JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/jacsau
Article

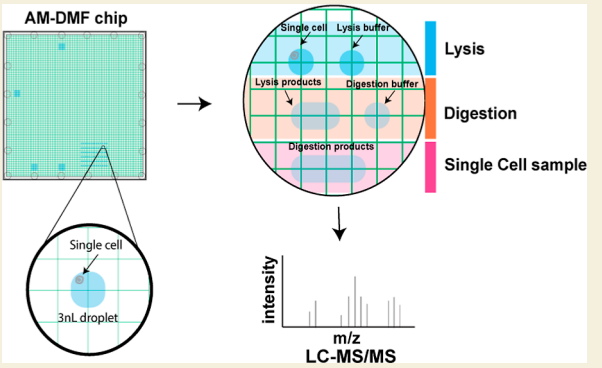
AM-DMF-SCP: Integrated Single-Cell Proteomics Analysis on an Active Matrix Digital Microfluidic Chip

Zhicheng Yang,[▽] Kai Jin,[▽] Yimin Chen,[▽] Qian Liu, Hongxu Chen, Siyi Hu, Yuqiu Wang, Zilu Pan, Fang Feng, Mude Shi, Hua Xie,* Hanbin Ma,* and Hu Zhou*

Cite This: *JACS Au* 2024, 4, 1811−1823
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Single-cell proteomics offers unparalleled insights into cellular diversity and molecular mechanisms, enabling a deeper understanding of complex biological processes at the individual cell level. Here, we develop an integrated sample processing on an active-matrix digital microfluidic chip for single-cell proteomics (AM-DMF-SCP). Employing the AM-DMF-SCP approach and data-independent acquisition (DIA), we identify an average of 2258 protein groups in single HeLa cells within 15 min of the liquid chromatography gradient. We performed comparative analyses of three tumor cell lines: HeLa, A549, and HepG2, and machine learning was utilized to identify the unique features of these cell lines. Applying the AM-DMF-SCP to characterize the proteomes of a third-generation EGFR inhibitor, ASK120067-resistant cells (67R) and their parental NCI-H1975 cells, we observed a potential correlation between elevated VIM expression and 67R resistance, which is consistent with the findings from bulk sample analyses. These results suggest that AM-DMF-SCP is an automated, robust, and sensitive platform for single-cell proteomics and demonstrate the potential for providing valuable insights into cellular mechanisms.



KEYWORDS: *single-cell proteomics, active matrix digital microfluidics (AM-DMF), integrated microfluidic platform, mass spectrometry, cellular heterogeneity*

INTRODUCTION

Single-cell proteomics provides us with a more in-depth knowledge map in research areas such as disease,[1] development,[2] and drug treatment,[3] helping us to comprehensively understand the heterogeneity of complex life systems. While single-cell proteomics has been proposed for almost 3 decades, it has not been as widely used as single-cell transcriptional profiling to tackle biological challenges.[4] Recent improvements in mass spectrometer,[5−8] which can now detect peptides at nanogram or even hundreds of picogram levels, have truly brought single-cell proteomics to the forefront. For single-cell proteomics, there are mainly two methods of acquisition: chemical labeling and label-free quantitation. Within the chemical labeling category, innovative approaches include the SCoPE[9]/SCoPE2[10] method, which utilizes Tandem Mass Tags (TMT) and carrier channels, and plexDIA,[11] a technique that combines mTRAQ labeling with data-independent acquisition (DIA). These methods have quantified over 1500 protein groups (PGs) in individual cells, and nearly 3000 PGs can be identified in all cells in an experiment.[8−13]

The main process of single-cell proteomics, comprising single-cell sorting/isolation, proteomic sample preparation (including lysis, reduction, alkylation, and digestion), and mass spectrometry data acquisition, faces various challenges. While single-cell sorting has advanced through techniques like flow cytometry, single-cell printing, and laser capture microdissection (LCM), these methods generally lack direct, independent control over individual cells. Additionally, preparing proteomic samples from single cells is inherently difficult. Key to improving this process is handling samples in small volumes, which effectively reduces analyte dilution and minimizes interaction with adsorptive surfaces, thus benefiting single-cell proteomics.[14] Preventing sample loss during preparation is critical, prompting the development of ultralow volume sample processing techniques, such as OAD,[15] nanoPOTS,[16,17] iPAD-1,[18] OCAM,[19] and certain multiwell plate methods,[8,20−22] which have shown promising results.

For many years, digital microfluidics (DMF) technology has been considered a powerful tool for proteomics sample processing due to its automation capabilities for handling microvolume samples.[13,23,24] The small reaction volumes

Received: January 8, 2024
Revised: March 8, 2024
Accepted: March 8, 2024
Published: March 26, 2024

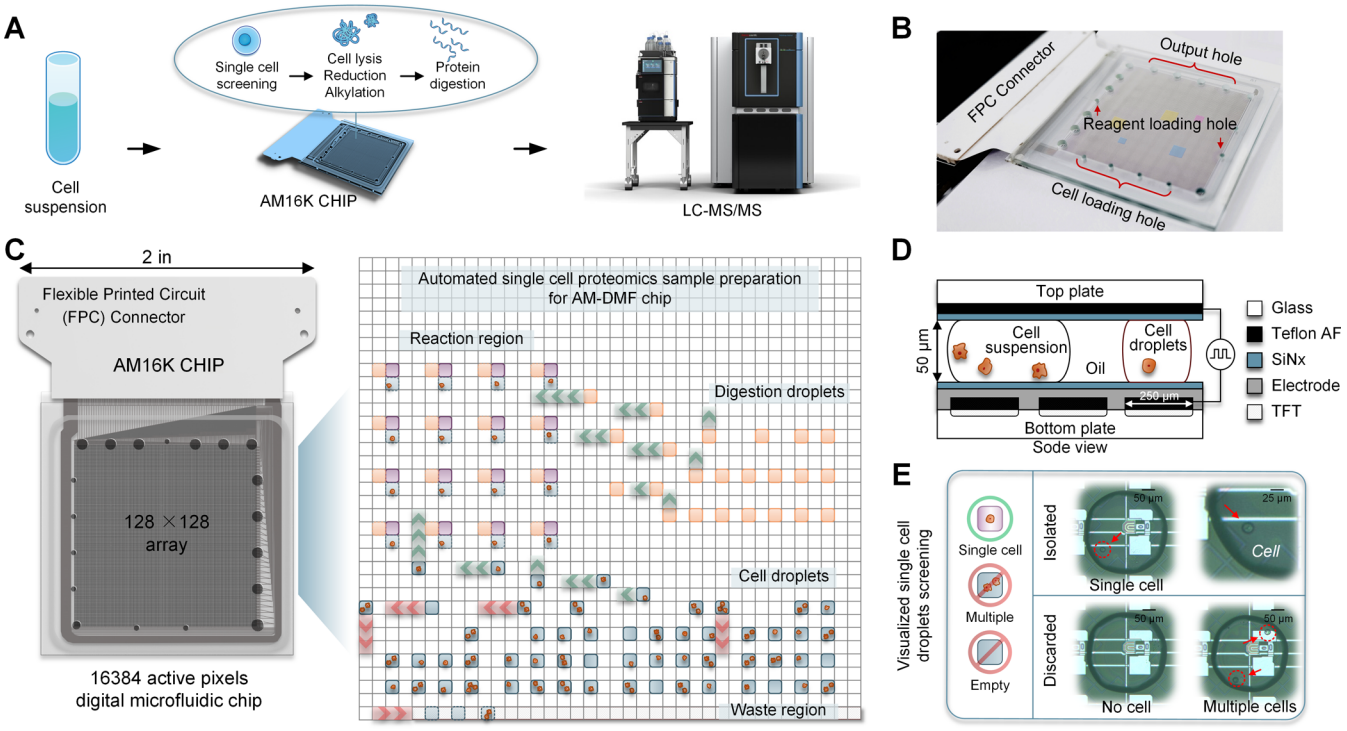


Figure 1. Single-cell proteomics sample preparation using active-matrix digital microfluidics (AM-DMF-SCP). (A) Workflow schematic for mass spectrometry-based single-cell proteomics analysis utilizing AM-DMF chip. (B) Photographic image of AM-DMF chip. (C) Top-view diagram of AM-DMF chip, illustrating the layout and workflow for single-cell sample preparation. The chip is equipped with 16384 active pixel electrodes for liquid samples manipulation. (D) Cross-section-view schematic of the chip showing the individual pixel electrodes, each 250 μm in diameter, which demonstrates the chip's capability to process minimal droplet volumes of about 3.125 nL within a 50 μm chip gap. (E) Visual representation of single-cell droplet selection. Displays the selective process for cell droplets, ensuring only one cell per droplet in three distinct scenarios.

characteristic of DMF systems offer a distinct advantage in reducing protein loss. By integrating laser cell lysis (LCL) into the Single Cells for -Omics (DISCO) platform, DMF has achieved multi-omics analysis including single-cell genomics, transcriptomics, and proteomics.[24] However, the throughput of single-cell analysis has been limited by the size and number of electrodes on traditional DMF chips, confining them to the separation and analysis of only a very small number of single cells and unable to meet the needs for large-scale single-cell sorting and SCP analysis.[24,25] Emerging active matrix digital microfluidic technology overcomes the limitations of electrode size and number on conventional DMF chips, featuring the capability to generate and manipulate picoliter-sized microdroplets. By employing an active matrix backplane to replace traditional electrodes on digital microfluidic (DMF) chips, this approach overcomes the limitations associated with the number of electrodes. It enables the creation of a large-scale electrode array that can be independently addressed.[26,27]

Here, we developed an active-matrix digital microfluidic (AM-DMF) platform providing integrated and easily used in-chip procedures with parallel sample preparations for single-cell proteomics, including single-cell isolation, cell lysis, and enzymatic digestion. First, we conducted single-cell isolation and sample preparation on the DMF chip, realizing a truly all-in-one sample preparation workflow. Second, sample reaction volumes were as low as 10 nL to minimize sample loss and reduce reagent consumption. In addition, all single cells within the DMF chip could be mixed and reacted with reagents simultaneously. This means that parallel sample processing effectively avoids individual differences caused by serial sample processing. We highlighted the distinctions between the AM-DMF-SCP platform and existing SCP solutions in a comparative table (Table S1).[20,28−32] The AM-DMF-SCP platform demonstrates clear advantages in minimizing reaction volumes, cutting costs, and enhancing the automation efficiency. These advancements significantly enhance single-cell proteomics performance. In our efforts to optimize cell lysis within the AM-DMF-SCP workflow, we evaluated five distinct single-pot methods and found that the *n*-dodecyl β-D-maltoside (DDM) method was chosen due to its proven effectiveness. Our comparisons revealed that AM-DMF-SCP outperforms conventional manual sample preparation by identifying more proteins. Utilizing this platform, we prepared samples from various cell types and successfully differentiated them. The analysis of NCI-H1975 cells and their drug-resistant variant 67R confirmed the consistency of AM-DMF-SCP with bulk proteomic results, demonstrating its reliability and efficacy in single-cell proteomics.

RESULTS AND DISCUSSION

Single-Cell Proteomics with the Active-Matrix Digital Microfluidics Chip

Our study addressed the challenges associated with sample preparation for mass-spectrometry-based single-cell proteomics (MS-based SCP) by introducing the AM-DMF-SCP platform. This platform provides researchers with a reliable and easy-to-use tool, simplifying sample preparation and making single-cell proteomics (SCP) more accessible. We developed an integrated, in-chip, parallel sample preparation method specifically designed for MS-based SCP (Figure 1A). This method enables the automation of the sample preparation

https://doi.org/10.1021/jacsau.4c00027
JACS Au 2024, 4, 1811−1823

within a digital microfluidic chip, from cell suspension to mass spectrometry analysis, including steps like single-cell sorting, lysis, and enzymatic digestion, effectively reducing protein loss due to sample transfer and improving the automation of single cell sample preprocessing.

The SCP preprocessing workflow relies on a large-scale automated fluid handling system using an AM-DMF instrument (Figure 1B−D). The top plate of the chip features microwells for reagent and cell loading and sample output. The bottom plate contains an array of 16384 active pixel electrodes, giving the chip its name, AM16K. Arranged in a grid of 128 rows and 128 columns, each measuring 250 μm in length, the chip can theoretically handle droplet volumes as small as 3.125 nL. The chip design ensures minimal droplet volume for single-cell samples, less than 10 nL, significantly reducing protein loss by decreasing the contact area. Additionally, we can adjust the gap between the top and bottom plates of the chip to select cells of different sizes and to adjust the reagent reaction system. The internal environment of the chip is filled with medium oil hexadecane, ensuring that all sample processing occurs within the oil phase. This oil phase environment effectively reduces the evaporation and loss of microvolume samples and proteins while minimizing cross-contamination risk. To assess residual contamination as different samples, pass through the same electrodes, we utilized a simplified model using 200 pg peptide (GDFQFNISR, derived from β-galactosidase of *E. coli*). We measured the intensity of this peptide in samples traversing the same path and compared it with the intensity of the peptide dissolved in 0.1% formic acid (Solution A). No detectable intensity of the peptide segment was observed in Solution A, indicating the absence of potential contamination risks in single-cell experiments.

This work's ability to automate sample preprocessing for all-in-one SCP is an important attribute. The chip's electrode array is strategically divided into zones for single-cell droplet generation and screening, mixing zones for lysis, reduction, and alkylation reagents, and a zone for enzymatic digestion droplet generation (Figure 1C). The chip can automatically complete the preprocessing by loading the sample onto the chip. For easy sample loading, the droplet generation area is designed around the edge of the chip. The chip's sample loading volume is about 200 nL, offering a significant advantage in handling trace and scarce samples. Since single-cell and enzyme digestion droplets must be moved to the reaction area, the droplet spacing is designed to be larger to aid droplet transport and mixing. We introduced optical microscopic imaging and artificial intelligence vision recognition to select droplets and obtain uniform sample volumes. By fully visualized droplet screening, we ensured the authenticity and effectiveness of single-cell droplets (Figure S1 and Movie S1). Notably, the AM-DMF technology enables in-chip visualization and the selection of single cells (Figure 1E). By injecting a specific concentration of cell suspension into the chip, resulting in the dispersion into numerous tiny droplets, there is a probabilistic chance of obtaining droplets containing exactly one cell. The generation of single-cell droplets follows the principles of Poisson distribution, with the probability of getting a droplet containing exactly one cell being around 30%.[33] Therefore, in theory, a single chip can mix and react with up to 64 single cells. However, in practice, due to uncontrollable factors such as initial cell concentration and cell viability, some droplet loss occurs during secondary cell selection. Typically, a single chip can yield no fewer than 24 single-cell samples. Through microscopic imaging, we can identify three scenarios: single-cell droplets, multicell droplets, and blank droplets, enabling the selection of single-cell droplets (Figure 1E). Besides, we use the capillaries to obtain the single-cell samples after enzyme digestion in the chip. Through a capillary sampling test (details in Material and Methods), we concluded that the success rate of extracting sample through capillary is around 95% (Table S3), indicating this manipulation is suitable for this workflow.

The novel AM-DMF-SCP workflow proposed in this study significantly improves sample preparation automation, reduces the reaction volume of samples, and achieves parallel sample processing. This provides an effective solution for SCP sample preparation and is expected to advance the progress of single-cell proteomics research.

Compare Single-Pot Sample Preparation Method

To optimize the sample preparation for AM-DMF-SCP, a comparative analysis was performed on five distinct single-pot methods was studied including three surfactant-based (DDM, Rapigest, SDC) and two organic solvents-based (TFE, ACN) approaches. Using principal component analysis (PCA), significant differences were observed among these methods (Figure S3A). Notably, within the surfactant category, DDM and Rapigest showed greater similarity compared to SDC, which yielded fewer proteins and peptides. This discrepancy could be attributed to post-SDC removal via standard acid precipitation.[34] Our evaluation was thorough, considering the total number of protein groups (PGs) and peptides identified by LC−MS/MS (Figure S3B,C). The results highlighted the superiority of the DDM-based method, particularly in terms of the sum of the peptide intensities, which we used as a proxy for the total intensity (Figure S3D). Furthermore, DDM stood out in terms of digestion efficiency, as assessed by misscleavage rates (Figure S3E). The peptides identified through the DDM method were characterized by increased hydrophobicity, as determined by their grand average of hydropathy (GRAVY) scores (Figure S3F), aligning with findings from previous research.[35,36] Conclusively, the DDM lysis buffer demonstrated superior performance across various metrics, including the number of PGs and peptides identified, overall intensity, efficiency of digestion, and ability to recover hydrophobic peptides.

Our comprehensive comparative analysis to optimize sample preparation for AM-DMF-SCP underscores the effectiveness of various single-pot methods, including both surfactant-based and organic solvent-based approaches. The significant variations among these methods highlight the impact of the chosen sample preparation technique on the final proteomic outcome. The results notably indicate that the superiority of the DDM-based method, especially in digestion efficiency and recovery of hydrophobicity peptides, positioned it as a particularly effective approach for AM-DMF-SCP sample preparation.

Comparison of the Proteome Results Obtained from In-Chip and Out-Chip Sample Processing

To evaluate the effectiveness of the in-chip workflow, a batch of samples was also prepared by using the out-chip method as a comparison. In the out-chip approach, single cells are isolated from the chip and transferred into low-absorption vials for subsequent sample preparation. The average number of identified PGs was significantly higher in the in-chip method

https://doi.org/10.1021/jacsau.4c00027
JACS Au 2024, 4, 1811−1823

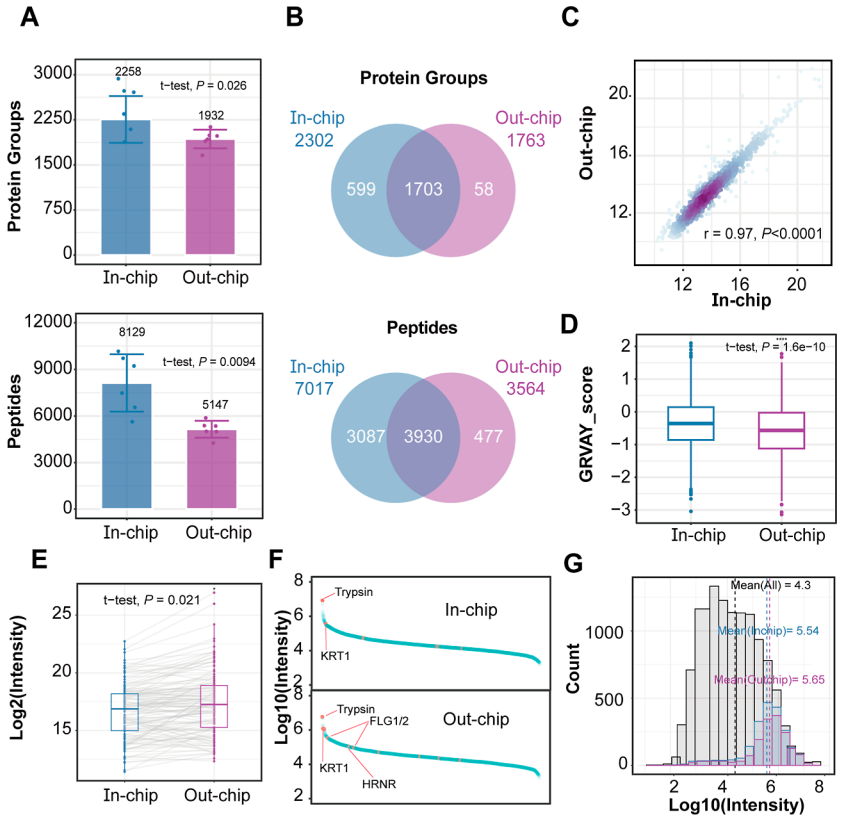


Figure 2. Comparison of in-chip and out-chip sample processing methods. (A) The bars and error bars show the resulting PGs or peptides and their standard deviations (SD). To estimate significance, an unpaired two-tailed Student's *t*-test was performed. The resulting *p* values are shown above, $n = 6$ for each condition. (B) Venn diagram showing the quantifiable PGs and peptides. (C) Pearson correlations of PGs intensity between two conditions. (D) Peptide hydrophobicity of differentially expressed and group-specific peptides was evaluated using the GRAVY score. (E) The log2-transformed intensities of peptides corresponding to commonly identified contaminant proteins were analyzed, and their significance was determined using an unpaired two-tailed Student's *t*-test. The resulting *p* values are shown above. (F) Rank plot comparing two conditions. Contaminant proteins were marked in red. (G) The copy number distribution of two conditions overlays the data using 20 bins with purple for out-chip and blue for in-chip.

($P < 0.05$; *t*-test), with 2288 PGs and 8129 peptides identified on average per sample) compared to the out-chip method (1932 PGs and 5147 peptides) (Figure 2A). Figure 2B shows that the inchip method can identify 541 extra PGs and 2610 extra peptides. After features with less than 50% missing values in each group were filtered out and common ones were retained, uniform manifold approximation and projection (UMAP) analysis was used to differentiate these two methods. Moreover, a correlation analysis was conducted on the intensity of the proteins jointly identified by both groups revealing a significant positive correlation ($r = 0.97$; $P < 0.0001$; Figure 2C). Peptides, which are differentially expressed and group-specific and identified by the inchip method, have higher hydrophobicity, as measured by their GRAVY scores (Figure 2D). To evaluate the presence of contaminating proteins in the experiment, we analyzed peptides derived from proteins listed in the contamination library from MaxQuant.[37] The results showed that the intensity of the contaminating peptides in the out-chip method was significantly higher than that in the in-chip methods ($P < 0.05$; *t*-test; Figures 2E and S4), and the number and rank of contaminating proteins were also higher in the out-chip (Figure 2F). Moreover, we compared the copy number distribution[38] of proteins identified by the two methods and found that the in-chip method exhibited a lower average copy number (5.44 versus 5.65), indicating that the in-chip method is more sensitive (Figure 2G). This heightened sensitivity may be attributed to the in-chip method's smaller digestion environment, resulting in fewer exogenous contaminants and lower absorption loss.

The comparison between in-chip and out-of-chip methods in our study offers significant insights into the effectiveness of these workflows in single-cell proteomics. The in-chip method demonstrated a distinct advantage, as evidenced by the significantly higher number of identified PGs and peptides, increased sensitivity, and reduced contamination. This suggests that the in-chip workflow facilitates more comprehensive proteome coverage, which is crucial for in-depth cellular analysis.

Reproducibility of Single-Cell Proteomic Sample Preparation with AM-DMF-SCP

We prepared single-cell proteomic samples from four distinct cell types by using three chips. Two chips processed 12 single cells each, while the third chip handled six single cells from each cell type (Table S4). This distribution ensured that the processing of each cell type was spread over two different chips, which is essential for assessing potential batch effects among the cell samples. To avoid possible background interference in MS analysis, a blank sample was analyzed before the samples to determine the presence of any pre-existing contamination in the system. Upon conducting a database search of this blank sample, we confirmed no identifiable precursors or PGs, affirming the absence of

https://doi.org/10.1021/jacsau.4c00027
JACS Au 2024, 4, 1811−1823

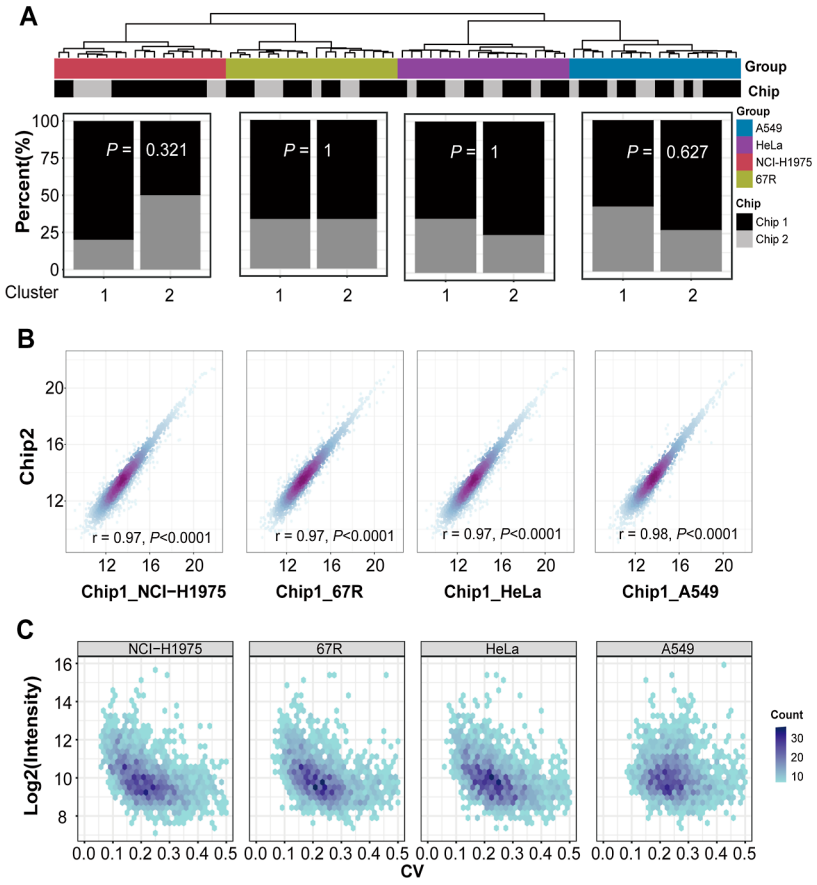


Figure 3. Reproducibility analysis of single-cell proteomic sample preparation with AM-DMF chip. (A) Top panel: cluster analysis plot. Bottom panel: stacked bar plot, where the significance of the *p*-values is annotated using Fisher's exact test. (B) Pearson correlations of PGs intensity between two chips for all cell lines. (C) Hexbin plots showing the log2-transformed intensity and CV distribution for quantified PGs.

contamination. Quality control (QC) samples, which is 200 pg Hela standard dilution, were integral during the sample runs to monitor the status of the instrumentation. The metadata of all samples are shown in Table S5 according to the recommendations of the single-cell white paper.[39]

The QC samples exhibited impressive consistency, with a correlation exceeding 0.97 (Figure S5A) and a median coefficient of variation (CV) below 20% (Figure S5B). Each sample was followed by a blank run to flush the whole system to eliminate residual peptides, resulting in a markedly lower intensity than that of the QC samples (Figure S5C). The blank sample yielded no identifiable results, confirming the absence of residual peptides and indicating that the single-cell samples were thoroughly analyzed within a 15 min gradient run. Cluster analysis of the collected data showed clear segregation with each cell type forming distinct clusters (Figure 3A), thereby confirming the efficacy of the AM-DMF-SCP workflow. To further investigate potential batch effects, we divided each cell line into two groups based on the chip used for sample preparation. We applied Fisher's exact test to evaluate the differences between chips and found that all *p*-values for the four cell lines exceeded 0.05 (Figure 3A). Moreover, UMAP analysis revealed that samples prepared on different chips were indistinguishable (Figure S6A). The proteomic intensity profiles derived from samples processed on different chips exhibited a significant positive correlation (Figure 3B). These results suggest the absence of considerable batch effects in our workflow. Furthermore, we investigated the variation across multiple cell lines by examining the CVs for each protein within each cell. This CV calculation was performed only when data for more than half the observations per protein in each cell were available. Notably, the median CV value was less than 0.3, with a trend showing that proteins of higher abundance had lower CVs (Figure 3C), consistent with the literature[40] that suggests lower abundance peptides are noisier.

The reproducibility of our single-cell proteomic sample preparation using AM-DMF-SCP is a crucial aspect of our study. By preparing samples from different cell types across different chips, we have demonstrated that AM-DMF-SCP can handle varied sample types effectively. Its capability to accurately process and analyze single-cell proteomes with minimal batch effects makes it a promising tool for advanced proteomic research.

Comparison of Three Human Cell Lines Based on AM-DMF-SCP

In our study, HeLa, A549, and HepG2 cells were utilized as models to assess AM-DMF-SCP's ability to detect cellular heterogeneity. On average, approximately 2300 proteins were identified in HeLa cells, with A549 and HepG2 cells yielding around 2000 proteins each (Figure 4A). This discrepancy in protein counts may be attributed to the slightly larger size of HeLa cells compared to the other two cell lines.[41,42] The UMAP analysis clearly delineates the three cell types into clusters (Figure 4B). A correlation heatmap further supports this segregation by demonstrating lower correlations between different cell lines compared to within the same cell line

https://doi.org/10.1021/jacsau.4c00027
JACS Au 2024, 4, 1811−1823

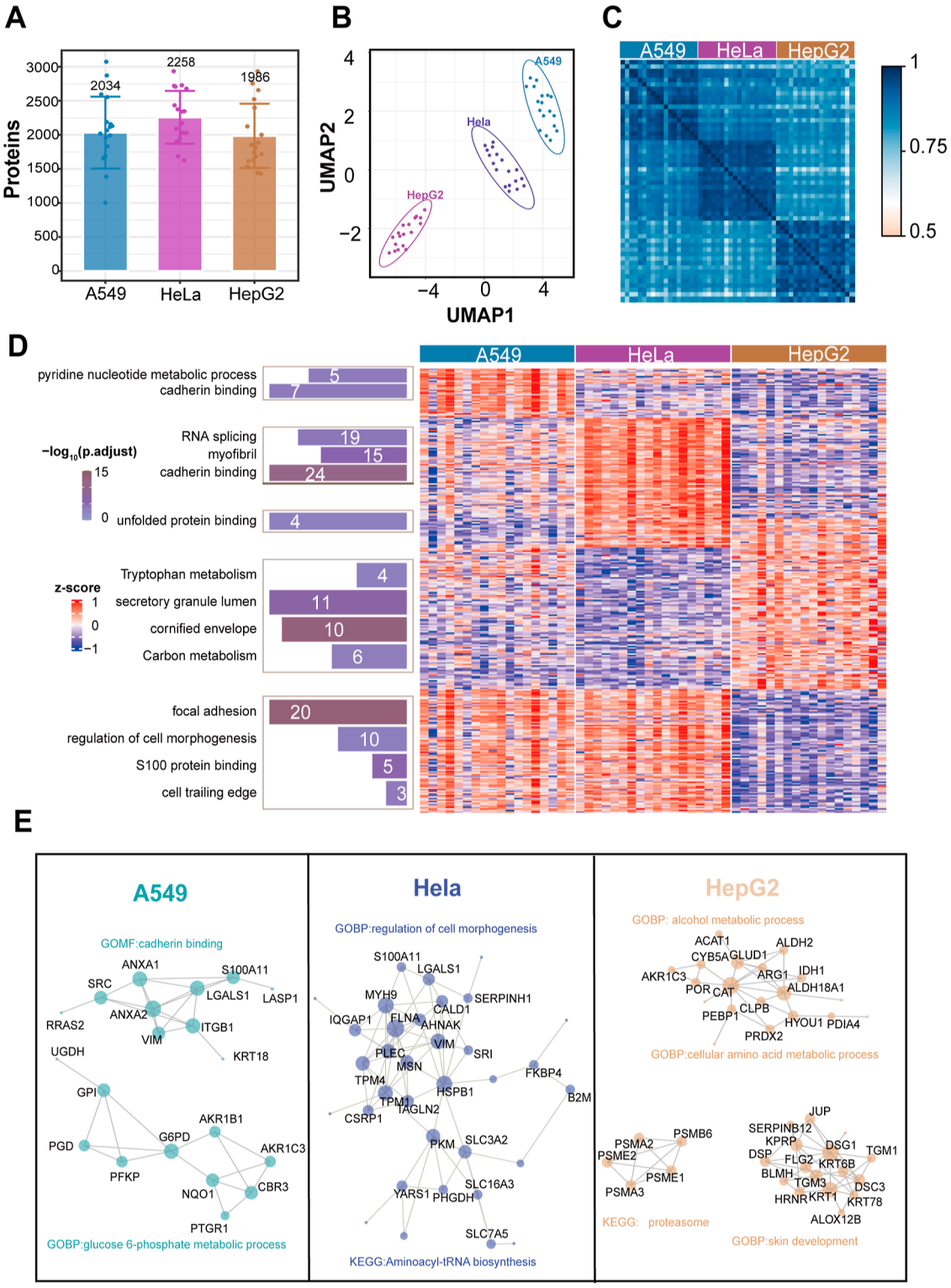


Figure 4. Comparison of single A549, HeLa, HepG2 cells using AM-DMF-SCP. (A) The bars and error bars showing the resulting PGs or peptides and their SD. N = 18 for each condition. (B) UMAP analysis plot of the samples. (C) Correlations of PGs intensity between cells. (D) Heatmap showing the differentially expressed PGs. The bars on the left represent Enriched GO terms. (E) STRING network of the genes upregulated in A549, HeLa, and HepG2, respectively.

(Figure 4C). These results confirm the capability of our workflow to distinguish these cell lines effectively. To identify the proteins contributing to the clustering of the three cell types, we conducted a comparison between one cell line and the other two using stringent criteria [(False discovery rate) FDR < 0.01, |Fold change (FC)| > 2]. Among the proteomes identified by filtering features with less than 50% missing values in each group and retaining those common across groups, 298 out of 1533 PGs showed significant differential abundance across the three cell types. Additionally, gene

https://doi.org/10.1021/jacsau.4c00027
JACS Au 2024, 4, 1811−1823

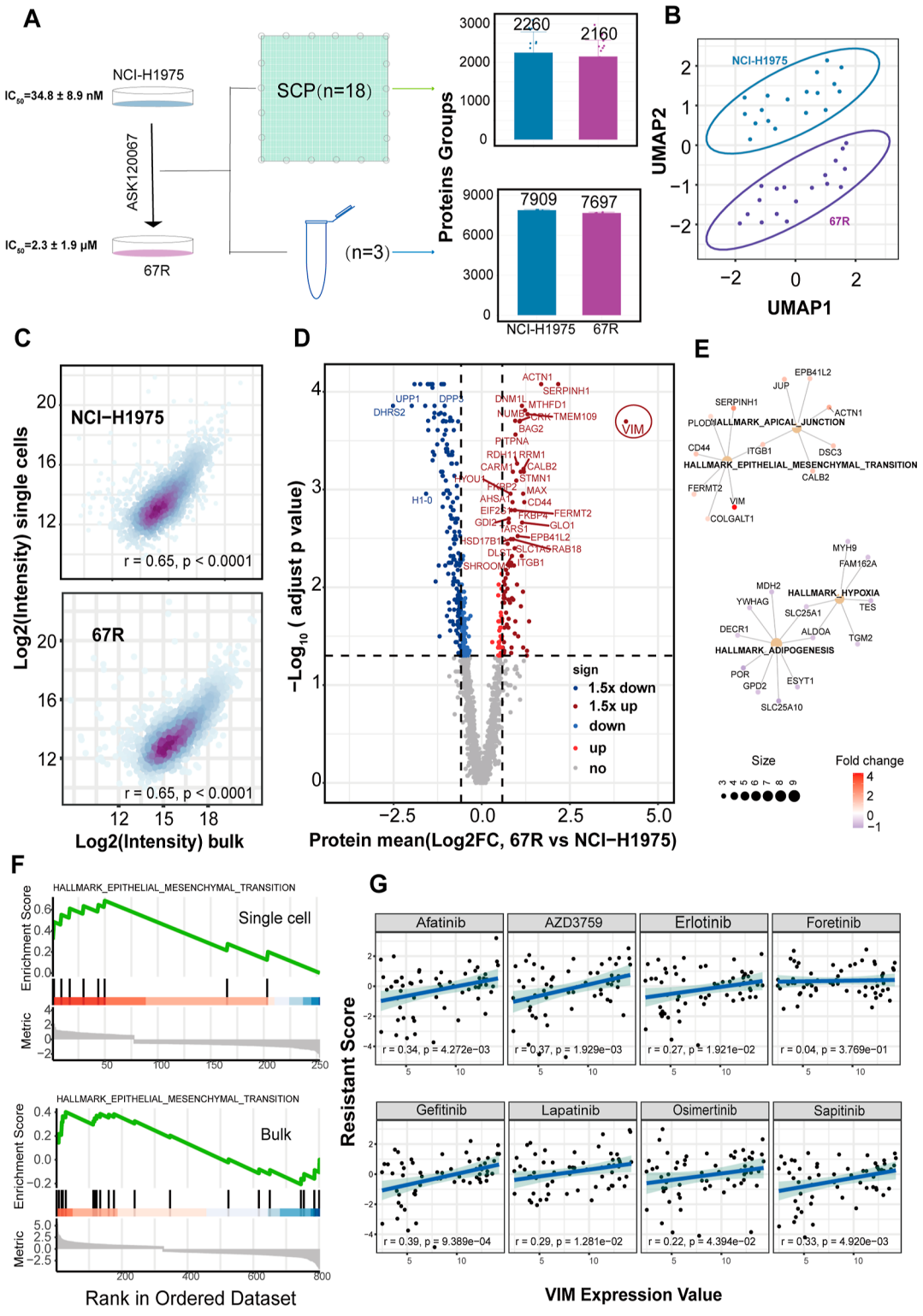


Figure 5. Comparison of 67R and NCI-H1975 using AM-DMF-SCP and bulk. (A) Experimental workflow and PGs identified. For bulk experiments, we conducted 3 biological replicates. In the case of single-cell experiments, 18 individual cells were analyzed. (B) UMAP plot of the samples. (C) Pearson correlations of PGs Intensity for all cell lines. (D) Volcano plot depicting differentially expressed PGs (Wilcoxon signed-rank test, BH adjusted). (E) Network plot of GESA results. (F) GSEA plot for EMT pathway. (G) Pearson correlations between VIM expression value and drug-resistant score.

ontology (GO) enrichment analysis[43] was performed to identify biological processes associated with proteins that were uniquely up- or down-regulated in each cell line. In A549 cells, there was an upregulation in cadherin binding and the pyridine nucleotide metabolic process and a downregulation in unfolded protein binding. HeLa cells showed upregulation in RNA splicing, myofibril, and cadherin binding, whereas HepG2 cells exhibited upregulation in carbon metabolism,

https://doi.org/10.1021/jacsau.4c00027
JACS Au 2024, 4, 1811−1823

tryptophan metabolism, secretory granule lumen, and the cornified envelope (Figure 4D). STRING-based protein−protein interaction (PPI) analysis[44] of cell-specific upregulated proteins revealed major PPI modules and were further annotated by GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis[45] (Figure 4E). Specifically, proteins in A549 cells were enriched in cadherin binding and glucose 6-phosphate metabolic processes. The major PPI modules in HeLa cells were associated with the regulation of cell morphogenesis and aminoacyl-tRNA biosynthesis. HepG2 cells exhibited enrichment in the alcohol metabolic process, cellular amino acid metabolic process, proteasome, and skin development. We also employed machine learning techniques using OmicLearn[46] (v1.3) for data analysis, model execution, and visualization. The training process identified the most distinct features of each cell line and their weight ranking (Figure S7A; Table S6). Notably, the abundance levels of characteristic proteins identified through machine learning aligned with the gene expression trends of the corresponding cell lines in the DepMap database, underscoring the reliability of single-cell proteomics data obtained through AM-DMF-SCP (Figure S7B).

The protein identification in cells using the AM-DMF-SCP platform is in line with recent reports,[21,47] but it is noteworthy that our approach utilized shorter sample processing times and less mass spectrometry runtime. AM-DMF-SCP has proven to be a powerful tool in the comparative analysis of different human cell lines. Its ability to provide detailed and reliable data paves the way for a deeper understanding of cellular diversity and function, offering significant implications for basic research and clinical applications. We have utilized a label-free quantitation approach to demonstrate our platform's capability for direct and effective analysis of single-cell proteomes. The motivation for this choice is the direct compatibility of label-free quantitation with the automated features of the AM-DMF-SCP platform. Despite this focus, we acknowledge the versatility required in proteomics research and the significance of adapting the results to various analytical methods. Therefore, we are planning to expand our platform's capabilities to integrate chemical labeling methods such as TMT, iTRAQ, and plexDIA in the future, and we will systematically evaluate these quantitative approaches.

Single-Cell Proteomics Analysis of Changes in Drug-Resistant Cell Lines

We applied the AM-DMF-SCP workflow to examine drug-resistant cell lines postresistance development changes. To corroborate the single-cell results, we concurrently conducted a bulk proteome analysis on the same cells. As models, we utilized the ASK120067-resistant cell lines 67R and its parental NCI-H1975 cells (Figure S8). Our SCP results revealed no significant difference in the average number of proteins identified between the two cell lines (Figure 5A). However, a distinct separation of the cell lines was observed in the UMAP analysis (Figure 5B), indicating underlying proteomic differences. A differential protein analysis based on SCP results between the two cell lines revealed that 44 proteins were upregulated, and 121 proteins were downregulated in the drug-resistant 67R cell line ($P < 0.01$, $|FC| \geq 1.5$, Figure 5D; Table S7). Notably, the upregulation of VIM, an essential protein in the epithelial−mesenchymal transition (EMT) process, was particularly striking. VIM belongs to the intermediate filament (IF) protein family and is widely found in normal mesenchymal cells.[48] It plays a significant role in maintaining cellular structure and conferring resistance to stress.[49] EMT is known for its critical contribution to cancer progression, metastasis, and drug resistance[50] and is considered a potential target for treating drug resistance. In our study, subsequent Gene Set Enrichment Analysis (GSEA) of the expression data indicated that the 67R cell line exhibited an upregulation in the EMT pathway and apical junction, along with a down-regulation in adipogenesis and hypoxia (Figure 5E). The finding that showed enrichment in the EMT pathway was in line with bulk proteome analysis (Figure 5F), suggesting that single-cell proteomics data can mirror some characteristics of bulk samples. ASK120067 is a third-generation EGFR inhibitor.[51] So we investigated the correlation between VIM protein abundance[52] and the resistance scores[53] of various EGFR inhibitors in non-small cell lung adenocarcinoma cell lines. For most drugs, a significant positive correlation was observed between drug resistance and VIM expression (Figure 5G). However, Foretinib displayed an insignificant positive correlation, likely due to its lower inhibitory effectiveness on EGFR. This finding further emphasizes the link between VIM expression and EGFR inhibitor drug resistance.

The platform's precision in detecting subtle differences in protein expression and pathways provides valuable insights into the mechanisms of drug resistance at a single-cell level. This capability has significant implications for clinical applications, particularly in personalized medicine, where understanding the specific protein profiles of cancer cells, for instance, can lead to more targeted and effective treatments. However, the single-cell approach did not uncover undetected novel elements or pathways in bulk analysis. While our single-cell proteomics approach provided a detailed view at the cellular level, its findings largely aligned with those of bulk proteome analysis in this study. It is important to highlight that on another level, single-cell data effectively captured the heterogeneity within cell populations without the need for amplification to bulk samples. This aspect is particularly crucial in complex biological scenarios such as drug resistance in cancer cells, where understanding individual cell responses can offer invaluable insights. Single-cell proteomics thus plays a vital role in complementing bulk data, offering a more nuanced understanding of cellular diversity and behavior.

CONCLUSIONS

Designed to simplify sample preparation, the AM-DMF-SCP platform makes single-cell proteomics more accessible. It reduces the need for expensive equipment and lowers reagent use, significantly cutting costs and making this advanced research feasible for more laboratories. Our study underscores the capability of AM-DMF-SCP in single-cell proteomics, with the in-chip method demonstrating superiority in identifying more proteins than the out-chip. This method enhances the sensitivity, reduces sample contamination, and distinguishes cellular heterogeneity. Its effectiveness in differentiating cell types was crucial in our drug resistance analysis in cell lines, providing deeper insights into the complex nature of cellular responses. Thus, the AM-DMF-SCP platform is a vital tool in proteomic research, enabling a comprehensive understanding of diverse biological processes at the single-cell level.

https://doi.org/10.1021/jacsau.4c00027
JACS Au 2024, 4, 1811−1823

MATERIAL AND METHODS

Reagents and Materials

The reagents and materials used are provided in Table 1.

Table 1. Reagents and Materials

Reagents	Vendor	Catalog Number
n-Dodecyl-*b*-D-maltoside (DDM)	Sigma-Aldrich	324355
Tris(2-carboxyethyl) phosphine (TCEP)	Sigma-Aldrich	75259
Chloracetamide (CAA)	Sigma-Aldrich	22790
Sodium deoxycholate (SDC)	Sigma-Aldrich	D5670
2,2,2-Trifluoroethanol (TFE)	Sigma-Aldrich	96924
Trypsin	Promega	V5113
HeLa protein digest standards	Thermal Fisher	PI88329
10% Pluronic F68 solution	Sigma-Aldrich	P5556
Teflon-AF	Chemours	1600
Hexadecane	Sigma-Aldrich	H6703
Water (LC−MS grade)	Fisher Chemical	219205
Trifluoroacetic acid (TFA)	Sigma-Aldrich	8.08260
Acetonitrile (ACN)	Fisher Chemical	220889
Isopropanol (ISO)	Fisher Chemical	210948
Formic acid (FA)	Sigma-Aldrich	5.33002
RapiGest	Waters	213051
Fused silica capillary	Polymicro	TSP100375

AM16K DMF Chip Design and Fabrication

AM16K Digital microfluidic chip was formed from bottom and top plates. The top plate of the chip is approximately 2 in. of indium tin oxide (ITO) coated glass. The top plate has laser-cut microholes (1 or 2 mm in diameter) for sample injection and output. The chip bottom plate is manufactured using a standard flat panel display process and is the same size as the top plate. The bottom glass substrate contains a 128-row ×128-column array of thin film transistor (TFT) square pixel electrodes (250 μm side length) with 300 nm SiN_x deposited as the dielectric layer. A gap of known height is fixed between the top plate and the bottom plate of the chip, and the cavity formed is used for liquid flow and control. The volume of the liquid is positively correlated with the liquid area and chip gap. The complete AM16K DMF chip package was fabricated in acxel's Class 10000 clean room (Foshan, China). It mainly consists of ultrasonic cleaning of the chip's top and bottom plates. Spin-coating and baking of Teflon-AF (1% w/w in Fluorinert FC-40), an in-chip hydrophobic coating. Defining the 50 μm chip gap and sealing the chip was done using UV curable adhesive or plastic spacers. Standard hot press process for binding FPC connectors to DMF chips, etc.

DMF Instrument and Operation

The full-process single-cell proteomics sample preprocessing is conducted on an integrated digital microfluidics platform known as BOXmini SCP. This platform comprises several critical components: electronic drive module: central to the system is the AM16K digital microfluidic chip, managed by an electronic drive module. This module includes a control panel with a microcontroller unit (MCU) and row-column address controllers, ensuring precise handling of the chip's functionalities. Flexible printed circuit board (FPC): the control panel connects to the AM-16K chip via a flexible printed circuit board, facilitating seamless interaction between the control module and the chip. Imaging module: the electrode array on the AM16K chip is monitored by a dual-magnification imaging module. The low-magnification lens offers an overview of the entire chip, while the high-magnification lens focuses on individual cells or droplets. Positioning of the imaging module is controlled by a three-axis module using lead screw stepper motors paired with a ring light source to ensure comprehensive imaging of the chip array without blind spots. Temperature control: the temperature module, consisting of a thermo electric cooler (TEC) and a DC cooling fan, provides a wide range of temperature settings from 0 to 100 °C, crucial for various stages of sample processing. Software interface: the hardware functions are integrated through a custom software interface connected to computer software. This set up allows us for closed-loop control of the chip, including chip management, imaging, and temperature regulation, streamlining the entire experimental workflow. For operation, researchers simply load the AM16K chip onto the BOXmini SCP platform and fill it with medium oil to minimize microvolume sample evaporation and reduce nonspecific protein adsorption. The computer software is preprogramed with standard single-cell proteomics sample preprocessing protocols, guiding users through the process. The only manual intervention required is the timely addition of reagents and samples to the chip using a pipet, following the software's predefined schedule.

Electrode Residue Experiment

Peptide GDFQFNISR, derived from β-galactosidase of *E. coli* with a purity greater than 90%, was diluted to 1.5 pmol/μL, resulting in an approximate peptide mass of 200 pg/sample during dispensing. Concurrently, 0.1% FA droplets of the same size were dispensed by following the same path for subsequent mass spectrometry analysis.

Capillary Sampling Test

We conducted 20 sampling tests and measured the lengths of the capillaries used. We applied the interquartile range (IQR) method to identify and exclude outliers as follows: first, we calculated the first (Q1) and third (Q3) quartiles of the data, then determined the IQR as Q3−Q1. We defined the boundaries for outliers as lower_bound = Q1−1.5 × IQR and upper bound = Q3 + 1.5 × IQR. By identifying and excluding outliers, we then defined a reasonable range as the mean ± two standard deviations and calculated the success rate based on this criterion.

Cell Culture

Lung adenocarcinoma cell line A549, human embryonic kidney cell line HEK-293T, liver cancer cell line HepG2, and cervical cancer cell line HeLa were cultured in DMEM basic (1×) (Thermo Fisher Scientific, REF: C11995500BT) were cultivated in Dulbecco's modified Eagle's medium (DMEM 4.5 g/L glucose) (Gibco) supplemented with 10% fetal bovine serum (FBS) (Gibco) and 1% penicillin−streptomycin (Sigma-Aldrich) in 10 cm dishes under 5% CO_2 at 37 °C. Lung adenocarcinoma cell line NCI-H1975 and its drug-resistant cell line 67R were provided by Xie Lab[39] (Shanghai Institute of Materia Medica, Chinese Academy of Sciences).

Cell Harvested

Cells were harvested at ∼70% confluency by a 1 min treatment with trypsin (Sigma-Aldrich) at 37 °C, followed by a 1:3 dilution with full media to stop the digest. Cells were pelleted by centrifugation (1 min at 1000*g*, 25 °C) and washed thrice with 1× phosphate-buffered saline (PBS). For single pot methods, aliquots of ∼10,000 HEK-293T cells were added to eight-strip PCR tubes and later lysed with different buffers (all of the lysis buffers were freshly prepared just before use). For single-cell sorting, cells were washed three times with ice-cold PBS and then diluted with PBS plus 0.05% F68 to a concentration of 3.5×10^5 cells/mL, a concentration determined based on our previous work.[33]

Organic Solvent-Based Methods

HEK-293T cells were added 50 μL basic lysis buffer (20 mM TCEP,80 mM CAA, 200 mM Tris−HCl pH 8.5), followed by adding 50 μL ACN/TFE. Incubate at 56 °C for 5 min and at 95 °C for 5 min for sufficient lysis, reduction, and alkylation. Next, a vacuum drying instrument was used to dry the sample until it is complete and 100 μL of digestion buffer (10 mM $CaCl_2$) containing 0.5 μL of Trypsin (Promega). Proteins were digested for 4 h at 47 °C in a thermal shaker (1500 rpm). Digests were stopped by adding 10% trifluoroacetic acid (TFA) to a final concentration of 1%, and all of the resulting peptide samples were desalted on two layers of SDB-RPS StageTips. In brief, about 100 μL of peptides was directly loaded into StageTips and passed through by centrifugation for 5 min at 1500*g*.

https://doi.org/10.1021/jacsau.4c00027
JACS Au 2024, 4, 1811−1823

Then 100 μL of wash buffer1 [1%(v/v) TFA in ISO] and 100 μL of wash buffer 2 [0.2%(v/v)TFA] were used to wash the peptides in sequence. Finally, use 60 μL elution buffer [80%ACN/1% ammonium hydroxide (NH_4OH)] to elute the peptides into a new collection tube. The eluted peptides were dried using a SpeedVac centrifuge and resuspended in 10 μL of buffer A [0.1% (v/v) FA].

Surfactant-Based Methods

HEK-293T cells were added 100 μL of lysis buffer (0.2% DDM/0.1% Rapigest/1% SDC,10 mM TCEP, 40 mM CAA, 100 mM Tris−HCl pH 8.5,10 mM $CaCl_2$). Incubate at 56 °C for 5 min and at 95 °C for 5 min for sufficient lysis, reduction, and alkylation. Add 0.5 μL Trypsin (Promega) to the protein lysate and enzymatically digest at 47 °C for 4 h in a thermal shaker (1500 rpm). Digests were stopped by adding 10% TFA to a final concentration of 1%, except DDM-based Method digests were stopped by adding 2%(v/v) TFA in isopropanol. All of the resulting peptide samples were desalted on two layers of SDB-RPS StageTips according to the method described above.

Single-Cell Proteomics Sample Preparation

The detailed procedure of the in-chip method is showcased in Movie S1. The process involves the sequential injection of lysis solution (0.2% DDM, 10 mM TCEP, 40 mM CAA, 100 mM Tris−HCl pH 8.5), cell suspension, and digestion solution (50 ng/μL Trypsin in 100 mM Tris−HCl pH 8.5) into the chip. These solutions are then segmented into individual droplets as per the programed sequence, followed by the selection of appropriate droplets by using optical microscopic imaging and artificial intelligence vision screening.

The specific steps are as follows: suitable lysis solution droplets are mixed with single-cell droplets. This mixture undergoes reduction and alkylation at 56 °C for 15 min, followed by mixing with the digestion solution droplets and enzymatic digestion at 37 °C for 120 min. The sample output process postdigestion is detailed in Figure S1. The sample is combined with an antievaporation droplet and overflows from the output hole. The chip's temperature is then adjusted to 10 °C to solidify the medium oil, and the sample is extracted using a capillary tube and transferred to a low-binding sample vial. For the out-chip method, the cell suspension is injected into the chip through the sample inlet and split into individual droplets according to the programed sequence. Droplets containing a single cell are selected and directly output to a low-binding sample vial. Each selected droplet is then mixed with 1 μL of lysis solution and 1 μL of digestion solution to undergo the same reaction steps. Finally, 2 μL of 1% FA is added to acidify the sample.

NanoLC−MS/MS Analysis

Two different MS systems were used in this study. To determine the optimum single pot method for sample preparation, the nanoElute (Bruker) coupled with a timsTOF Pro mass spectrometer (Bruker) via nanoelectrospray ion source (Captive spray, Bruker) was used to measure samples. 1 μL (corresponding to ∼1000 cells) peptides were loaded on a 20 cm × 75 μm homemade column [ReproSil Saphir $100C_{18}$, 1.5 μm resin (Dr. Maisch GmbH, Germany)] The column was heated to 50 °C using an in-house-manufactured oven. The gradient time is 60 min, and the gradient was set as 2−4% B in 1 min, 4−26% B in 47 min, 26−32% B in 5 min, 32−90% B in 2 min at a flow rate of 250 nL/min, followed by 90% B maintained for 5 min at a flow rate of 300 nL/min. Source capillary voltage was set to 1500 V in positive ion mode and dry gas flow to 3 L/min at 180 °C. For data-independent acquisition, samples were acquired with a diaPASEF[54] method consisting of 14 cycles, including a total of 28 mass-width windows (25 Da width, from 452 to 1152 Da) with 4 mobility windows each, making a total of 56 windows (Table S8) covering the ion mobility range ($1/K_0$) from 0.76 to 1.29 V s/cm^2. The acquisition time of each PASEF scan was set to 100 ms with a nearly 100% duty cycle, which led to a total cycle time of around 1.59 s.

The single-cell proteomic analysis was performed on an Orbitrap Astral mass spectrometer (Thermo Fisher Scientific) and a Vanquish Neo UHPLC (Thermo Fisher Scientific) with the FAIMS Pro interface (Thermo Fisher Scientific) using a compensation voltage set to −48 V. Mobile phases A, and B were 0.1% formic in water and 0.1% formic acid/80% acetonitrile, respectively. The column was heated to 50 °C using an in-house-manufactured oven. The gradient was set for 15 min, starting with an increase from 4 to 8% B in 0.1 min at 750 nL/min, then 8 to 18% B over 1.8 min at the same flow rate, briefly shifting to 18 to 18.1% B in 0.1 min while reducing flow to 200 nL/min, followed by 18.1 to 28% B over 3 min at 200 nL/min, then escalating from 28 to 48% B in 2.5 min while decreasing flow to 180 nL/min, sharply rising from 48 to 99% B in 0.4 min, and finally holding at 90% B for 5 min at a flow rate of 180 nL/min. A source voltage of 1500 V was used for all of the experiments. For DIA experiments on the Orbitrap Astral MS, MS1 spectra were acquired with the Orbitrap at a resolution of 120,000 and a scan range of 400−800 m/z with normalized automatic gain control (AGC) target of 500% and maximum injection time of 100 ms. Fragmentation of precursor ions was performed by using higher energy collisional dissociation (HCD) using a normalized collision energy (NCE) of 25%. AGC target was set to 800% and the maximum injection time was 40 ms. The isolation widths were 20 Th.

Data Analysis

For all experiments, DIA-NN 1.8.1[55] was used to process raw MS files. Bruker timsTOF data was searched with a library-free search against the *Homo sapiens* Swissprot reference proteomes database (release 2021_09, 20588 entries). The false discovery rate (FDR) was set to 1% at the precursor and protein levels. N-terminal methionine excision was enabled. Cysteine carbamidomethylation was set as fixed, and methionine oxidation and N-terminal acetylation were used as variable modifications. Max misscleavage was set to 2, and the maximum variable modification of a peptide was set to 5. MBR mode was turned off. The rest of the parameters were kept as a default. Astral raw data was converted from .raw to .dia format. All the parameters were the same as the timsTOF data, except the precursor m/z range was set from 400 to 800.

The peptide data was imported into *R* using the "diann" package (https://github.com/vdemichev/diann-rpackage), specifically from the "report.tsv" file, employing the MaxLFQ algorithm for analysis. Protein Groups data was imported into *R* from the "···_pg_matrix.tsv" file. Missing quantitative values were imputed using the *K*-nearest neighbor (KNN) algorithm with $k = 5$. Proteins detected in less than half of the samples of the compared group were removed from the downstream analysis. The CV values were calculated for each feature based on the formula CV = SD/mean. The GRAVY value is calculated by the sum of hydropathy values of all amino acids divided by the protein length based on Kyte and Doolittle methods.[56] Pearson correlations were determined using the log-2 transformed mean abundances of proteins that were quantifiable and common among the groups. *P* values for differentially expressed proteins were calculated by using a Wilcoxon test. Functional and pathway enrichment analyses were performed using ClusterProfiler.[57] The networks of cell-specific genes were constructed by *R* package STRINGdb[32] and ggraph.

The cell line protein-coding gene expression levels are publicly available at the DepMap portal (https://depmap.org/portal/download/all/, 23Q2 version of the "OmicsExpressionProteinCodingGenesTPMLogp1"). The cell line proteomics data are publicly available at ProCan-DepMapSanger. Drug response measurements are publicly available at Genomics of drug sensitivity in cancer (GDSC, https://www.cancerrxgene.org/downloads/drug_data).

OmicLearn (https://github.com/OmicEra/OmicLearn, v1.3) was used to characterize cell-specific characters and perform data analysis, model execution, and creation of plots and charts. Machine learning was done in Python (3.8.1). Feature tables were imported via the Pandas package (1.5.0) and manipulated by using the Numpy package (1.23.3). The machine learning pipeline was employed using the scikit-learn package (1.1.2). The Plotly (5.9.0) library was used for plotting. No normalization of the data was performed. The data set contained no missing values; hence, no imputation was performed. Features were selected using an ExtraTrees ($n\_trees = 100$) strategy with a maximum number of 20 features. During training, normalization, and feature selection were individually performed using the

https://doi.org/10.1021/jacsau.4c00027
JACS Au 2024, 4, 1811−1823

data of each split. For classification, we used an AdaBoost-Classifier (random_state = 23, *n*_estimators = 100, learning_rate = 1.0). We used a repeated (*n*_repeats = 10), stratified cross-validation (RepeatedStratifiedKFold, *n*_splits = 5) approach to classify one vs another cellline, and we achieved a receiver operating characteristic (ROC) with an average AUC (area under the curve).

■ ASSOCIATED CONTENT

Data Availability Statement

The raw mass spectrometry data and diann-search results have been deposited to the integrated proteome resources (iProX)[58] (https://www.iprox.cn/) via the PRIDE partner repository with the identifier IPX0007681000. Code Availability: The code for data analysis is available at https://github.com/youngbee12/AM-DMF-SCP.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.4c00027.

Partial results of system optimization evaluation and outcomes of biological relevance analysis (PDF)

Complete sample processing procedure of AM-DMF-SCP (ZIP)

■ AUTHOR INFORMATION

Corresponding Authors

Hua Xie − *Division of Antitumor Pharmacology, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; University of the Chinese Academy of Sciences, Beijing 100049, China; Zhongshan Institute for Drug Discovery, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Zhongshan 528400, China*; Email: hxie@simm.ac.cn

Hanbin Ma − *CAS Key Laboratory of Bio-Medical Diagnostics, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, China; Guangdong ACXEL Micro & Nano Tech Co. Ltd., Foshan, Guangdong Province 528000, China*; orcid.org/0000-0002-7629-2287; Email: mahb@sibet.ac.cn

Hu Zhou − *Department of Analytical Chemistry, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; University of the Chinese Academy of Sciences, Beijing 100049, China; Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China*; orcid.org/0000-0001-7006-4737; Email: zhouhu@simm.ac.cn

Authors

Zhicheng Yang − *Department of Analytical Chemistry, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; University of the Chinese Academy of Sciences, Beijing 100049, China*; orcid.org/0000-0001-7477-2984

Kai Jin − *CAS Key Laboratory of Bio-Medical Diagnostics, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, China*

Yimin Chen − *Department of Analytical Chemistry, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; University of the Chinese Academy of Sciences, Beijing 100049, China*

Qian Liu − *Department of Analytical Chemistry, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China*

Hongxu Chen − *School of Chinese Materia Medica, Nanjing University of Chinese Medicine, Nanjing, Jiangsu 210023, China*

Siyi Hu − *CAS Key Laboratory of Bio-Medical Diagnostics, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, China*; orcid.org/0000-0002-0686-5182

Yuqiu Wang − *Department of Analytical Chemistry, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China*; orcid.org/0009-0001-8931-8853

Zilu Pan − *Division of Antitumor Pharmacology, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China*

Fang Feng − *Division of Antitumor Pharmacology, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China*

Mude Shi − *Guangdong ACXEL Micro & Nano Tech Co. Ltd., Foshan, Guangdong Province 528000, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.4c00027

Author Contributions

▽Z.Y., K.J., and Y.C. contributed equally to this work. CRediT: **Z.Y.** data curation, formal analysis, investigation, methodology, validation, visualization, writing-original draft; **K.J.** formal analysis, funding acquisition, investigation, methodology, visualization, writing-original draft; **Y.C.** data curation, formal analysis, methodology, visualization, writing-original draft; **Q.L.** methodology, writing-review & editing; **H.C.** data curation, methodology; **S.H.** funding acquisition, methodology; **Y.W.** investigation, writing-review & editing; **Z.P.** validation; **F.F.** validation; **M.S.** writing-review & editing; **H.X.** conceptualization, funding acquisition, resources, supervision, writing-review & editing; **H.M.** conceptualization, funding acquisition, resources, supervision, writing-review & editing; **H.Z.** conceptualization, funding acquisition, resources, supervision, writing-review & editing.

Notes

The authors declare no competing financial interest.

Data Visualization: Plots were generated using the ggplot2 (v3.4.2) package unless otherwise indicated.

■ ACKNOWLEDGMENTS

This work was financially supported by the National Key Research and Development Program of China (2022YFA1302902 and 2023YFF0721500), Shanghai Municipal Science and Technology Major Project, the Shanghai Young Excellent Academic Leader Program (20XD1424900), the China Postdoctoral Science Foundation (2022M722338), the National Natural Science Foundation of China (82273948 and 62374102), the Innovation and Entrepreneurship Team of Jiangsu Province (JSSCTD202145), the High-level Innovative Research Institute (2021B0909050003), and the Science and Technology Innovation Project of Foshan, Guangdong Province, China (1920001000047). We thank the Mass Spectrometry Service Department, Institutional Center for Shared Technologies and Facilities of the Shanghai Institute of

Materia Medica, Chinese Academy of Sciences, and Shanghai Bioprofile Technology Co. Ltd for mass spectrometry analysis.

REFERENCES

(1) Mund, A.; Coscia, F.; Kriston, A.; Hollandi, R.; Kovacs, F.; Brunner, A. D.; Migh, E.; Schweizer, L.; Santos, A.; Bzorek, M.; Naimy, S.; Rahbek-Gjerdrum, L. M.; Dyring-Andersen, B.; Bulkescher, J.; Lukas, C.; Eckert, M. A.; Lengyel, E.; Gnann, C.; Lundberg, E.; Horvath, P.; Mann, M. Deep Visual Proteomics defines single-cell identity and heterogeneity. *Nat. Biotechnol.* **2022**, *40*, 1231−1240.

(2) Dang, Y.; Zhu, L.; Yuan, P.; Liu, Q.; Guo, Q.; Chen, X.; Gao, S.; Liu, X.; Ji, S.; Yuan, Y.; Lian, Y.; Li, R.; Yan, L.; Wong, C. C. L.; Qiao, J. Functional profiling of stage-specific proteome and translational transition across human pre-implantation embryo development at a single-cell resolution. *Cell Discovery* **2023**, *9*, 10.

(3) Vegvari, A.; Rodriguez, J. E.; Zubarev, R. A. Single-Cell Chemical Proteomics (SCCP) Interrogates the Timing and Heterogeneity of Cancer Cell Commitment to Death. *Anal. Chem.* **2022**, *94*, 9261−9269.

(4) Rosenberger, F. A.; Thielert, M.; Strauss, M. T.; Schweizer, L.; Ammar, C.; Madler, S. C.; Metousis, A.; Skowronek, P.; Wahle, M.; Madden, K.; Gote-Schniering, J.; Semenova, A.; Schiller, H. B.; Rodriguez, E.; Nordmann, T. M.; Mund, A.; Mann, M. Spatial single-cell mass spectrometry defines zonation of the hepatocyte proteome. *Nat. Methods* **2023**, *20*, 1530−1536.

(5) Heil, L. R.; Damoc, E.; Arrey, T. N.; Pashkova, A.; Denisov, E.; Petzoldt, J.; Peterson, A. C.; Hsu, C.; Searle, B. C.; Shulman, N.; Riffle, M.; Connolly, B.; MacLean, B. X.; Remes, P. M.; Senko, M. W.; Stewart, H. I.; Hock, C.; Makarov, A. A.; Hermanson, D.; Zabrouskov, V.; et al. Evaluating the Performance of the Astral Mass Analyzer for Quantitative Proteomics Using Data-Independent Acquisition. *J. Proteome Res.* **2023**, *22*, 3290−3300.

(6) Stewart, H. I.; Grinfeld, D.; Giannakopulos, A.; Petzoldt, J.; Shanley, T.; Garland, M.; Denisov, E.; Peterson, A. C.; Damoc, E.; Zeller, M.; Arrey, T. N.; Pashkova, A.; Renuse, S.; Hakimi, A.; Kuhn, A.; Biel, M.; Kreutzmann, A.; Hagedorn, B.; Colonius, I.; Schutz, A.; Stefes, A.; Dwivedi, A.; Mourad, D.; Hoek, M.; Reitemeier, B.; Cochems, P.; Kholomeev, A.; Ostermann, R.; Quiring, G.; Ochmann, M.; Mohring, S.; Wagner, A.; Petker, A.; Kanngiesser, S.; Wiedemeyer, M.; Balschun, W.; Hermanson, D.; Zabrouskov, V.; Makarov, A. A.; Hock, C. Parallelized Acquisition of Orbitrap and Astral Analyzers Enables High-Throughput Quantitative Analysis. *Anal. Chem.* **2023**, *95*, 15656−15664.

(7) Petrosius, V.; Aragon-Fernandez, P.; Arrey, T. N.; Üresin, N.; Furtwängler, B.; Stewart, H.; Denisov, E.; Petzoldt, J.; Peterson, A. C.; Hock, C.; Damoc, E.; Makarov, A.; Zabrouskov, V.; Porse, B. T.; Schoof, E. M. Evaluating the capabilities of the Astral mass analyzer for single-cell proteomics. *bioRxiv* **2023**, 2023.06.06.543943.

(8) Brunner, A.-D.; Thielert, M.; Vasilopoulou, C.; Ammar, C.; Coscia, F.; Mund, A.; Hoerning, O. B.; Bache, N.; Apalategui, A.; Lubeck, M.; Richter, S.; Fischer, D. S.; Raether, O.; Park, M. A.; Meier, F.; Theis, F. J.; Mann, M. Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *Mol. Syst. Biol.* **2022**, *18*, No. e10798.

(9) Budnik, B.; Levy, E.; Harmange, G.; Slavov, N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* **2018**, *19*, 161.

(10) Specht, H.; Emmott, E.; Petelski, A. A.; Huffman, R. G.; Perlman, D. H.; Serra, M.; Kharchenko, P.; Koller, A.; Slavov, N. Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome Biol.* **2021**, *22*, 50.

(11) Derks, J.; Leduc, A.; Wallmann, G.; Huffman, R. G.; Willetts, M.; Khan, S.; Specht, H.; Ralser, M.; Demichev, V.; Slavov, N. Increasing the throughput of sensitive proteomics by plexDIA. *Nat. Biotechnol.* **2023**, *41*, 50−59.

(12) Peng, J.; Chan, C.; Zhang, S.; Sklavounos, A. A.; Olson, M. E.; Scott, E. Y.; Hu, Y.; Rajesh, V.; Li, B. B.; Chamberlain, M. D.; Zhang, S.; Peng, H.; Wheeler, A. R. All-in-One digital microfluidics pipeline for proteomic sample preparation and analysis. *Chem. Sci.* **2023**, *14*, 2887−2900.

(13) Ctortecka, C.; Hartlmayr, D.; Seth, A.; Mendjan, S.; Tourniaire, G.; Udeshi, N. D.; Carr, S. A.; Mechtler, K. An Automated Nanowell-Array Workflow for Quantitative Multiplexed Single-Cell Proteomics Sample Preparation at High Sensitivity. *Mol. Cell. Proteomics* **2023**, *22*, 100665.

(14) Sun, B.; Kumar, S. Protein Adsorption Loss−The Bottleneck of Single-Cell Proteomics. *J. Proteome Res.* **2022**, *21*, 1808−1815.

(15) Li, Z. Y.; Huang, M.; Wang, X. K.; Zhu, Y.; Li, J. S.; Wong, C. C. L.; Fang, Q. Nanoliter-Scale Oil-Air-Droplet Chip-Based Single Cell Proteomic Analysis. *Anal. Chem.* **2018**, *90*, 5430−5438.

(16) Zhu, Y.; Piehowski, P. D.; Zhao, R.; Chen, J.; Shen, Y.; Moore, R. J.; Shukla, A. K.; Petyuk, V. A.; Campbell-Thompson, M.; Mathews, C. E.; Smith, R. D.; Qian, W. J.; Kelly, R. T. Nanodroplet processing platform for deep and quantitative proteome profiling of 10−100 mammalian cells. *Nat. Commun.* **2018**, *9*, 882.

(17) Cong, Y.; Motamedchaboki, K.; Misal, S. A.; Liang, Y.; Guise, A. J.; Truong, T.; Huguet, R.; Plowey, E. D.; Zhu, Y.; Lopez-Ferrer, D.; Kelly, R. T. Ultrasensitive single-cell proteomics workflow identifies > 1000 protein groups per mammalian cell. *Chem. Sci.* **2021**, *12*, 1001−1006.

(18) Shao, X.; Wang, X.; Guan, S.; Lin, H.; Yan, G.; Gao, M.; Deng, C.; Zhang, X. Integrated Proteome Analysis Device for Fast Single-Cell Protein Profiling. *Anal. Chem.* **2018**, *90*, 14003−14010.

(19) He, Y.; Yuan, H.; Liang, Y.; Liu, X.; Zhang, X.; Ji, Y.; Zhao, B.; Yang, K.; Zhang, J.; Zhang, S.; Zhang, Y.; Zhang, L. On-capillary alkylation micro-reactor: a facile strategy for proteo-metabolome profiling in the same single cells. *Chem. Sci.* **2023**, *14*, 13495−13502.

(20) Matzinger, M.; Muller, E.; Durnberger, G.; Pichler, P.; Mechtler, K. Robust and Easy-to-Use One-Pot Workflow for Label-Free Single-Cell Proteomics. *Anal. Chem.* **2023**, *95*, 4435−4445.

(21) Sanchez-Avila, X.; Truong, T.; Xie, X.; Webber, K. G. I.; Johnston, S. M.; Lin, H.-J. L.; Axtell, N. B.; Puig-Sanvicens, V.; Kelly, R. T. Easy and Accessible Workflow for Label-Free Single-Cell Proteomics. *J. Am. Soc. Mass Spectrom.* **2023**, *34*, 2374−2380.

(22) Masuda, T.; Inamori, Y.; Furukawa, A.; Yamahiro, M.; Momosaki, K.; Chang, C. H.; Kobayashi, D.; Ohguchi, H.; Kawano, Y.; Ito, S.; Araki, N.; Ong, S. E.; Ohtsuki, S. Water Droplet-in-Oil Digestion Method for Single-Cell Proteomics. *Anal. Chem.* **2022**, *94*, 10329−10336.

(23) Leipert, J.; Steinbach, M. K.; Tholey, A. Isobaric Peptide Labeling on Digital Microfluidics for Quantitative Low Cell Number Proteomics. *Anal. Chem.* **2021**, *93*, 6278−6286.

(24) Lamanna, J.; Scott, E. Y.; Edwards, H. S.; Chamberlain, M. D.; Dryden, M. D. M.; Peng, J.; Mair, B.; Lee, A.; Chan, C.; Sklavounos, A. A.; Heffernan, A.; Abbas, F.; Lam, C.; Olson, M. E.; Moffat, J.; Wheeler, A. R. Digital microfluidic isolation of single cells for -Omics. *Nat. Commun.* **2020**, *11*, 5632.

(25) Zeng, X.; Guo, X.; Jiang, S.; Yang, X.; Zhong, Z.; Liu, S.; Zhu, Z.; Song, J.; Yang, C. Digital-scRRBS: A Cost-Effective, Highly Sensitive, and Automated Single-Cell Methylome Analysis Platform via Digital Microfluidics. *Anal. Chem.* **2023**, *95*, 13313−13321.

(26) Anderson, S.; Hadwen, B.; Brown, C. Thin-film-transistor digital microfluidics for high value in vitro diagnostics at the point of need. *Lab Chip* **2021**, *21*, 962−975.

(27) Ma, H.; Shi, S.; Jin, K.; Wang, D.; Hu, S.; Su, Y.; Zhang, Y.; Li, J.; Liu, Z.; Jiang, C.; Feng, L.; Guo, X.; Nathan, A. *2020 IEEE International Electron Devices Meeting (IEDM), 12−18 Dec 2020*, 2020.

(28) Wang, Y.; Guan, Z. Y.; Shi, S. W.; Jiang, Y. R.; Zhang, J.; Yang, Y.; Wu, Q.; Wu, J.; Chen, J. B.; Ying, W. X.; Xu, Q. Q.; Fan, Q. X.; Wang, H. F.; Zhou, L.; Wang, L.; Fang, J.; Pan, J. Z.; Fang, Q. Pick-up single-cell proteomic analysis for quantifying up to 3000 proteins in a Mammalian cell. *Nat. Commun.* **2024**, *15*, 1279.

(29) Liang, Y.; Acor, H.; McCown, M. A.; Nwosu, A. J.; Boekweg, H.; Axtell, N. B.; Truong, T.; Cong, Y.; Payne, S. H.; Kelly, R. T. Fully Automated Sample Processing and Analysis Workflow for Low-Input Proteome Profiling. *Anal. Chem.* **2021**, *93*, 1658−1666.

https://doi.org/10.1021/jacsau.4c00027
JACS Au 2024, 4, 1811−1823

(30) Woo, J.; Williams, S. M.; Markillie, L. M.; Feng, S.; Tsai, C. F.; Aguilera-Vazquez, V.; Sontag, R. L.; Moore, R. J.; Hu, D.; Mehta, H. S.; Cantlon-Bruce, J.; Liu, T.; Adkins, J. N.; Smith, R. D.; Clair, G. C.; Pasa-Tolic, L.; Zhu, Y. High-throughput and high-efficiency sample preparation for single-cell proteomics using a nested nanowell chip. *Nat. Commun.* **2021**, *12*, 6246.
(31) Leduc, A.; Huffman, R. G.; Cantlon, J.; Khan, S.; Slavov, N. Exploring functional protein covariation across single cells using nPOP. *Genome Biol.* **2022**, *23*, 261.
(32) Gebreyesus, S. T.; Siyal, A. A.; Kitata, R. B.; Chen, E. S.; Enkhbayar, B.; Angata, T.; Lin, K. I.; Chen, Y. J.; Tu, H. L. Streamlined single-cell proteomics by an integrated microfluidic chip and data-independent acquisition mass spectrometry. *Nat. Commun.* **2022**, *13*, 37.
(33) Hu, S.; Ye, J.; Shi, S.; Yang, C.; Jin, K.; Hu, C.; Wang, D.; Ma, H. Large-Area Electronics-Enabled High-Resolution Digital Microfluidics for Parallel Single-Cell Manipulation. *Anal. Chem.* **2023**, *95*, 6905−6914.
(34) Leon, I. R.; Schwammle, V.; Jensen, O. N.; Sprenger, R. R. Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. *Mol. Cell. Proteomics* **2013**, *12*, 2992−3005.
(35) Nie, S.; O'Brien Johnson, R.; Livson, Y.; Greer, T.; Zheng, X.; Li, N. Maximizing hydrophobic peptide recovery in proteomics and antibody development using a mass spectrometry compatible surfactant. *Anal. Biochem.* **2022**, *658*, 114924.
(36) Tsai, C.-F.; Zhang, P.; Scholten, D.; Martin, K.; Wang, Y.-T.; Zhao, R.; Chrisler, W. B.; Patel, D. B.; Dou, M.; Jia, Y.; Reduzzi, C.; Liu, X.; Moore, R. J.; Burnum-Johnson, K. E.; Lin, M.-H.; Hsu, C.-C.; Jacobs, J. M.; Kagan, J.; Srivastava, S.; Rodland, K. D.; Steven Wiley, H.; Qian, W.-J.; Smith, R. D.; Zhu, Y.; Cristofanilli, M.; Liu, T.; Liu, H.; Shi, T. Surfactant-assisted one-pot sample preparation for label-free single-cell proteomics. *Commun. Biol.* **2021**, *4*, 265.
(37) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26*, 1367−1372.
(38) Orsburn, B. C. Evaluation of the Sensitivity of Proteomics Methods Using the Absolute Copy Number of Proteins in a Single Cell as a Metric. *Proteomes* **2021**, *9* (3), 34.
(39) Gatto, L.; Aebersold, R.; Cox, J.; Demichev, V.; Derks, J.; Emmott, E.; Franks, A. M.; Ivanov, A. R.; Kelly, R. T.; Khoury, L.; Leduc, A.; MacCoss, M. J.; Nemes, P.; Perlman, D. H.; Petelski, A. A.; Rose, C. M.; Schoof, E. M.; Van Eyk, J.; Vanderaa, C.; Yates, J. R.; Slavov, N. Initial recommendations for performing, benchmarking and reporting single-cell proteomics experiments. *Nat. Methods* **2023**, *20*, 375−386.
(40) Petrosius, V.; Aragon-Fernandez, P.; Üresin, N.; Kovacs, G.; Phlairaharn, T.; Furtwängler, B.; Op De Beeck, J.; Skovbakke, S. L.; Goletz, S.; Thomsen, S. F.; Keller, U. a. d.; Natarajan, K. N.; Porse, B. T.; Schoof, E. M. Exploration of cell state heterogeneity using single-cell proteomics through sensitivity-tailored data-independent acquisition. *Nat. Commun.* **2023**, *14*, 5910.
(41) Arzumanian, V. A.; Kiseleva, O. I.; Poverennaya, E. V. The Curious Case of the HepG2 Cell Line: 40 Years of Expertise. *Int. J. Mol. Sci.* **2021**, *22* (23), 13135.
(42) Jiang, R.-D.; Shen, H.; Piao, Y.-J. The morphometrical analysis on the ultrastructure of A549 cells. *Rom. J. Morphol. Embryol.* **2010**, *51*, 663−667.
(43) Thomas, P. D.; Ebert, D.; Muruganujan, A.; Mushayahama, T.; Albou, L.-P.; Mi, H. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* **2022**, *31*, 8−22.
(44) Szklarczyk, D.; Gable, A. L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N. T.; Morris, J. H.; Bork, P.; Jensen, L. J.; Mering, C. V. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607−d613.
(45) Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **2016**, *44*, D457−D462.
(46) Torun, F. M.; Virreira Winter, S.; Doll, S.; Riese, F. M.; Vorobyev, A.; Mueller-Reif, J. B.; Geyer, P. E.; Strauss, M. T. Transparent Exploration of Machine Learning for Biomarker Discovery from Proteomics and Omics Data. *J. Proteome Res.* **2023**, *22*, 359−367.
(47) Johnston, S. M.; Webber, K. G. I.; Xie, X.; Truong, T.; Nydegger, A.; Lin, H. J. L.; Nwosu, A.; Zhu, Y.; Kelly, R. T. Rapid, One-Step Sample Processing for Label-Free Single-Cell Proteomics. *J. Am. Soc. Mass Spectrom.* **2023**, *34*, 1701−1707.
(48) Franke, W. W.; Schmid, E.; Osborn, M.; Weber, K. Different intermediate-sized filaments distinguished by immunofluorescence microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5034−5038.
(49) Satelli, A.; Li, S. Vimentin in cancer and its potential as a molecular target for cancer therapy. *Cell. Mol. Life Sci.* **2011**, *68*, 3033−3046.
(50) De Las Rivas, J.; Brozovic, A.; Izraely, S.; Casas-Pais, A.; Witz, I. P.; Figueroa, A. Cancer drug resistance induced by EMT: novel therapeutic strategies. *Arch. Toxicol.* **2021**, *95*, 2279−2297.
(51) Zhang, T.; Qu, R.; Chan, S.; Lai, M.; Tong, L.; Feng, F.; Chen, H.; Song, T.; Song, P.; Bai, G.; Liu, Y.; Wang, Y.; Li, Y.; Su, Y.; Shen, Y.; Sun, Y.; Chen, Y.; Geng, M.; Ding, K.; Ding, J.; Xie, H. Discovery of a novel third-generation EGFR inhibitor and identification of a potential combination strategy to overcome resistance. *Mol. Cancer* **2020**, *19*, 90.
(52) Goncalves, E.; Poulos, R. C.; Cai, Z.; Barthorpe, S.; Manda, S. S.; Lucas, N.; Beck, A.; Bucio-Noble, D.; Dausmann, M.; Hall, C.; Hecker, M.; Koh, J.; Lightfoot, H.; Mahboob, S.; Mali, I.; Morris, J.; Richardson, L.; Seneviratne, A. J.; Shepherd, R.; Sykes, E.; Thomas, F.; Valentini, S.; Williams, S. G.; Wu, Y.; Xavier, D.; MacKenzie, K. L.; Hains, P. G.; Tully, B.; Robinson, P. J.; Zhong, Q.; Garnett, M. J.; Reddel, R. R. Pan-cancer proteomic map of 949 human cell lines. *Cancer Cell* **2022**, *40*, 835−849.e8.
(53) Yang, W.; Soares, J.; Greninger, P.; Edelman, E. J.; Lightfoot, H.; Forbes, S.; Bindal, N.; Beare, D.; Smith, J. A.; Thompson, I. R.; Ramaswamy, S.; Futreal, P. A.; Haber, D. A.; Stratton, M. R.; Benes, C.; McDermott, U.; Garnett, M. J. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **2012**, *41*, D955−D961.
(54) Meier, F.; Brunner, A.-D.; Frank, M.; Ha, A.; Bludau, I.; Voytik, E.; Kaspar-Schoenefeld, S.; Lubeck, M.; Raether, O.; Bache, N.; Aebersold, R.; Collins, B. C.; Röst, H. L.; Mann, M. diaPASEF: parallel accumulation-serial fragmentation combined with data-independent acquisition. *Nat. Methods* **2020**, *17*, 1229−1236.
(55) Demichev, V.; Messner, C. B.; Vernardis, S. I.; Lilley, K. S.; Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **2020**, *17*, 41−44.
(56) Kyte, J.; Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105−132.
(57) Wu, T.; Hu, E.; Xu, S.; Chen, M.; Guo, P.; Dai, Z.; Feng, T.; Zhou, L.; Tang, W.; Zhan, L.; Fu, X.; Liu, S.; Bo, X.; Yu, G. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2021**, *2*, 100141.
(58) Ma, J.; Chen, T.; Wu, S.; Yang, C.; Bai, M.; Shu, K.; Li, K.; Zhang, G.; Jin, Z.; He, F.; Hermjakob, H.; Zhu, Y. iProX: an integrated proteome resource. *Nucleic Acids Res.* **2019**, *47*, D1211−D1217.

https://doi.org/10.1021/jacsau.4c00027
JACS Au 2024, 4, 1811−1823