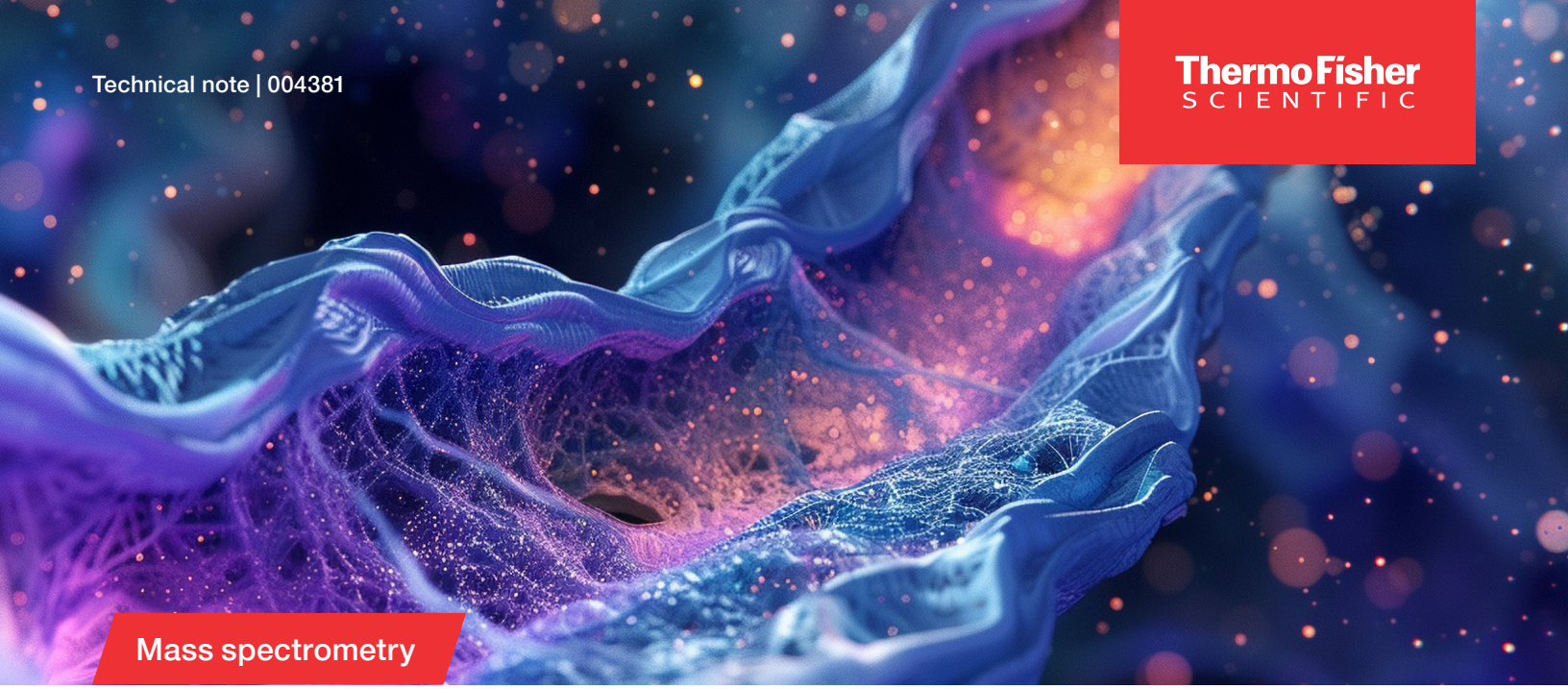

Technical note | 004381

Mass spectrometry

Unlocking the archived proteome: High-throughput, deep FFPE proteome profiling using the Orbitrap Astral mass spectrometer

Authors

Sudipa Maity[1], Amarjeet Flora[2], Bhavin Patel[2], William Comstock[1], Katherine Walker[1], Amirmansoor Hakimi[1], Ellen Casavant[1], Tonya Pekar Hart[1]

[1]Thermo Fisher Scientific, San Jose, CA

[2]Thermo Fisher Scientific, Rockford, IL

Keywords

Orbitrap Astral, OptiSpray ion source, formalin-fixed, paraffin-embedded, FFPE, cancer, lung tumor

Goal

To demonstrate a streamlined, rapid formalin-fixed, paraffin-embedded (FFPE) proteomics workflow using the Thermo Scientific™ Orbitrap™ Astral™ Mass Spectrometer and multiple liquid chromatography (LC) systems, enabling robust and sensitive protein quantitation and scalable elucidation of biological differences with minimal analysis time.

Introduction

FFPE tissue specimens are foundational to clinical pathology and translational cancer research. FFPE preservation enables long-term storage of patient samples while maintaining tissue morphology required for routine histological evaluation and molecular testing. As a result, FFPE tissues represent a uniquely valuable resource for retrospective studies, biomarker discovery, and precision oncology, particularly when paired with rich clinical annotations such as diagnosis, treatment history, and patient outcomes. Compared to freshly collected tissues which are often difficult to obtain, costly to process, and rarely available at scale, FFPE samples are widely accessible through biobanks and archives, offering a cost-effective and clinically relevant foundation for proteomic analyses aligned with real-world oncology workflows.[1]

Despite their clinical importance, proteomic analysis of FFPE tissues presents methodological challenges that can impact reproducibility and sensitivity if not systematically addressed. Formaldehyde fixation introduces protein crosslinking and chemical modifications that complicate protein extraction and enzymatic digestion, while essential preparatory steps such as paraffin removal and crosslink reversal add processing complexity. Traditional FFPE workflows often involve multiple manual steps and extended processing times, increasing variability in peptide recovery and limiting throughput. Recent advances in sample preparation have created new opportunities to overcome these limitations.[2] However, even well-processed FFPE samples can contain residual contaminants that accumulate over time and require frequent instrument cleaning, limiting robustness and scalability.

Recent advances in FFPE-compatible sample preparation and mass spectrometry technologies are helping to overcome these barriers. Here, we implemented a streamlined FFPE proteomics workflow that simplifies tissue processing, protein recovery, and digestion, reducing sample handling while improving peptide yield and reproducibility. Using FFPE lung tissue sections as a representative cancer model, we leveraged the novel Thermo Scientific™ OptiSpray™ Ion Source with the Orbitrap Astral mass spectrometer with two different LC platforms to enable confident detection of low-abundance peptides across a wide input range (20–200 ng). When combined with fast and robust LC methods and high acquisition speeds of 60, 180, and 500 samples per day (SPD), this integrated workflow enables deep, scalable proteomic profiling of FFPE cancer tissues. In addition, the automated OptiSpray ion source helps keep the mass spectrometer cleaner, enabling reliable, deep, and scalable proteomic analysis of archived clinical tissues. Together, these advances facilitate and maintain robust detection of biologically meaningful differences across large sample cohorts, supporting biomarker discovery and translational cancer research at scale.

Experimental

Consumables

- Fisher Chemical™ Optima™ LC/MS Grade Water with 0.1% Formic Acid (v/v) (Part No. LS118-500)
- Fisher Chemical™ Optima™ LC/MS 80% Acetonitrile (ACN), 20% Water with 0.1% Formic Acid (Part No. LS122500)
- Fisher Chemical™ Optima™ LC/MS Grade Formic Acid, 99.0+% (Part No. A117-50)
- Thermo Scientific™ EasyPep™ Mini MS Sample Prep Kit
- Invitrogen™ DDM (n-dodecyl β-D-maltoside) (Part No. BN2005)
- Thermo Scientific™ OptiSpray™ µPAC™ Neo 50 cm Nano Cartridge (Part No. OS-UPAC050NAN)
- Thermo Scientific™ OptiSpray™ µPAC™ Neo High Throughput Cartridge (Part No. OS-UPAC005CAP)
- Evosep™ EV1182 Performance Column
- Evosep™ EV2018 Evotip Pure™ Tip Column

UHPLC system

- Thermo Scientific™ Vanquish™ Neo UHPLC System
- Evosep™ Eno

Mass spectrometer

- OptiSpray ion source (Part No. B51004132)
- Orbitrap Astral mass spectrometer

Software

- Biognosys™ Spectronaut™ software

Methodology

Lung tumor and lung normal samples were de-paraffinized using xylene and sequential ethanol washes. The optimized protein extraction protocol and EasyPep Mini MS sample prep kit were used to prepare the digest samples from FFPE sections of normal and tumor lung samples. Protein concentration was measured using the Thermo Scientific™ Pierce™ Rapid Gold BCA kit. The peptides were quantified using the Thermo Scientific™ Pierce™ fluorometric peptide assay kit before LC-MS/MS analysis.

LC-MS analysis

Samples were injected onto an OptiSpray µPAC Neo 50 cm nano cartridge and OptiSpray µPAC Neo high throughput cartridge and separated using 20 min (60 SPD) and 6.8 min (180 SPD) gradients, respectively, in direct injection mode at 55 °C on a Vanquish Neo UHPLC system. For 500 SPD (2.3 min) separation, samples were processed on the Evosep Eno using an Evosep EV1182 Performance Column at 40 °C. Detailed LC and MS parameters are provided in Tables 1 and 2.

Table 1. Vanquish Neo UHPLC system conditions for direct injection mode.

60 SPD (20 min, OptiSpray µPAC Neo 50 cm cartridge, Part No. OS-UPAC050NAN)			
	Time	%B	Flow (µL/min)
Gradient	0	4	0.75
	0.3	8	0.75
	14.1	30	0.4
	17.5	50	0.4
	18	99	0.75
	20	99	0.75
LC parameters	Column temperature	55 °C	
	Fast loading/equilibration	Pressure control	
	Pressure for loading	450 bar	
	Equilibration factor	3	
	Sampler temperature	7 °C	
180 SPD (6.8 min, OptiSpray µPAC Neo high throughput cartridge, Part No. OS-UPAC005CAP)			
	Time	%B	Flow (µL/min)
Gradient	0	4	2.5
	0.1	8	2.5
	0.9	12.5	2.5
	1	12.5	1.4
	4.5	28.5	1.4
	5.8	50	1.4
	6.2	99	2.5
	6.8	99	2.5
LC parameters	Column temperature	55 °C	
	Fast loading/equilibration	Pressure Control	
	Pressure for loading	450 bar	
	Equilibration factor	3	
	Sampler temperature	7 °C	

Table 2. Mass spectrometer parameters.

MS parameters					
	Vanquish Neo UHPLC system (60 & 180 SPD)			Evosep Eno (500 SPD)	
	Sample amount (ng)	**200**	**20**	**200**	**20**
MS^1	Resolution	240,000	240,000	240,000	240,000
	Scan range	380–980	380–920	400–800	400–800
	AGC	500	800	500	500
	Max-IT	5	100	3	3
MS^2	Scan range	150–2,000	150–2,000	150–2,000	150–2,000
	Isolation window	2	5	4	4
	AGC	500	800	500	500
	Max-IT	3	10	3	3
	HCD	25	25	25	25

Data analysis

The raw DIA files from both the labeled and unlabeled samples were analyzed together using Spectronaut software in directDIA™ mode. Protein identification was conducted using the Pulsar search engine, and the reference database employed was UniProtKB/Swiss-Prot, specifically the human proteome database (Homo sapiens: 20,423 entries) with default search settings. All results were processed and filtered with a 1% precursor and 1% protein group false discovery rate (FDR). Exported output files were imported to RStudio (2023.09.0 Build 463) with R (v4.3.1) for downstream data analysis and visualization.

Results and discussion

Fast and scalable proteomics for clinical FFPE specimens

A streamlined workflow was established enabling deep proteomic profiling of FFPE tissues in less than 1 day (Figure 1). The protocol integrates rapid deparaffinization and protein extraction (~3 h), efficient peptide digestion and cleanup using the EasyPep Mini MS sample prep kit (3–4 h, dry time ~4 h), fast LC–MS analysis (<25 min per sample), and automated data processing (~20 min). By minimizing sample handling and total processing time, the workflow addresses key bottlenecks traditionally associated with FFPE proteomics.

This workflow was successfully applied to FFPE lung tissues, including both normal lung and lung cancer samples (Figure 2A). FFPE tissue sections were stained with hematoxylin and eosin (H&E) to verify tumor and normal status (Figure 2B), followed by preparation of the tissue for mass spectrometry-based proteomic analysis. Efficient deparaffinization and lysis enabled effective reversal of formalin-induced crosslinks and robust protein recovery (Figure 2C), which is critical for downstream enzymatic digestion. The resulting peptides were readily compatible with rapid LC-MS analysis, demonstrating that extensive fractionation or long chromatographic gradients are not required to achieve comprehensive proteome coverage from FFPE material.

Notably, the use of short LC gradients combined with high-resolution mass spectrometry supports high-throughput analysis while maintaining sensitivity and reproducibility. This is particularly advantageous for studies involving large FFPE cohorts or comparative analyses between normal and tumor tissues, where throughput and consistency are essential.

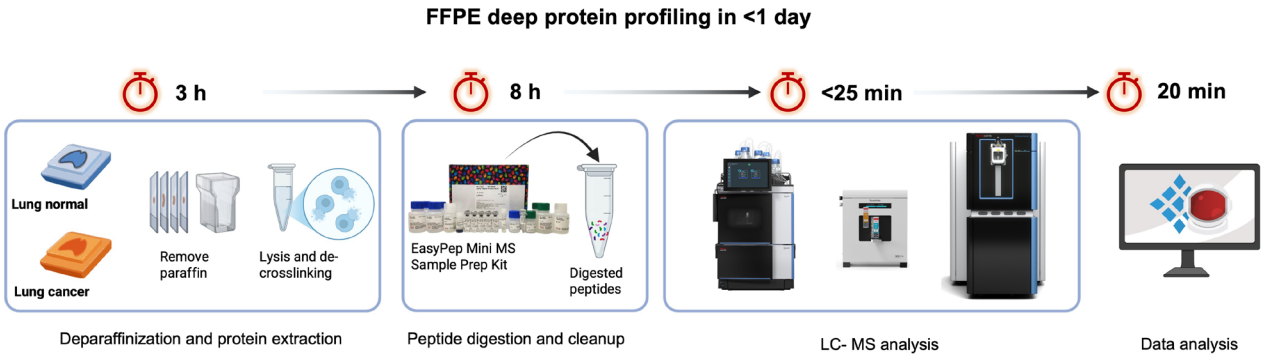

Figure 1. Workflow for FFPE sample preparation and mass spectrometry data acquisition to enable deep protein profiling.

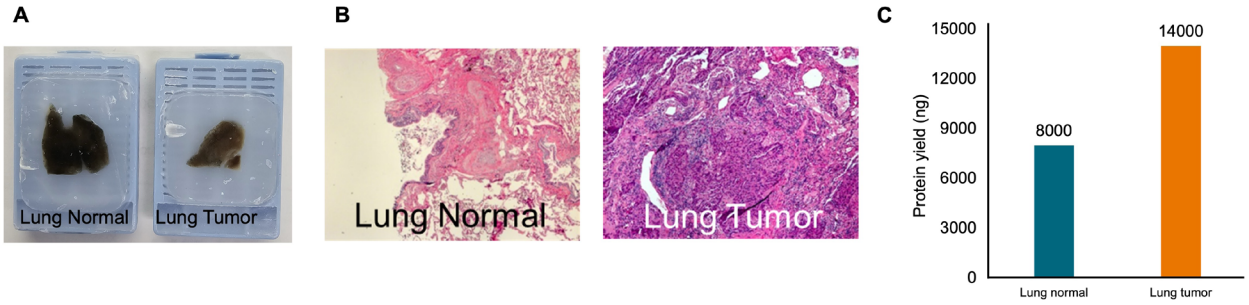

Figure 2. FFPE lung tissue and protein yield. (A) Lung normal and lung tumor tissue in FFPE blocks. (B) H&E staining of lung normal and lung tumor tissue to confirm normal and tumor status. (C) Protein recovery after deparaffinization and lysis.

Protein and peptide identifications across acquisition speeds and input amounts

Across normal and tumor FFPE lung tissues, high proteome coverage was maintained across all acquisition speeds, with only gradual reductions in protein group and peptide identifications as throughput increased (Figure 3). At the deepest protein profiling setting (60 SPD; ~20 min gradient), up to ~8,600 protein groups and ~105,000 peptides were identified from 200 ng input, representing the maximum proteome depth achieved in this study.

Increasing throughput from 60 to 180 SPD (~3× higher SPD; ~7 min gradient), at 200 ng load, ~6,100 protein groups and ~62,000 peptides were identified, corresponding to ~70% of protein groups and ~60% of peptides relative to 60 SPD. Notably, even at 20 ng input, ~5,400 protein groups and ~57,000 peptides were identified at 180 SPD, demonstrating robust performance at low sample amounts.

Further increasing throughput to 500 SPD (~8× higher SPD relative to 60 SPD; ~2 min gradient) still enabled identification of ~3,700–3,900 protein groups and ~25,000–27,000 peptides at 200 ng load. This corresponds to ~45% of protein groups and ~25% of peptides compared to the deepest acquisition, despite more than an 8-order-of-magnitude increase in sample throughput. Similar proportional trends were observed at 20 ng input, indicating that depth retention with increasing SPD is largely independent of sample load.

Importantly, normal and tumor FFPE lung tissues exhibited comparable scaling behavior across all SPDs and input amounts, indicating consistent proteome sampling and workflow robustness. Collectively, these results demonstrate that while deeper coverage is achieved at lower SPDs, high-throughput acquisition preserves a large fraction of proteome information. This favorable depth–throughput trade-off enables flexible experimental design, allowing users to prioritize either maximal proteome depth or rapid, large-scale FFPE cohort analysis without compromising data quality.

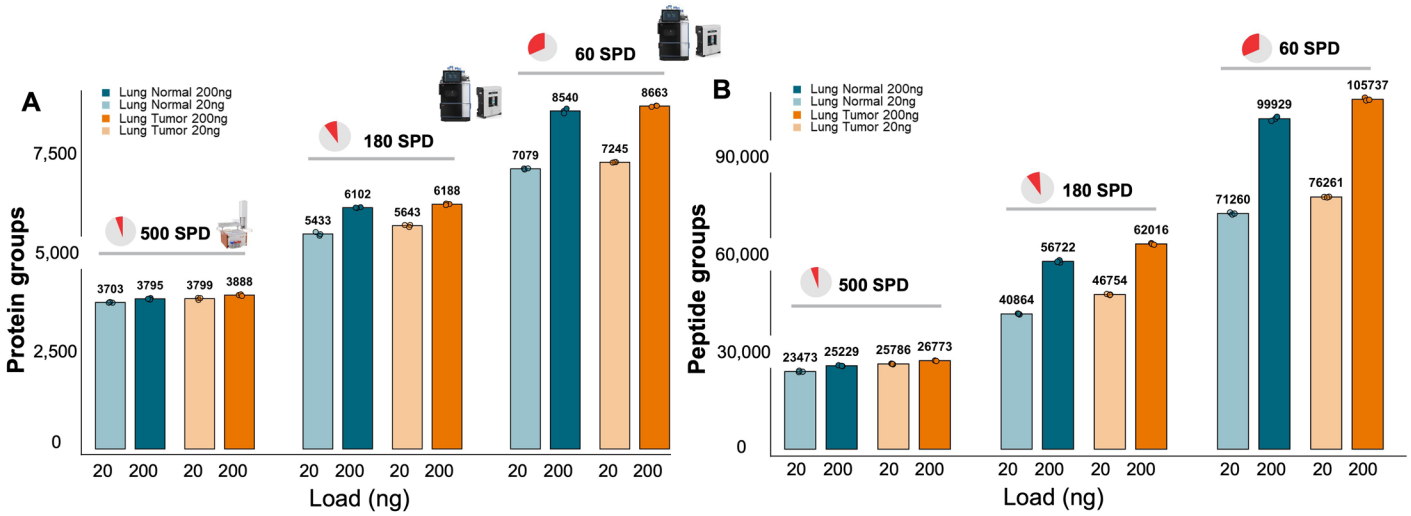

Figure 3. (A) Unique number of identified proteins and (B) peptides from three replicates of FFPE tissue across on-column load (20, 200ng) and gradient length (500, 180, and 60 SPD). Bar plots demonstrate unique identifications for each replicate in the given condition (color and shade identifiers in figure legend). 500 SPD is equivalent to a 2-min gradient, 180 SPD is equivalent to a 7-min gradient, and 60 SPD is equivalent to a 20-min gradient.

High quantitative reproducibility maintained across acquisition speeds

Quantitative reproducibility was consistently high across all acquisition speeds, sample types, and input amounts (Figure 4). Median protein abundance coefficients of variations (CVs) across triplicate injections remained low for all SPDs, ranging from ~7–10% at 500 SPD and 180 SPD to ~5–6% at 60 SPD (Figure 4A). Importantly, even at the highest throughput (500 SPD), CVs remained within 15% for quantitative proteomics, indicating that accelerated gradients do not substantially compromise measurement precision. Comparable CV distributions were observed for both normal and tumor FFPE lung tissues and across 20 ng and 200 ng loads, demonstrating robustness across biologically and analytically distinct conditions.

Overlap analysis further showed strong consistency in protein identifications across gradient lengths (Figure 4B). A large proportion of proteins identified at 500 SPD were also detected at 180 and 60 SPD, with the highest-throughput method retaining 45% of the core proteome observed at 60 SPD. This indicates that increased throughput primarily affects lower-abundance, gradient-dependent identifications while preserving reproducible detection of the main proteome.

Finally, fold-change measurements between lung tumor and normal tissues were highly concordant across SPDs (Figure 4C). Fold changes quantified at 180 SPD and 500 SPD showed strong correlation with those obtained at 60 SPD (Pearson r = 0.93 and 0.87, respectively), demonstrating that biological trends are consistently captured even under highly accelerated acquisition conditions using two different LC conditions. Further, it demonstrates that high-throughput acquisition preserves biologically relevant proteome observations.

Collectively, these results show that increasing throughput by up to ~8-fold has minimal impact on quantitative reproducibility and biological consistency. The workflow therefore supports reliable, high-throughput FFPE proteomics without sacrificing quantitative accuracy, enabling scalable analysis of large clinical cohorts.

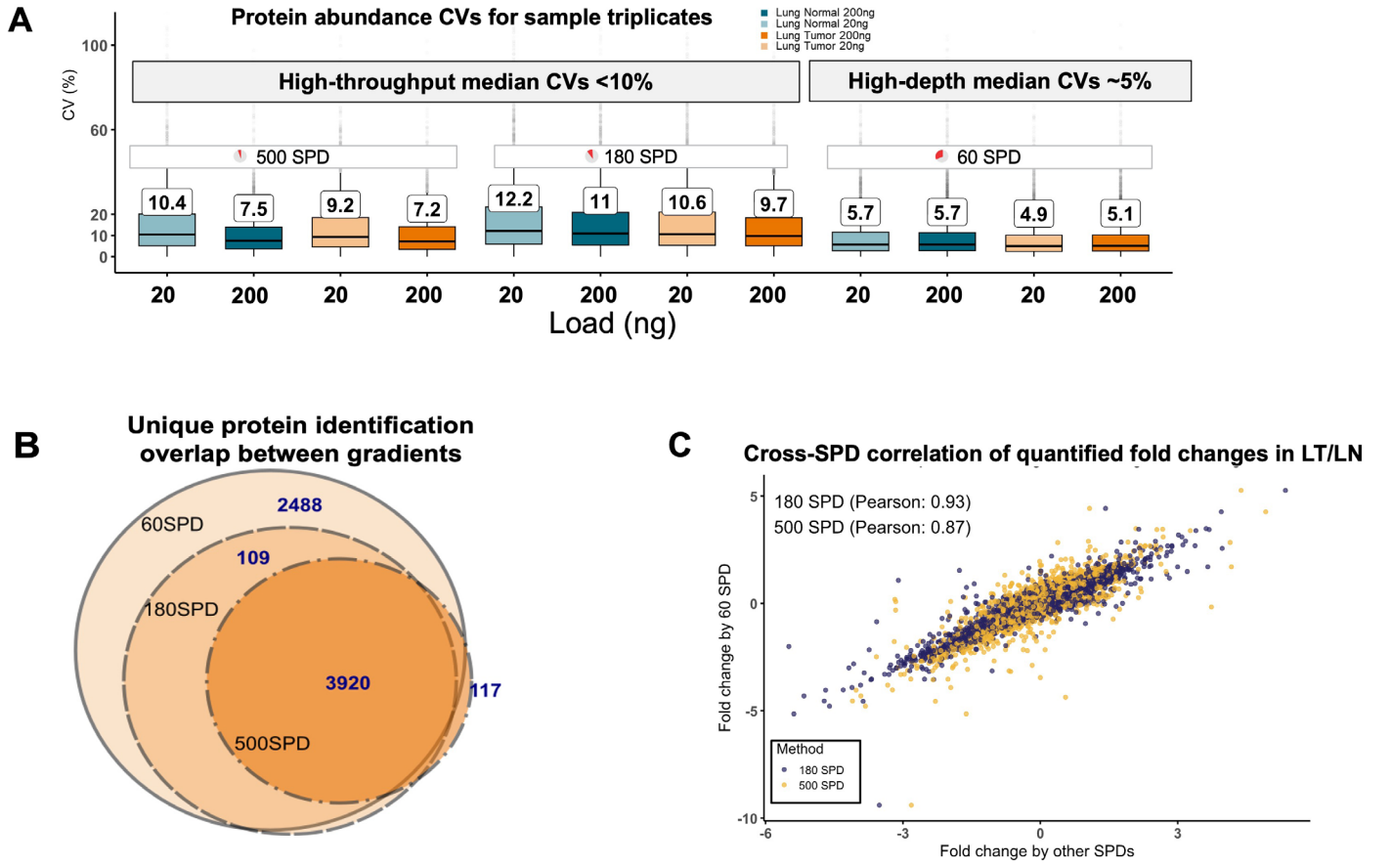

Figure 4. Reproducibility assessment across sample type, on-column amount, and gradient lengths. (A) CVs of peptide abundances for triplicate sample replicates across all FFPE on-column loads and gradients. Median CV marked with black line and labeled in the figure. (B) Unique protein identifications from 200 ng sample load, overlap analysis between gradient lengths. Circles labeled according to gradient length. (C) Fold change calculations between lung normal and lung tissue samples, then correlation between fold changes were plotted between 180 SPD vs. 60 SPD (blue, r = 0.93) and 500 SPD vs. 60 SPD (yellow, r = 0.87).

Differential protein expression and pathway alterations between lung tumor and normal FFPE tissues

Differential expression analysis of 200 ng of sample loads at 60 SPD revealed widespread proteomic alterations between lung tumor and matched normal FFPE tissues (Figure 5 left). About 1,500 protein groups were significantly differentially expressed ($|\log_2 FC| \geq 1$, $p \leq 0.05$), indicating robust detection of tumor-associated proteomic remodeling. The clear separation between significantly regulated and non-significant proteins highlights the sensitivity of the workflow for identifying biologically meaningful differences from FFPE sections.

Ranking proteins by median intensity showed that differentially upregulated proteins spanned a wide dynamic range and were not limited to only the most abundant species (Figure 5 center). While many high-intensity proteins exhibited tumor-associated upregulation, a substantial number of regulated proteins were detected at mid- to low-abundance levels, demonstrating that the workflow captures both dominant and subtler proteomic changes. This depth is essential for uncovering signaling and metabolic pathways that may be masked in more shallow analyses.

Pathway enrichment analysis of upregulated proteins identified key biological processes associated with tumor progression and the tumor microenvironment (Figure 5 right). Enriched pathways include neutrophil extracellular trap formation, transcriptional dysregulation in cancer, glycolysis/gluconeogenesis, HIF-1 signaling, ECM–receptor interaction, and central carbon metabolism in cancer. The prominence of metabolic reprogramming, hypoxia signaling, and extracellular matrix remodeling is consistent with known hallmarks of lung cancer biology, 3,4 while enrichment of immune- and phagosome-related pathways suggests contributions from tumor–immune interactions captured within FFPE tissue sections.

Together, these results demonstrate that the rapid FFPE proteomics workflow not only provides deep and reproducible proteome coverage but also enables robust detection of biologically relevant differential expression and pathway-level alterations. This supports its utility for translational studies aimed at interrogating disease mechanisms directly from archived clinical specimens.

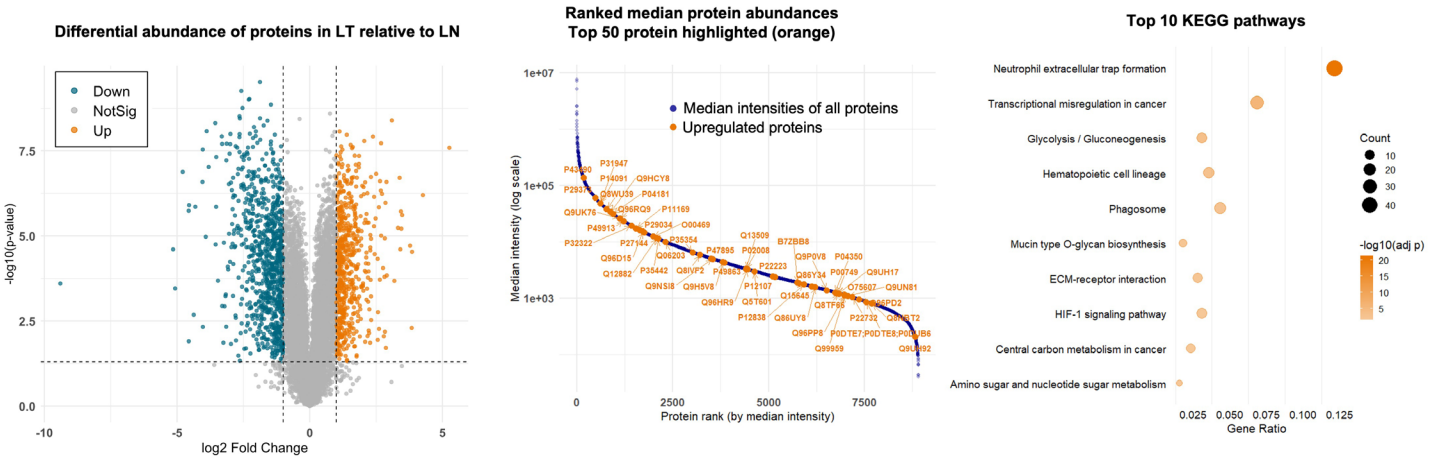

Figure 5. Biological insights for FFPE lung tissue, 200 ng, 60 SPD gradient. (A) Volcano plot: Proteins significantly enriched in lung tumor compared to lung normals samples highlighted in orange, and proteins significantly enriched in lung normal compared to lung tumor highlighted in blue. (B) Rank order plot of median intensities for all proteins, with proteins significantly enriched in lung tumor highlighted in orange. (C) Top 10 KEGG pathways identified from proteins found to be significantly enriched in lung tumor samples. Bubble size indicates numbers of proteins found in the pathway, shade indicates significance.

Summary

- End-to-end FFPE proteomics is completed in under 1 day, integrating rapid sample preparation, fast LC–MS, and streamlined data analysis
- Deep proteome coverage is achieved from FFPE lung tissue, with up to ~8,600 protein groups and >100,000 peptides identified from 200 ng input in 20 min
- Increasing throughput by up to ~8× (60 to 500 SPD) with moderate loss in depth, retaining ~70% of proteins at 180 SPD and ~45% at 500 SPD
- Robust quantitative performance is maintained at low sample input (20 ng) and across acquisition speeds, with median CVs ≤10%
- The workflow enables reliable differential expression and pathway analysis at scale, demonstrating its utility for biological insight into lung tumor and normal tissues

References

1. Haines, M. et al. High-Throughput Proteomic and Phosphoproteomic Analysis of Formalin-Fixed Paraffin-Embedded Tissues. *Molecular & Cellular Proteomics* 24, (2025).
2. Zhu, Y. et al. High-throughput proteomic analysis of FFPE tissue samples facilitates tumor stratification. Mol Oncol 13, 2305–2328 (2019).
3. Wang, Y. an et al. Effects of tumor metabolic microenvironment on regulatory T cells. Molecular Cancer vol. 17 Preprint at https://doi.org/10.1186/s12943-018-0913-y (2018).
4. Liu, Y., Vandekeere, A., Xu, M., Fendt, S. M. & Altea-Manzano, P. Metabolite-derived protein modifications modulating oncogenic signaling. *Frontiers in Oncology* vol. 12 Preprint at https://doi.org/10.3389/fonc.2022.988626 (2022).

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

TN 004381 0226