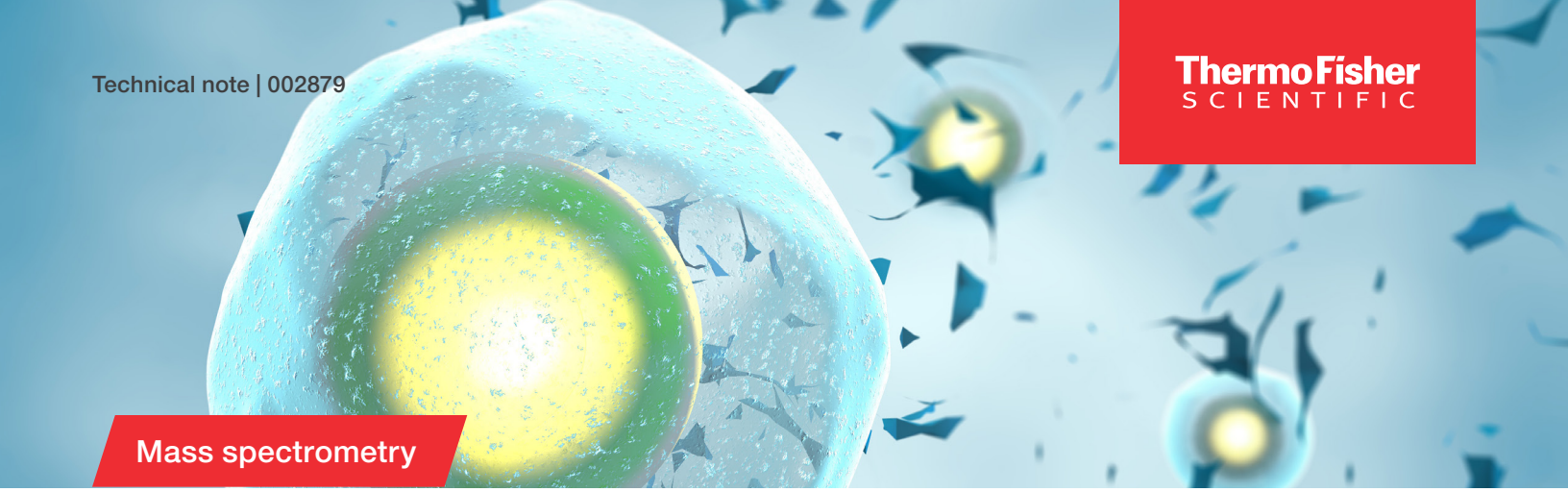
Technical note | 002879



Mass spectrometry

Deeper proteome coverage and faster throughput for low-input and single cell samples on the Orbitrap Ascend MultiOmics Tribrid mass spectrometer

Authors

Fernanda Salvato[1], Bernard Delanghe[2], Julia Kraegenbring[2], David Hartlmayr[3], Anjali Seth[3], Amirmansoor Hakimi[1], Tonya Pekar Hart[1]

[1]Thermo Fisher Scientific, San Jose, CA, USA

[2]Thermo Fisher Scientific, Bremen, Germany

[3]Cellenion, Lyon, France

Keywords

Orbitrap Ascend MultiOmics mass spectrometer, Vanquish Neo UHPLC, FAIMS Pro interface, data-independent acquisition, proteomics, Tribrid

Goal

To assess proteome coverage and sample throughput for low input and single cell samples using a library-free data-independent acquisition (DIA) method on the Thermo Scientific™ Orbitrap™ Ascend MultiOmics Tribrid™ mass spectrometer.

Introduction

In recent years, there has been a shift to analyzing smaller sample quantities, driven by an interest in profiling the proteome of individual cells. This increased focus on LC-MS analysis of limited sample amounts requires the highest possible sensitivity. At the same time, throughput, chromatographic performance, quantitative accuracy, and precision must be preserved to generate sufficiently high quality data from large data sets to draw meaningful biological insights. This trend toward analyzing smaller sample quantities reflects the need to understand the heterogeneous nature of biological systems through the dissection of complex systems into their individual parts, or specifically, individual cells. In contrast to the traditional proteomics approaches that read out the average of a sample, single-cell proteomics provides insights into the exact nature of each of the components in such samples.

The Orbitrap Ascend MultiOmics MS offers versatility and performance on a single platform. The improved front end makes the instrument the most sensitive and fastest Tribrid mass spectrometer on the market. The already impressive sensitivity is further boosted by the combination with the Thermo Scientific™ FAIMS Pro™ interface, which removes background ions with ease. The Thermo Scientific™ Vanquish™ Neo UHPLC system delivers maximum performance in terms of precise low flow rates and gradient formation, as well as low injection volumes, an important benefit for single-cell proteomics and limited sample amounts.

In this technical note, we present a low-input proteomics sample workflow (Figure 1) using the Thermo Scientific™ Pierce™ HeLa protein digest standard in nanogram and picogram loads on the Orbitrap Ascend MultiOmics MS to demonstrate sensitivity and depth of proteome coverage using a library-free DIA approach.

Experimental

Recommended consumables

- Fisher Scientific™ LC-MS grade water with 0.1% formic acid (FA) (P/N LS118-500)
- Fisher Scientific™ LC-MS grade 80% acetonitrile (ACN) with 0.1% formic acid (P/N LS122500)
- Fisher Scientific™ LC-MS grade formic acid (P/N A117-50)
- Fisher Scientific™ Optima™ LC-MS grade water (P/N 10505904)
- Fisher Scientific™ Optima™ LC-MS acetonitrile (P/N A955-1)
- Fisher Scientific™ Optima™ LC-MS isopropanol (P/N A461-212)
- DDM (Dodecyl β D maltoside) (P/N D4641 – Sigma)

Samples

- Pierce HeLa protein digest standard (P/N 88328)
- HeLa cells prepared on Cellenion cellenONE™

LC columns

- IonOpticks Aurora Ultimate™ TS 25 cm (AUR3-25075C18-TS)

HPLC system

- Vanquish Neo UHPLC system including:
  - Vanquish Neo Pump/Autosampler (P/N VN-S10-A-01)
  - Column Compartment (P/N VN-C10-A-01)

Mass spectrometer

- Orbitrap Ascend MultiOmics MS
- FAIMS Pro interface
- Thermo Scientific™ EasySpray™ ion source

Data analysis software

- Thermo Scientific™ Proteome Discoverer™ 3.1 software with CHIMERYS™ intelligent search algorithm by MSAID
- Spectronaut™ 18 software (Biognosys AG)

HeLa standard

All proteomics experiments were performed using the Pierce HeLa protein digest standard. 200 µL of resuspension buffer (0.015% DDM prepared in 0.1% formic acid) was added to the vial containing 20 µg of protein digest. The vial was then sonicated at room temperature for 5 minutes, making a final concentration of 100 ng/µL. To the autosampler vial, 95 µL of resuspension buffer and 5 µL of 100 ng/µL HeLa digest were added to make the final concentration 5 ng/µL. This solution was vortexed for 30 s. All injections were done from the same vial.

HeLa cells

HeLa cells were sorted and prepared using cellenONE and Cellenion proteoChip LF 48 and transferred manually to the wells of a 384 well-plate.

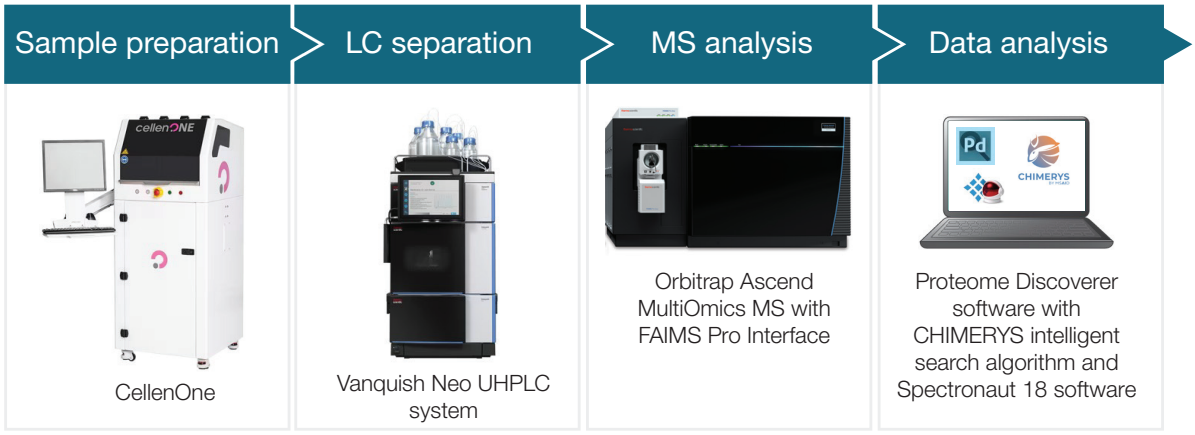


Figure 1. Single-cell proteomics workflow with the Orbitrap Ascend MultiOmics MS for new standards in sensitivity and depth of coverage for single-cell samples

LC conditions

To evaluate the performance of our workflow, a 50 sample per day (SPD) method was employed consisting of a 22-minute gradient and 3-minute washing, equilibration, and injection steps for a total injection-to-injection cycle time of 25 minutes.

HPLC conditions are described in Table 1, with the 50 SPD method gradient details in Table 2.

Table 1. HPLC conditions

HPLC method parameters	
Mobile phase A	0.1% formic acid (FA) in water
Mobile phase B	0.1% FA in 80% acetonitrile (ACN)
Flow rate	0.2 µL/min
Column	Aurora Ultimate TS 25 cm
Colum temperature	50 °C
Autosampler temperature	7 °C
Injection wash solvents	Strong wash: 0.1% FA in 80% ACN Weak wash: 0.1% FA in water
Needle wash	Enabled after-draw

Table 2. 50 SPD method

	Time (min)	Duration (min)	%B	Flow rate (µL/min)
	Run			
Active gradient	0.0	0.0	1.0	0.45
	0.1	0.1	4.0	0.45
	1.9	1.8	12.0	0.45
	2.0	0.1	12.0	0.20
	12.0	10.0	22.5	0.20
	19.5	7.5	40.0	0.20
	Column wash			
	22.0	2.5	99.0	0.3
	25.0	3.0	99.0	0.3
	Stop run			
	Column equilibration			

MS parameters

The Orbitrap Ascend MultiOmics MS was operated with the parameters shown in Table 3.

Table 3. MS parameters

Source parameters	
Spray voltage	1.9 kV
Capillary temperature	275 °C
FAIMS CV	–50
Total carrier gas flow (L/min)	3.5
Orbitrap MS full scan parameters	
Resolution	120 k
Normalized AGC target	300%
Maximum IT	Auto
RF lens	45%
Scan range	400–800 *m/z*
Orbitrap DIA MS^2 scan parameters	
Precursor mass range	400–800 *m/z*
Resolution	60,000
Isolation window (*m/z*)	40 (>1 ng load) or 50 (<1 ng load)
Number of scan events	10
HCD normalized collision energy	28%
First mass	120 *m/z*
RF lens	60%
Normalized AGC target	1,000%
Loop control	All
Max IT time (ms)	118

Data processing parameters

The HeLa protein digest standard dilution data contained triplicates of 50 pg, 100 pg, 250 pg, 500 pg, 1 ng, 2.5 ng, 5 ng, and 10 ng. DIA runs were processed using triplicate runs and a library-free approach using the directDIA+™ workflow on Spectronaut 18 software or Proteome Discoverer 3.1 software with CHIMERYS intelligent search algorithm. The Human UniProt protein database (20,607 FASTA entries) was used for the library-free searches. False-discovery rate (FDR) of 1% was applied at the precursor, peptide, and protein levels. Library-based searches were processed with Spectronaut 18 software. Spectral libraries were generated using DIA data with the Pulsar search engine in Spectronaut 18 software against the Human UniProt Protein database.

Results and discussion

High-throughput and high protein coverage from low sample amounts

We created a dilution series from 50 pg to 10 ng HeLa digest and analyzed the digests with a 50 SPD workflow using the optimized LC-MS methods (see Tables 1 to 3). We identified >1,200 protein groups and >4,300 peptides from 50 pg HeLa digest and >3,200 protein groups and 14,400 peptides from 250 pg HeLa digest in a library-free workflow. Increasing the sample amount to 10 ng, over 5,500 protein groups and 33,300 peptides were identified (Figure 2). Recent studies show increased protein and peptide identifications for DIA data using library-free approaches, incorporating machine learning-based in-silico library predictions from a protein database. For low sample amounts, it was observed that a project-specific library generated from limited sample amounts yielded better results compared to a library-free approach. This could be due to the reduced spectral complexity in low sample amounts that negatively impacts the library-free data analysis approach.

To evaluate the impact of different library sizes, we generated a DIA library from triplicate runs of 1 ng, 2.5 ng, 5 ng, and 10 ng and a combined library using 5 ng and 10 ng HeLa digests. The 250 pg DIA runs were then searched against the respective libraries, as shown in Figure 3. The DIA library generated from ≥5 ng DIA runs yielded >4,800 proteins from 250 pg HeLa digest triplicate runs. The library, composed of 5 and 10 ng files, resulted in the highest numbers of protein and peptide identifications (Figure 3).

Using a spectral library-based approach with a DIA library comprising 3 x 5 ng and 3 x 10 ng DIA runs (38,698 precursors) resulted in the identification of 2,557 and 4,998 protein groups from 50 pg and 250 pg HeLa protein digests, respectively, and >5,500 proteins were identified from loads above 1 ng (Figure 4).

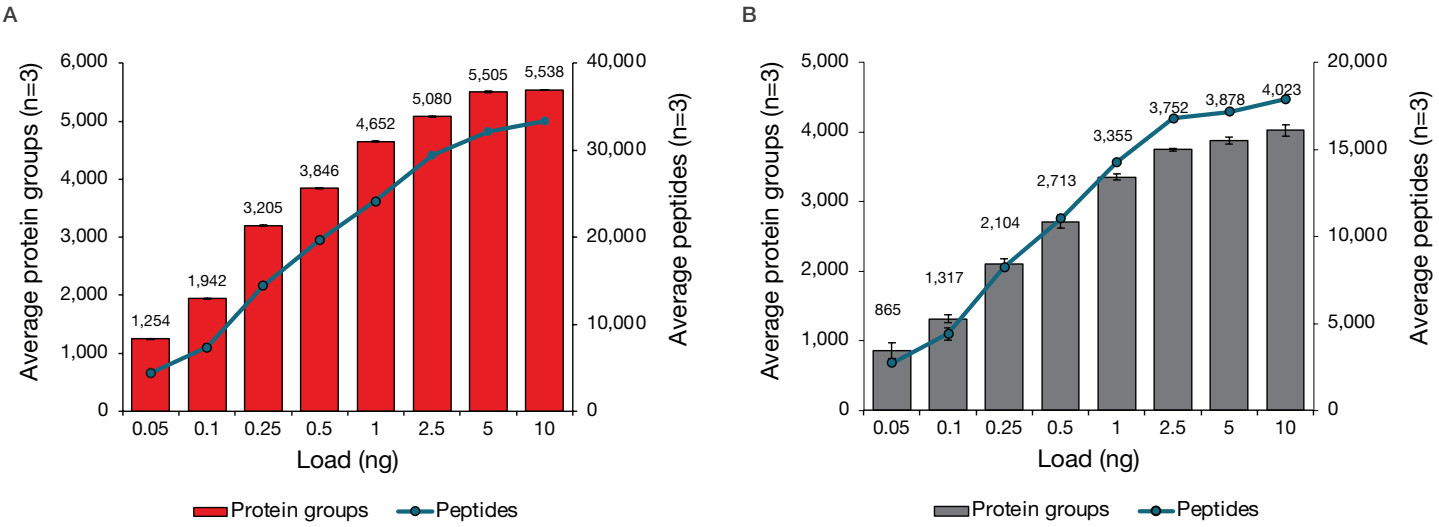


Figure 2. Protein groups and peptides identified using a 50 SPD method with library-free search using Spectronaut 18 software (A) and Proteome Discoverer 3.1 software with CHIMERYS intelligent search algorithm (B)

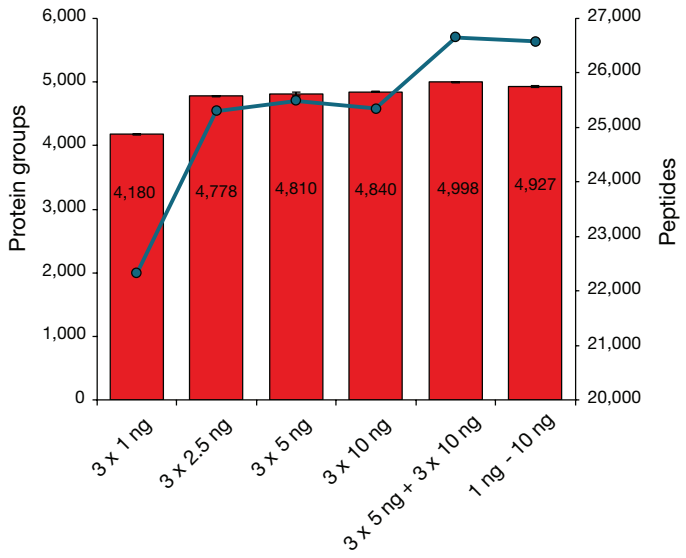


Figure 3. Effect of DIA library size on 250 pg HeLa protein digest standard (n = 3) runs. Files were searched against DIA libraries generated with varying amounts of HeLa digests (x-axis) on Spectronaut 18 software.

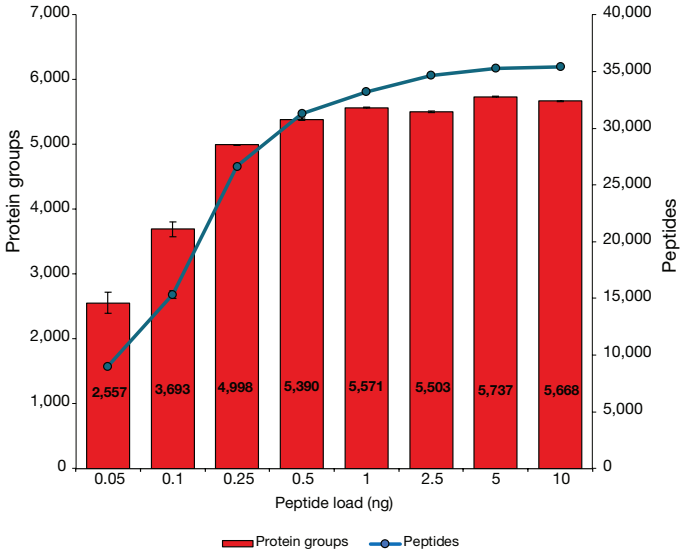


Figure 4. The average number of protein groups and peptides identified using a 50 SPD method with a DIA library-based approach using Spectronaut 18 software

High accuracy and precision of quantitation

In addition to a list of protein identifications, quantitative data is necessary to classify single cells. The quantitative data must be precise and accurate to be able to discern differences between individual cells. We observed that ~93% of identified proteins in the 250 pg HeLa digest could be quantified with <20% CV, and 76% of proteins were quantified with <10% CV using the library-free approach (Figure 5). Using a library-based approach, 79% of identified proteins had <20% CV, and 59% of proteins had <10% CV for a 250 pg HeLa digest (Figure 6).

Median CVs of protein abundances across triplicate runs from 50 pg to 10 ng were <8% and <12% with library-free and library-based searches (Figure 7), respectively. For 250 pg HeLa digest (equivalent to the protein amount found in a single HeLa cell), the median CV was 5.2% and 7.4% with library-free and library-based searches, respectively. This clearly shows highly reproducible measurements spanning a broad concentration range (Figure 7).

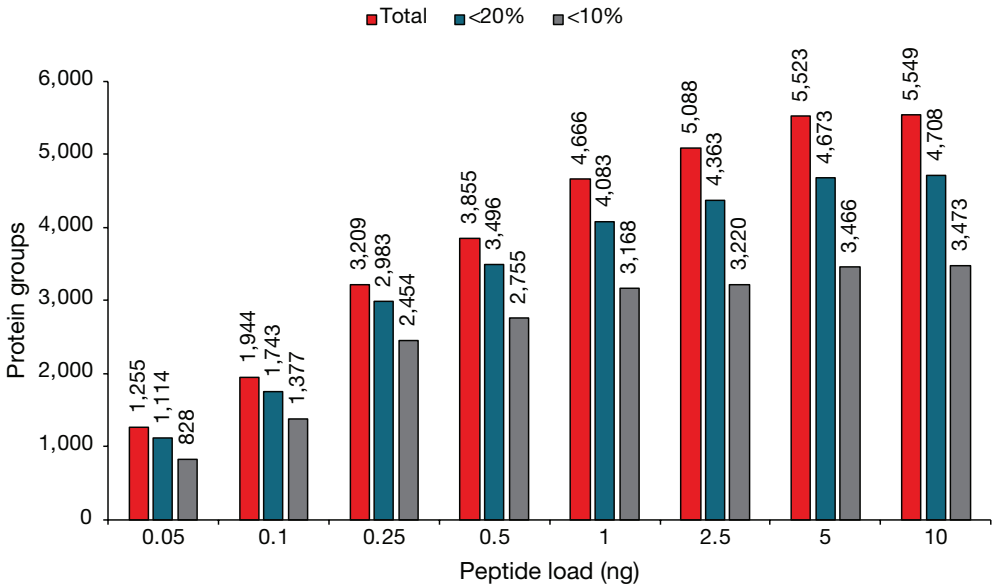


Figure 5. Total protein groups identified in triplicate runs (red) with protein groups CVs <20% (blue) and <10% (gray) analyzed using library-free search using Spectronaut 18 software

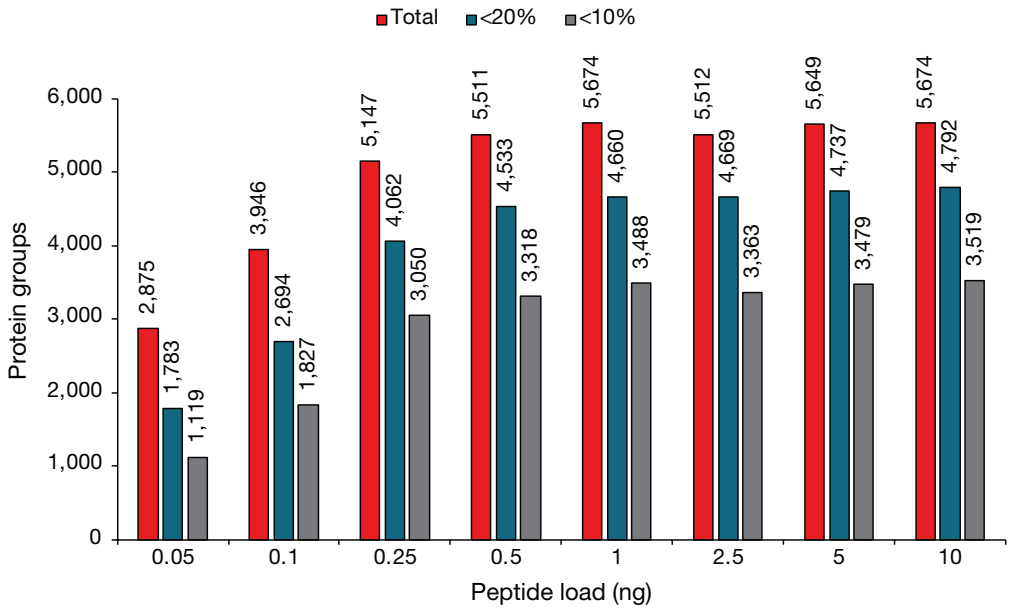


Figure 6. Total protein groups identified in triplicate runs (red) with protein groups CVs <20% (blue) and <10% (gray) analyzed using library-based search using Spectronaut 18 software

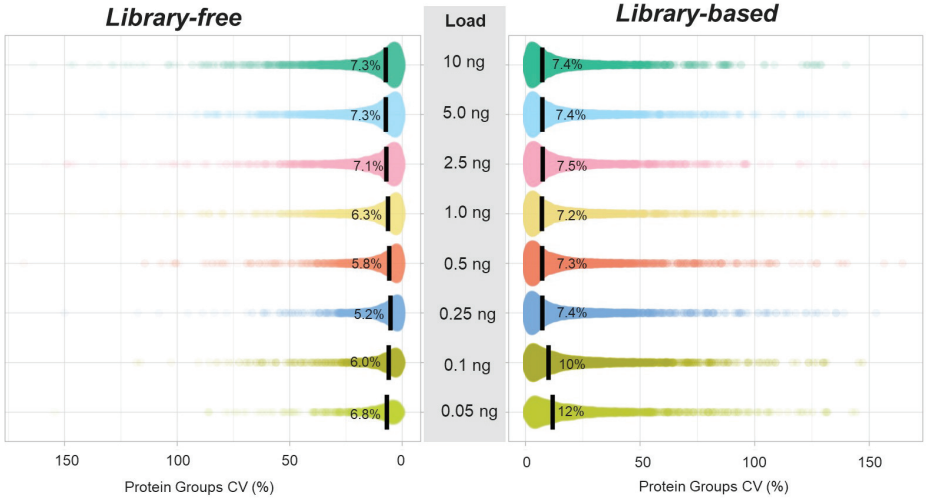


Figure 7. Violin plots indicate the %CV of protein groups identified for each dilution data point (n=3) of HeLa digest using library-free (on the left) and library-based searches (on the right). Black bars with numbers labeled in the figure represent the median %CVs for each load. Data processed in Spectronaut 18 software.

Analyzing individual single cells

To demonstrate the applicability of our DIA workflow for the analysis of individual cells, we utilized HeLa cells sorted and prepared on CellenONE. On average (n=6), 3,762 protein groups were identified from single cells, while 4,131 and 4,634 protein groups were identified from 5 and 10 cells, respectively, using the library-free search (Figure 8). We also demonstrated different DIA search strategies that can be employed when searching single cells or low-load input data to improve identification. Figure 9 summarizes the impact of different search strategies on protein group and peptide identification. In addition to the above-described search strategies—with or without a library—we also included two other strategies using three raw files of 5 cells and/or 10 cells to boost protein group identification. We see a large increase in protein identification when the raw files are processed all together (3,762 protein groups) compared to when files are processed individually (3,101 protein groups). Another boost in identification is observed when 10 cells and 5 cells raw files (three files from each) are added to the library-free search (4,322 and 4,406 protein groups, respectively). A library-based approach was also demonstrated by building a library from 5 and 10 cells raw files (three from each) on Pulsar (Figure 9).

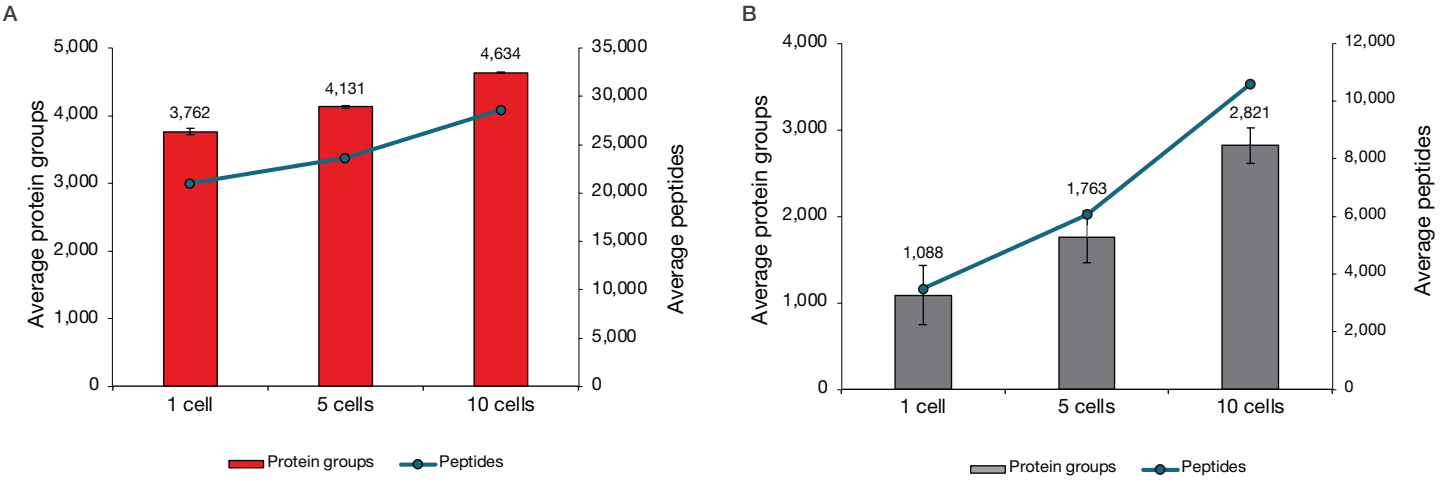


Figure 8. Number of protein groups and peptides identified across single, 5 and 10 cells using the 50 SPD method and library-free approach. Replicates from the same load condition were searched together on Spectronaut 18 software (A) and Proteome Discoverer 3.1 software with CHIMERYS intelligent search algorithm (B).

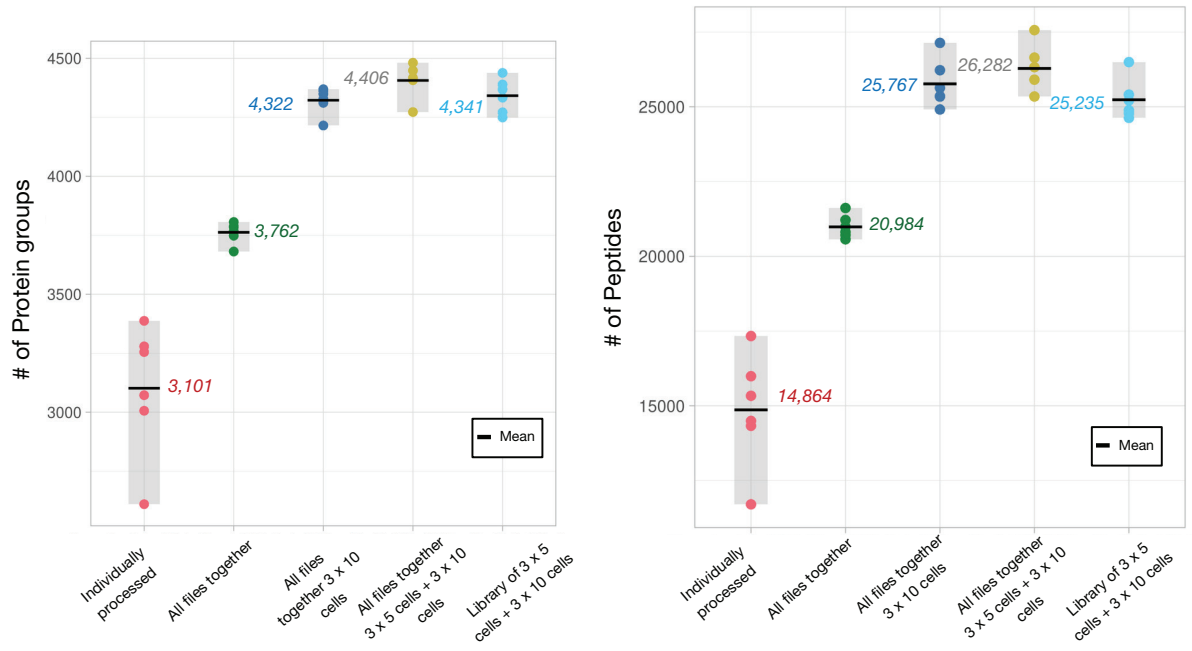


Figure 9. Impact of different DIA searching strategies on protein and peptide identification that can be employed for processing single cell datasets. Black bars represent the mean values of six replicates. Grey bars represent the range of observed data.

High quantitative precision and accuracy were achieved with CVs <7% in the protein groups analyzed in a library-free search at each dilution point. For a 250 pg HeLa digest, CVs were <10% for protein groups quantified using either a library-free or spectral library search.

Summary

- The Orbitrap Ascend MultiOmics MS enables robust, reproducible, rapid, and sensitive deep proteome coverage from low sample amounts using DIA.
- The FAIMS Pro interface improves the signal-to-noise (S/N) ratio with decreased background ions, enhancing spectral quality for improved protein and peptide identifications for low sample amounts.
- Reproducible injections and separations were achieved with the Vanquish Neo UHPLC system, with consistent performance from 50 pg to 10 ng sample loads.

Learn more at **thermofisher.com/singlecellproteomics**