U.PORTO

INTEGRATED MASTER'S DEGREE

MEDICINE

Study of the epigenetic silencing of the tumor suppressor gene *SPINT2* in lung cancer

Marta Sofia de Carvalho Ribeiro Viana Pereira

M

2019

Study of the epigenetic silencing of the tumor suppressor gene *SPINT2* in lung cancer

Marta Sofia de Carvalho Ribeiro Viana Pereira

martavianap@gmail.com

Integrated Master's Degree in Medicine 2018/2019

Institute of Biomedical Sciences Abel Salazar

University of Porto, Portugal

Supervisor: Rui Manuel Ferreira Henrique, MD, PhD

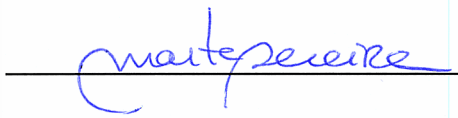
Director of the Department of Pathology at *Instituto Português de Oncologia (IPO) do Porto.* Invited Full Professor at Department of Pathology and Molecular Immunology, Institute of Biomedical Sciences Abel Salazar, University of Porto, Portugal.

Co-supervisor: Carmen de Lurdes Fonseca Jerónimo, PhD

Coordinator of the Cancer Biology & Epigenetics Group at Research Centre of IPO-Porto. Invited Associate Professor at Department of Pathology and Molecular Immunology, Institute of Biomedical Sciences Abel Salazar, University of Porto, Portugal.

May, 2019

Author: Marta Sofia de Carvalho Ribeiro Viana Pereira



Supervisor: Rui Manuel Ferreira Henrique, MD, PhD

Co-supervisor: Carmen de Lurdes Fonseca Jerónimo, PhD

May, 2019

Acknowledgements

I would like to acknowledge all the people that, in many ways, have given me support during the elaboration of this work and throughout this Master's Degree.

A special regard my supervisors at IPO-Porto, Professor Rui Henrique and Professor Carmen Jerónimo, who so nicely welcomed me in their lab. Thank you for trusting my work and for the scientific support and advices. It was a great opportunity to work and learn with you.

To the University of Porto, the Institute of Biomedical Sciences Abel Salazar, *Centro Hospitalar do Porto* and *Instituto Português de Oncologia* (IPO) Porto, especially to the all the Professors that directly or indirectly helped throughout this Master, I would like to express my gratitude.

To the people directly involved in this work and my lab colleagues, in particular, to Daniela Silva (at IPO) and Sónia Celeiro (at ICVS), thank you for all the support.

To all my friends and colleagues with whom I have shared so many great moments for the past years.

Finally, a special word to my family, my guiding star, my mum, Valter and, in particular, Margarida, you mean the world to me.

Resumo

Introdução: O cancro do pulmão é a neoplasia com maior taxa de mortalidade e uma das mais incidentes a nível global. Há dois grupos principais, o carcinoma de pequenas células e o carcinoma de não pequenas células, que se subdivide, em adenocarcinoma, carcinoma epidermoide e outros subtipos mais raros. O *SPINT2* é um inibidor de proteases que regula moléculas como matriptase e prostasina, e tem sido descrito maioritariamente como gene supressor tumoral. A hipermetilação do gene está descrita como principal responsável pela diminuição dos níveis de proteína e associada a pior prognóstico. Porém, em algumas neoplasias, o seu papel parece ser pró-tumoral, dependendo do tipo de proteases expressas nos tumores. Em carcinoma pulmonar, o SPINT2 foi reportado como supressor tumoral.

Objetivos: Estando a importância clínica do SPINT2 no cancro do pulmão pouco explorada, este trabalho pretendeu estudar este tópico, avaliando a expressão proteica e níveis de metilação do gene e correlacionando com os dados clinico-patológicos dos doentes.

Metodologia: Uma série de 183 carcinomas pulmonares e 19 tecidos pulmonares normais adjacentes a tumor, foram avaliados para metilação do gene *SPINT2* por PCR-quantitativo de metilação em tempo-real. Níveis de expressão da proteína foram avaliados em 170 casos por imunohistoquímica e os de RNA, em 74 amostras, por PCR-quantitativo de transcrição-reversa em tempo-real.

Resultados: Cerca de 74% das amostras apresentaram baixa expressão de SPINT2, havendo uma correlação entre os níveis de RNA e proteína. No entanto, níveis elevados de SPINT2 estavam associados a progressão tumoral e metastização, bem como a uma menor sobrevivência específica de doença. Os níveis de metilação do *SPINT2* estavam mais elevados no tecido normal adjacente do que nos tumores, sendo que menores níveis de metilação estavam associados a um fenótipo mais agressivo. Não se observou uma correlação entre os níveis de metilação de *SPINT2* e a expressão de RNA ou proteína.

Conclusões: Globalmente, os resultados revelam um papel pró-tumoral do SPINT2 em carcinoma pulmonar. Estes resultados necessitam de validação e uma análise *in silico* independente, e uma avaliação da expressão do SPINT2 em tecido

pulmonar não tumoral seriam importantes. Observou-se que a expressão de SPINT2 poderá ter um papel prognóstico, propondo-se a avaliação da sua utilidade clínica como biomarcador em plasma de doentes. Ensaios funcionais para avaliação das proteases reguladas pelo SPINT2 e caracterização do seu papel funcional na tumorigénese e resposta à terapia, seriam fundamentais para um conhecimento abrangente desta proteína em cancro do pulmão.

Palavras-Chave: Biomarcador; Metilação do DNA; Cancro; Pulmão; Inibidor de Proteases

Abstract

Introduction: Worldwide, lung cancer is the leading cause of cancer-related death and one of the most incident. There are two major groups of lung cancer, small cell lung cancer and non-small cell lung cancer, which, in turn, comprises adenocarcinoma and squamous-cell carcinoma, as well as other rarer subtypes. SPINT2 is a protease inhibitor that regulates molecules such as matriptase and prostasin, and has been described mostly as a tumor suppressor gene. The hypermethylation of the gene is described as main cause for the decrease on protein levels and is associated with worse prognosis. However, in some neoplasms, SPINT2 role seems to be pro-tumoral, depending on the type of proteases expressed in the tumors. In lung cancer, a single recent article reports SPINT2 as a tumor suppressor.

Aims: Since the role of SPINT2 in lung cancer is poorly studied, this work aimed to explore the clinical relevance of SPINT2 in this tumor type by evaluating the protein expression and methylation levels of the gene and correlating it with the clinical and pathological data of the patients.

Methods: A series of 183 lung carcinomas and 19 normal lung tissues adjacent to the tumor were evaluated for *SPINT2* methylation levels by real-time quantitative methylation-specific PCR. Protein expression levels were evaluated in 170 cases by immunohistochemistry and RNA in 74 samples using real-time quantitative reverse-transcription-PCR.

Results: About 74% of the samples disclosed low SPINT2 expression, with a correlation between RNA and protein levels. Higher SPINT2 levels were associated with tumor progression and metastasis, as well as lower disease-specific survival. *SPINT2* methylation levels were higher in adjacent normal tissue than in tumors, and lower levels of *SPINT2* methylation were associated with a more malignant phenotype. No correlation was observed between *SPINT2* methylation levels and the expression of RNA or protein.

Conclusions: Overall, our results reveal a pro-tumoral role of SPINT2 in lung cancer. These results require validation and an independent *in silico* analysis, and the evaluation of SPINT2 expression in non-tumoral lung tissue would also be important. It was seen that SPINT2 expression may have a prognostic role and the

evaluation of its clinical utility as a biomarker in plasma of patients, would be of interest. Functional assays for the evaluation of proteases regulated by SPINT2 and characterization of SPINT2 functional role in tumorigenesis and response to therapy would be critical to a comprehensive understanding of this protein´s role in lung cancer.

Keywords: Biomarker; DNA Methylation; Cancer; Lung; Protease inhibitor

Abbreviations and acronyms

AJCC	American Joint Committee on Cancer
ALK	Anaplastic lymphoma kinase
BRAF	Serine/threonine-protein kinase B-Raf
CpG	Cytosine-phosphate-Guanine
DSS	Disease specific survival
EGFR	Epidermal growth factor receptor
EMT	Epithelial-to-mesenchymal transition
FFPE	Formalin-fixed paraffin-embedded
GUSB	Glucuronidase, Beta
HAI-2	Hepatocyte growth factor activator inhibitor type-2
HGF	Hepatocyte growth factor
HGFA	Hepatocyte growth factor-activator
KOP	Kunitz domain-containing protein overexpressed in pancreatic cancer
LCa	Lung Cancer
MET	MET proto-oncogene
mRNA	messenger RNA
MSP	Methylation Specific PCR
NSCLC	Non-small cell lung cancer
PCR	Polymerase chain reaction
PD-1	Programed cell death protein 1
PD-L1	Programed cell death-ligand 1
PRSS8	Serine protease 8 or Prostasin
qMSP	Quantitative methylation specific PCR
qRT-PCR	Quantitative Reverse Transcriptase PCR
RAF	Rapidly Accelerated Fibrosarcoma
ROS	Proto-oncogene tyrosine-protein kinase ROS
SCC	Squamous cell carcinoma
SCLC	Small cell lung cancer
SPINT2	Serine protease inhibitor Kunitz type 2
TCGA	The cancer genome atlas
TKi	Tyrosine kinase inhibitor
TTF1	Thyroid transcription factor 1
WHO	World health organization

Index

Introduction

Worldwide, lung cancer (LCa) is the most commonly diagnosed cancer, in both sexes combined, and the leading cause of cancer related death[1]. In 2018, in developed countries, including Southern Europe, among males it also represented the second most frequently diagnosed cancer, after prostate cancer and the leading cause of cancer related death[1,2]. In females, it is the third most frequently diagnosed cancer, after breast and colorectal cancer[1]. The overall 5-year survival rate for LCa is low (18% when including all cancer stages) and has presented only minimal changes over decades[3]. In spite of this, the recent advances on surgical techniques, and available targeted therapies and immunotherapies might improve these numbers in future reports[3].

The most frequently diagnosed cancers and the leading cause of cancer-related deaths differ across countries and within regions of the same country, depending on the economic developmental status and are highly associated with social and life style factors[1]. In fact, the major risk factor for LCa is behavioral, specifically through smoking habits and it is known that declines in smoking prevalence reduce LCa incidence and mortality. Furthermore, environmental and genetic factors may play a role in tumor development, influencing also individual capacity for therapeutical and immunological response[4-7].

There are two major groups of LCa according to the microscopic aspect of the cells: small-cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC), which also exhibit distinct clinical behavior[8]. NSCLC is more frequent, accounting for about 85% of cases of LCa, whereas the proportion of patients with SCLC has decreased to 13% in Western world[9,10].

SCLC originates from neuroendocrine-cell precursors, mostly staining positive for neuroendocrine markers, such as chromogranin A, synaptophysin, and CD56[11]. Patients are typically old men and virtually all have a history of tobacco use[11]. These tumors disclose a central and bulky location on chest radiography, rapid tumor growth, high initial response rates to chemotherapy and ultimately development of treatment resistance, as well as high rates of tumor recurrence and metastases[10,11]. Unlike NSCLC, these tumors are not associated with a specific somatic mutation.

As such, the standard regimen for SCLC has not evolved much in the past years, consisting of chemotherapy with a platinum agent (cisplatin or carboplatin) combined with etoposide[11]. Recently, the addition of the antibody against programed cell death-ligand 1 (PD-L1), atezolizumab, to chemotherapy as first-line therapy for advanced SCLC has significantly improved median survival, from 10.3 to 12.3 months, and progression-free survival[12].

NSCLC usually grows and disseminates more slowly compared to SCLC, nevertheless it represents the leading cause of cancer-related death[2,9]. NSCLC is further subtyped as adenocarcinoma or squamous cell carcinoma (SCC), according to the morphologic features or immunohistochemical specific staining that favor the classification in one subtype (p40 or p63 to predict SCC and thyroid transcription factor 1 (TTF1) to predict adenocarcinoma)[13], as well as other less common subtypes. Complete histological diagnosis and molecular testing currently allows for individual treatment decisions. In fact, NSCLC can be categorized according to molecular changes that predict sensitivity to selective pathway-directed systemic therapies (targeted therapies) such as tyrosine kinase inhibitors (TKi)[14].

Targeted therapies have revolutionized the treatment and prognosis of metastatic NSCLC. Key oncogenic alterations (driver mutations), have been identified in NSCLC, with some of them, *epidermal growth factor receptor* (*EGFR*) and *Serine/threonine-protein kinase B-Raf* (*BRAF*) mutations, *Anaplastic lymphoma kinase* (*ALK*) and *Proto-oncogene tyrosine-protein kinase ROS* (*ROS1*) rearrangements determining approved targeted therapies[14-18]. Also, immunotherapy has been approved for advanced NSCLC, with Nivolumab and Pembrolizumab targeting programed cell death protein 1 (PD-1) and Atezolizumab and Durvalumab targeting PD-L1, with promising results[19].

Several drugs targeting EGFR have been approved, namely Gefitinib, Erlotinib, Afatinib, Osimertinib, Dacomitinib and Necitumumab; for *BRAFV600E* mutations Dabrafenib and Trametinib have been approved; Alectinib, Ceritinib, Brigatinib and Loratinib approved for *ALK* translocations and Crizotinib approved for *ALK* and ROS1 rearrangements; however, resistance to these drugs is still a major concern, hampering the survival of these patients[14]. Amplification of *MET proto-oncogene* (*MET*) in NSCLC arises primarily as a mechanism of resistance to treatment directed

to EGFR[16]. Tumors that show *MET* amplification are initially sensitive to inhibition of this signaling pathway, but eventually develop resistance to MET-targeted therapies, and the combination of EGFR and MET-targeted therapies has demonstrated limited efficacy[20,21]. Tumor microenvironment may play an important role in the acquisition of resistance to the targeted therapies[16]. In melanoma, the secretion of hepatocyte growth factor (HGF) by surrounding stromal cells, was demonstrated as a mechanism of resistance to rapidly accelerated fibrosarcoma (RAF) inhibitors[22].

Serine protease inhibitor Kunitz type 2 (*SPINT2*), also known as *hepatocyte growth factor activator inhibitor type-2* (*HAI-2*) or bikiunin, has been mainly reported as a tumor suppressor gene, with hypermethylation of its promoter being the main inactivation mechanism described in the literature. Decreased expression of SPINT2 and hypermethylation of its promoter are usually associated with cancer progression and poor prognosis[23-33], although some evidence suggests that SPINT2 role in cancer might be tissue specific[34].

SPINT2 protein inhibits serine proteases, many of them, involved in the degradation of the extracellular matrix, such as hepsin, matriptase, trypsin and plasmin, hindering tumor progression and implicating SPINT2 deregulation in epithelial-to-mesenchymal transition (EMT)[25,35-39]. EMT facilitates cell detachment, invasion and metastasis, hampering cancer patients' survivals. SPINT2 is also known to inhibit HGF-activator (HGFA), thereby regulating the HGF/MET signaling pathway. This signaling pathway plays a crucial role in angiogenesis, growth, motility and cell invasion. Pro-HGF is secreted in the inactive form and its activation into HGF depends on cleavage by the serine protease HGFA, the activity of which is regulated by SPINT2[23,39-42]. As previously stated, tumor microenvironment may play a role in the acquisition of resistance to the targeted therapies through HGF secretion from stromal cells. SPINT2, as an inhibitor of HGF activation, may have a role in this regulation. A molecule developed to inhibit HGF activation used in combination with MET inhibitors, had overcome resistance to the latter in NSCLC cell lines with *MET* amplification. This might indicate SPINT2 as a potential biomarker of response to therapy in these tumors[43].

The role of SPINT2 in LCa is scarcely explored. A single recent article has shown that, also in LCa, SPINT2 under-expression was associated with tumor progression

and that SPINT2 has a role on repressing NSCLC cell motility via plasmin inhibition[44]. Still, these results need validation and it remains to explore whether *SPINT2* hypermethylation is present in NSCLC.

Objectives

Considering that SPINT2 plays a role in the genesis and progression of several solid tumors, and that SPINT2 dysregulation in LCa is still scarcely studied, this study aimed to further explore the clinical relevance of SPINT2 in LCa. Specifically, the main goals of this study were:

1) Characterize SPINT2 protein expression in a cohort of LCa samples using immunohistochemistry;

2) Assess the methylation status of *SPINT2* gene promoter in LCa and correlate with transcript and protein expression;

3) Correlate deregulation of SPINT2 with standard clinical and pathological parameters to assess a putative biomarker value.

Material and Methods

Tissue samples

Tissue samples and data were collected from 183 patients with LCa diagnosed and treated at IPO-Porto, Portugal, between 2001 and 2017. All samples were reviewed and histologically classified according to the World Health Organization (WHO) guidelines[13] and tumor staged according to the American Joint Committee on Cancer (AJCC) classification[45]. This series comprised 173 NSCLC and 10 SCLC samples. Tumor adjacent normal tissue was also obtained from 19 patients. While DNA was available for methylation study of all samples, RNA was available only from 74 NSCLC which were diagnosed or with recurrence after 2015 and slides for immunohistochemistry were available for 170 NSCLC samples.

In order to preserve patients' anonymity, a numerical code was assigned to each patient, with researchers not having access to the identification of any of the study participants. The information collected in the database constructed referred to a set of clinical data, preoperative examinations, surgical reports, pathological data and follow-up of patients. The study protocol was approved by the Ethics Committee of IPO Porto (CES. 150/018).

Nucleic acids isolation and DNA bisulfite treatment

DNA and RNA were previously isolated from the tissue samples. Briefly, DNA and RNA were extracted from formalin-fixed paraffin-embedded (FFPE) tissue sections using Norgen's FFPE RNA/DNA Purification Plus Kit (Norgen, Canada), according to manufacturer's instructions. After elution, DNA and RNA concentrations were measured by NanoDrop Spectrophotometer (NanoDrop Technologies, USA) and the RNA quality measured in a 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA).

Treatment of 50ng of DNA with sodium bisulfite was performed using the EZ DNA Methylation Golf Kit (Zymo Research Corporation, Irvine, California, USA) according to the manufacturer's instructions.

Real-time quantitative polymerase chain reaction

For the real-time quantitative methylation-specific polymerase chain reaction (qMSP), primers were specifically designed to bind the bisulfite converted DNA of the CpG islands of *SPINT2* gene. Primers were designed using Methyl Primer Express 1.0, checked using NetPrimer (http://www.premierbiosoft.com) and purchased from Sigma-Aldrich (St. louis, MO, USA). For quantitative reverse-transcriptase-PCR qRT-PCR primers were already available in the laboratory. Primers sequence and annealing temperatures are provided in Table I.

PCR reactions were performed in triplicates in a volume of 10μL, consisting of 5μL of Xpert Fast SYBR 2X-Mastermix (Grisp, Portugal); 0.4μL of 10μM forward and reverse primers; 2.6μL of water, and 2μL of bisulfite modified DNA, as a template. Amplifications were carried out at: 95˚C for 10 minutes, followed by 45 cycles at 95˚C for 15 seconds and 60˚C for 1 minute, in a 384-well plate using Roche LightCycler 480 II (Roche, Germany). Each plate included five serial dilutions (dilution factor of 5) of methylated bisulfite modified universal DNA control (*in vitro* methylated human DNA, Chemicon) or multiple non-template controls and 5 serial dilutions (1:10) of a cDNA from Human Reference Total RNA (Agilent Technologies), for qMSP and qRT-PCR, respectively, in order to construct a standard curve and multiple water blanks, as negative controls. The PCR product samples were run on a 3% agarose gel, to confirm the size of the PCR product. The relative level of methylated DNA or RNA of *SPINT2* in each sample were determined using β-*Actin* for qMSP and Glucuronidase, Beta (*GUSB*) for qRT-PCR internal reference genes.

Table I - Primer sequences and PCR product size.

	Forward	Reverse	Size (bp)
SPINT2_qMSP	5'-GTAAGAACGGGGTGAGGAGTC-3'	5´-TTCCGAATCTAAAACGACCAC-3'	143
β-ACTIN_qMSP	5'-TGGTGATGGAGGAGGTTTAGTAAGT-3'	5´-AACCAATAAAACCTACTCCTCCCTTAA-3'	130
SPINT2_qRT-PCR	5'-AACAGCAATAATTACCTGACC-3'	5'-AAGGATGCACGGCAAGGC-3'	218
GUSB_qRT-PCR	5'-CTCATTTGGAATTTTGCCGATT-3'	5´-CCGAGTGAAGATCCCCTTTTTA-3'	81

bp-base pairs;

Immunohistochemistry

Tissue sections (3-µm-thick) were cut from 170 NSCLC samples for the immunohistochemical study. Normal placenta was used as positive control and negative control consisted on omission of the primary antibody. Immunohistochemistry was performed using the anti-human SPINT2 polyclonal antibody (clone HPA011101, 1:500 dilution, Sigma-Aldrich) and the streptavidin-biotin peroxidase complex system (UltraVision Large Volume Detection System Anti-Polyvalent, HRP; LabVision Corporation, Fremont, CA), according to the manufacturer's instructions.

The intensity of SPINT2 staining was semi-quantitated by intensity scoring: 0-negative; 1-weak; 2-moderate; 3-strong. Tumor tissues disclosing scores 0 or 1+ were grouped and classified as low SPINT2 immunoexpression and samples with scores 2+ and 3+ as high SPINT2 immunoexpression levels.

Statistical analysis

All analyses were performed using SPSSv25 (IBM SPSS Inc., Chicago, IL). The relationship between methylation ratios or expression levels and standard clinical-pathological variables were evaluated using Chi-Square ($\chi 2$) and Mann-Whitney (pairwise comparisons) or Kruskal-Wallis (multiple groups) nonparametric correlation tests, when appropriate. Analysis of disease-specific survival (DSS) were performed using the Kaplan-Meier method and the log-rank test. The threshold used for statistical significance was set at $p<0.05$.

Results

Characterization of the clinical samples

A total of 202 samples were used in this work, 183 from LCa samples, diagnosed between 2001 and 2017, and 19 from "normal" adjacent tissues. Among the 183 tumor samples, 173 (94.5%) were from NSCLC and 10 (5.5%) from SCLC.

The clinical-pathological data of the patients, at the time of diagnosis, is summarized on Table II. Of the patients, 142 (77.6%) were male and 41 (22.4%) were female, with a median age of 64 years (ranging from 29 to 88 years of age). All SCLC were diagnosed at an advanced stage of disease (AJCC, stage III or IV), whereas in NSCLC the pathological stage at diagnosis was more uniformly distributed. Furthermore, SCLC were larger at diagnosis (median of 8cm) when compared to NSCLC (median between 3.3 and 4.5 cm). Not surprisingly, the large majority (76.7%) of LCa patients included smoking habits.

Table II – Clinical and Histopathological characterization of the Lung Cancer patients.

Clinicopathological characteristics	Lung Cancer Patients	NSCLC		SCLC
		Adenocarcinoma	SCC	
Patients, n	183	92	81	10
Gender, n (%)	183	92	81	10
Male	142 (77.6%)	57 (62%)	79 (97.5%)	6 (60%)
Female	41 (22.4%)	35 (38%)	2 (2.5%)	4 (40%)
Median Age, years (range)	64 (29-88)	65 (42-84)	63 (29-88)	65 (51-72)
Pathological stage, n (%)	179	91	81	7
Stage I	64 (35.8%)	36 (39.6%)	28 (34.6%)	-
Stage II	38 (21.2%)	13 (14.3%)	25 (30.9%)	-
Stage III	35 (19.6%)	17 (18.7%)	17 (21.0%)	1 (14.3%)
Stage IV	42 (23.5%)	25 (27.5%)	11 (13.6%)	6 (85.7%)
Differentiation, n (%)	113	56	57	n.a
Well	14 (12.4%)	12 (21.4%)	2 (3.5%)	n.a
Moderate	73 (64.6%)	33 (58.9%)	40 (70.2%)	n.a
Poor	26 (23%)	11 (19.6%)	15 (26.3%)	n.a
Smoking, n (%)	181	92	79	9
(median package/year, range)	(46, 0-140)	(37.5, 0-120)	(60, 0-140)	(n.a)
Yes	138 (76.2%)	56 (60.9%)	74 (93.7%)	7 (77.8%)
No	43 (23.8%)	36 (39.1%)	5 (6.3%)	2 (22.2%)
Tumor size, n	177	92	80	5
Median (range)	3.8 (1.0-12.5)	3.3 (1.0-10.0)	4.5 (1.0-12.5)	8.0 (6-9.9)

NSCLC – Non-Small Cell Lung Cancer; SCC – Squamous Cell Carcinoma; SCLC – Small Cell Lung Cancer; n – number of samples; n.a – not available;

The molecular and targeted therapy characterization of the NSCLC is summarized in Table III. The characterization of *EGFR* mutations, *ALK* and *ROS1* rearrangements was not performed in all samples, with the largest characterization completed for *EGFR*, with a total of 108 samples studied, 78 adenocarcinoma and 30 SCC. Only adenocarcinoma samples presented alterations in these genes, no alterations were observed on the 54 samples analyzed for *ROS1*; *EGFR* mutations were present in 22 patients (28.2%) and *ALK* rearrangements in 3/52 (5.5%) patients. Targeted therapies were used in 15/172 (8.7%) of the patients, 13 adenocarcinomas with *EGFR* mutation were treated with erlotinib or gefitinib and one patient with adenocarcinoma and *ALK* rearrangement was treated with crizotinib.

Table III – Molecular characterization and targeted therapies of the NSCLC patients.

Molecular and Targeted Therapies Characteristics	NSCLC (total)	Adenocarcinoma	SCC
Patients, n	173	92	81
EGFR mutation, n (%)	108	78	30
Yes	22 (20.4%)	22 (28.2%)	0 (0%)
No	86 (79.6%)	56 (71.8%)	30 (100%)
ALK rearrangements, n (%)	55	52	3
Yes	3 (5.5%)	3 (5.8%)	0 (0%)
No	52 (94.5%)	49 (94.2%)	3 (100%)
ROS1 rearrangements, n (%)	54	51	3
No	54 (100%)	51 (100%)	3 (100%)
Targeted therapy, n (%)	172	92	80
No	157 (91.3%)	78 (84.8%)	79 (98.8%)
Erlotinib	13 (7.6%)	12 (13%)	1 (1.3%)
Gefitinib	1 (0.6%)	1 (1.1%)	0 (0%)
Crizotinib	1 (0.6%)	1 (1.1%)	0 (0%)
Anti-PD1 treatment, n (%)	172	92	80
No	165 (95.9%)	86 (93.5%)	79 (98.8%)
Pembrolizumab	6 (3.5%)	5 (5.4%)	1 (1.3%)
Atezolizumab	1 (0.6%)	1 (1.1%)	0 (0%)

NSCLC – Non-Small Cell Lung Cancer; SCC – Squamous Cell Carcinoma; n – number of samples

SPINT2 expression levels in NSCLC

SPINT2 expression was evaluated in 170 NSCLC samples by immunohistochemistry. A total of 126 (74.1%) samples displayed low SPINT2 expression and 44 (25.9%) presented high SPINT2 levels. Representative examples of each staining level are represented on Figure 1. The staining was mostly present on the cellular membrane

and/or cytoplasm (Figure 1).

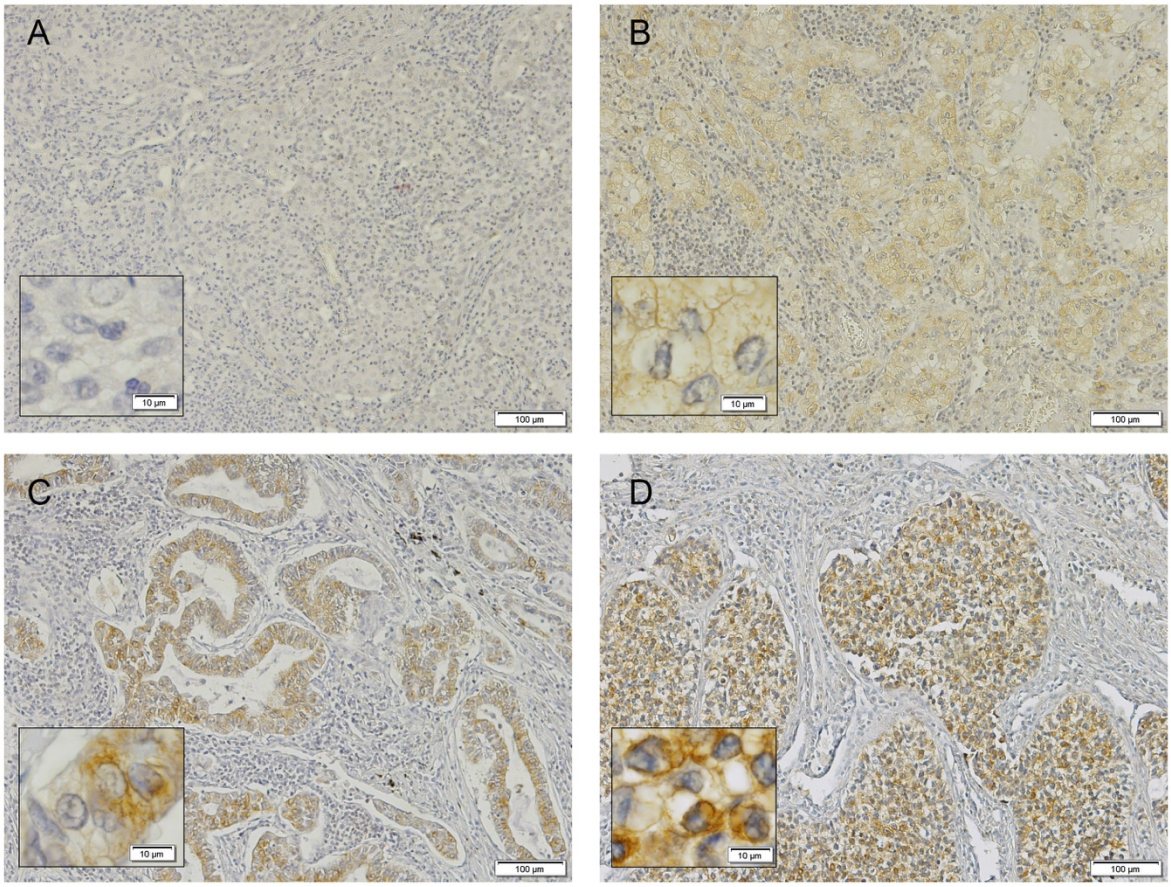


Figure 1. Representative staining for SPINT2 protein expression in NSCLC tissues showing A) SPINT2 no expression (intensity score 0); B) SPINT2 low expression (intensity score 1+); C) SPINT2 moderate expression (intensity score 2+) and D) SPINT2 high expression (intensity score 3+) samples. Original magnification 100x (Scale 100µm); insets 600x (Scale 10µm).

In NSCLC, higher SPINT2expression was associated with adenocarcinoma histological subtype (p=0.009) (Table IV). Moreover, high SPINT2 levels were associated with tumor progression, particularly with higher AJCC pathological stages when taking into account all four stages (p=0.036) or when comparing initial (I+II) with advanced (III+IV) stages (p=0.012), and also with the presence of metastasis (p=0.004), in particular brain metastasis (p=0.040) (Table IV). When evaluating NSCLC subtypes, adenocarcinoma presented significant associations with high SPINT2 levels associating with tumor progression, namely pathological stages, when considering four stages (p=0.013) or when comparing initial vs. advanced stages (p=0.014), and with the presence of metastasis at diagnosis (p=0.001), particularly liver metastasis (p=0.043) (Table IV). No other associations were disclosed, including with patients' gender or smoking habits, tumor stage, presence of lymph node invasion, differentiation grade or recurrence (Table IV).

Also, no associations were found with other metastatic locations at diagnosis, particularly, bone, pleural or adrenal metastasis (Supplementary Table I).

Table IV – Associations between SPINT2 expression and Clinicopathological features of NSCLC

Clinicopathological features	NSCLC (total)		*P* value	Adenocarcinoma		*P* value	SCC		*P* value
	Low SPINT2	High SPINT2		Low SPINT2	High SPINT2		Low SPINT2	High SPINT2	
Patients, n	126	44		60	31		66	13	
Subtype, n (%)	126	44							
Adenocarcinoma	60 (65.9%)	31 (34.1%)	**0.009**	-	-		-	-	
SCC	66 (83.5.4%)	13 (16.5%)		-	-		-	-	
Gender, n (%)	126	44		60	31		66	13	
Male	100 (75.2%)	33 (24.8%)	0.546	36 (64.3%)	20 (35.7%)	0.675	64 (83.1%)	13 (16.9%)	1.000[a]
Female	26 (70.3%)	11 (29.7%)		24 (68.6%)	11 (32.4%)		2 (100%)	0 (0.0)	
Smoking, n (%)	124	44		60	31		64	13	
Yes	94 (74.0%)	33 (26.0%)	0.915	35 (63.6%)	20 (36.4%)	0.568	59 (81.9%)	13 (18.1%)	0.582[a]
No	30 (73.2%)	11 (26.8%)		25 (69.4%)	11 (30.6%)		5 (100%)	0 (0.0%)	
Tumor stage, n (%)	125	44		60	31		65	13	
T1+T2	90 (75.6%)	29 (24.4%)	0.446	50 (70.4%)	21 (29.6%)	0.089	40 (83.3%)	8 (16.7%)	1.000
T3+T4	35 (70.0%)	15 (30.0%)		10 (50.0%)	10 (50.0%)		25 (83.3%)	5 (16.7%)	
Lymph node invasion, n (%)	125	44		60	31		65	13	
N0	71 (74.0%)	25 (26.0%)	0.998	33 (68.8%)	15 (31.3%)	0.549	38 (79.2%)	10 (20.8%)	0.212
N1+N2+N3	54 (74.0%)	19 (26.0%)		27 (62.8%)	16 (37.2%)		27 (90.0%)	3 (10.0%)	
Distant metastasis, n (%)	122	44		60	31		62	13	
M0	103 (78.6%)	28 (21.4%)	**0.004**	50 (75.8%)	16 (24.2%)	**0.001**	53 (81.5%)	12 (18.5%)	1.000[a]
M1	19 (54.3%)	16 (45.7%)		10 (40.0%)	15 (60.0%)		9 (90.0%)	1 (10.0%)	
Brain metastasis, n (%)	125	44		60	31		65	13	
No	114 (76.5%)	35 (23.5%)	**0.040**	52 (70.3%)	22 (29.7%)	0.069	62 (82.7%)	13 (17.3%)	1.000[a]
Yes	11 (55.0%)	9 (45.0%)		8 (47.1%)	9 (52.9%)		3 (100%)	0 (0.0%)	
Liver metastasis, n (%)	125	44		60	31		65	13	
No	118 (75.2%)	39 (24.8%)	0.303[a]	58 (69.0%)	26 (31.0%)	**0.043**[a]	60 (82.2%)	13 (17.8%)	0.583[a]
Yes	7 (58.3%)	5 (41.7%)		2 (28.6%)	5 (71.4%)		5 (100%)	0 (0.0%)	
Pathological stage, n (%)	125	44		59	31		66	13	
Stage I	48 (77.4%)	14 (22.6%)		26 (74.3%)	9 (25.7%)		22 (81.5%)	5 (18.5%)	
Stage **II**	33 (86.8%)	5 (13.2%)	**0.036**	11 (84.6%)	2 (15.4%)	**0.013**	22 (88.0%)	3 (12.0%)	0.088[b]
Stage **III**	23 (69.7%)	10 (30.3%)		12 (70.6%)	5 (29.4%)		11 (68.8%)	5 (31.3%)	
Stage IV	21 (58.3%)	15 (41.7%)		10 (40.0%)	15 (60.0%)		11 (100%)	0 (0.0%)	
Advanced stages, n (%)	125	44		59	31		66	13	
Stages I+II	81 (81.0%)	19 (19.0%)	**0.012**	37 (77.1%)	11 (22.9%)	**0.014**	44 (84.6%)	8 (15.4%)	0.755[a]
Stages III+IV	44 (63.8%)	25 (36.2%)		22 (52.4%)	20 (47.6%)		22 (81.5%)	3 (18.5%)	
Differentiation, n (%)	93	17		45	10		48	7	
Well	10 (76.9%)	3 (23.1%)		8 (72.7%)	3 (27.3%)		2 (100%)	0 (0.0%)	
Moderate	65 (89.0%)	8 (11.0%)	0.201[b]	30 (90.9%)	3 (9.1%)	0.094	35 (87.5%)	5 (12.5%)	0.732[b]
Poor	18 (75.0%)	6 (25.0%)		7 (63.6%)	4 (36.4%)		11 (84.6%)	2 (15.4%)	
Recurrence, n (%)	126	44		60	31		66	13	
Yes	36 (78.3%)	10 (21.7%)	0.453	15 (75.0%)	5 (25.0%)	0.333	21 (80.8%)	5 (19.2%)	0.749[a]
No	90 (72.6%)	34 (27.4%)		45 (63.4%)	26 (36.6%)		45 (84.9%)	8 (15.1%)	

NSCLC – Non-Small Cell Lung Cancer; SCC – Squamous Cell Carcinoma; n – number of samples (% of SPINT2 expression)
P value based on Pearson $\chi 2$ test, except [a] Fisher's exact test and [b] Likelihood Ratio, due to violation of $\chi 2$ assumptions.

The correlation between patients' DSS and SPINT2 expression in NSCLC was evaluated using Kaplan-Meier curves (Figure 2). Interestingly, high SPINT2 levels were correlated with lower DSS on adenocarcinoma patients, log-rank $p<0.001$ (Figure 2A), whereas the curves for SCC demonstrated the opposite trend, low SPINT2 levels associating with lower DSS, although the differences were not significant, log-rank p=0.284 (Figure 2B).

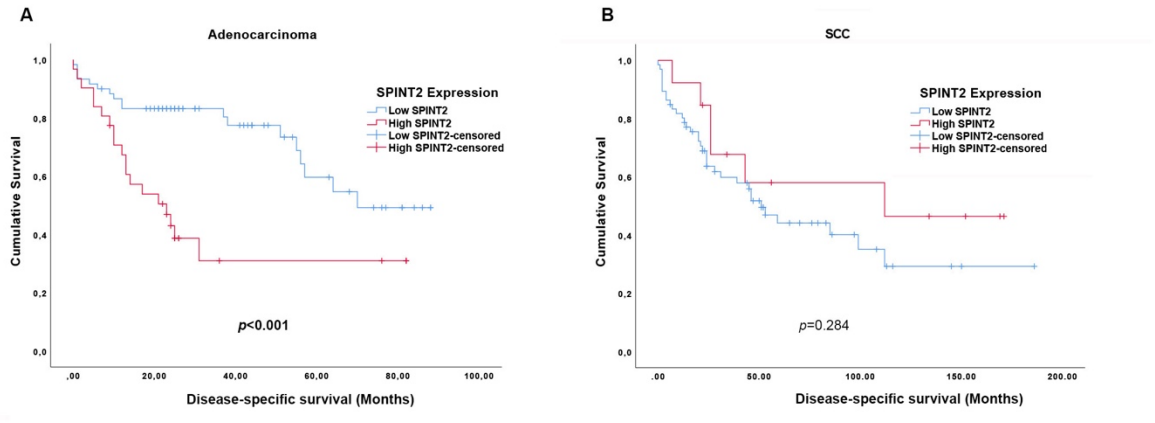


Figure 2. Disease-specific survival according to SPINT2 protein levels for a) Adenocarcinoma and B) Squamous Cell Carcinoma (SCC). Presented *p* values from log-rank tests.

The evaluation of *SPINT2* expression at RNA level, was also possible in a subset of NSCLC patients, namely those diagnosed or with recurrence after 2015. We found a positive association between RNA and protein levels on NSCLC (Figure 3, p=0.012).

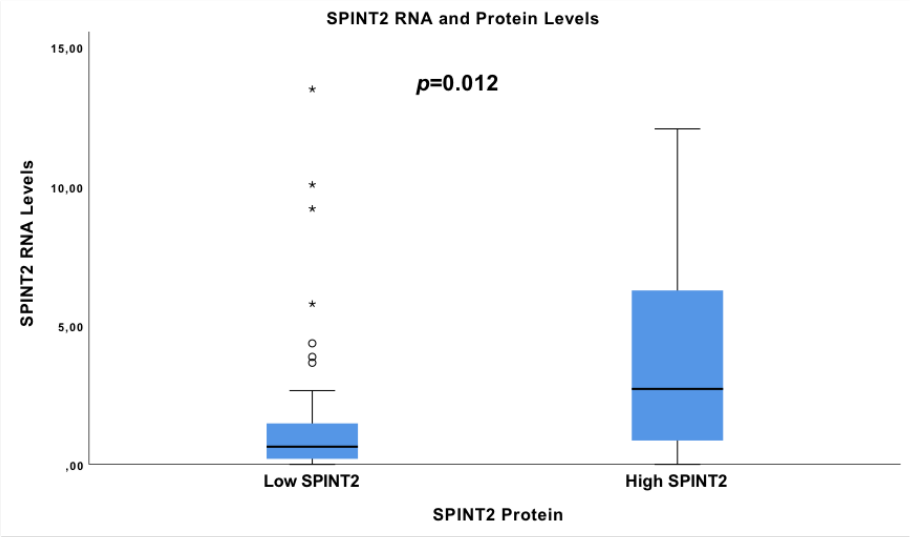


Figure 3. Boxplots of *SPINT2* RNA and protein levels for NSCLC. Presented *p* value from Mann-Whitney Test.

Evaluation of *SPINT2* gene methylation levels in LCa Patients

The levels of *SPINT2* gene methylation were significantly higher in controls compared to LCa (p<0.001, Figure 4A). When comparing LCa subtypes (NSCLC vs SCLC), there was a tendency for higher methylation levels in NSCLC (p=0.073, Figure 4B).

Significant associations were also found between lower *SPINT2* promoter methylation levels and more aggressive clinicopathologic features, specifically tumor stage (p<0.000), distant metastasis (p=0.001), pathological stage (p=0.002) and when comparing initial (I+II) with advanced (III+IV) stages (p=0.005) (Figure 5).

Similar results were obtained in the analysis of NSCLC, with lower *SPINT2* promoter methylation levels correlating with more aggressive tumor features, namely tumor stage (p=0.001), distant metastasis (p=0.005), pathological stage (p=0.009) and comparing initial (I+II) with advanced (III+IV) stages (p=0.012) (Figure 6).

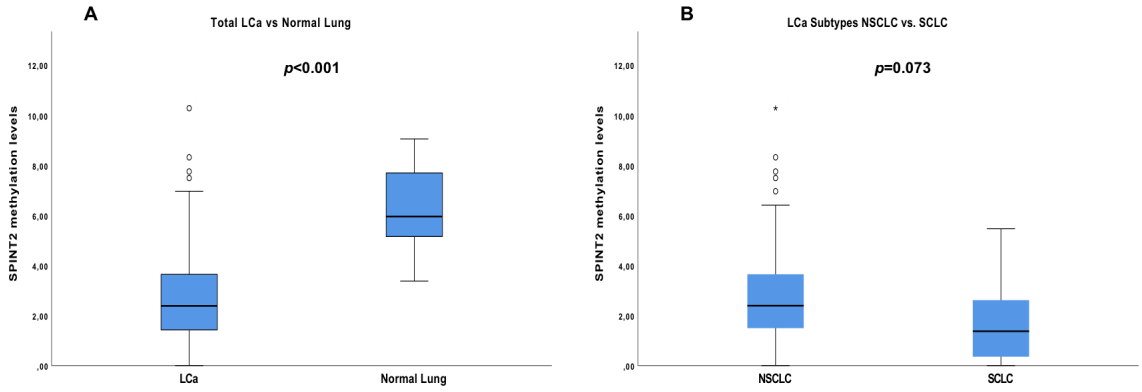


Figure 4. Boxplots of *SPINT2* methylation levels for A) LCa and Normal Lung and B) NSCLC and SCLC. Presented p values from Mann-Whitney Test.

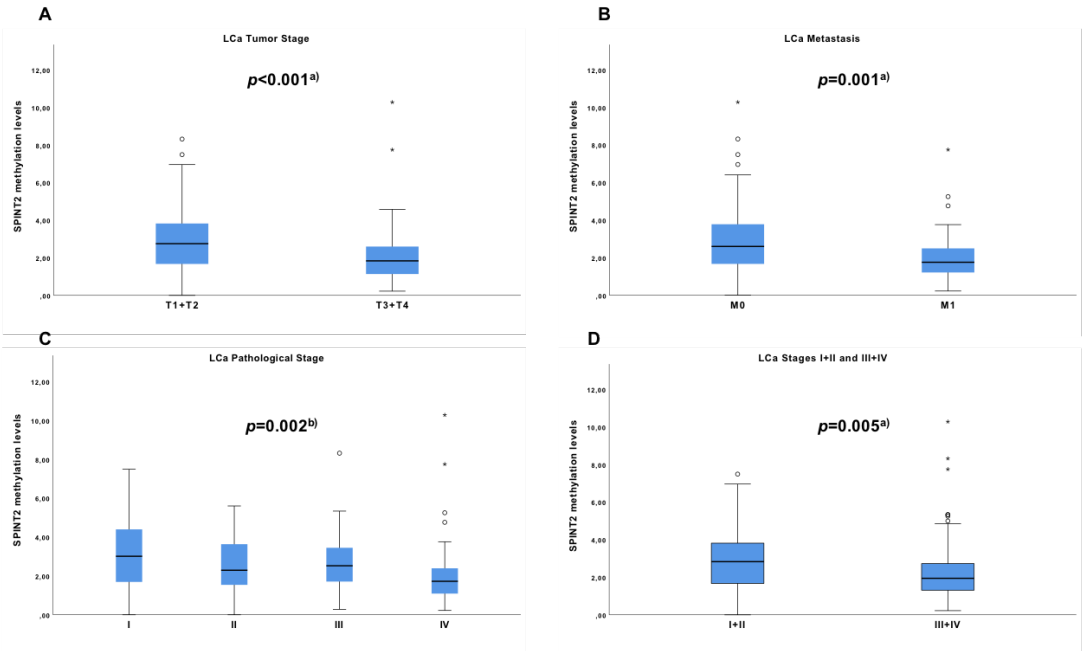


Figure 5. Boxplots of association between *SPINT2* promoter methylation levels and A) Tumor Stage B) Metastasis C) Pathological Stage in four and D) two groups in LCa. Presented p values from [a)]Mann-Whitney or [b)]Kruskal-Wallis Tests.

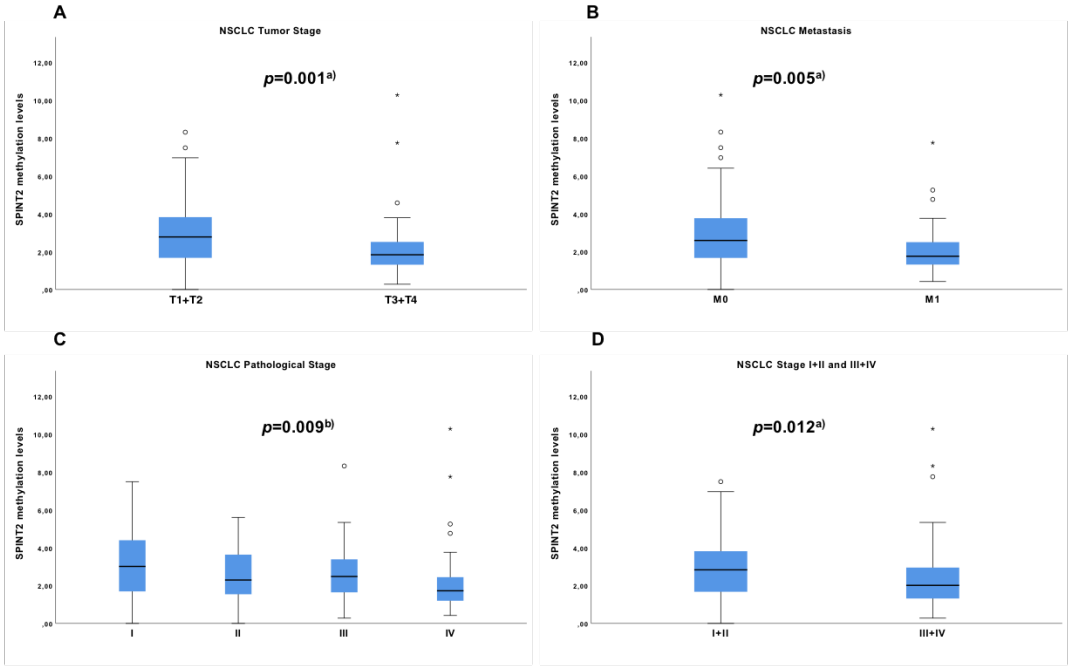


Figure 6. Boxplots of association between *SPINT2* promoter methylation levels and A) Tumor Stage B) Metastasis C) Pathological Stage in four and D) two groups in NSCLC. Presented *p* values from [a)]Mann-Whitney or [b)]Kruskal-Wallis Tests.

When evaluating NSCLC histological subtypes, for adenocarcinoma, an association between lower *SPINT2* promoter methylation levels and higher pathological stage (I, II, III and IV) was observed (p=0.031, Figure 7A), whereas for SCC there was an association between lower *SPINT2* promoter methylation levels and higher tumor stage (p<0.001), distant metastasis (p=0.026), in particular brain metastasis (p=0.033) and more advanced pathologic stages (I-IV p=0.034 and I+II compared with III+IV, p=0.005) (Figure 7B-F). No significant association was found between *SPINT2* promoter methylation levels and other clinicopathologic features, including gender, lymph node metastasis, tumor differentiation grade or smoking habits.

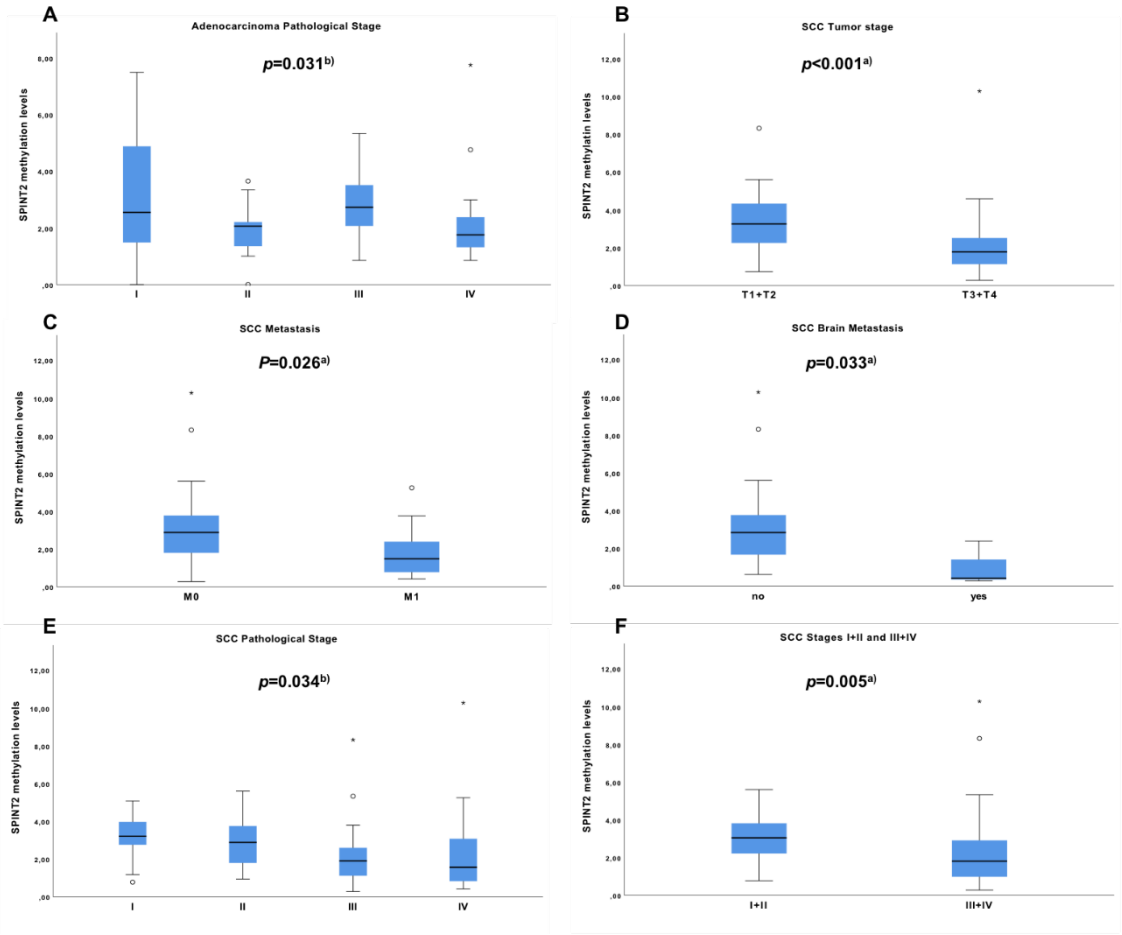


Figure 7. Boxplots of association between *SPINT2* methylation levels and A) Pathological Stage in adenocarcinoma; B) Tumor Stage C) Metastasis D) Brain metastasis E) Pathological Stage in four and F) two groups in NSCLC. Presented *p* values from [a)]Mann Whitney or [b)]Kruskal-Wallis Tests.

We further evaluated whether *SPINT2* promoter methylation was the mechanism responsible for RNA and protein downregulation. Assessing the correlation between mRNA and methylation levels, the linear correlation Spearman r was -0.009654 (p=0.9382) (Figure 8A) and r^2=0.008969. The correlation between gene promoter methylation and protein levels was also not significant, although there was a trend for association between higher levels of gene methylation and lower levels of protein (p=0.081) (Figure 8B), and therefore our results do not suggest an epigenetic regulation of SPINT2 expression through gene hypermethylation.

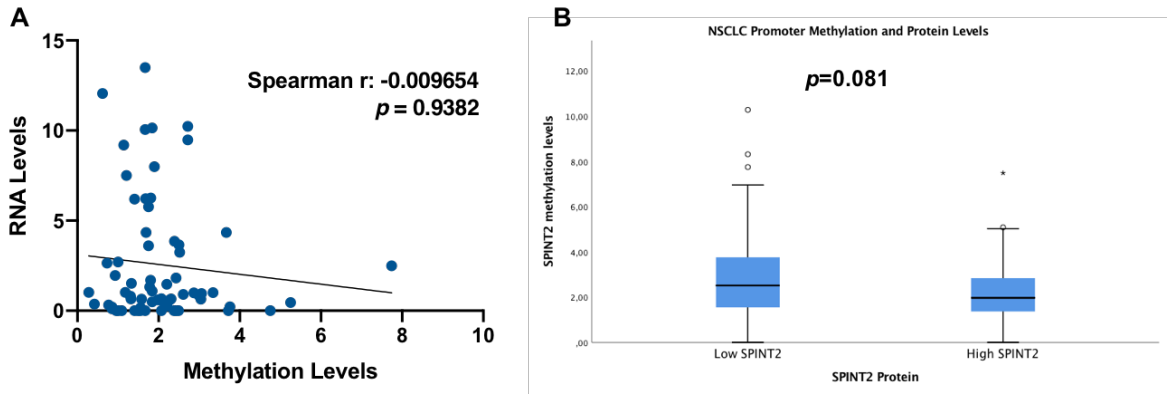


Figure 8. Association between methylation status and A) mRNA expression of *SPINT2* in NSCLC, based on nonparametric Spearman correlation or B) protein levels, *p* values from Mann Whitney Test.

For statistical purposes, *SPINT2* promoter methylation levels were dichotomized into percentiles, with percentile 75 used as threshold value on Figure 9. No significant associations were observed between *SPINT2* methylation and patients' survival, however the median survival of patients with methylation P≤75 was 51 months and the median survival of patients with methylation P>75 was 85 months (p=0.093, log-rank test).

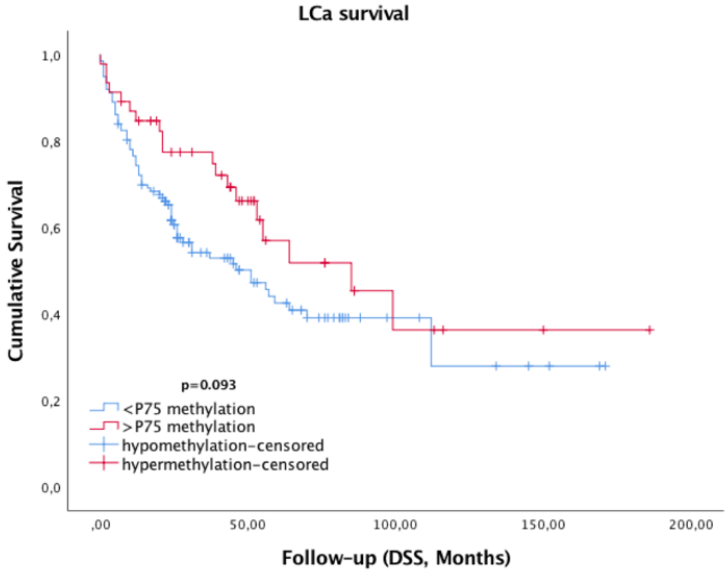


Figure 9. Disease-specific survival according to *SPINT2* methylation levels dichotomized using the percentile 75. Presented *p* value from log-rank test.

Discussion

Lung cancer is a major health issue worldwide, representing one of the most frequently diagnosed cancers and being the leading cause of cancer-related death. LCa is more frequently found in male than in female gender, as we also observed in this study (77.6% male vs. 22.4% female)[1]. As typically described in the literature[9,10], SCLC were larger at diagnosis (SCLC, median of 8cm; adenocarcinoma, median of 3.3cm; and SCC, median of 4.5cm) in our cohort. Also, as expected, the great majority (76.7%) of LCa patients had previous smoking habits, in particular those with SCC (93.7%) and SCLC (77.8%), whereas these numbers were lower for adenocarcinoma (60.9%)[9,10].

The molecular signatures of LCa are complex, particularly in patients with smoking habits, as the acquisition of several mutations promotes tumor complexity and heterogeneity[46]. The understanding of these molecular signatures is essential for the design of efficient targeted therapies and have boosted the treatment of LCa, in particular NSCLC. Nowadays, drugs targeting EGFR, BRAF, ALK and ROS1 molecular alterations are used in the clinics. Some of these molecules are frequently mutated in LCa and their molecular alterations routinely evaluated in these tumors, in particular adenocarcinoma. *EGFR* mutations are one of the most frequently found in NSCLC patients (10–20% of patients or in about 40% if patients are from east Asian origin), mainly in adenocarcinoma, younger women, and non-smokers[14]. In fact, 20% of the NSCLC of our cohort of tested samples, presented EGFR mutations and all mutated samples were adenocarcinoma. *ALK* rearrangements are observed in 2-7% of NSCLCs, younger at diagnosis (median age around 50 years) and mostly men, particularly in never or light smokers with adenocarcinomas[14]. Accordingly, *ALK* was positive in 5.8% of the adenocarcinomas evaluated. *ALK* fusions are thought to be mutually exclusive with molecular alterations on *EGFR*, *ROS1*, and *KRAS*, which was also seen in our samples. *ROS1* rearrangements are present in 1–2% of NSCLC, mostly in adenocarcinomas and predominantly in female, never or light smokers, with a median age at diagnosis around 50 years of age[14]. None of the samples tested for *ROS1* was positive in our series. Targeted therapies against these molecules are used in advanced NSCLC, and among our samples, 13 cases (7.6%) were treated with erlotinib and one with gefitinib Most of these samples were *EGFR*-mutated adenocarcinomas. One of the

patients treated with erlotinib had a SCC although according to our database this sample was not tested for *EGFR*.

Immunotherapeutic agents, in particular monoclonal antibodies directed to the PD-1 receptor (nivolumab, pembrolizumab) or its ligand PD-L1 (atezolizumab, durvalumab, avelumab) have shown promising responses in 14–20% of patients with advanced NSCLC, with atezolizumab also in use for SCLC[12,14]. On our NSCLC series, five adenocarcinomas were treated with pembrolizumab and one with atezolizumab, and one SCC was treated with pembrolizumab, as well.

SPINT2 has been described as a tumor suppressor gene in cancer[33]. A recently published study had also described SPINT2 as a tumor suppressor gene in LCa and described the functional mechanism through which SPINT2, by inhibiting plasmin, a serine protease involved in the extracellular matrix remodeling, could suppress the development of NSCLC invasion and metastasis. Those authors have used a tissue microarray to study SPINT2 expression through immunohistochemistry and a Web resource (SurvExpress) to evaluate the prognostic significance of SPINT2 in adenocarcinoma, showing that lower levels of SPINT2 expression were related with poor prognosis and also with tumor progression[44]. Nevertheless, that study has some limitations, namely immunohistochemistry was performed in adenocarcinoma only and in a small number of samples, 64 cases, with only two of those samples presenting distant metastasis. Moreover, authors presented results for only one dataset (comprising 255 adenocarcinoma) among the 24 available for LCa on SurvExpress, and therefore their results need validation.

Our data points towards a different direction and do not validate Wu *et al*. [44] results. We have shown that reduced expression of SPINT2 was frequently observed in NSCLC (74.1% of the samples), with SPINT2 lower levels more frequently seen in SCC, when compared to adenocarcinoma. In adenocarcinoma, higher levels of SPINT2 were related with aggressive features, namely the presence of distant metastasis, in particular liver metastasis and more advanced pathological stage, and also with reduced DSS. On the other hand, in SCC, there was a trend for correlation between lower levels of SPINT2 and more advanced pathological stage and survival analysis revealed that tumors with low SPINT2 levels disclosed reduced DSS, although the differences were not significant. Although these results need further validation, they point towards a further molecular differentiation between

adenocarcinoma and SCC. The study of a larger series of SCC would be important to clarify this subject.

Because the baseline hypothesis of this study was that SPINT2 would be a tumor suppressor gene in LCa, we investigated whether promoter methylation might be responsible for its downregulation. Unexpectedly, gene promoter methylation levels were higher in normal tissue than in LCa and there was a trend for higher methylation levels in NSCLC compared to SCLC. Lower methylation levels were associated with more aggressive features of LCa, in particular NSCLC, like advanced tumor stage, distant metastasis and in SCC with brain metastasis, reinforcing a putative oncogenic role for SPINT2 in LCa.

Finally, we aimed to understand whether *SPINT2* gene methylation was responsible for RNA or protein downregulation in LCa, but no significant associations were found between these parameters. The CpG-rich region of *SPINT2* comprises over 100 CpG islands and in this study, we have designed new primers appropriate for using on qMSP that include a total of four CpG islands. As we have not found a positive association between gene methylation levels and protein expression and there are no reports describing whether methylation on these specific CpG islands correlate with SPINT2 expression, a bisulfite sequencing of the *SPINT2* CpG-enriched region should be performed in LCa, for a better characterization of *SPINT2* methylation in this tumor type.

Cancer develops from consecutive genetic or epigenetic alterations that lead to gain of function in oncogenes and loss of function in tumor-suppressor genes. However, expression of several genes in cancers may differ based on tissue type and tumor site, with these molecules having multifactorial effects on tumorigenesis. *SPINT2* has mainly been described as a tumor suppressor gene in cancer with gene promoter methylation being associated with aggressive features of several solid tumor types[23-33], nevertheless a few exceptions were also reported in the literature. SPINT2 was named as Kunitz domain-containing protein overexpressed in pancreatic cancer (KOP), and described as being overexpressed in pancreatic cancer as compared to normal pancreas or chronic pancreatitis tissues[47]. Also, in breast cancer, higher levels of SPINT2 correlated with HER2 expression, malignant tumor features and poor treatment response[48], in addition, high levels of SPINT2 were present in primary and metastatic breast cancer cell

lines, with higher levels present in the metastatic lines[49]. Recently, SPINT2 was shown to promote tumor growth and invasion on oral squamous cell carcinoma through inhibition of prostasin and increased levels of SPINT2 expression were seen throughout neoplastic progression of the oral epithelium[34]. Prostasin (*PRSS8*) is poorly characterized in LCa, a single study reports low levels of this protease serine in NSCLC cell lines and that its ectopic expression inhibited tumor growth, migration and invasion of NSCLC cells[50]. Therefore, the roles of SPINT2 in different cancers may be cell or tissue type-specific and might depend on the target proteases co-expressed in the same cells or on the cellular location of SPINT2, membrane and/or cytoplasm locations, and that will require additional investigations[39,51,52].

Overall, our results point towards an oncogenic role of SPINT2 in LCa, yet these results need validation. An independent *in silico* analysis evaluating *SPINT2* expression and methylation levels in normal lung and cancer tissues would be important in this context, as well as the evaluation of SPINT2 expression on a series of non-tumoral lung tissue would be important to better understand the pathological role of SPINT2.

Given that our results point towards a prognostic role of SPINT2 expression in LCa, it would be interesting to isolate the circulating cancer cells from patients' plasma and assess *SPINT2* expression levels to evaluate its role as a prognostic marker.

In addition, functional experiments, with cell lines, to further characterize the proteases regulated by SPINT2 in LCa, in particular prostasin, as well as its tumorigenic role and the influence on therapeutic response, would also be important steps to fully understand the functional role of SPINT2 in LCa.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* 2018; **68**(6): 394-424.

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* 2018; **68**(1): 7-30.

3. Boloker G, Wang C, Zhang J. Updated statistics of lung and bronchus cancer in United States (2018). *J Thorac Dis* 2018; **10**(3): 1158-61.

4. Ma J, Jemal A, Fedewa SA, et al. The American Cancer Society 2035 challenge goal on cancer mortality reduction. *CA Cancer J Clin* 2019.

5. Akhtar N, Bansal JG. Risk factors of Lung Cancer in nonsmoker. *Curr Probl Cancer* 2017; **41**(5): 328-39.

6. Nasim F, Sabath BF, Eapen GA. Lung Cancer. *Med Clin North Am* 2019; **103**(3): 463-73.

7. de Groot PM, Wu CC, Carter BW, Munden RF. The epidemiology of lung cancer. *Transl Lung Cancer Res* 2018; **7**(3): 220-33.

8. World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart. Lyon: IARCPress; 2004.

9. Goldstraw P, Ball D, Jett JR, et al. Non-small-cell lung cancer. *Lancet* 2011; **378**(9804): 1727-40.

10. van Meerbeeck JP, Fennell DA, De Ruysscher DK. Small-cell lung cancer. *Lancet* 2011; **378**(9804): 1741-55.

11. Fruh M, De Ruysscher D, Popat S, Crino L, Peters S, Felip E. Small-cell lung cancer (SCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013; **24 Suppl 6**: vi99-105.

12. Horn L, Mansfield AS, Szczesna A, et al. First-Line Atezolizumab plus Chemotherapy in Extensive-Stage Small-Cell Lung Cancer. *N Engl J Med* 2018; **379**(23): 2220-9.

13. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. *J Thorac Oncol* 2015; **10**(9): 1243-60.

14. Hirsch FR, Scagliotti GV, Mulshine JL, et al. Lung cancer: current therapies and new targeted treatments. *Lancet* 2017; **389**(10066): 299-311.

15. Rosell R, Bivona TG, Karachaliou N. Genetics and biomarkers in personalisation of lung cancer treatment. *Lancet* 2013; **382**(9893): 720-31.

16. Gao J, Li HR, Jin C, Jiang JH, Ding JY. Strategies to overcome acquired resistance to EGFR TKI in the treatment of non-small cell lung cancer. *Clin Transl Oncol* 2019.

17. Yang Z, Yang N, Ou Q, et al. Investigating Novel Resistance Mechanisms to Third-Generation EGFR Tyrosine Kinase Inhibitor Osimertinib in Non-Small Cell Lung Cancer Patients. *Clin Cancer Res* 2018; **24**(13): 3097-107.

18. Halliday PR, Blakely CM, Bivona TG. Emerging Targeted Therapies for the Treatment of Non-small Cell Lung Cancer. *Curr Oncol Rep* 2019; **21**(3): 21.

19. Doroshow DB, Sanmamed MF, Hastings K, et al. Immunotherapy in Non-Small Cell Lung Cancer: Facts and Hopes. *Clin Cancer Res* 2019.

20. Gherardi E, Birchmeier W, Birchmeier C, Vande Woude G. Targeting MET in cancer: rationale and progress. *Nat Rev Cancer* 2012; **12**(2): 89-103.

21. Drilon A, Cappuzzo F, Ou SI, Camidge DR. Targeting MET in Lung Cancer: Will Expectations Finally Be MET? *J Thorac Oncol* 2017; **12**(1): 15-26.

22. Straussman R, Morikawa T, Shee K, et al. Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. *Nature* 2012; **487**(7408): 500-4.

23. Kongkham PN, Northcott PA, Ra YS, et al. An epigenetic genome-wide screen identifies SPINT2 as a novel tumor suppressor gene in pediatric medulloblastoma. *Cancer Res* 2008; **68**(23): 9945-53.

24. Lee EJ, Rath P, Liu J, et al. Identification of Global DNA Methylation Signatures in Glioblastoma-Derived Cancer Stem Cells. *J Genet Genomics* 2015; **42**(7): 355-71.

25. Nakamura K, Hongo A, Kodama J, Hiramatsu Y. The role of hepatocyte growth factor activator inhibitor (HAI)-1 and HAI-2 in endometrial cancer. *Int J Cancer* 2011; **128**(11): 2613-24.

26. Nakamura K, Abarzua F, Kodama J, et al. Expression of hepatocyte growth factor activator inhibitors (HAI-1 and HAI-2) in ovarian cancer. *Int J Oncol* 2009; **34**(2): 345-53.

27. Nakamura K, Abarzua F, Hongo A, et al. Hepatocyte growth factor activator inhibitor-2 (HAI-2) is a favorable prognosis marker and inhibits cell growth through the apoptotic pathway in cervical cancer. *Ann Oncol* 2009; **20**(1): 63-70.

28. Pereira MS, de Almeida GC, Pinto F, Viana-Pereira M, Reis RM. SPINT2 Deregulation in Prostate Carcinoma. *J Histochem Cytochem* 2016; **64**(1): 32-41.

29. Bergum C, List K. Loss of the matriptase inhibitor HAI-2 during prostate cancer progression. *Prostate* 2010; **70**(13): 1422-8.

30. Morris MR, Gentle D, Abdulrahman M, et al. Tumor suppressor activity and epigenetic inactivation of hepatocyte growth factor activator inhibitor type 2/SPINT2 in papillary and clear cell renal cell carcinoma. *Cancer Res* 2005; **65**(11): 4598-606.

31. Dong W, Chen X, Xie J, Sun P, Wu Y. Epigenetic inactivation and tumor suppressor activity of HAI-2/SPINT2 in gastric cancer. *Int J Cancer* 2010; **127**(7): 1526-34.

32. Hwang S, Kim HE, Min M, et al. Epigenetic Silencing of SPINT2 Promotes Cancer Cell Motility via HGF-MET Pathway Activation in Melanoma. *J Invest Dermatol* 2015; **135**(9): 2283-91.

33. Roversi FM, Olalla Saad ST, Machado-Neto JA. Serine peptidase inhibitor Kunitz type 2 (SPINT2) in cancer development and progression. *Biomed Pharmacother* 2018; **101**: 278-86.

34. Yamamoto K, Kawaguchi M, Shimomura T, et al. Hepatocyte growth factor activator inhibitor type-2 (HAI-2)/SPINT2 contributes to invasive growth of oral squamous cell carcinoma cells. *Oncotarget* 2018; **9**(14): 11691-706.

35. Friis S, Sales KU, Schafer JM, Vogel LK, Kataoka H, Bugge TH. The protease inhibitor HAI-2, but not HAI-1, regulates matriptase activation and shedding through prostasin. *J Biol Chem* 2014; **289**(32): 22319-32.

36. Kirchhofer D, Peek M, Lipari MT, Billeci K, Fan B, Moran P. Hepsin activates pro-hepatocyte growth factor and is inhibited by hepatocyte growth factor activator inhibitor-1B (HAI-1B) and HAI-2. *FEBS Lett* 2005; **579**(9): 1945-50.

37. Tsai CH, Teng CH, Tu YT, et al. HAI-2 suppresses the invasive growth and metastasis of prostate cancer through regulation of matriptase. *Oncogene* 2014; **33**(38): 4643-52.

38. Nonboe AW, Krigslund O, Soendergaard C, et al. HAI-2 stabilizes, inhibits and regulates SEA-cleavage-dependent secretory transport of matriptase. *Traffic* 2017; **18**(6): 378-91.

39. Kataoka H, Kawaguchi M, Fukushima T, Shimomura T. Hepatocyte growth factor activator inhibitors (HAI-1 and HAI-2): Emerging key players in epithelial integrity and cancer. *Pathol Int* 2018; **68**(3): 145-58.

40. Kawaguchi T, Qin L, Shimomura T, et al. Purification and cloning of hepatocyte growth factor activator inhibitor type 2, a Kunitz-type serine protease inhibitor. *J Biol Chem* 1997; **272**(44): 27558-64.

41. Kawaguchi M, Kataoka H. Mechanisms of hepatocyte growth factor activation in cancer tissues. *Cancers (Basel)* 2014; **6**(4): 1890-904.

42. Parr C, Jiang WG. Hepatocyte growth factor activation inhibitors (HAI-1 and HAI-2) regulate HGF-induced invasion of human breast cancer cells. *Int J Cancer* 2006; **119**(5): 1176-83.

43. Owusu BY, Thomas S, Venukadasula P, et al. Targeting the tumor-promoting microenvironment in MET-amplified NSCLC cells with a novel inhibitor of pro-HGF activation. *Oncotarget* 2017; **8**(38): 63014-25.

44. Wu SR, Lin CH, Shih HP, et al. HAI-2 as a novel inhibitor of plasmin represses lung cancer cell invasion and metastasis. *Br J Cancer* 2019; **120**(5): 499-511.

45. AJCC. Lung. In: Amin MB, Edge S, Greene F, et al., eds. American Joint Committee on Cancer: Cancer Staging Manual. 8th ed. New York: Springer; 2017: 431-56.

46. Massion PP, Zou Y, Chen H, et al. Smoking-related genomic signatures in non-small cell lung cancer. *Am J Respir Crit Care Med* 2008; **178**(11): 1164-72.

47. Muller-Pillasch F, Wallrapp C, Bartels K, et al. Cloning of a new Kunitz-type protease inhibitor with a putative transmembrane domain overexpressed in pancreatic cancer. *Biochim Biophys Acta* 1998; **1395**(1): 88-95.

48. Generali D, Fox SB, Berruti A, et al. Regulation of hepatocyte growth factor activator inhibitor 2 by hypoxia in breast cancer. *Clin Cancer Res* 2007; **13**(2 Pt 1): 550-8.

49. Yancy HF, Mason JA, Peters S, et al. Metastatic progression and gene expression between breast cancer cell lines from African American and Caucasian women. *J Carcinog* 2007; **6**: 8.

50. Ma C, Ma W, Zhou N, Chen N, An L, Zhang Y. Protease Serine S1 Family Member 8 (PRSS8) Inhibits Tumor Growth In Vitro and In Vivo in Human Non-Small Cell Lung Cancer. *Oncol Res* 2017; **25**(5): 781-7.

51. Lee SP, Kao CY, Chang SC, et al. Tissue distribution and subcellular localizations determine in vivo functional relationship among prostasin, matriptase, HAI-1, and HAI-2 in human skin. *PLoS One* 2018; **13**(2): e0192632.

52. Chang HH, Xu Y, Lai H, et al. Differential subcellular localization renders HAI-2 a matriptase inhibitor in breast cancer cells but not in mammary epithelial cells. *PLoS One* 2015; **10**(3): e0120489.

Annex I

Supplementary data

Supplementary Table I – Associations between SPINT2 expression and Clinicopathological features of NSCLC

Clinicopathological features	NSCLC (total)		*P* value	Adenocarcinoma		*P* value	SCC		*P* value
	Low SPINT2	High SPINT2		Low SPINT2	High SPINT2		Low SPINT2	High SPINT2	
Patients, n	126	44		60	31		66	13	
Bone metastasis, n (%)	125	44		60	31		65	13	
No	108 (75.0%)	36 (25.0%)	0.462	47 (67.1%)	23 (32.9%)	0.657	61 (82.4%)	13 (17.6%)	1.000[a)]
Yes	17 (68.0%)	8 (32.0%)		13 (61.9%)	8 (38.1%)		4 (100.0%)	0 (0.0%)	
Pleural metastasis, n (%)	125	44		60	31		65	13	
No	116 (74.8%)	39 (25.2%)	0.361[a)]	55 (67.1%)	27 (32.9%)	0.484[a)]	61 (83.6%)	12 (16.4%)	1.000[a)]
Yes	9 (64.3%)	5 (35.7%)		5 (55.6%)	4 (44.4%)		4 (80.0%)	1 (20.0%)	
Adrenal metastasis, n (%)	125	44		60	31		65	13	
No	120 (75.5%)	39 (24.5%)	0.129[a)]	56 (68.3%)	26 (31.7%)	0.264[a)]	64 (83.1%)	13 (16.9%)	1.000[a)]
Yes	5 (50.0%)	5 (50.0%)		4 (44.4%)	5 (55.6%)		1 (100%)	0 (0.0%)	

NSCLC – Non-Small Cell Lung Cancer; SCC – Squamous Cell Carcinoma; n – number of samples (% of SPINT2 expression)
P value based on Pearson $\chi 2$ test, except [a)] Fisher's exact test, due to violation of $\chi 2$ assumptions.