The application of high-throughput proteomics in cytopathology

Ilias P. Nikas¹, Han Suk Ryu²,³

¹School of Medicine, European University Cyprus, Nicosia, Cyprus; ²Department of Pathology, Seoul National University College of Medicine, Seoul; ³Department of Pathology, Seoul National University Hospital, Seoul, Korea

High-throughput genomics and transcriptomics are often applied in routine pathology practice to facilitate cancer diagnosis, assess prognosis, and predict response to therapy. However, the proteins rather than nucleic acids are the functional molecules defining the cellular phenotype in health and disease, whereas genomic profiling cannot evaluate processes such as the RNA splicing or post-translational modifications and gene expression does not necessarily correlate with protein expression. Proteomic applications have recently advanced, overcoming the issue of low depth, inconsistency, and suboptimal accuracy, also enabling the use of minimal patient-derived specimens. This review aims to present the recent evidence regarding the use of high-throughput proteomics in both exfoliative and fine-needle aspiration cytology. Most studies used mass spectrometry, as this is associated with high depth, sensitivity, and specificity, and aimed to complement the traditional cytomorphologic diagnosis, in addition to identify novel cancer biomarkers. Examples of diagnostic dilemmas subjected to proteomic analysis included the evaluation of indeterminate thyroid nodules or prediction of lymph node metastasis from thyroid cancer, also the differentiation between benign and malignant serous effusions, pancreatic cancer from autoimmune pancreatitis, non-neoplastic from malignant biliary strictures, and benign from malignant salivary gland tumors. A few cancer biomarkers—related to diverse cancers involving the breast, thyroid, bladder, lung, serous cavities, salivary glands, and bone marrow—were also discovered. Notably, residual liquid-based cytology samples were suitable for satisfactory and reproducible proteome analysis. Proteomics could become another routine pathology platform in the near future, potentially by using validated multi-omics protocols.

Key Words: Cytology; Fine-needle aspiration; Mass spectrometry; Cancer biomarker; Proteomics

Since next-generation sequencing (NGS) technologies were introduced, sequencing data output significantly increased and brought unprecedented revolution into cancer genomic profiling [1,2]. In addition, the affordable cost of NGS technologies has made their clinical application feasible, as well as their use in the research setting [2,3]. Comprehensive genetic profiling of tumor samples has driven the construction of The Cancer Genome Atlas (TCGA), comprising enormous genomic landscapes across various cancer types. Notably, NGS-based gene panel tests have put genomic sequencing into routine clinical practice as diagnostic tools enabling precision medicine [4]. In addition to surgical pathology, NGS has been extensively used in the field of cytology, utilizing both exfoliative and fine-needle aspiration (FNA) samples [5-9].

However, the number of transcripts does not necessarily correlate with that of the translated proteins, which are the actual functional molecules defining the cellular phenotype in health and disease. Multiple splicing variants could be formed from each transcript during RNA maturation [10-12], while more than 400 different types of post-translational modifications such as acetylation, phosphorylation, glycosylation, methylation, and peptide cleavage might change the properties of the final protein product [12-14]. Furthermore, it may be difficult to define which mutations are the driver and passenger ones while analyzing nucleic acids. All these may limit our understanding of the complexity of cancer and our quest for optimal diagnostic, prognostic, and therapeutic biomarkers, especially when counting solely on data derived from genomics and/or transcriptomics [15]. Thus, the integration of multi-omic approaches, including genomics, epigenomics, transcriptomics, proteomics, and/or metabolomics,
could combine the strengths of each high-throughput application, enhancing cancer diagnosis, prognosis, and therapy [16,17].

In the past, classic analytical methods to detect proteins struggled due to the structural instability of proteins, which are sensitive to degradation by proteases [12,18]. Proteins cannot be amplified, similar to the nucleic acids via the polymerase chain reaction. Thus, analyzing small amounts of proteins was challenging and a large amount of proteins per sample was needed for quality assurance and successful proteomic analysis [12]. However, since mass spectrometry (MS) has been established as the modern technology of choice for proteomics, it has provided researchers with high depth, improved accuracy, and unbiased quality [15,19]. Recent technological improvements have allowed the analysis of large-scale proteomes and improved the speed of analysis with short turnaround times [19]. Such technical advances have succeeded in the detection of almost entire proteomes in clinical as well as research samples [20,21]. Furthermore, the enhanced sensitivity and specificity of mass spectrometry, enabling the measurement of minute amount of proteins, has allowed the consideration of proteomics application into future routine clinical practice [22,23].

**BASIC PRINCIPLES OF PROTEOMICS**

The general aims of proteomic approaches are as follows: (1) identification of specific proteome groups, (2) analysis (e.g., expression levels) of differentially expressed protein signatures from two or more samples, (3) bioinformatic analysis, including the study of protein-protein interactions and gene set enrichment, and (4) study of post-translational modifications in a variety of samples including cell lines, tissue biopsies, and cytology [24,25].

There are two types of proteomic approaches based on the analytic platform used, the protein microarrays and MS-based techniques [26-28]. Regarding the former, there are three types of protein arrays: the analytic microarrays, functional microarrays, and reverse-phase protein microarrays [29]. These arrays have been used to detect differentially expressed protein landscapes, identifying the presence of altered proteins or molecular interactions in certain diseases [30]. However, the restricted number of suitable antibodies needed for such analysis, which could also result in non-specific antigen-antibody interactions, is considered as their main limitation for its use in research or the clinical laboratories [18,28].

During the last years, MS has been significantly improved and emerged as the next generation technology of proteomics, due to its capacity to analyze large-scale proteomes with high sensitivity and specificity [19]. This advanced technique has made protein sequencing possible through three major steps; protein ionization, separation of the ionized analytes based on their own m/z (mass-to-charge) ratio, and detection of the analytes. Finally, the mass spectrum displays the relative abundance of charged analytes vs. their m/z ratios [31,32]. Due to the aforementioned highly accurate and unbiased proteomic analysis through MS, a recent typical proteomic workflow is a mass spectrometry-based one.

**THE HISTORY OF PROTEOMIC APPLICATION IN CYTOLOGY**

Since the 2000s, numerous studies have utilized high-throughput proteomics in cytology, most of which have been conducted on breast and thyroid specimens (Table 1). In the early days, the two-dimensional gel electrophoresis (2D-GE) was being used for proteomics analysis [33,34], yet this lacked the reproducibility and accuracy of the newer proteomic applications [18]. In this technique, the proteins are initially separated based on their charge and molecular weight with gel electrophoresis. Subsequently, the areas containing the target proteins are excised from the gel and then identified with MS [35]. Through the matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS), the cytologic samples are mixed with the substrates, followed by their crystallization within the matrix on a metal plate. Then, the laser energy is absorbed in the matrix generating analyte ions, which are then accelerated into a mass spectrometer [36,37]. In the surface-enhanced laser desorption/ionization time-of-flight mass spectrometry, which is considered as an extended technique of the MALDI-TOF-MS method, the ionized proteins can be directly identified in an electric field by mass spectrometry, without involving protein separation on a 2D gel [38,39]. Over the last decade, electrospray ionization tandem mass spectrometry analysis has become one of the most advanced analytical proteomics methods [40] and has also been applied in cytologic specimens [41].

Regarding breast cancer, most published cytology-based proteomics studies utilized nipple aspirate fluid (NAF), whereas a smaller number FNA samples (Table 1). A few reported significant proteomic profile differences between the NAF of patients with breast cancer compared to non-malignant controls [39,42-44]. In a breast FNA-based study performed by Franzen et al. [45], expression levels of several immune-related proteins differed between cancer and controls, while a few were associated with estrogen receptor, Ki-67 status, and tumor grading. Of interest, liquid-based cytology samples, stored in the methanol-based Pre-
Proteomic profile differences (40 proteins) were found in the NAF from DCIS and invasive breast cancer (stages 3 and 2) compared to controls. Whereas no proteomic profile differences were found in the NAF from breast cancer (stages 1, 2, and 3 from breast cancer; paired and unilateral samples from 21 patients with breast cancer; 13 controls). The in situ proteomics-based model was able to predict the classification derived from the FNA morphologic evaluation and malignant lesions, identify PTC, also distinguish neoplastic and malignant lesions, and correctly triaged indeterminate FNA lesions as either benign or malignant.

Samples from 6 patients with cancer compared to controls, and invasive cancer compared to DCIS. The RPPM technology successfully identified and quantified selected proteins in FNA samples. Vitamin D binding protein precursor was overexpressed in the NAF of patients with early-stage breast cancer compared to controls (Continued to the next page)

Table 1. Studies utilizing high-throughput proteomics on cytology samples received from various organ sites

| Study | Sample type | No. of samples | High-throughput proteomics approach | Key findings |
|-------|-------------|----------------|-------------------------------------|-------------|
| Breast | NAF         | 18 from breast cancer (stages I and II); 4 controls | ICAT LC-MS/MS | Vitamin D binding protein precursor was overexpressed in the NAF of patients with early-stage breast cancer compared to controls |
|        | NAF         | 23 from breast cancer (stages I and II); 5 controls | SELDI-MS | Significant proteomic profile differences were found in the NAF of patients with early-stage breast cancer compared to controls |
|        | NAF         | 27 from breast cancer; 87 controls | SELDI-MS | Proteomic profile differences were found in the NAF of patients with DCIS compared to controls, and invasive cancer compared to DCIS |
|        | NAF         | 52 from DCIS and invasive cancer; 53 controls | 2D PAGE and MALDI-MS | GCDFP-15 was significantly underexpressed and AAG overexpressed in the breast cancer samples tested |
|        | NAF         | 20 from breast cancer; 13 controls | SELDI-MS | Proteomic profile differences (5 proteins) were found in the NAF of patients with cancer compared to controls |
|        | NAF         | 9 from breast cancer; 4 controls | LC-MS/MS | Proteomic profile differences (40 proteins) were found in the NAF of patients with cancer compared to controls |
|        | NAF         | 3 from breast cancer; 3 controls | LC-MS/MS | More than 900 proteins were discovered, as part of the NAF proteome |
|        | NAF         | Paired samples from 21 patients with breast cancer; paired and unilateral samples from 44 controls | SELDI-MS | Whereas no proteomic profile differences were found in the NAF received from the breast with cancer compared to the contralateral healthy one, significant differences were identified between women with cancer (in both cancerous and healthy breasts) and healthy controls |
|        | FNA         | 24 (benign and malignant lesions) | SELDI-MS | Liquid-based cytology samples, stored in the methanol-based PresenCyt, were suitable for satisfactory and reproducible proteomic analysis |
|        | FNA         | 25 from breast cancer, 32 controls | PEA | Expression levels of several immune-related proteins differed between cancer and controls, while a few were associated with ER, Ki-67 status, and tumor grading |
|        | FNA         | 63 (50 with cancer) from 21 patients | RPPM | The RPPM technology successfully identified and quantified selected proteins in FNA samples |
|        | FNA         | Samples from 6 patients (3 non-neoplastic, 1 Hurthle cell adenoma, 1 PTC, 1 MTC) | MALDI-MS | Proteomic profile differences were identified between diverse thyroid lesions sampled with FNA |
|        | FNA         | Samples from 7 patients (non-neoplastic and neoplastic) | MALDI-MS | In situ proteomic analysis could differentiate between non-neoplastic and malignant lesions, identify PTC, also distinguish PTC cases carrying the BRAF V600E mutation |
|        | FNA         | Samples from 43 patients (non-neoplastic and neoplastic; training and validation cohorts) | MALDI-MS | In situ proteomic analysis distinguished Hashimoto thyroiditis from hyperplastic nodules and PTCs |
|        | FNA         | 36 (13 benign, 10 indeterminate, 13 PTCs) | MALDI-MS | In situ proteomic analysis distinguished benign thyroid lesions from PTCs and correctly triaged indeterminate FNA lesions as either benign or malignant |
|        | FNA         | 17 suspicious and malignant thyroid lesions | 2D-GE and MALDI-MS | Several proteins were identified, involved in various cell processes (e.g., metabolism, apoptosis, motility) |
|        | FNA         | 13 PTCs | 2D-GE and MALDI-MS | 17 proteins were overexpressed in thyroid cancer patients compared to controls; proteomic profile differences were also identified between classic and tall cell PTC variants |
|        | FNA         | 240 (internal and external validation cohorts) | MALDI-MS | Whereas the diagnostic accuracy of the in situ proteomics-based classification model was inferior in the external than internal validation cohort, this was improved when sample cellularity was adequate |
|        | FNA         | 212 (benign, intermediate, suspicious for malignancy, and malignant) | 2D-GE and LC-ESI-MS/MS | Proteomic profile differences (25 proteins) were found between benign and malignant lesions; ROC curve analysis showed the combination of ENO1, ANXA1, DJ1, SOD, CRNN protein levels had the best discriminatory capacity |
|        | FNA         | 18 (12 PTCs, 6 benign) | LC-MS/MS | Several actin cytoskeleton proteins (e.g., Arp 2/3 complex overexpression) were altered in PTC; IQGAP1 was upregulated in CV-PTC, while IQGAP2 in FV-PTC, at significant levels, respectively |
|        | FNA         | 28 (benign, intermediate, and malignant; training and validation cohorts) | MALDI-MS | The in situ proteomics-based model was able to predict the classification derived from the FNA morphologic evaluation of the thyroid lesions |
|        | FNA         | 120 PTMCs (60 with LN metastasis, and 60 without) | TMT and LC-MS/MS | ISG15 levels distinguished PTMC patients developing LN metastasis from the ones that did not |
One patient with serous
Proteomic analysis of LBC samples revealed moesin as a biomarker
1D PAGE and
Pleural
No. of samples
LC-MS/MS
655 (non-malignant and
54 cancer, and 46 controls
Proteomic analysis revealed AHNAK as a biomarker differentiating
LBC is suitable for high-throughput proteomic analysis to identify
In situ proteomics analysis was able to correctly assign most lesions
among their original cytologic classification group.

75

Bone marrow
Chen et al. (2021) [41]
Bone marrow aspirate
5 RRMM, 5 NDDM
TMT-MS/MS
Overexpression of the biomarker SERPINB9 was found in RRMM,
compared to NDDM

Table 1. Continued

| Study | Sample type | No. of samples | High-throughput proteomics approach | Key findings |
|-------|-------------|----------------|------------------------------------|-------------|
| Park et al. (2023) [70] | Urine (LBC cytology) | 16 (6 NIBUC, 5 SIBUC, and 5 MIBUC) | LC-MS/MS | Proteomic analysis of LBC samples revealed moesin as a biomarker predicting bladder urothelial cancer invasion |
| Yang et al. (2011) [71] | Urine | 54 cancer, and 46 controls | LC-MS/MS | Overexpression of A1AT was associated with the presence of bladder urothelial cancer, at a significant level |
| Theodorescu et al. (2006) [72] | Urine | 655 (non-malignant and malignant) | CE-MS | The model predicted the presence of urothelial cancer in urine samples with high diagnostic accuracy |
| Lee et al. (2018) [73] | Urine (LBC cytology) | 20 (10 bladder cancer; 10 controls) | LC-MS/MS | Proteomic analysis revealed AHNAK as a biomarker differentiating bladder cancer from controls in LBC cytology samples |
| Schwamborn et al. (2011) [54] | Pap test | 32 (18 with LSIL or higher; 14 NILM) | MALDI-MS | In situ proteomics analysis was able to correctly assign most lesions into their original cytologic classification group |
| Boylan et al. (2014) [74] | Pap test | 100, all with normal cytology | 1D PAGE and LC-MS/MS | The core proteome of normal Pap test, comprising 153 proteins, was created by proteomics analysis of residual LBC samples |
| Boylan et al. (2021) [75] | Pap test | One patient with serious ovarian cancer | LC-MS/MS | LBC is suitable for high-throughput proteomic analysis to identify ovarian cancer biomarkers |
| Schwamborn et al. (2019) [55] | Pleural effusion | 24 with serious ovarian cancer, 19 with non-ovarian cancers | MALDI-MS | In situ proteomics analysis was able to differentiate among diverse cancer types in effusions |
| Perzanowska et al. (2018) [66] | Pleural effusion | 69 malignant, 49 benign (controls) | LC/MMR-MS | Multiplex proteomic analysis was able to differentiate between benign and malignant effusions, besides among lung cancer histologic subtypes (SCC, AC, SqCC) |
| Li et al. (2016) [57] | Pleural effusion | 83 malignant (lung ACs), 60 benign (training and validation cohorts) | MALDI-MS | The model was able to differentiate between benign and malignant effusions with high diagnostic accuracy; CARD9 was downregulated in malignant effusions |
| Liu et al. (2015) [76] | Pleural effusion | 405 malignant and benign effusions (discovery and validation cohorts) | 1D-PAGE and LC-MS/MS | Overexpression of MET, DPP4, and PTPR6 identified metastatic lung adenocarcinomas in effusion samples with high diagnostic accuracy |
| Li et al. (2015) [77] | Pleural effusion | 6 (3 NSOCL, 3 TB) | 1D-PAGE and LC-MS/MS | Proteomic analysis was able to differentiate NSCLC from TB effusions; IL1A was overexpressed in NSCLC compared to TB effusions |
| Hegmans et al. (2009) [78] | Pleural effusion | 89 (mesothelioma, metastatic carcinoma, benign effusions) | SELDI-MS | SMRP was identified as a diagnostic biomarker of mesothelioma in pleural effusions |
| Inoue et al. (2022) [58] | EUS-FNA | 40 PDAC, 6 AIP | LC-MS/MS | Expression of several EV proteins differed between PDAC and AIP patients |
| Lee et al. (2012) [59] | EUS-FNA | 5 BD-IPMNs, 5 inflammatory cysts | Cytokine microarray | HGF and GM-CSF differentiated inflammatory cysts from BD-IPMNs |
| Navaneethan et al. (2015) [60] | Bile | 24 (PDAC, CCA, PSC, other non-neoplastic) | SDS-PAGE and LC-MS/MS | Expression of several proteins differed between malignant and non-neoplastic biliary strictures |
| Seccia et al. (2020) [61] | FNA | 20 MSGTs, 37 PAs, 14 WTs | 2D-GE and LC-ESI-MS/MS | Overexpression of 4 proteins (annexin-5, collagen-1, peptidyl-prolyl-cis-trans-isomerase-A, and F-actin-capping-alpha-1) differentiated MSGTs from benign aspirates |
| Bone marrow | Chen et al. (2021) [41] | Bone marrow aspirate | 5 RRMM, 5 NDDM | TMT-MS/MS | Overexpression of the biomarker SERPINB9 was found in RRMM, compared to NDDM |

NAF, nasal aspirate fluid; ICAT, isotope-coded affinity tag; LC-MS/MS, liquid chromatography-tandem mass spectrometry; SELDI-MS, surface-enhanced laser desorption/ionization-mass spectrometry; DCIS, ductal carcinoma in situ; 2D PAGE, two-dimensional polyacrylamide gel electrophoresis; MALDI, matrix-assisted laser desorption/ionization; GCDFP, gross cystic disease fluid protein; AAG, alpha-1-acid glycoprotein; FNA, fine-needle aspiration; PEA, proximity extension assay; ER, estrogen receptor; RPMP, reverse-phase protein microarrays; PTC, papillary thyroid carcinoma; MTC, medullary thyroid carcinoma; MSI, microsatellite instability; 2D-GE, two-dimensional gel electrophoresis; LC-ESI-MS/MS, liquid chromatography electrospray ionization tandem mass spectrometry; ROC, receiver operating characteristic; ENO1, enolase 1; ANXAI, annexin A1; DJ1, protein DJ-1; SOD, superoxide dismutase; OV-PTC, classic variant PTC; RV-PTC, follicular variant PTC; PTMC, papillary thyroid microcarcinoma; LN, lymph node; TMT, tandem mass tags; ISG15, interferon-stimulated gene 15 protein; LBC, liquid-based cytology; NIBUC, non-invasive bladder urothelial carcinoma; SIBUC, stromal-invasive bladder urothelial carcinoma; MIBUC, muscle-invasive bladder urothelial carcinoma; A1AT, alpha 1 antitrypsin; CE-MS, capillary electrophoresis coupled to mass spectrometry; LSIL, low-grade squamous intraepithelial lesion; NILM, normal for intraepithelial lesion or malignancy; PAP, Papaincloula; MRI, magnetic resonance imaging; SOC, small cell carcinoma; AC, adenocarcinoma; SquCC, squamous cell carcinoma; CAR9, caspase recruitment domain family member 9; DPP4, dipeptidyl peptidase-4; PTPR6, protein tyrosine phosphatase receptor type F; NSCLC, non-small cell lung cancer; TB, tuberculosis; SMRP, soluble mesothelin-related protein; EUS-FNA, endoscopic ultrasound-guided fine-needle aspiration; PDAC, pancreatic adenocarcinoma; AIP, autoimmune pancreatitis; EV, extracellular vesicles; BD-IPMNs, branch duct intraductal papillary mucinous neoplasms; HGF, hepatocyte growth factor; GM-CSF, granulocyte-macrophage colony-stimulating factor; CCA, cholangiocarcinoma; PSC, primary sclerosing cholangitis; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; MSGTs, malignant salivary gland tumors; PAs, pleomorphic adenomas; WTs, Warthin tumors; RRMM, recurrent and relapsed multiple myeloma; NDDM, newly diagnosed multiple myeloma; TMT-MS, tandem mass tag-mass spectrometry.
Proteomics in cytology

To complement the morphologic evaluation of FNA in the evaluation of thyroid lesions, especially the ones with indeterminate interpretations, a few studies utilized in situ proteomics, more specifically the MALDI–mass spectrometry imaging (MSI) technique [48-53]. For instance, MALDI-MSI distinguished benign thyroid lesions from papillary thyroid carcinomas (PTCs) and correctly triaged indeterminate FNA lesions as either benign or malignant [51], while it also distinguished Hashimoto thyroiditis from hyperplastic nodules and PTC in another study [50]. Notably, except differentiating between non-neoplastic lesions from PTC, MALDI-MSI was also able to identify PTC cases carrying the \( \text{BRAF} \) V600E mutation [49]. Furthermore, Schwamborn et al. applied MALDI-MSI aiming to facilitate Papanicolaou (Pap) test and serous effusion cytologic diagnoses; in situ proteomics was able to correctly assign most lesions into their original cervical cytology classification group and differentiate among diverse cancer types in serous effusions, respectively [54,55].

Apart from breast and thyroid cytology, high-throughput proteomics have additionally been applied in urine cytology, Pap tests, serous effusions, pancreatobiliary samples, salivary FNAs, and bone marrow aspirates (Table 1) with the goal to either improve morphologic diagnosis or identify novel cancer biomarkers. Diagnostic dilemmas in cytology subjected to proteomic analysis have been the differentiation between benign and malignant serous effusions [56,57], pancreatic cancer from autoimmune pancreatitis in FNAs of solid pancreatic lesions [58], inflammatory pancreatic cysts from branch duct intraductal papillary mucinous neoplasms while evaluating cystic pancreatic lesions (BD-IPMNs) [59], non-neoplastic from malignant biliary strictures [60], and benign from malignant salivary gland FNAs [61].

**BIOMARKERS DISCOVERED USING CYTOLOGY SPECIMENS THROUGH HIGH-THROUGHPUT PROTEOMICS**

Fig. 1 gives a general proteomic workflow used to discover a successful cancer biomarker with cytologic specimens. With the recent advances of MS-based proteomics, even small protein amounts are detectable, while the discovery of biomarker candidates via proteomics has been presented in several studies using cytologic material (Table 2).

Regarding breast cancer, NAF has mainly been used to identify potential breast cancer biomarkers, besides suggesting several proteomic profiles that might have value in assessing the risk of breast cancer (Tables 1, 2). Alexander et al. [33] identified 41 different proteins through 2D-GE and MALDI-MS and suggested two candidate biomarkers, gross cystic disease fluid protein (GCDFP)-15 and alpha1-acid glycoprotein (AAG), testing 52 NAFs from breast cancer patients (in situ and invasive) and 53 controls. GCDFP-15 was found significantly underexpressed, whereas AAG overexpressed in the breast cancer samples [33]. In another study, Pawlik et al. [62] reported that vitamin D binding...
Table 2. Examples of novel cancer biomarkers discovered by utilizing high-throughput proteomics on cytology samples

| Study            | Cancer type/sample type | Novel biomarker(s)                                      | Expression status in cancer                        |
|------------------|-------------------------|--------------------------------------------------------|-----------------------------------------------------|
| Pawlik et al.    | Breast/NAF              | Vitamin D-binding protein precursor                    | Vitamin D-binding protein precursor:                |
|                  |                         |                                                        | ↓ in breast cancer                                   |
|                  |                         |                                                       | ↓ in breast cancer                                   |
|                  |                         |                                                       | ↓ in breast cancer                                   |
| Alexander et al. | Breast/NAF              | GCDFP-15, AAG                                         | GCDFP-15: ↓ in breast cancer                         |
|                  |                         |                                                        | GCDFP-15: ↓ in breast cancer                         |
| Ciregia et al.   | Thyroid/Thyroid FNA, serum, saliva | ANXA1, AAG                                           | ANXA1: ↑ in thyroid cancer                           |
|                  |                         |                                                        | ANXA1: ↑ in thyroid cancer                           |
| Ucal et al.      | Thyroid/FNA             | IQGAP1, IQGAP2                                        | IQGAP1: ↑ in CV-PTC                                   |
|                  |                         |                                                        | IQGAP2: ↑ in FV-PTC                                   |
| Lin et al.       | Thyroid/Thyroid FNA     | ISG15                                                  | ISG15: ↑ in PTMC patients with metastasis to cervical lymph nodes (prognostic biomarker) |
| Giusti et al.    | Thyroid/FNA             | TTR, FLC, proteasome activator complex subunit 1 and 2, alpha-1-antitrypsin precursor, GAPDH, LDH-B, Apo-A1, annexin A1, DJ-1 protein and coffin-1 | TTR, FLC, proteasome activator complex subunit 1 and 2, alpha-1-antitrypsin precursor, GAPDH, LDH-B, Apo-A1, annexin A1, DJ-1 protein and coffin-1: ↑ in PTC |
| Park et al.      | Bladder/Urine           | Meso-1                                                | Meso-1: ↑ in invasive bladder cancer                 |
| Yang et al.      | Bladder/Urine           | A1AT                                                  | A1AT: ↑ in bladder cancer                            |
| Lee et al.       | Bladder/Urine           | AHNAK                                                 | AHNAK: ↑ in bladder cancer                           |
| Li et al.        | Lung/Effusions          | CAPD9                                                 | CAPD9: ↓ in malignant effusions                      |
| Liu et al.       | Lung/Effusions          | MET, DPP4, and PTTPR                                   | MET, DPP4, and PTTPR: ↑ in malignant effusions      |
| Li et al.        | Lung/Effusions          | IL1A                                                  | IL1A: ↑ in malignant effusions                       |
| Hiegmans et al.  | Mesotheioma/Effusions   | SMRP                                                  | SMRP: ↑ in mesothelioma                              |
| Seccia et al.    | MGMTs/FNA               | Annexin-5, coffin-1, peptidyl-prolyl-cis-trans-isomerase-A and F-actin-capping-alpha-1 | Annexin-5, coffin-1, peptidyl-prolyl-cis-trans-isomerase-A and F-actin-capping-alpha-1: ↑ in MGMTs |
| Chen et al.      | MM/Bone marrow aspirate | SERPIN9                                               | SERPIN9: ↑ in RRMM (prognostic biomarker)            |

NAF, nipple aspirate fluid; GCDFP, gross cystic disease fluid protein; AAG, alpha-1-acid glycoprotein; FNA, fine-needle aspiration; ANXA1, annexin A1; IQGAP1, IQ motif containing GTPase activating protein 1; CV-PTC, classic variant PTC; FV-PTC, follicular variant PTC; ISG15, interferon-stimulated gene 15 protein; PTMC, papillary thyroid microcarcinoma; TTR, transthyretin; FLC, feritin light chain; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; LDH-B, lactate dehydrogenase chain B; Apo-A1, apolipoprotein A1 precursor; A1AT, alpha 1 antitrypsin; CAPD9, caspase recruitment domain family member 9; DPP4, dipeptidyl peptidase-4; PTTPR, protein tyrosine phosphatase receptor type F; IL1A, interleukin 1A; SMRP, soluble mesothelin-related protein; MGMTs, malignant salivary gland tumors; MM, multiple myelomas; RRMM, recurrent and relapsed multiple myelomas.
reporting the proteome of normal cervical cytology, which was composed of 153 proteins. Regarding serous effusions, caspase recruitment domain family member 9 was found downregulated in malignant effusions [57]. Overexpression of MET, dipeptidyl peptidase-4, and protein tyrosine phosphatase receptor type F identified metastatic lung adenocarcinomas [76], interleukin 1A was overexpressed in non–small cell lung cancer compared to tuberculosis effusions [77], and serum soluble mesothelin-related protein was identified as a diagnostic biomarker of mesothelioma in pleural effusions [78]. Notably, hepatocyte growth factor and granulocyte-macrophage colony-stimulating factor differentiated inflammatory cysts from BD-IPMNs [59], whereas the overexpression of four proteins (annexin-5, cofilin-1, peptidyl-prolyl cis–trans-isomerase-A, and F-actin-capping-alpha-1) differentiated malignant from benign salivary gland FNAs [61].

In two recent studies, our group applied MS-based proteomics on liquid-based urine cytology samples obtained from urothelial carcinoma patients, and reported potential diagnostic and predictive biomarkers through several validation test layers. The latter included cross validation with TCGA, tumor cell lines with gene editing techniques, and immunocytochemistry in independent patient cohorts [70,73]. Lee et al. [73] selected 112 differentially expressed proteins altered in urothelial carcinoma and validated neuroblast differentiation-associated protein AHNAK (AHNAK) as a new cancer biomarker, able to differentiate between urothelial carcinoma and benign urothelial cytology. TCGA also identified AHNAK as a candidate biomarker along with EPPK1, MYH14, and OLFM4. Furthermore, Park et al. [70] found moesin (MSN) as a potential biomarker predicting the presence of invasive urothelial carcinoma in urine cytology. Of interest, MSN knockdown using siRNA led to inhibition of tumor invasion in urothelial carcinoma cell lines. Also, immunocytochemistry consistently showed that MSN is a crucial biomarker predicting invasion when applied in urine cytology [70].

**PERSPECTIVES**

High-throughput proteomic applications have recently advanced, enabling the use of minimal patient-derived specimens and overcoming the issue of low depth, inconsistency, and suboptimal accuracy. These technical advances are applicable to cytology samples, especially the ones processed with liquid-based cytology, providing reproducible results and revealing a few candidate biomarkers of diagnostic, prognostic, and therapeutic value (Table 2). Most published studies have utilized breast and thyroid cytology samples, showing the potential to help pathologists solve various diagnostic dilemmas and avoid common pitfalls. Such dilemmas comprise the evaluation of indeterminate thyroid nodules while examining thyroid FNAs, the detection of malignant serous effusions, also the differential diagnosis of a few entries in the challenging field of pancreatobiliary cytology, including pancreatic cancer from autoimmune pancreatitis, non-neoplastic from neoplastic pancreatic cysts, and non-neoplastic from malignant biliary strictures. Proteomic profiling of NAF breast samples may identify early-stage breast cancers, also differentiate between in situ and invasive breast cancers and provide information related to prognosis and therapy. Notably, according to the literature, in situ proteomics has exhibited the capacity to triage indeterminate thyroid FNAs thus prevent unnecessary surgeries and reduce healthcare costs, besides provide prognostic information through identifying PTGs carrying the BRAF V600E mutation and predicting the presence of lymph node metastasis or PTC histology associated with a more aggressive behavior (e.g., the tall cell variant) (Table 1). Indeed, proteomic profiling could complement traditional morphologic evaluation and ancillary testing used to examine various exfoliative and FNA cytopathology samples in routine practice or even constitute a stand-alone diagnostic modality in specific settings. However, evidence is still primitive, mostly resulting from studies with small sample size. Apart from the shortage of high-quality evidence, the demands of highly-skilled laboratory personnel, also the cost of analytic equipment, have prohibited the routine application of such approaches and limited them in the research setting. To implement high-throughput proteomics into everyday clinical practice, well-designed prospective studies and randomized controlled trials involving large patient cohorts should be used, aiming to evaluate the proteomics benefits and limitations compared to already established cytomorphologic and ancillary approaches, also their potential implementation in diagnostic algorithms used in cytopathology. Most importantly, cytopathologists and researchers should validate these methods in different sample preparations, and assess their clinical utility in diverse diagnostic scenarios. In conclusion, proteomics could become another diagnostic platform—along with genomics, transcriptomics and/or metabolomics—in the near future, potentially by using validated multi-omics approaches.

**Ethics Statement**

Not applicable.

**Availability of Data and Material**

Data sharing not applicable to this article as no datasets were generated or analyzed during the study.
Code Availability
Not applicable.

ORCID
Ilias P. Nikas https://orcid.org/0000-0001-3825-2556
Han Suk Ryu https://orcid.org/0000-0003-2308-9796

Author Contributions
Conceptualization: HSR. Project administration: HSR. Supervision: HSR. Writing—original draft: IPN, HSR. Writing—review & editing: IPN, HSR. Approval of final manuscript: all authors.

Conflicts of Interest
The authors declare that they have no potential conflicts of interest.

Funding Statement
No funding to declare.

References
1. Mardis ER. Next-generation sequencing platforms. Annu Rev Anal Chem (Palo Alto Calif) 2013; 6: 287-303.
2. Hou YC, Neidich JA, Duncavage EJ, Spencer DH, Schroeder MC. Clinical whole-genome sequencing in cancer diagnosis. Hum Mutat 2022; 43: 1519-30.
3. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: an overview. Hum Immunol 2021; 82: 80-111.
4. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015; 372: 793-5.
5. Esagian SM, Grigoriadou G, Nikas IP, et al. Comparison of liquid-based to tissue-based biopsy analysis by targeted next generation sequencing in advanced non-small cell lung cancer: a comprehensive systematic review. J Cancer Res Clin Oncol 2020; 146: 2051-66.
6. Grigoriadou G, Esagian SM, Ryu HS, Nikas IP. Molecular profiling of malignant pleural effusions with next generation sequencing (NGS): evidence that supports its role in cancer management. J Pers Med 2020; 10: 206.
7. Nikas IP, Mountzios G, Sydnee GI, Ioakim KJ, Won JK, Papageorisis P. Evaluating pancreatic and biliary neoplasms with small biopsy-based next generation sequencing (NGS): doing more with less. Cancers (Basel) 2022; 14: 397.
8. Tsamis KI, Sakkas H, Giannakis A, Ryu HS, Gartzonika C, Nikas IP. Evaluating infectious, neoplastic, immunological, and degenerative diseases of the central nervous system with cerebrospinal fluid-based next-generation sequencing. Mol Diagn Ther 2021; 25: 207-29.
9. Roy-Chowdhuri S, Pisapia P, Salto-Tellez M, et al. Invited review—next-generation sequencing: a modern tool in cytopathology. Virchows Arch 2019; 475: 3-11.
10. Liu Y, Beyer A, Aebersold R. On the dependency of cellular protein levels on mRNA abundance. Cell 2016; 165: 535-50.
11. Nagaraj N, Wisniewski JR, Geiger T, et al. Deep proteome and transcriptome mapping of a human cancer cell line. Mol Syst Biol 2011; 7: 548.
12. Kageyama S, Isono T, Iwaki H, et al. Proteome research in urothelial carcinoma. Int J Urol 2015; 22: 621-8.
13. Banks RE, Dunn MJ, Hochstrasser DF, et al. Proteomics: new perspectives, new biomedical opportunities. Lancet 2000; 356: 1749-56.
14. Ramazi S, Allahverdi A, Zahiri J. Evaluation of post-translational modifications in histone proteins: a review on histone modification defects in developmental and neurological disorders. J Biosci 2020; 45: 135.
15. Doll S, Gnadt F, Mann M. The case for proteomics and phosphoproteomics in personalized cancer medicine. Proteomics Clin Appl 2019; 13: e1800113.
16. Menyhart O, Gyoryffy B. Multi-omics approaches in cancer research with applications in tumor subtyping, prognosis, and diagnosis. Comput Struct Biotechnol J 2021; 19: 949-60.
17. Singla D, Sangha MK. Multi-omic approaches to improve cancer diagnosis, prognosis, and therapeutics. In: Raza K, ed. Computational intelligence in oncology: applications in diagnosis, prognosis and therapeutics of cancers. Singapore: Springer Singapore, 2022: 411-33.
18. Chandramouli K, Qian P Y. Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. Hum Genomics Proteomics 2009; 2009: 239204.
19. Aebersold R, Mann M. Mass-spectrometric exploration of proteome structure and function. Nature 2016; 537: 347-55.
20. Choudhary C, Weinert BT, Nishida Y, Verdin E, Mann M. The growing landscape of lysine acetylation links metabolism and cell signaling. Nat Rev Mol Cell Biol 2014; 15: 536-50.
21. Bekker-Jensen DB, Kelstrup CD, Barth TS, et al. An optimized shotgun strategy for the rapid generation of comprehensive human proteomes. Cell Syst 2017; 4: 587-99.
22. Zhu Y, Piehowski PD, Zhao R, et al. Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. Nat Commun 2018; 9: 882.
23. Altealaf AE, Heck AJ. Trends in ultrasensitive proteomics. Curr Opin Chem Biol 2012; 16: 206-13.
24. Anderson NL, Anderson NG. Proteome and proteomics: new technologies, new concepts, and new words. Electrophoresis 1998; 19: 1853-61.
25. Blackstock WP, Weir MP. Proteomics: quantitative and physical mapping of cellular proteins. Trends Biotechnol 1999; 17: 121-7.
26. Melton L. Protein arrays: proteomics in multiplex. Nature 2004; 429: 101-7.
27. Mueller LN, Rinner O, Schmidt A, et al. SuperHirn—a novel tool for high resolution LC-MS-based peptide/protein profiling. Proteomics 2007; 7: 3470-80.
28. Ghose A, Gullapalli SV, Chohan N, et al. Applications of proteomics in ovarian cancer: dawn of a new era. Proteomes 2022; 10: 16.
29. Hall DA, Ptacek J, Snyder M. Protein microarray technology. Mech Ageing Dev 2007; 128: 161-7.
30. Seekumar A, Nyati MK, Varambally S, et al. Profiling of cancer cells age 2007; 128: 161-7.
31. Domon B, Aebersold R. Mass spectrometry and protein analysis. Science 2006; 312: 212-7.
32. Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem 2001; 70: 437-73.
33. Alexander H, Stengner AL, Wagner-Mann C, Du Bois GC, Alexander S, Sauter ER. Proteomic analysis to identify breast cancer biomarkers in nipple aspirate fluid. Clin Cancer Res 2004; 10: 7500-10.
34. Ghose A, Gullapalli SV, Chohan N, et al. Profiling of cancer cells using protein microarrays: discovery of novel radiation-regulated proteins. Cancer Res 2001; 61: 7585-93.
35. Krause K, Jessneriter B, Fuhrer D. Proteomics in thyroid tumor re-
search. J Clin Endocrinol Metab 2009; 94: 2717-24.
36. Duncan M, DeMarco ML. MALDI-MS: Emerging roles in pathology and laboratory medicine. Clin Mass Spectrom 2019; 13: 1-4.
37. Kriegsmann J, Kriegsmann M, Casadonte R. MALDI TOF imaging mass spectrometry in clinical pathology: a valuable tool for cancer diagnostics (review). Int J Oncol 2015; 46: 893-906.
38. Wright GI, Jr. SELDI protechip® Ms: a platform for biomarker discovery and cancer diagnosis. Expert Rev Mol Diagn 2002; 2: 549-63.
39. Sauter ER, Shan S, Hewett JE, Speckman P, Du Bois GC. Proteomic analysis of nipple aspirate fluid using SELDI-TOF-MS. Int J Cancer 2005; 114: 791-6.
40. Leito I, Herodes K, Huopolainen M, et al. Towards the electrospary ionization mass spectrometry ionization efficiency scale of organic compounds. Rapid Commun Mass Spectrom 2008; 22: 379-84.
41. Chen Y, Quan L, Jia C, et al. Proteomics-based approach reveals the involvement of SERPINB9 in recurrent and relapsed multiple myeloma. J Proteome Res 2021; 20: 2673-86.
42. Pawlik TM, Fritsche H, Coombs KR, et al. Significant differences in nipple aspirate fluid protein expression between healthy women and those with breast cancer demonstrated by time-of-flight mass spectrometry. Breast Cancer Res Treat 2005; 89: 149-57.
43. Sauter ER, Zhu W, Fan XJ, Wassell RP, Chervoneva I, Du Bois GC. Proteomic analysis of nipple aspirate fluid to detect biologic markers of breast cancer. Br J Cancer 2002; 86: 1440-3.
44. George AL, Shaheed SU, Sutton CW. High-throughput proteomic profiling of nipple aspirate fluid from breast cancer patients compared with non-cancer controls: a step closer to clinical feasibility. J Clin Med 2021; 10: 2243.
45. Franzen B, Alexeyenko A, Kamali-Moghaddam M, et al. Protein profiling of fine-needle aspirates reveals subtype-associated immune signatures and involvement of chemokines in breast cancer. Mol Oncol 2019; 13: 376-91.
46. Fowler LJ, Lovell MO, Iacconi P, Ciregia F, et al. Fine-needle aspiration of thyroid lesions: preliminary report. Cytotherapy 2015; 26: 318-24.
47. Pavlou MP, Kulasingam V, Sauter ER, Klie thermes B, Diamandis EP. Nipple aspirate fluid proteome of healthy females and patients with breast cancer. Clin Chem 2010; 56: 848-55.
48. Noble JL, Dua RS, Coulton GR, Isacke CM, Gui GP. A comparative proteomic analysis of nipple aspirate fluid from women with early-stage breast cancer using isotope-coded affinity tags and tandem mass spectrometry reveals differential expression of vitamin D binding protein. BMC Cancer 2006; 6: 68.
49. Li H, Yang N, Feng S, Shedden K, et al. Urinary glycoprotein biomarker of breast cancer. J Proteome Res 2021; 20: 2673-86.
50. Schwamborn K, Krieg RC, Uhlig S, Ikenberg H, Wellmann A. MALDI imaging as a specific diagnostic tool for routine cervical cytology specimens. Int J Mol Med 2011; 27: 417-21.
51. Schwamborn K, Weirich G, Steiger K, et al. Discerning the primary carcinoma in malignant peritoneal and pleural effusions using imaging mass spectrometry: a feasibility study. Proteomics Clin Appl 2019; 13: e1800064.
52. Perzanowska A, Fatalaska A, Wojtas G, et al. An MRM-based cyto-keratin marker assay as a tool for cancer studies: application to lung cancer pleural effusions. Proteomics Clin Appl 2018; 12: 1700884.
53. Li H, Yang N, Feng S, Cui S, Jiang W. Proteomic study of benign and malignant pleural effusion. J Cancer Res Clin Oncol 2016; 142: 1191-200.
54. Inoue H, Eguchi A, Kobayashi Y, et al. Extracellular vesicles from pancreatic ductal adenocarcinoma endoscopic ultrasound-fine needle aspiration samples contain a protein barcode. J Hepatobiliary Pancreat Sci 2022; 29: 394-403.
55. Lee LS, Banks PA, Bellizzi AM, et al. Inflammatory protein profiling of pancreatic cyst fluid using EUS-FNA in tandem with cytokine microarray differentiates between branch duct IPMN and inflammatory cysts. J Immunol Methods 2012; 382: 142-9.
56. Navaneethan U, Lourdusamy V, Gk Venkatesh P, Willard B, Sanaka MR, Parsi MA. Bile proteomics for differentiation of malignant from benign biliary strictures: a pilot study. Gastroenterol Rep (Oxf) 2015; 3: 136-43.
57. Seccia V, Navari E, Donadie E, et al. Proteomic investigation of malignant major salivary gland tumors. Head Neck Pathol 2020; 14: 362-73.
58. Rapkiewicz A, Espina V, Zujewski JA, et al. The needle in the haystack: application of breast fine-needle aspirate samples to quantitative protein microarray technology. Cancer 2007; 111: 173-84.
59. Noble JL, Dua RS, Coulton GR, Isacke CM, Gui GP. A comparative proteomic analysis of nipple aspirate fluid from women with early-stage breast cancer using isotope-coded affinity tags and tandem mass spectrometry reveals differential expression of vitamin D binding protein. BMC Cancer 2006; 6: 68.
60. Pavlou MP, Kulasingam V, Sauter ER, Klie thermes B, Diamandis EP. Nipple aspirate fluid proteome of healthy females and patients with breast cancer. Clin Chem 2010; 56: 848-55.
61. Noble JL, Dua RS, Coulton GR, Isacke CM, Gui GP. A comparative proteomic analysis of nipple aspirate fluid from healthy women and women with breast cancer. Eur J Cancer 2007; 43: 2315-20.
62. Giusti L, Iacconi P, Ciregia F, et al. Proteomic analysis of human thyroid fine needle aspiration fluid. J Endocrinol Invest 2007; 30: 865-9.
63. Giusti L, Iacconi P, Ciregia F, et al. Fine-needle aspiration of thyroid nodules: proteomic analysis to identify cancer biomarkers. J Proteome Res 2008; 7: 4079-88.
64. Ciregia F, Giusti L, Molinaro A, et al. Proteomic analysis of fine-needle aspiration in differential diagnosis of thyroid nodules. Transl Res 2016; 176: 81-94.
65. Ucal Y, Ercaci M, Tokat F, Duren M, Ince U, Ozpinar A. Proteomic analysis reveals differential protein expression in variants of papillary thyroid carcinoma. EurPA Open Proteom 2017; 17: 1-6.
66. Liu P, Yao Z, Sun Y, et al. Deciphering novel biomarkers of lymph node metastasis of thyroid papillary microcarcinoma using proteomic analysis of ultrasound-guided fine-needle aspiration biopsy samples. J Proteomics 2019; 204: 103414.
67. Park JH, Lee C, Han D, et al. Moesin (MSN) as a novel proteome-based diagnostic marker for early detection of invasive bladder urothelial carcinoma in liquid-based cytology. Cancers (Basel) 2020; 12: 1018.
discovery for bladder cancer detection using LC/MS-MS and label-free quantification. Clin Cancer Res 2011; 17: 3349-59.

72. Theodorescu D, Wittke S, Ross MM, et al. Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis. Lancet Oncol 2006; 7: 230-40.

73. Lee H, Kim K, Woo J, et al. Quantitative proteomic analysis identifies AHNAK (neuroblast differentiation-associated protein AHNAK) as a novel candidate biomarker for bladder urothelial carcinoma diagnosis by liquid-based cytology. Mol Cell Proteomics 2018; 17: 1788-802.

74. Boylan KL, Afuni-Zadeh S, Geller MA, et al. A feasibility study to identify proteins in the residual Pap test fluid of women with normal cytology by mass spectrometry-based proteomics. Clin Proteomics 2014; 11: 30.

75. Boylan KL, Afuni-Zadeh S, Geller MA, Argenta PA, Griffin TJ, Skubitz APN. Evaluation of the potential of Pap test fluid and cervical swabs to serve as clinical diagnostic biospecimens for the detection of ovarian cancer by mass spectrometry-based proteomics. Clin Proteomics 2021; 18: 4.

76. Liu PJ, Chen CD, Wang CL, et al. In-depth proteomic analysis of six types of exudative pleural effusions for nonsmall cell lung cancer biomarker discovery. Mol Cell Proteomics 2015; 14: 917-32.

77. Li Y, Lian H, Jia Q, Wan Y. Proteome screening of pleural effusions identifies IL1A as a diagnostic biomarker for non-small cell lung cancer. Biochem Biophys Res Commun 2015; 457: 177-82.

78. Hegmans JP, Veltman JD, Fung ET, et al. Protein profiling of pleural effusions to identify malignant pleural mesothelioma using SELDI-TOF MS. Technol Cancer Res Treat 2009; 8: 323-32.

79. Ucal Y, Tokat F, Duren M, Ince U, Ozpinar A. Peptide profile differences of noninvasive follicular thyroid neoplasm with papillary-like nuclear features, encapsulated follicular variant, and classical papillary thyroid carcinoma: an application of matrix-assisted laser desorption/ionization mass spectrometry imaging. Thyroid 2019; 29: 1125-37.

80. Torres-Cabala C, Bibbo M, Panizo-Santos A, et al. Proteomic identification of new biomarkers and application in thyroid cytology. Acta Cytol 2006; 50: 518-28.