



# Biomarker identification through spatial proteomics for the characterization of indeterminate thyroid nodules

Giulia Capitoli <sup>1,2</sup> · Antonio Maria Alviano <sup>3</sup> · Nicole Monza <sup>4</sup> · Lisa Pagani <sup>4</sup> · Isabella Piga <sup>4,9</sup> · Davide Paolo Bernasconi <sup>1</sup> · Angela Greco <sup>3</sup> · Davide Leni <sup>5</sup> · Alice Maggioni <sup>3</sup> · Andrea-Valer Gatti <sup>6</sup> · Fausto Maffini <sup>7</sup> · Nicola Fusco <sup>7,8</sup> · Mattia Garancini <sup>6</sup> · Fulvio Magni <sup>4</sup> · Stefania Galimberti <sup>1,2</sup> · Fabio Pagni <sup>3</sup> · Vincenzo L'Imperio <sup>3</sup> · Vanna Denti <sup>4</sup>

Received: 31 May 2025 / Accepted: 30 July 2025  
© The Author(s) 2025

## Abstract

**Purpose** The identification of novel molecular biomarkers may assist in the characterization of indeterminate thyroid nodules, which pose significant diagnostic challenges. Here, we aimed to explore the potential of proteomic analyses to support biomarker discovery in challenging thyroid lesions.

**Methods** Linear Discriminant Analysis (LDA) was applied to Matrix-Assisted Laser Desorption Ionization Mass Spectrometry Imaging (MALDI-MSI) data from 44 thyroid neoplasms to select the most impactful molecular features for the classification of different tumor histologies, as well as for the distinction between NRAS-mutant (mNRAS) and NRAS-wild-type (wtNRAS) tumors. Relevant peaks were subsequently identified through nanoscale liquid chromatography electrospray ionization tandem mass spectrometry (nLC-ESI-MS/MS).

**Results** The LDA selected nine relevant molecular markers distinguishing noninvasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTPs) from other tumor histologies (balanced accuracy = 73%), as well as 19 relevant markers able to identify mNRAS cases (balanced accuracy = 84%). Nine differentially expressed proteins were putatively identified: among them, ATP-dependent RNA helicase DDX42 showed a similar distribution between NIFTPs and papillary thyroid carcinomas (PTCs) / follicular variant PTCs (FVPTCs), while the distribution of the Histone H4 signal was similar between NIFTPs and follicular adenomas (FAs). In addition, Protein disulfide-isomerase A1 and Complement C4-B were overexpressed in wtNRAS compared to mNRAS cases, regardless of histology.

**Conclusion** The LDA-selected features enable to distinguish NIFTPs from morphologically similar lesions and to discriminate between mNRAS and wtNRAS cases. The identified markers might complement genetic analyses and provide insights into the distinct pathogenic drivers behind the development of mNRAS compared to wtNRAS lesions.

---

These authors contributed equally: Giulia Capitoli, Antonio Maria Alviano.

---

These authors jointly supervised this work: Vincenzo L'Imperio, Vanna Denti.

---

✉ Giulia Capitoli  
giulia.capitoli@unimib.it

<sup>1</sup> Bicocca Bioinformatics Biostatistics and Bioimaging Research Centre - B4, Department of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy

<sup>2</sup> Biostatistics and Clinical Epidemiology, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy

<sup>3</sup> Pathology, Fondazione IRCCS San Gerardo dei Tintori and Center of Digital Medicine, Department of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy

<sup>4</sup> Proteomics and Metabolomics Unit, Department of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy

<sup>5</sup> Department of Radiology, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy

<sup>6</sup> Endocrine and Metabolic Surgery Unit, Department of Surgery, IRCCS Fondazione San Gerardo dei Tintori, Monza, Italy

<sup>7</sup> Division of Pathology, European Institute of Oncology IRCCS, Milan, Italy

<sup>8</sup> Department of Oncology & Hemato-Oncology, University of Milan, Milan, Italy

<sup>9</sup> Present address: Independent Researcher, Cagliari, Italy

**Keywords** Thyroid cancer · Fine-needle aspiration · Cytology · Digital pathology · Computational pathology · MALDI-MSI

## Introduction

Given the high incidence of thyroid nodules and the associated risk of malignancy, an accurate characterization of such lesions is of paramount importance [1]. Despite being the gold standard for risk stratification purposes, fine-needle aspiration cytology (FNAC) yields indeterminate results in up to 30% of cases, many of which undergo diagnostic surgery [2]. To prevent this, various ancillary classifications based on radiology have been proposed [3, 4], possibly complemented by genetic analyses, with still perfectible discriminative capabilities for the different follicular-patterned lesions [5]. The recently updated histological terminologies (e.g. Noninvasive Follicular Thyroid Neoplasm with Papillary-like nuclear features, NIFTP) have added additional complexity to the diagnostic workup of borderline thyroid tumors, also considering their marked heterogeneity. In this context, the application of Matrix Assisted Laser Desorption Ionization Mass Spectrometry Imaging (MALDI-MSI) has shown promise in the characterization of the molecular fingerprints of different thyroid lesions [1, 5], demonstrating encouraging preliminary results even in the identification of NIFTP molecular signatures [6]. Moreover, statistical classification algorithms have emerged as a promising tool to support the diagnostic process of tumors [7], including in the field of thyroid cytology to distinguish some challenging entities (e.g. NIFTP) from more aggressive or indolent lesions [10, 11]. Here, we tested a MALDI-MSI-based approach for biomarker discovery on a challenging series of thyroid neoplasms to assess whether the classification ability of proteomics can be improved by statistical classification algorithms, with a specific focus on the distinction between NIFTPs and histologically similar neoplasms (e.g., follicular adenoma [FA], papillary thyroid carcinoma [PTC], and invasive encapsulated and infiltrative follicular variants of PTC [iE-FVPTC and I-FVPTC, respectively]). Within follicular-patterned tumors, we also evaluated the performance of this combined approach in the distinction between NRAS-mutant (mNRAS) and NRAS-wild-type (wtNRAS) lesions.

## Materials and methods

### Cases

The study cohort was composed of 44 patients who underwent thyroid surgery for different thyroid tumors at

the IRCCS Fondazione San Gerardo dei Tintori, Monza, Italy. Two different areas were selected for each tumor sample from the original formalin-fixed paraffin-embedded (FFPE) tissue block by an experienced thyroid pathologist (FP) to build a tissue microarray (TMA) using the semi-automatic arrayer ISE Galileo TMA CK 4500 through the ISE Galileo TMA R4.30 software (Integrated Systems Engineering, Milan, Italy). The TMA was tested for the NRAS Q61R mutation by immunohistochemistry (rabbit monoclonal antibody, clone RST-NRAS, dilution 1:20) on DAKO Omnis (Agilent, Santa Clara, CA, USA).

For each case, demographic (age and sex), clinical (nodule laterality, nodule diameter), and cytological characteristics were extracted. Cytological diagnoses were reported according to the TIR classification of the Italian Society of Pathology (SIAPEC) [12]. Approval was obtained from the local ethics committee (FINAL-TIR PU 3581/21), and all subjects enrolled in the study signed an informed consent. The study was not part of a clinical trial.

### MALDI-MSI analysis

Sections from the obtained TMA block were treated as previously described for MALDI-MSI preparation [6]. The mass spectra were acquired in the reflectron positive mode, within the range of  $m/z$  700 to 3000, using a rapifleX MALDI TissueTyper (Bruker Daltonics, Bremen, Germany) MALDI-TOF/TOF MS equipped with a Smartbeam 3D laser operating at 2 kHz frequency. The MALDI-MSI images were acquired with a single-spot laser setting of 10  $\mu\text{m}$  and a scan range of 6  $\times$  6  $\mu\text{m}$ . A mixture of standard peptides within the mass range of  $m/z$  750 to 3150 (PepMix I, Bruker Daltonics) was used for the external calibration, directly applied on the glass slide (mass accuracy <10 ppm). FlexControl 4.0 (Bruker Daltonics) was used to set up the instrument parameters for the acquisition method, and FlexImaging 5.0 (Bruker Daltonics) for the MALDI-MSI analysis visualization. Next, the matrix was removed and the slides were stained with hematoxylin and eosin (H&E). Finally, the slides were converted to a digital format using a MIDI II digital scanner (3DHISTECH, Budapest, Hungary), allowing the integration of the proteomic and morphological data. Within each core, an expert thyroid pathologist (FP) annotated different regions of interest (ROIs), separating normal tissue (thyroid parenchyma and stroma) from lesional areas, as reported in Fig. S1, for a total of 144 ROIs of lesional areas. The annotated regions were then imported into SCiLS Lab 2024 Pro software (Bruker, Bremen,

Germany) and the corresponding average mass spectrum was extracted from each ROI.

### nLC-ESI-MS/MS sample preparation and analysis

After the MALDI-MSI analysis, a nanoscale liquid chromatography electrospray ionization tandem mass spectrometry (nLC-ESI-MS/MS) analysis for peptide identification was performed, as previously described with some modifications [6]. The matrix collected from each slide was dried with a vacuum centrifugal evaporator (Hetovac, Savant) and resuspended in 50  $\mu$ L of phase A (98:2:0.1; water/acetonitrile/trifluoroacetic acid). The nLC-ESI MS/MS analysis was performed using a Dionex UltiMate 3000 rapid separation (RS) LC nano system coupled with an Impact HD UHR-QqToF (Bruker Daltonics). Sample desalting and concentration were carried out using a pre-column (Dionex, Acclaim PepMap 100 C18, cartridge, 300  $\mu$ m), and the peptides were separated with a 50 cm column (Dionex, ID 0.075 mm, Acclaim PepMap100, C18) at 40 °C, using a 240 min gradient from 96% to 2% of phase A (0.1% formic acid), whilst phase B was 0.08% formic acid:acetonitrile (80:20). The MS was operating in data-dependent acquisition mode. Compass DataAnalysis v4.1 software (Bruker Daltonics) was used to calibrate, deconvolute and convert the acquired raw data prior to the protein identification and quantification. The Peaks Studio X-Plus (Bioinformatics Solutions Inc., Waterloo, ON, CA) was used for protein identification. The parameters were set as follows: trypsin as the digestive enzyme, no fixed modifications, oxidation (M) and FFPE + 12 and FFPE + 30 as the variable modifications. The precursor mass error and the fragment mass error tolerances were set at 20 ppm and 0.05 Da, respectively. The identification engine used the Swissprot database, selecting the homo sapiens taxonomy. For identification, a peptide false discovery rate (FDR) of  $\leq 1\%$  was applied. The proteins were considered identified if they had at least one unique significant peptide ( $-\log P \geq 20$ ). The signals of interest present within the MALDI-MSI dataset were correlated with the positively identified peptide sequences obtained using the nLC-ESI-MS/MS. An identification was putatively assigned to a signal if an error of  $< 100$  ppm was noted between the two measured  $m/z$  values.

### Statistical analysis

The database containing the average spectra of each ROI underwent a pre-processing workflow. This process included baseline subtraction (TopHat algorithm), normalization (Total Ion Current algorithm), spatial denoising, alignment, and peak picking. As a result, 373  $m/z$  molecular features indicating the molecular

abundance contained in each of the 144 ROIs were detected. Linear Discriminant Analysis (LDA) was applied for feature selection and classification [13, 14]. LDA aims to reduce dimensionality by finding a lower-dimensional subspace that maximizes the separability of the samples [15], assuming that the features might be dependent. The optimal number of features was selected controlling for FDR. The usual performance metrics, i.e. sensitivity (Sn), specificity (Sp), positive and negative predictive values (PPV and NPV, respectively) were calculated on the entire cohort, considering only the selected features. Balanced accuracy was used instead of generic accuracy to account for differences in sample size between diagnostic classes. Features selected by LDA were compared for each task of the study: discriminating between tumor histotypes and differentiating between mNRAS and wtNRAS samples. Pairwise Wilcoxon-tests were used to quantify the differences between lesions and samples with and without positivity to NRAS Q61R, accompanied by the associated fold-change, defined as the ratio of the mean values between two classes. All the analyses were performed using R software version 4.2.2 (The R Foundation For Statistical Computing, Vienna, Austria) with the MixOmics library (<http://mixomics.org/install/>).

## Results

### Cases

The study cohort consisted of 44 cases, corresponding to 88 TMA cores (two cores were collected for each patient) and a total of 144 pathologist-annotated ROIs. The clinicopathological characteristics of enrolled subjects are summarized in Table 1. The majority of patients were female ( $n = 36$ , 82%), with a median age of 51 years (interquartile range [IQR] 45–63). The median diameter of the nodules was 1.25 cm (IQR 0.88–1.80), and most nodules were located in the right lobe ( $n = 19$ , 43.2%), with significantly larger nodules belonging to the FA category (median size = 3 cm, IQR 2.8–3.13,  $p = 0.002$ ). The final histological diagnoses included 18 cases (44 ROIs) of conventional PTC, 7 (23 ROIs) of I-FVPTC, 3 (5 ROIs) of iE-FVPTC, 4 (8 ROIs) of FA, and 12 (34 ROIs) of NIFTP. Given the small number of ROIs within the I-FVPTC and iE-FVPTC classes, these PTC variants were grouped into a single category (i.e., FVPTC) for subsequent analyses. A higher prevalence of the NRAS Q61R mutation was observed in follicular-patterned lesions (50% of FAs, 28.6% of I-FVPTCs, and 50% of NIFTPs) compared to classic PTCs, where it was found in only one case (6.2%,  $p = 0.082$ ).

**Table 1** Clinical and pathological features of patients enrolled in the study

Characteristic	Overall, n = 44	FA, n = 4	I-FVPTC, n = 7	iE-FVPTC, n = 3	NIFTP, n = 12	PTC, n = 18	p-value <sup>c</sup>
Age (years) <sup>a</sup>	51 (45, 63)	45 (30, 60)	58.1 (50, 64)	52 (43, 64)	56 (48, 66)	48 (44, 50)	0.208
Sex <sup>b</sup>							0.464
F	36 (82)	4 (100)	4 (57.1)	3 (100)	10 (83)	15 (83)	
M	8 (18)	0 (0)	3 (42.9)	0 (0)	2 (17)	3 (17)	
NRAS mutational status <sup>b</sup>							0.082
MUT	11 (26.2)	2 (50)	2 (28.6)	0 (0)	6 (50)	1 (6.2)	
WT	31 (73.8)	2 (50)	5 (71.4)	3 (100)	6 (50)	15 (93.8)	
N/A	2	0	0	0	0	2	
Nodule laterality <sup>b</sup>							0.941
I	2 (4.5)	1 (25)	0 (0)	0 (0)	1 (8.3)	0 (0)	
L	15 (34.1)	1 (25)	3 (42.9)	1 (33.3)	4 (33.4)	6 (33.4)	
MF	8 (18.2)	0 (0)	1 (14.2)	1 (33.3)	2 (16.7)	4 (22.2)	
R	19 (43.2)	2 (50)	3 (42.9)	1 (33.3)	5 (41.6)	8 (44.4)	
Nodule size (cm) <sup>a</sup>	1.25 (0.88, 1.80)	3.00 (2.80, 3.13)	0.99 (0.60, 1.40)	0.60 (0.55, 0.70)	1.70 (0.98, 2.05)	1.15 (0.93, 1.50)	0.004
Cytological classification <sup>b</sup>							0.140
Tir1	1 (3.5)	0 (0)	0 (0)	0 (0)	1 (12.5)	0 (0)	
Tir2	8 (27.6)	1 (50)	1 (25)	0 (0)	4 (50)	2 (14.3)	
Tir3	4 (13.8)	1 (50)	0 (0)	1 (100)	2 (25)	0 (0)	
Tir4	7 (24.1)	0 (0)	1 (25)	0 (0)	1 (12.5)	5 (35.7)	
Tir5	9 (31)	0 (0)	2 (50)	0 (0)	0 (0)	7 (50)	
N/A	15	2	3	2	4	4	
Lymphovascular invasion <sup>b</sup>							0.055
Absent	37 (84.1)	4 (100)	7 (100)	2 (66.7)	12 (100)	12 (67)	
Present	7 (15.9)	0 (0)	0 (0)	1 (33.3)	0 (0)	6 (33)	
Capsular invasion <sup>b</sup>							0.001
Absent	16 (84.2)	4 (100)	N/A	0 (0)	12 (100)	N/A	
Present	3 (15.8)	0 (0)	N/A	3 (100)	0 (0)	N/A	
N/A	25	0	7	0	0	18	

n Number; F Female; M Male; MUT Mutant; WT Wild-type; I Isthmus; L Left; MF Multifocal; R Right; N/A Not assessed; FA Follicular adenoma; I-FVPTC Infiltrative follicular variant papillary thyroid carcinoma; iE-FVPTC invasive Encapsulated follicular variant papillary thyroid carcinoma; NIFTP Noninvasive follicular thyroid neoplasm with papillary-like nuclear features; PTC Papillary thyroid carcinoma; IQR Interquartile range

<sup>a</sup>Median (IQR)

<sup>b</sup>n (%)

<sup>c</sup>Kruskal-Wallis rank sum test; Fisher's exact test

## Combined LDA and MALDI approach for the classification of thyroid nodules

After appropriate filtering and selection of MALDI-MSI features from the original 373 *m/z* signals, the LDA identified nine relevant molecular markers achieving a balanced accuracy of 73% in distinguishing NIFTPs from the other thyroid lesions. The model also performed well in identifying the FA and PTC classes ( $S_n = 75\%$ ), showing the highest PPV

(87%) in the PTC class, followed by the FA and NIFTP categories (PPV = 67% and 62%, respectively) (Table 2A). The distribution of the nine LDA-selected features among the different tumor classes revealed interesting relationships of NIFTPs with both malignant (e.g., *m/z* 1542.777 and 1586.813 signals in FVPTC/PCT) and benign (e.g., *m/z* 1094.621 and 1325.645 signals in FA) lesions in terms of protein expression patterns, highlighting the controversial biological behavior of this neoplasm (Fig. 1 and Table S1).

**Table 2** Performance and predicted classes of the LDA models developed to discriminate among the different lesions (A) and according to NRAS mutational status (B)

A										
Class	n	Predicted as FA	Predicted as FVPTC	Predicted as NIFTP	Predicted as PTC	Sn	Sp	PPV	NPV	Balanced Accuracy
FA	8	6	0	2	0	0.75	0.97	0.67	0.98	0.86
FVPTC	28	1	18	7	2	0.64	0.83	0.55	0.88	0.73
NIFTP	34	2	8	21	3	0.62	0.84	0.62	0.84	0.73
PTC	44	0	7	4	33	0.75	0.93	0.87	0.86	0.84
B										
Class	n	TN	FP	TP	FN	Sn	Sp	PPV	NPV	Balanced Accuracy
WT vs. MUT	62 (32 vs. 30)	23	9	29	1	0.97	0.72	0.76	0.96	0.84

*n* Number; *FA* Follicular adenoma; *FVPTC* Follicular variant papillary thyroid carcinoma; *NIFTP* Noninvasive follicular thyroid neoplasm with papillary like-nuclear features; *PTC* Papillary thyroid carcinoma; *Sn* Sensitivity; *Sp* Specificity; *PPV* Positive predictive value; *NPV* Negative predictive value; *WT* Wild-type; *MUT* Mutant; *TN* True negatives; *FP* False positives; *TP* True positives; *FN* False negatives

### Combined LDA and MALDI approach to discriminate between mNRAS and wtNRAS cases

Comparing mNRAS and wtNRAS cases, the LDA allowed a significant reduction of relevant signals, selecting 167 and 52 *m/z* peaks for FVPTC and NIFTP, respectively. 51 of these signals were overexpressed in wtNRAS compared to mNRAS cases, regardless of histology (Table S2). Based on the 19 most relevant molecular markers (Table S2), the model achieved a balanced accuracy of 84%, correctly classifying 29 out of 30 mNRAS ROIs and 23 out of 32 wtNRAS cases, with a Sn of 97% and a Sp of 72% (Table 2B). The visual inspection of the distribution of some of these signals highlighted clear differences between mNRAS and wtNRAS cases, independently of histology (e.g., *m/z* 838.5 and 1293.7 signals, both overexpressed in wtNRAS cases. Figure 2).

### Signal identification through nLC-ESI-MS/MS

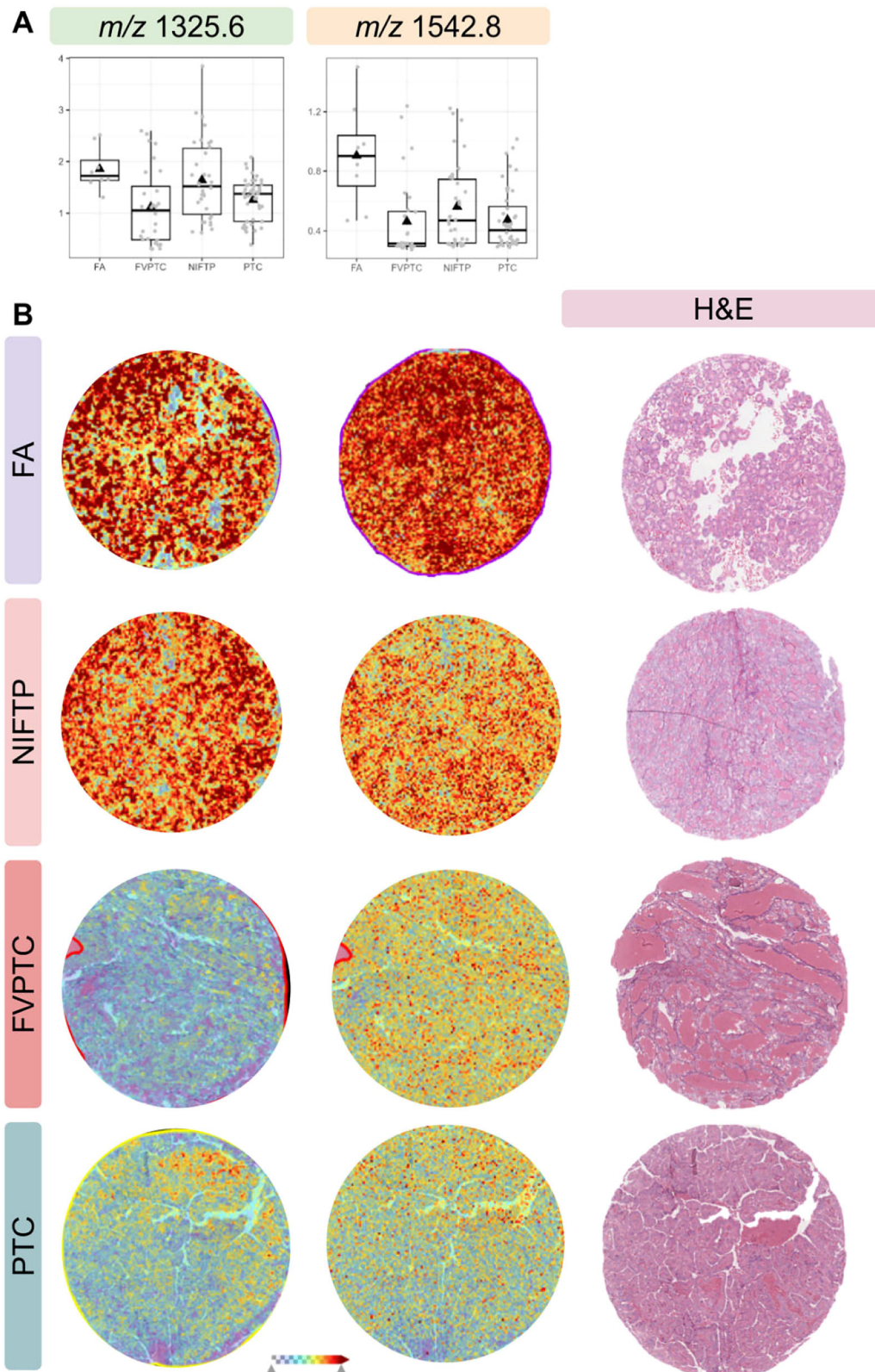
By means of nLC-ESI-MS/MS, nine of the proteins with different expression patterns among the various classes of thyroid neoplasms were putatively identified. Among them, ATP-dependent RNA helicase DDX42 (*m/z* 1542.777, DDX42) showed a similar distribution between NIFTPs and FVPTCs/PTCs, while the distribution of the Histone H4 signal (*m/z* 1325.645, H4) was similar between NIFTPs and FAs. In addition, Complement C4-B (*m/z* 838.5, CO4B) and Protein disulfide-isomerase, also known as cellular thyroid hormone-binding protein (*m/z* 1293.7, PDIA1), were both overexpressed in wtNRAS cases independently of histology (a complete list of the identified proteins is reported in Table S3). These proteins are involved in the organization of chromatin structure (H4) and regulation of

mRNA splicing (DDX42), as well as in the degradation of misfolded proteins and modulation of thyroid hormone receptor function (PDIA1) [16–18], suggesting that a disruption in these key cellular processes might play a role in the pathogenesis of the lesions studied.

### Discussion

Extensive research efforts have focused on identifying novel biomarkers for challenging thyroid lesions, with the aim to improve diagnostic accuracy and define useful prognostic indicators to guide patient management [19]. The introduction of innovative molecular markers able to fill the existing accuracy gap of the combined ultrasound-FNAC approach would more precisely identify patients who are truly candidates for invasive therapeutic procedures. Follicular-patterned thyroid lesions represent a particularly challenging group of neoplasms, encompassing entities with distinct behaviors and different malignant potential. Among these, NIFTPs have been recently “downgraded” as indolent neoplasms [20], although their biology is still far from being fully understood. Moreover, the presence of “atypical” cases, often bearing RAS mutations, further complicates efforts to distinguish NIFTPs from the other follicular-patterned lesions [21].

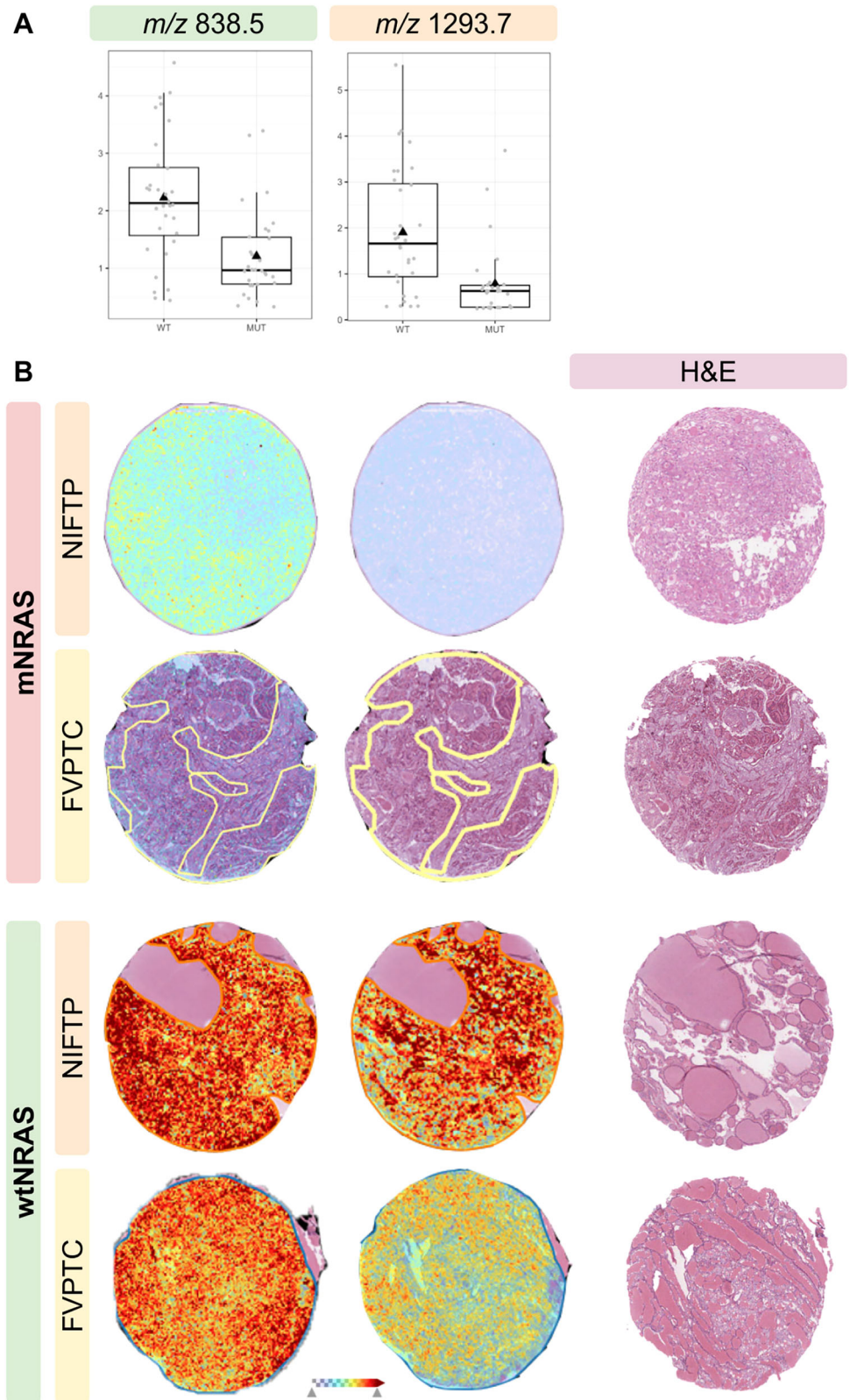
MALDI-MSI has emerged as a promising tool for biomarker discovery, enabling the identification of the specific molecular signals associated with the analyzed lesions [22]. Given the considerable amount of information deriving from MALDI-MSI analysis, we built an LDA-supervised approach for feature selection and classification based on proteomic data obtained from a group of challenging thyroid lesions (i.e., FAs, NIFTPs, FVPTCs, and PTCs).



**Fig. 1** **A** Molecular distribution of the H4\_HUMAN ( $m/z$  1325.6) and DDX42 ( $m/z$  1542.8) signals by category of thyroid lesions. The black triangle represents the mean, while the dots represent the single observations. **B** MALDI-MSI images showing the spatial localization

of the signals. FA Follicular adenoma, FVPTC Follicular variant papillary thyroid carcinoma, NIFTP Noninvasive follicular thyroid neoplasm with papillary-like nuclear features, PTC Papillary thyroid carcinoma

**Fig. 2** **A** Molecular distribution of the CO4B ( $m/z$  838.5) and PDIA1 ( $m/z$  1293.7) signals by NRAS mutational status. The black triangle represents the mean, while the dots represent the single observations. **B** MALDI-MSI images showing the spatial localization of the signals. FVPTC Follicular variant papillary thyroid carcinoma, NIFTP Noninvasive follicular thyroid neoplasm with papillary-like nuclear features, mNRAS NRAS mutant, wtNRAS NRAS-wild-type



Among these, FAs were significantly larger at diagnosis compared to the other neoplasms, reflecting their indolent course and the subsequent late diagnosis. As expected,

NRAS mutations were significantly more frequent in follicular-patterned lesions (i.e., NIFTP and FVPTC) compared to classic PTCs. Our approach enabled the

identification of specific molecular markers that correspond to follicular-patterned lesions, including NIFTPs and FVPTCs, while also providing valuable insights into the complex behavioral trajectories of NIFTPs with respect to morphologically similar neoplasms. This, in turn, could be useful for discerning whether such behavior more closely resembles that of the malignant class (PTC, FVPTC) as opposed to the benign category (FA). Specifically, the similar downregulation of DDX42 in NIFTPs and FVPTCs/PTCs compared to FAs may suggest that the former lesions share similar alterations in pre-mRNA processing. Splicing abnormalities could thus represent important mechanisms underlying the development and progression of these neoplasms, similar to what has been observed in other tumor types [23]. On the other hand, the comparable distribution of the Histone H4 signal between NIFTPs and FAs and its underexpression in FVPTCs/PTCs may indicate a higher burden of chromatin structural aberrations in the latter category of lesions, which is a hallmark of a malignant behavior. Indeed, downregulation of conventional histone proteins may reflect the increased expression of modified histone variants, which have been identified as key contributors to tumor progression and invasiveness in several malignancies, including breast and colorectal cancer [24].

Apart from their pathobiological significance, the identified signals may be valuable for differentiating between these challenging thyroid lesions in routine pathological practice. Indeed, previous work by our group has shown a significant downregulation of the DDX42 and Histone H4 proteins in RAS-mutant compared to RAS-wild type NIFTPs from a separate cohort of patients, encouraging further validation of these markers as ancillary tools to assist in the differential diagnosis of borderline thyroid neoplasms [25]. This would require assessing the expression of the identified markers through immunohistochemistry in a separate and, possibly, larger sample of FAs, PTCs, FVPTCs, and NIFTPs. Specifically, the commercial availability of antibodies for the immunoenzymatic detection of the DDX42 and Histone H4 proteins may warrant the recruitment of additional subjects to establish a validation set of sufficient size for statistically robust analyses.

Finally, the comparison of wtNRAS vs. mNRAS cases among follicular-patterned lesions (i.e., FVPTC and NIFTP) enabled the identification of molecular signals associated with NRAS Q61R positivity, potentially providing insights into the distinct pathogenic pathways through which these lesions develop. Among the differentially expressed signals, we were able to putatively identify the Complement C4-B and PDIA1 proteins, which we found to be overexpressed in wtNRAS compared to mNRAS cases. Altered expression of both of these proteins has already been implicated in the biology of thyroid neoplasms: PTC cells have been shown to escape host immunosurveillance through the synthesis of

anti-idiotypic immunoglobulin G (IgG) molecules that bind to and neutralize host IgGs directed against tumor antigens. The subsequent formation of immune complexes leads to the activation of the classical complement pathway, explaining the observed co-localization of IgGs and complement proteins (including C4) in PTC tissues [26]. Similarly, members of the protein disulfide isomerase family, to which PDIA1 belongs, have been identified as pivotal factors in promoting the progression and dissemination of several cancer types, including PTC [27]. PDIs, which are redox proteins bearing thioredoxin-like domains [28], have pleiotropic roles in regulating thyroid hormone receptor signaling, misfolded protein degradation, and progression through the cell cycle [29]. Therefore, aberrant PDIA1 expression may contribute to tumor development and progression by disrupting intracellular protein homeostasis, as well as by altering the transcriptional responses elicited by the binding of thyroid hormones to their receptors. In addition, reduced expression of another member of the PDI family, namely PDIA3, has been found to correlate with lymph node metastasis and poor prognosis in PTC patients [30]. Overall, these observations support the hypothesis that mNRAS and wtNRAS lesions are intrinsically distinct entities that rely on different pathways for tumor growth and immune evasion. Whether these biological differences lead to a distinct clinical behavior of mNRAS compared to wtNRAS tumors, regardless of histology, is still uncertain, warranting prospective clinical studies with large patient cohorts.

The wealth of discriminative features identified in our study further confirms the remarkable potential of MALDI-MSI to streamline biomarker discovery in the field of thyroid pathology, particularly when machine learning algorithms are applied to support data analysis. However, the implementation of this tool into the routine diagnostic practice is currently hampered by its relatively high costs, limited availability, and technical constraints regarding sample preparation for proteomic analyses. Thus, future advancements to improve the accessibility and scalability of this methodology are highly anticipated. This may be achieved by defining a restricted panel of relevant markers and testing their expression through multiplex spatial imaging approaches, such as MALDI-HiPLEX-IHC, which allows to precisely map the distribution of proteins of interest in FFPE tissues [31]. As shown by our group using clear cell renal cell carcinoma models, the integration of MALDI-HiPLEX-IHC with untargeted spatial proteomics enables the spatial localization of antigens of interest as well as the characterization of the tumor immune microenvironment, requiring only a single FFPE tissue section [32]. Based on these promising results, we are planning to extend the applicability of our workflow also to challenging thyroid lesions.

Although promising, our preliminary application of the LDA model to distinguish between histologically similar thyroid lesions has some limitations: due to the small number of subjects enrolled in our study, we decided to analyze the full patient cohort without splitting our dataset, as this would have resulted in underpowered training folds. For similar reasons, our proteomic analyses were performed considering I-FVPTCs and iE-FVPTCs as a single neoplastic category (i.e., FVPTC). Thus, the full generalizability of our model still requires further validation with external cohorts.

Moreover, the main focus of our study was to investigate putative biomarkers to facilitate the differential diagnostic process of borderline thyroid lesions. Whether our methodology can also identify markers correlated with relevant histopathological features (e.g., the presence of capsular and/or lymphovascular invasion) remains to be explored. Nevertheless, since the identified proteins may aid in the distinction between NIFTPs and histologically similar invasive neoplasms, they might serve as indirect indicators of tumor invasiveness irrespective of histology. Again, these preliminary findings warrant further validation studies with larger patient cohorts.

Finally, our analyses were carried out on a relatively limited set of neoplasms, so additional efforts are needed to extend the applicability of our approach to other challenging entities.

## Conclusion

Biomarker discovery is crucial for diagnostic and prognostic purposes, particularly in the context of thyroid lesions. Our work showed promise in identifying characteristic protein patterns associated with specific thyroid tumors, including NIFTPs, also considering their NRAS mutational status. The generalizability of our findings will require further validation on external cohorts and on a broader and more heterogeneous group of thyroid neoplasms.

## Data availability

The data that support the findings of this study are not publicly available as they contain information that could compromise the privacy of research participants, but will be made available by the corresponding author (GC) upon reasonable request.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1007/s12020-025-04383-9>.

**Author contributions** GC performed the statistical analysis, wrote the original manuscript and prepared the figures. AMA wrote the original manuscript and prepared the figures. NM, LP and IP

performed the MALDI-MSI and nLC-ESI-MS/MS analyses, revised the manuscript and prepared the figures. DPB supervised the statistical analysis. AG, DL, AM, AVG, FMaf, NF and MG provided clinical information. FMag and SG critically revised the manuscript in its final version. FP and VL provided the funding acquisition and administrative support and critically revised the manuscript in its final version. VD performed the MALDI-MSI and nLC-ESI-MS/MS analyses, revised the manuscript and prepared the figures. All authors had final approval of the submitted version of the manuscript.

**Funding** This work was funded by the National Plan for NRRP Complementary Investments (PNC, established with the decree-law 6 May 2021, n. 59, converted by law n. 101 of 2021) in the call for the funding of research initiatives for technologies and innovative trajectories in the health and care sectors (Directorial Decree n. 931 of 06-06-2022) - project n. PNC0000003 - AdvanCed Technologies for Human-centrEd Medicine (project acronym: ANTHEM). FP, SG, GC, FMag and VL participated in the Italian Ministry of the University MUR Dipartimenti di Eccellenza 2023–2027 IMPACT MEDICINE Project (IMPACT MEDICINE, l. 232/2016, art. 1, commi 314–337). This work reflects only the authors' views and opinions, neither the Ministry for University and Research nor the European Commission can be considered responsible for them. Funds to NF: Italian Ministry of Health through Ricerca Corrente 5×1000 funds; the Italian Ministry of Innovations via the Sustainable Growth Fund – Innovation Agreements under the Ministerial Decree of December 31, 2021, and the Director's Decree of November 14, 2022 (2nd Call), Project No.: F/350104/01-02/X60; and the Italian Ministry of University and Research (MUR) 2023 through the “Future Artificial Intelligence Research – FAIR” program, PE0000013, CUP D53C22002380006, within the National Recovery and Resilience Plan (PNRR), Mission 4, Component 2, Investment 1.3 – funded by the European Union – NextGenerationEU. Project: “AIDH – Adaptive AI Methods for Digital Health”.

## Compliance with ethical standards

**Ethics approval and consent for participation** This study protocol was reviewed and approved by the local ethical committee (Comitato Etico Brianza, via Pergolesi, 33, 20900 Monza, Italy; approval number FINAL-TIR PU 3581/21; approval date: January 14, 2021). All subjects enrolled in the study signed an informed consent. The study was not part of a clinical trial (clinical trial number: not applicable).

**Conflict of interest** The authors declare no competing interests.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- M. Galli, F. Pagni, G. De Sio et al., Proteomic profiles of thyroid tumors by mass spectrometry-imaging on tissue microarrays. *Biochim. Biophys. Acta Proteins Proteom.* **1865**(7), 817–827 (2017). <https://doi.org/10.1016/j.bbapap.2017.03.015>.
- B.R. Haugen, E.K. Alexander, K.C. Bible et al., 2015 American thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American thyroid association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid* **26**(1), 1–133 (2016). <https://doi.org/10.1089/thy.2015.0020>.
- D. Seminati, G. Capitoli, D. Leni et al., Use of diagnostic criteria from ACR and EU-TIRADS systems to improve the performance of cytology in thyroid nodule triage. *Cancers (Basel)* **13**(21), 5459 (2021). <https://doi.org/10.3390/cancers13215439>.
- D. Leni, D. Seminati, D. Fior et al., Diagnostic performances of the ACR-TIRADS system in thyroid nodules triage: a prospective single center study. *Cancers (Basel)* **13**(9), 2230 (2021). <https://doi.org/10.3390/cancers13092230>.
- V. Mainini, F. Pagni, M. Garancini et al., An alternative approach in endocrine pathology research: MALDI-IMS in papillary thyroid carcinoma. *Endocr. Pathol.* **24**(4), 250–253 (2013). <https://doi.org/10.1007/s12022-013-9276-9>.
- I. Piga, V. L'Imperio, L. Principi et al., Spatially resolved molecular approaches for the characterisation of non-invasive follicular tumours with papillary-like features (NIFTPs). *Int. J. Mol. Sci.* **24**(3), 2567 (2023). <https://doi.org/10.3390/ijms24032567>.
- G. Capitoli, I. Piga, S. Galimberti et al., MALDI-MSI as a complementary diagnostic tool in cytopathology: a pilot study for the characterization of thyroid nodules. *Cancers (Basel)* **11**(9), 1377 (2019). <https://doi.org/10.3390/cancers11091377>.
- Pergamon, Unsupervised neural networks as a support tool for pathology diagnosis in MALDI-MSI experiments: a case study on thyroid biopsies. *Expert. Syst. Appl.* **215**, 119296 (2023). <https://doi.org/10.1016/j.eswa.2022.119296>.
- G. Capitoli, I. Piga, V. L'Imperio et al., Cytomolecular classification of thyroid nodules using fine-needle washes aspiration biopsies. *Int. J. Mol. Sci.* **23**(8), 4156 (2022). <https://doi.org/10.3390/ijms23084156>.
- V. L'Imperio, V. Coelho, G. Cazzaniga et al., Machine learning streamlines the morphometric characterization and multi-class segmentation of nuclei in different follicular thyroid lesions: everything in a NUTSHELL. *Mod. Pathol.* **37**, 100608 (2024). <https://doi.org/10.1016/j.mod.2024.100608>.
- I. Piga, G. Capitoli, F. Clerici et al., Molecular trait of follicular-patterned thyroid neoplasms defined by MALDI-imaging. *Biochim. Biophys. Acta Proteins Proteom.* **1868**(11), 140511 (2020). <https://doi.org/10.1016/j.bbapap.2020.140511>.
- F. Nardi, F. Basolo, A. Crescenzi et al., Italian consensus for the classification and reporting of thyroid cytology. *J. Endocrinol. Invest.* **37**(6), 593–599 (2014). <https://doi.org/10.1007/s40618-014-0062-0>.
- M. Mourad, S. Moubayed, A. Dezube et al., Machine learning and feature selection applied to SEER data to reliably assess thyroid cancer prognosis. *Sci. Rep.* **10**(1), 5176 (2020). <https://doi.org/10.1038/s41598-020-62078-2>.
- H. Pang, T. Tong, M. Ng, Block-diagonal discriminant analysis and its bias-corrected rules. *Stat. Appl. Genet. Mol. Biol.* **12**(3), 347–359 (2013). <https://doi.org/10.1515/sagmb-2012-0084>.
- P. Xanthopoulos, P.M. Pardalos, T.B. Trafalis Linear discriminant analysis. In: *Robust Data Mining*. Springer; 2013:27-33. [https://doi.org/10.1007/978-1-4419-9878-1\\_3](https://doi.org/10.1007/978-1-4419-9878-1_3).
- S. Martire, L.A. Banaszynski, The roles of histone variants in fine-tuning chromatin organization and function. *Nat. Rev. Mol. Cell Biol.* **21**(9), 522–541 (2020). <https://doi.org/10.1038/s41580-020-0257-1>.
- F. Yang, T. Bian, X. Zhan et al., Mechanisms of the RNA helicases DDX42 and DDX46 in human U2 snRNP assembly. *Nat. Commun.* **14**(1), 897 (2023). <https://doi.org/10.1038/s41467-023-36536-0>.
- S. Hashimoto, S. Imaoka, Protein-disulfide isomerase regulates the thyroid hormone receptor-mediated gene expression via redox factor-1 through thiol reduction-oxidation. *J. Biol. Chem.* **288**(3), 1706–1716 (2013). <https://doi.org/10.1074/jbc.M112.420281>.
- S. Agarwal, A. Bychkov, C.-K. Jung, Emerging biomarkers in thyroid practice and research. *Cancers (Basel)* **14**(1), 204 (2021). <https://doi.org/10.3390/cancers14010204>.
- R.R. Seethala, Z.W. Baloch, J.A. Barletta et al., Noninvasive follicular thyroid neoplasm with papillary-like nuclear features: a review for pathologists. *Mod. Pathol.* **31**(1), 39–55 (2018). <https://doi.org/10.1038/modpathol.2017.108>.
- K. Zajkowska, J. Kopczyński, S. Góźdz et al., Noninvasive follicular thyroid neoplasm with papillary-like nuclear features: a problematic entity. *Endocr. Connect.* **9**(3), R47–R58 (2020). <https://doi.org/10.1530/EC-19-0522>.
- A.L. Swan, A. Mobasher, D. Allaway et al., Application of machine learning to proteomics data: classification and biomarker identification in postgenomics biology. *OMICS* **17**(12), 595–610 (2013). <https://doi.org/10.1089/omi.2013.0083>.
- Y. Zhang, J. Qian, C. Gu et al., Alternative splicing and cancer: a systematic review. *Signal. Transduct. Target. Ther.* **6**(1), 78 (2021). <https://doi.org/10.1038/s41392-021-00484-z>.
- C. Vardabasso, D. Hasson, K. Ratnakumar et al., Histone variants: emerging players in cancer biology. *Cell Mol. Life Sci.* **71**(3), 379–404 (2014). <https://doi.org/10.1007/s00018-013-1331-5>.
- V. Denti, A. Greco, A.M. Alviano et al., Spatially resolved molecular characterization of Noninvasive Follicular Thyroid Neoplasms with Papillary-Like Nuclear Features (NIFTPs) identifies a distinct proteomic signature associated with RAS-mutant lesions. *Int. J. Mol. Sci.* **25**, 13115 (2024). <https://doi.org/10.3390/ijms252313115>.
- Y. Qiu, C. Korteweg, Z. Chen et al., Immunoglobulin G expression and its colocalization with complement proteins in papillary thyroid cancer. *Mod. Pathol.* **25**(1), 36–45 (2012). <https://doi.org/10.1038/modpathol.2011.133>.
- N.S.A. Rahman, S. Zahari, S.E. Syafruddin et al., Functions and mechanisms of protein disulfide isomerase family in cancer emergence. *Cell Biosci.* **12**(1), 129 (2022). <https://doi.org/10.1186/s13578-022-00883-5>.
- H. Ali Khan, B. Mutus, Protein disulfide isomerase: a multi-functional protein with multiple physiological roles. *Front. Chem.* **2**, 70 (2014). <https://doi.org/10.3389/fchem.2014.00070>.
- E.Y. Bakker, M. Fujii, M. Krstic-Demonacos et al., Protein disulfide isomerase A1-associated pathways in the development of stratified breast cancer therapies. *Int. J. Oncol.* **60**(2), 16 (2022). <https://doi.org/10.3892/ijo.2021.5284>.
- S. Kure, T. Chiba, A. Ebina et al., Correlation between low expression of protein disulfide isomerase A3 and lymph node metastasis in papillary thyroid carcinoma and poor prognosis: a clinicopathological study of 1,139 cases with long-term follow-up. *Endocr. J.* **69**(3), 273–281 (2022). <https://doi.org/10.1507/endocrj.EJ21-0568>.
- M.J. Lim, G. Yagnik, C. Henkel et al., MALDI HiPLEX-IHC: multiomic and multimodal imaging of targeted intact proteins in tissues. *Front. Chem.* **11**, 1182404 (2023). <https://doi.org/10.3389/fchem.2023.1182404>.
- G. Bindi, N. Monza, G. Santos de Oliveira et al., Sequential MALDI-HiPLEX-IHC and untargeted spatial proteomics mass spectrometry imaging to detect proteomic alterations associated with tumour infiltrating lymphocytes. *J. Proteome Res.* **24**(2), 871–880 (2025). <https://doi.org/10.1021/acs.jproteome.4c00914>.