

Article

Tumor transcriptome in patients with lower pectoralis muscle area reveals IL-8 as a prognostic biomarker in non-small cell lung cancer

Sarah Santiloni Cury¹, Diogo de Moraes¹, Paula Paccielli Freire ¹, Grasieli de Oliveira¹, Douglas Venâncio Pereira Marques¹, Geysson Javier Fernandez¹, Maeli Dal-Pai-Silva¹, Érica Nishida Hasimoto², Patricia Pintor dos Reis^{2,3}, Silvia Regina Rogatto⁴, Robson Francisco Carvalho^{*1}.

¹ São Paulo State University (UNESP), Institute of Biosciences, Department of Morphology, Botucatu São Paulo, Brazil. santiloni.cury@unesp.br, demoraesdiogo2017@gmail.com, paula.freire@unesp.br, oliveira.grase@gmail.com, oliveira.grase@gmail.com, jasonfergar@hotmail.com, maeli.dal-pai@unesp.br, robson.carvalho@unesp.br

² São Paulo State University (UNESP), Department of Surgery and Orthopedics, Faculty of Medicine, Botucatu, São Paulo, Brazil. ericanh80@gmail.com

³ São Paulo State University (UNESP), Experimental Research Unity, Faculty of Medicine, Botucatu, São Paulo, Brazil. patpreis@gmail.com

⁴ Department of Clinical Genetics, Vejle Hospital, Institute of Regional Health Research, University of Southern Denmark, Denmark. Silvia.Regina.Rogatto@rsyd.dk

* Correspondence: robson.carvalho@unesp.br; Tel.: +55 14 3880 0473; Address: Prof. Dr. Antonio Celso Wagner Zanin, 250, Botucatu, SP, Brazil

Abstract: Cachexia is a syndrome characterized by an ongoing loss of skeletal muscle mass associated with poor patient prognosis in non-small cell lung cancer (NSCLC). However, prognostic cachexia biomarkers in NSCLC are unknown. Here, we analyzed computed tomography (CT) images and tumor transcriptome data to identify potentially secreted cachexia biomarkers (PSCB) in NSCLC patients with low-muscularity. We integrated radiomics features (pectoralis muscle, sternum, and T10 vertebra) from CT of 89 NSCLC patients, which allowed us to identify an index for screening muscularity. Next, a tumor transcriptomic-based secretome analysis from these patients (discovery set) was evaluated to identify potential cachexia biomarkers in patients with low-muscularity. The prognostic value of these biomarkers for predicting recurrence and survival outcome was confirmed using expression data from eight lung cancer datasets (validation set). Finally, C2C12 myoblasts differentiated into myotubes were used to evaluate the ability of the selected biomarker, IL-8, in inducing muscle cell atrophy. We identified 75 over-expressed transcripts in patients with low-muscularity, which included *IL6*, *CSF3*, and *IL8*. Also, we identified *NCAM1*, *CNTN1*, *SCG2*, *CADM1*, *IL8*, *NPTX1*, and *APOD* as PSCB in the tumor secretome. These PSCB were capable of distinguishing worse and better prognosis (recurrence and survival) in NSCLC patients. *IL8* was confirmed as a predictor of worse prognosis in all validation sets. *In vitro* assays revealed that IL-8 promoted C2C12 myotube atrophy. Tumors from low-muscularity patients presented a set of upregulated genes encoding for secreted proteins, including pro-inflammatory cytokines that predict worse overall survival in NSCLC. Among these up-regulated genes, *IL8* expression in NSCLC tissues was associated with worse prognosis and the recombinant IL-8 was capable of triggering atrophy in C2C12 myotubes.

Keywords: Secretome; Computed tomography; Interleukin-8; Tumor-derived factor; C2C12 cells; Cachexia

1. Introduction

Lung cancer is the most prevalent cancer type worldwide and responsible for an estimated 1.8 million deaths, each year [1]. Most patients (~ 85%) develop non-small cell lung cancer (NSCLC) [2], which is frequently diagnosed in an advanced stage, and consequently has an unfavorable prognosis [3]. Cancer cachexia is a syndrome that affects a considerable proportion of NSCLC patients [4]. It is characterized by an ongoing loss of skeletal muscle mass (with or without loss of fat mass) that cannot be fully reversed by conventional nutritional support and is associated with significant functional impairments [5].

The loss of skeletal muscle mass in cancer cachexia may lead to substantial weight loss and decreased body mass index (BMI), which are associated with worse outcome in NSCLC patients [6–8]. Studies using computed tomography (CT) images have revealed occult muscle depletion in NSCLC patients, regardless of overall body weight [4,9]. In addition, both the detection of muscle depletion or low muscle mass by CT images have been associated with shorter time to tumor progression, increased risk of chemotherapy toxicity, and shorter survival in NSCLC patients [4,9–14]. Skeletal muscle depletion detected by CT images in these patients also negatively affects their functional status and quality of life [15,16]. To our knowledge, tumor-secreted factors with the prognostic value associated with low muscle mass as detected by CT in NSCLC are unknown.

Several studies have highlighted that macromolecules secreted from cancer cells and cells within the tumor microenvironment (secretome), including many pro-inflammatory cytokines, act systemically leading to muscle wasting in cancer cachexia [17–19]. However, the secretome complexity and differences found in distinct lung cancer and cells lines [20–22] illustrates the need to apply global approaches, in order to identify tumor-specific secreted molecules associated with skeletal muscle depletion. Moreover, previous “omics” studies of cancer secretome in cachexia have focused on the analysis of cachectic conditioned media of single cancer cells lines to identify mediators of the syndrome [23–25]. However, *in vitro* systems ignore the contributions of the host-tumor microenvironment and the tumor heterogeneity as well as provide no insight into the disease progression [22]. These findings emphasize the importance of cancer cachexia studies in exploring the tumor secretome. Thus, we hypothesized that a tumor transcriptome-based secretome analysis in NSCLC patients with low muscularity is a strategy capable of identifying prognostic biomarkers and mediators of cancer-associated muscle loss.

Herein, we analyzed a cohort of NSCLC patients with CT images, clinical findings, and tumor expression microarrays data from a previous study that decoded tumor radiomics features associated with gene expression levels [26]. For these patients, we compared the pectoralis muscle area (PMA) with muscle normalizations based on different radiomics features to select an approach for screening muscularity. Next, we identified genes predicted to be secreted in patients with low muscularity and assessed their prognostic value as tumor markers of recurrence-free survival and overall survival. Finally, we demonstrated the potential of IL-8 as a putative secreted marker capable of inducing atrophy in C2C12 myotubes.

2. Results

The workflow of the integrative analyses of CT images and tumor transcriptome used to identify potentially secreted cachexia mediators and biomarkers in NSCLC patients with low-muscularity is depicted in Figure 1.

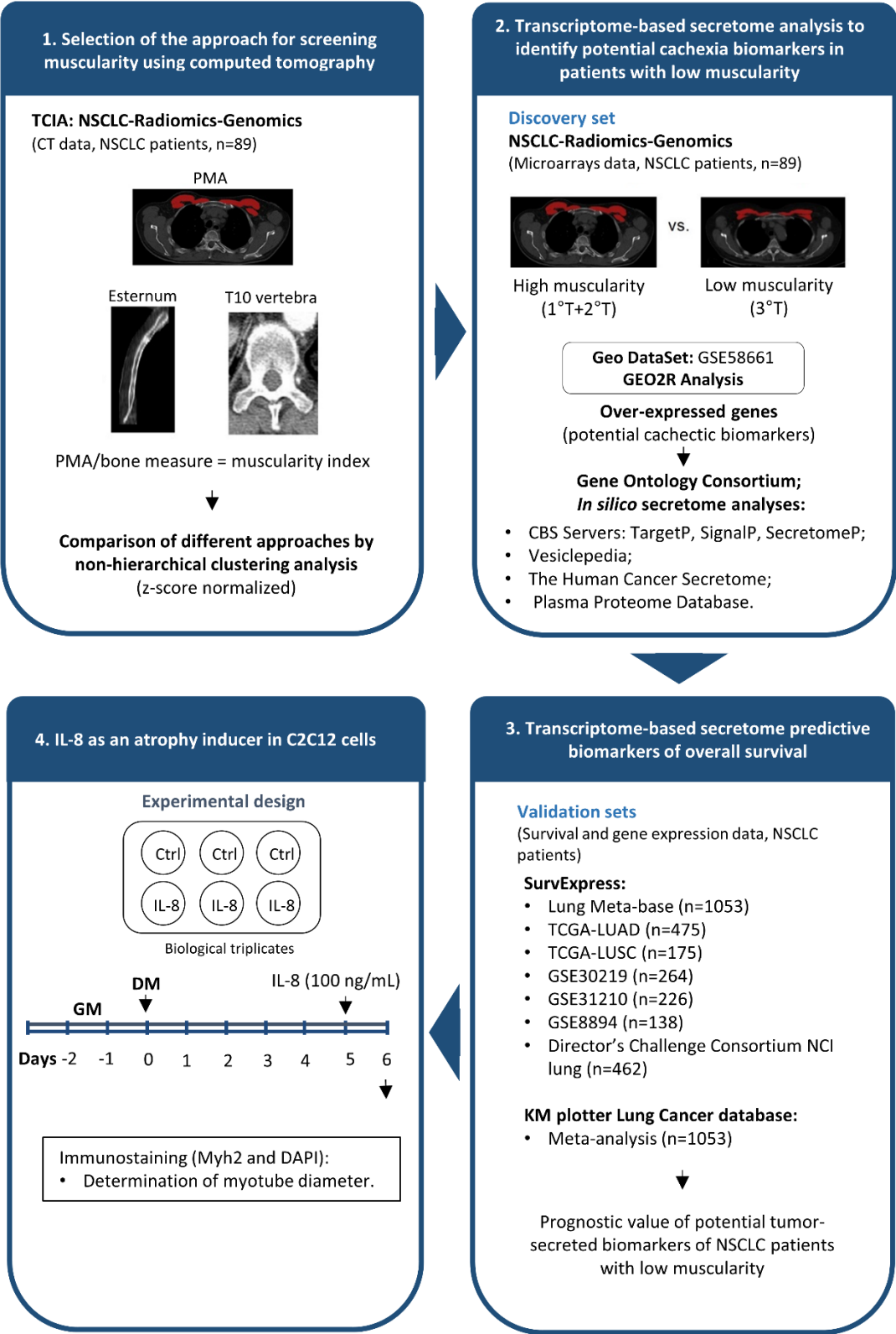


Figure 1. The workflow of the integrative analyses of CT images and tumor transcriptome data to identify secreted cachexia biomarkers in NSCLC patients with low muscularity. (1) We selected computed tomographies (CT) from 89 patients with non-small cell lung cancer (NSCLC) from

“NSCLC-Radiomics-Genomics” collection [27]. A total of 12 CTs features, including pectoralis muscle area (PMA), manubrium and sternum body lengths, six T10 different measures, and an anteroposterior length was used to determine an approach for screening muscularity. (2) This analysis revealed that PMA allows the identification of NSCLC patients with low (third tercile, 3rd T) and high muscularity (first and second terciles, 1st T + 2nd T). These groups were compared by using a tumor transcriptomic-based secretome analysis (discovery set; microarray data; GSE58661) to identify potential cachexia biomarkers (over-expressed genes) in patients with low muscularity. Transcripts with increased expression were further analyzed to identify enriched terms by Gene Ontology Consortium and to predict potentially secreted proteins using secretome and microvesicle databases. (3) The performance of these transcripts as tumor biomarkers able to determine patients’ survival outcome was validated in multiple independent lung cancer validation sets. (4) C2C12 myoblast differentiated into myotubes were used to evaluate the ability of the selected biomarker (IL-8) in inducing atrophy. C2C12 mouse myoblasts were cultured in a growth medium (GM) for two days. Myoblasts with 80% to 90% of confluence were induced to differentiate in a differentiation medium (DM) for five days when the cells were treated with recombinant IL-8 (100 ng/mL for 24h). n: number of patients; Ctrl: Control; Myh2: myosin heavy chain 2; DAPI: 4',6-Diamidine-2'-phenylindole dihydrochloride nuclear staining.

2.1. Study population

CT images, clinical, and microarrays data of 89 NSCLC patients, with an average age of 65.2 ± 8.7 years, were included in this study. The most common NSCLC histological type was adenocarcinoma (47.2%), and 20.2% of the patients were diagnosed with advanced stage cancers (stages III or IV). Adenocarcinoma was prevalent in women, while squamous cell carcinoma was more frequent in men. The muscle measurements revealed differences between sexes, with men and women presenting PMA of 42.5 ± 9.3 and 27 ± 6.0 cm², respectively. Based on this finding, the sex-specific categorical variable was taken into consideration for further analyses. Table 1 summarizes the clinical, histopathological, and muscle measurements in this cohort of NSCLC patients.

Table 1 Clinical findings and skeletal muscle parameters of patients with NSCLC with low and high muscularity defined by pectoralis muscle cross-sectional area assessed by computed tomography.

Characteristics	All	Men	Women	p-value*
Number of patients	89	60	29	
Age	65.2 ± 8.7	66.9 ± 7.3	61.8 ± 10.3	0.011 ^a
Cancer Stage (%)				
Early Stages (I-II)	79.7	74.6	90	0.08 ^b
Advanced Stages (III-IV)	20.2	25.4	10	
Histological Type (%)				
Adenocarcinoma	47.2	37.3	66.6	0.02 ^b
Squamous Cell Carcinoma	40.4	49.2	23.3	
Other Subtypes	12.4	13.5	10.1	
PMA (cm²)	37.3 ± 11.1	42.5 ± 9.3	27 ± 6.0	<0.001 ^a
HM (N)	59	40	19	
LM (N)	30	20	10	
LM PMA (cm²)	28.6 ± 6.5 ^{#c}	32.3 ± 4.3 ^{#c}	21 ± 1.4 ^{#c}	<0.001 ^a
HM PMA (cm²)	41.7 ± 10.2	47.5 ± 6.3	30 ± 5.2	<0.001 ^a

N: number of patients; PMA: Pectoralis Muscle Area; LM: low muscularity patients; HM: high-muscularity patients. The data represent the mean ± standard deviation. a: Student’s t-test; b: Chi-squared test; c: Mann-

Whitney’ U-test; * comparisons between men and women; # statistical difference between patients with low- and high-muscularity ($P < 0.001$).

2.2. PMA distinguishes NSCLC patients with low and high muscularity

Considering that CTs from NSCLC patients have information that goes beyond the tumor, we integrated different radiomics features to determine an approach to be used for screening muscularity (Supplementary Figure 1). The non-hierarchical, unsupervised clustering analysis of the PMA and its normalization by 11 CTs features (z-score normalized) revealed a similar pattern of patients’ distribution according to all muscularity indexes. The clustering analysis also revealed three subgroups of patients according to the muscularity indexes as depicted in the dendrogram on Supplementary Figure 2a. Applying k-means analysis ($k\text{-means} = 3$) resulted in a cluster composed of 34 patients with low muscularity (Figure 2a). Next, we used a descending PMA order based on gender as a sex-specific categorical variable. Finally, we segregated into terciles to generate two groups of study based on the patients’ muscularity. The low muscularity group includes patients within the third tercile, while the high muscularity group includes patients within the first and second terciles, regardless of the patient gender (Figure 2a). We highlight that PMA can be used to select potential NSCLC low muscularity patients; moreover, we suggest cut-offs values of $PMA < 32.2 \text{ cm}^2$ and $< 21 \text{ cm}^2$ (for men and women, respectively), as demonstrated by median values in the scatter dot plot on Supplementary figure 2b. The mean PMA differed significantly between the high and low muscularity groups considering all patients or comparing male and female patients (Table 1). We further compared high and low muscularity patients with other clinical variables using patient demographic information (Figure 2b). The comparison between these groups (high and low muscularity) revealed that muscularity seems to be related to tumor type and tumor stage, rather than age and tumor size (Figure 2b).

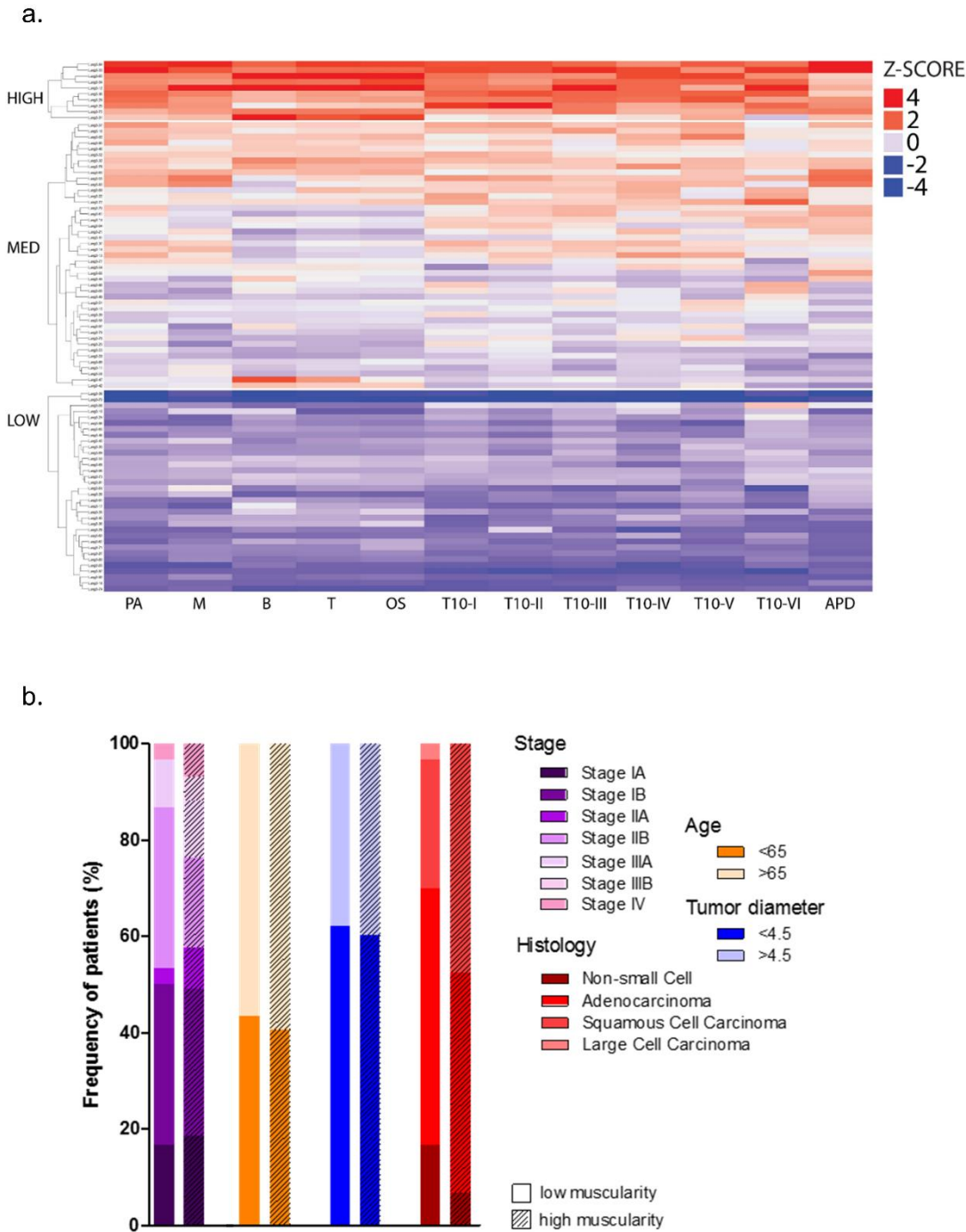


Figure 2. Pectoralis muscle area as an approach for screening muscularity. (a) Heatmap showing patients stratification into high, medium, and low muscularity by non-hierarchical k-means clustering analysis of the pectoralis muscle area (PA) and its normalization by eleven computed tomographies (CT) feature that includes: manubrium and sternum body lengths, six T10 different measures, and an anteroposterior length. Manubrium length (M); sternum Body length (B); T (M+B); total sternum length (OS); T10 body vertical length (T10-I); distance between T10 body and spinous process (T10-II); T10 body horizontal length (T10-III); distance between T10 pedicles (T10-IV); distance between T10 transverse processes (T10-V); T10 body area (T10-VI); anteroposterior distance (APD). (b) Bar graphs comparing the percentage of patients for clinical prognostic variables between high and low muscularity groups. These groups were generated based on the ordination of patients according to their PMA in descending order using a sex-specific categorical variable followed by segregation into terciles (high muscularity group: 1st and 2nd terciles; low muscularity group: 3rd tercile).

2.3. Patients with low muscularity upregulate tumor genes previously associated with cachexia

Considering that mediators released from cancer cells and cells within the tumor microenvironment have been associated with cachexia in lung cancers, we hypothesized that the identification of tumor deregulated genes in NSCLC patients with low muscularity could reveal potential factors associated with cachexia. Thus, an analysis using differential gene expression between patients with low and high muscularity revealed 105 genes exclusively deregulated (adj. p-value ≤ 0.05 and fold change ≥ 1.5) in patients with low muscularity, of which 75 and 30 were over- or down-expressed, respectively (Supplementary Table 1). Gene ontology and KEGG pathway analyses of the over-expressed transcripts highlighted cytokine activity and cytokine-receptor interaction activity as the most enriched categories in low muscularity patients. (Figure 3a). PPI analysis identified the interactions among these proteins (Figure 3b), including the pro-inflammatory cytokines IL-6, IL-8, and CSF3, which have been previously implicated in the development of cancer cachexia [28–31].

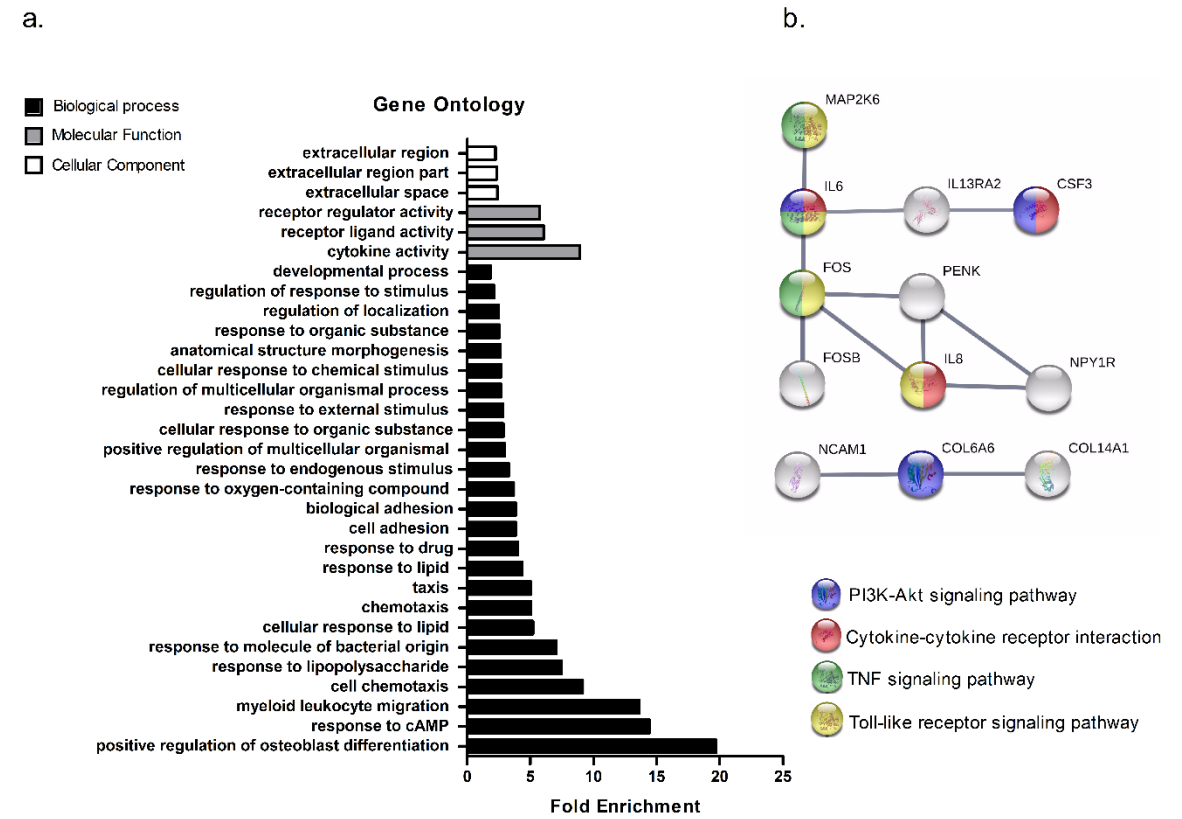


Figure 3. Over-expressed genes in tumors from patients with low muscularity (a) Enriched terms in gene ontology analysis of the 75 transcripts up-regulated in patients with low muscularity. (b) Protein-protein interaction (PPI) network of 75 up-regulated transcripts in patients with low muscularity generated by STRING using a high confidence interaction score (0.700).

2.4. Secretome-related genes with prognostic value in NSCLC

We then investigated whether these 75 up-regulated transcripts in the tumors from low muscularity patients are translated into secreted proteins. The intersection of the secretome databases CBS Servers, Vesiclepedia, Human Cancer Secretome Database, and Plasma Proteome Database showed seven overlapping proteins: IL-8, SCG2, NCAM1, CNTN1, CADM1, NPTX1, and APOD (Figure 4a). The microvesicle databases revealed that the predicted proteins in Evpedia (LPL, APOD, and COL14A1) were also identified in the Vesiclepedia dataset. However, the Exocarta did not show

184 any of these proteins in lung cancer samples, possibly due to the limited number of exosomes studies
185 in lung cancers deposited in this database.

186 The prognostic value related to the worse prognosis of *IL8*, *SCG2*, *NCAM1*, *CNTN1*, *CADM1*,
187 *NPTX1*, and *APOD* transcripts were evaluated in seven lung cancer transcriptome datasets
188 (validation set). Notably, these biomarkers were capable of distinguishing worse and better prognosis
189 (recurrence and survival) in seven NSCLC cohorts from the SurvExpress database (Figure 4c).
190 Interestingly, only IL8 was found with increased expression in the high-risk group in all NSCLC
191 validation set (Figure 4d and Supplementary Figure 3).

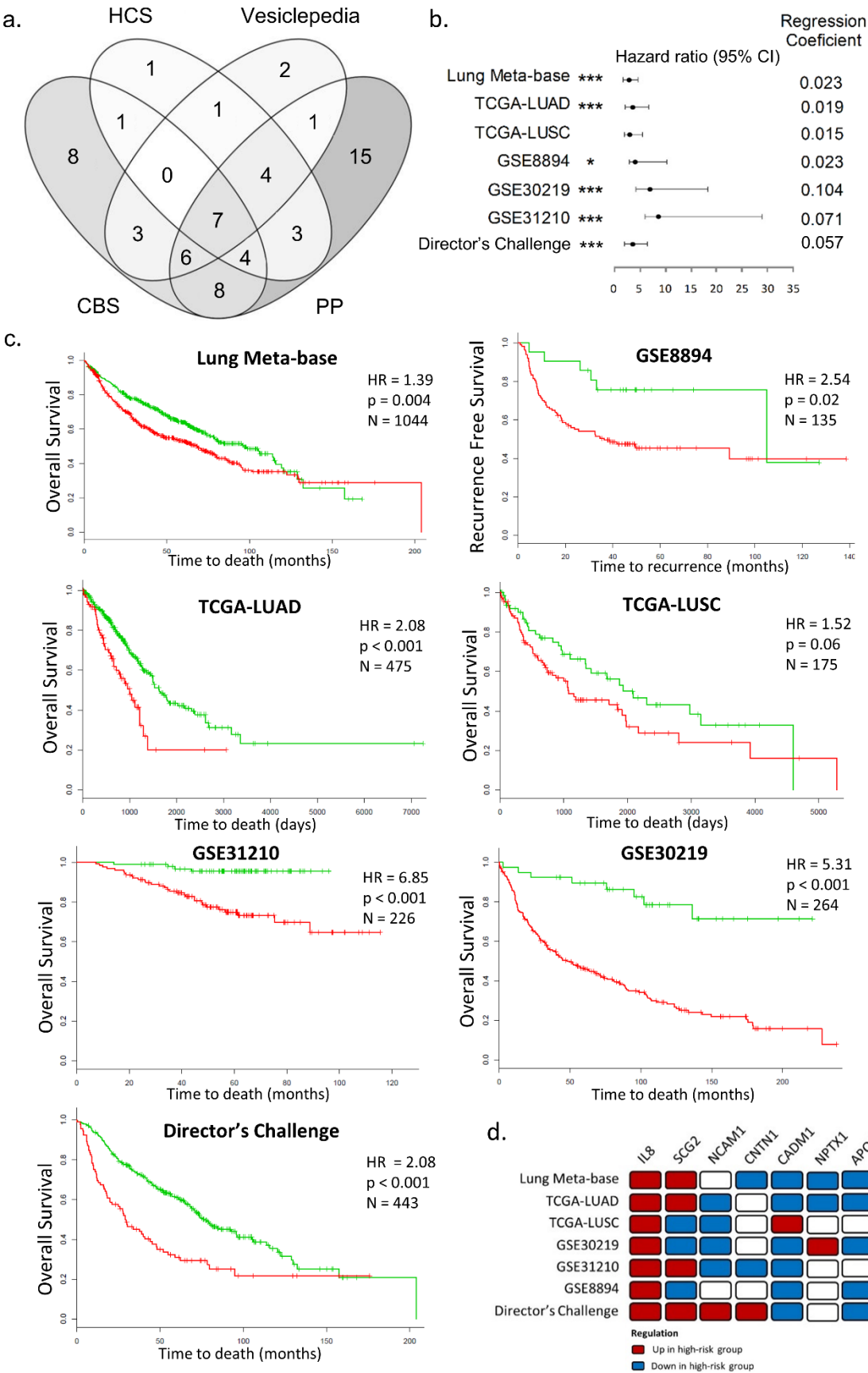


Figure 4. Prognostic values of the potentially secreted proteins in tumors from patients with low muscularity. (a) The intersection of databases used for prediction of secreted proteins revealed seven overlapped proteins: IL8, SCG2, NCAM1, CNTN1, CADM1, NPTX1, and APOD. (b) Forest plots representing the set of the seven potentially secreted biomarkers in each validation set. The horizontal

axis represents confidence intervals estimated by using a Cox proportional hazards model. The asterisks represent the statistical significance in patient survival outcome (** $p < 0.001$ and * $p < 0.05$, log-rank P-value). (c) Kaplan-Meier plots generated in SurvExpress [73] database for the NSCLC datasets (gene expression and survival or time to recurrence): Lung Meta-base, TCGA-LUAD, TCGA-LUSC, GSE30219, GSE31210, and Director's Challenge Consortium NCI. The Kaplan-Meier plot generated using the dataset GSE8894 was based on gene expression and time to recurrence data. (d) The direction of expression for the seven biomarkers in each validation set. N: number of patients; HR: adjusted hazard ratio; p: log-rank p-value determined from univariate Cox regression analyses (green curve: low-risk group; red curve: high-risk group); TCGA: The Cancer Genome Atlas; LUAD: Lung Adenocarcinoma; LUSC: Lung Squamous Cell Carcinoma

2.5. High IL8 expression in tumor tissues is associated with poor prognosis in NSCLC

All seven potential biomarkers were individually analyzed in the KM plotter server using gene expression and survival data of lung cancer patients available on the database (N = 1053), and IL8 proved to be a strong predictor of poor survival (Figure 5a). Moreover, as IL6 is a key regulator of muscle mass during cachexia [32] and has been associated with worse prognosis in lung cancer patients [33,34], we compared the prognostic value of IL8 with IL6 using KM plotter server. Notably, both IL8 and IL6 tumor transcripts presented similar prognostic values (IL8: HR = 1.28, 95% CI = 1.12-1.45; IL6: HR = 1.32, 95% CI = 1.16-1.5). These results demonstrate the upregulation of IL8 as a new biomarker associated with poor prognosis in lung cancer patients.

The 75 over-expressed transcripts were carefully evaluated in patients with low muscularity in KM plotter to detect additional potential cachexia biomarkers associated with poor prognosis in lung cancer patients. Nine genes (IL6, IL8, IL1R2, CEMIP, CLEC4E, FCGR3B, HAL, MAP2K6, and KIF1A) were validated as over-expressed in patients with worse overall survival (Supplementary Figure 4). Importantly, IL6, IL8, IL1R2, CEMIP, FCGR3B, and KIF1A are predicted as a potentially secreted protein in at least two secretome databases (Supplementary Table 2). Collectively, these results emphasize that IL8 is highly expressed in tumors from NSCLC patients with low muscularity and is associated with poor prognosis in this cancer type.

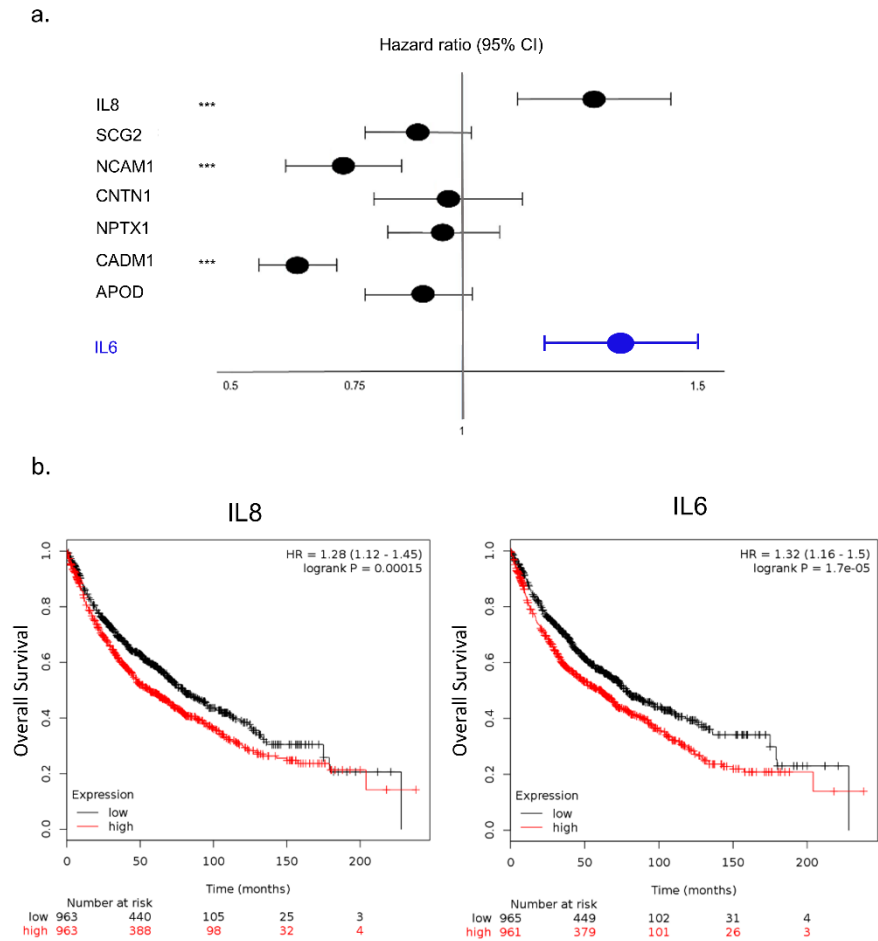


Figure 5. IL8 is associated with poor prognosis in NSCLC. (a) Forest plot for each tumor biomarkers (IL8, SCG2, NCAM1, CNTN1, CADM1, NPTX1, and APOD) in NSCLC patients from the dataset available on KM plotter[35] database. The hazard ratio (HR) with 95% confidence intervals (CI) determined by Cox proportional hazards model is represented in the horizontal axis. *** represent the statistical significance in NSCLC patient survival outcome ($p < 0.001$; log-rank p-value). (b) Kaplan-Meier overall survival curves for IL8 or IL6 in NSCLC patients from the dataset available on KM plotter [35] database. The resulting p-values for the log-rank test are shown.

2.6. IL-8 treatment induces in vitro myotube atrophy

The ability of IL-8 in inducing muscle atrophy was evaluated by treating C2C12 myotubes with different concentrations of this cytokine (10, 100, and 1000 ng/mL). The myotubes treated with the supraphysiological dose of 100 ng/mL presented a significant decrease in diameter compared to the control group after 24 hours (Supplementary Figure 5). The C2C12 myotubes treated with 100 ng/mL of IL-8 for 24 hours were evaluated by Myh2 immunostaining, which confirmed the significant decrease in myotubes diameter (Figure 6a, b). Myotubes treated with IL-8 also presented a higher number of myotubes with $< 10 \mu\text{m}$ of diameter compared to the control group. Conversely, a higher number of controls myotubes with $> 35 \mu\text{m}$ of diameter compared to those myotubes treated with IL-8 was observed (Figure 6c).

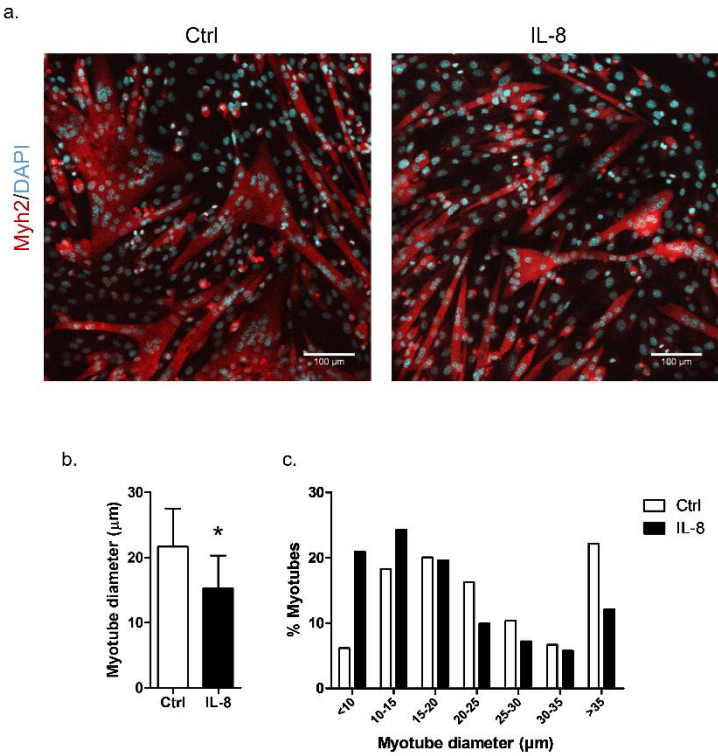


Figure 6. IL-8 induces atrophy in C2C12 myotubes. (a) Immunofluorescence of C2C12 myotubes treated with recombinant IL-8 (100 ng/mL for 24 hours) immunostained for Myh2 (red) and the nuclei was counterstained with DAPI. (b) Myotube diameter (μm) quantification using ImageJ software. (c) Determination of the frequency of myotubes according to the diameter classes. The data represents the mean ± standard deviation from three independent experiments. The statistical significance was analyzed using the Student's T-test. * p-value < 0.05. Ctrl: control myotubes; IL-8: myotubes treated with recombinant interleukin 8

3. Discussion

Using a tumor transcriptome-based secretome analysis in NSCLC patients with low muscularity, we aimed to identify potential cancer biomarkers of prognostic value and mediators of cancer-associated muscle loss. This strategy revealed increased expression levels of cachexia-related genes predicted to be secreted in NSCLC from patients with lower PMA. These genes were further associated with shorter recurrence-free survival and decreased overall survival in different validation sets of patients with NSCLC. Importantly, increased expression levels of *IL8* was detected in the high-risk group in all NSCLC validation sets, and IL-8 was sufficient to trigger atrophy in C2C12 myotubes.

Muscle depletion or low muscle mass in NSCLC patients identified by CT images has been extensively associated with poor outcome [4,9–13,15,16]. Previous studies using the same methodology to ours - the objective assessment of the PMA on CT scans -reported lower PMA associated with worse overall survival in NSCLC patients or cases with chronic obstructive pulmonary disease, despite normalization for BMI and performance status [14,36]. Teigen *et al.* reported that the PMA divided by height (used to standardize for body size) is a powerful predictor of outcome after left ventricular assist device implantation [37]. Unfortunately, the height in our cohort of PMA CT-based analysis was not available. However, the high quality of these CT images previously allowed the identification of new tumor radiomics features with prognostic value in NSCLC patients [26]. Thus, we hypothesized that the comparison of the PMA with muscle normalizations based on different radiomics features aiming the standardization for body size could reveal new approaches for screening muscularity in NSCLC patients. Interestingly, PMA distinguished NSCLC patients with low and high muscularity similarly in all muscle normalizations

tested. Considering that CTs images of lung cancer patients are preferentially performed in the thoracic region, our data additionally confirm that PMA is a feasible measurement easily applied to the clinical practice to distinguish NSCLC patients with different muscularity.

Although a large range of changes in body composition has been associated with tumor-derived factors, including many pro-inflammatory cytokines [17–19,38], only few NSCLC studies associated CT-derived body composition with systemic inflammatory response [39,40]. These studies showed that lower muscularity was associated with systemic inflammatory response (IL-6, C-reactive protein, and albumin blood levels, and neutrophil-to-lymphocyte ratio). However, the specific tumor-derived factors that induce muscle loss in NSCLC patients are still unknown. Using the tumor transcriptome analysis of NSCLC patients with low muscularity, we found 105 deregulated genes, of which 75 were up-regulated and 30 down-regulated. The functional enrichment analysis revealed up-regulated genes related to cytokine activity (*CSF3*, *IL8*, *IL6*, *BMP6*, *SCG2*, *CCL8*, *BMP2*) and extracellular space (*CSF3*, *FLRT2*, *IL8*, *PLA2G3*, *IL6*, *ATP1B1*, *COL14A1*, *LPL*, *HBB*, *ADAMTS4*). These results suggest that tumor of patients with low muscularity possibly secrete cachexia associated-factors.

The *in-silico* analysis confirmed that a set of over-expressed genes are translated into proteins presented in plasma or secretome of NSCLC patients. Seven of these predicted proteins (*NCAM1*, *CNTN1*, *SCG2*, *CADM1*, *IL-8*, *NPTX1*, and *APOD*) were identified in five databases (SignalP 4.1, SecretomeP 2.0, Vesiclepedia, Human Cancer Secretome, and Plasma Proteome), giving support to their relevance in NSCLC. Although not all NSCLC patients with low muscularity are cachectic, the tumor gene expression profile identified molecules, such as *IL6* and *IL8*, consistently linked to inflammation and cancer cachexia pathogenesis [28,30,41–48]. The low muscle mass detected by CT images can occur in the absence of systemic inflammation in other malignancies such as colorectal cancer, but the proportion of patients with low muscularity is substantially greater in the presence of systemic inflammation [49]. In cases where the inflammation coexists with low muscle mass, the prognosis is especially poor [50]. Taken together, we identified a specific set of up-regulated genes coding for secreted proteins that may constitute potential mediators of muscle loss in NSCLC.

Based on the fact that circulating levels of tumor-derived factors were correlated with cachexia development and predicted outcome in cancer [28,30,41–48], we also investigated the predictive potential of seven transcripts (*NCAM1*, *CNTN1*, *SCG2*, *CADM1*, *IL8*, *NPTX1*, and *APOD*). All of them were associated with shorter overall survival and recurrence-free survival for the predicted high-risk groups in the NSCLC validation set. However, only *IL8* was over-expressed in the high-risk group in all cohorts of our NSCLC validation set. We further confirmed that high *IL8* expression level in tumor tissue is a strong predictive biomarker significantly associated with worse survival (validation cohort of 1053 NSCLC patients). In agreement with our results, IL-8 expression in tumor tissues was recently associated with cachectic status and outcome in pancreatic cancer; cachectic patients with high IL-8 expression in tumor tissues had shorter overall survival or disease-free survival [30]. Importantly, these authors also showed that IL-8 expression level in tumor specimen paired with serum sample from the same patients was associated with tumor size.

We demonstrated that IL-8 directly induce myotube atrophy, reinforcing its potential as a new mediator of cancer cachexia. Muscle wasting in cancer cachexia has been attributed to the combinatorial action of mediators from host and tumor microenvironment [17–19,38]. Also, tumor expression and serum levels of IL-8 have been associated with muscle wasting in patients with different tumor types [28,30,41–45]. The potential direct effect of IL-8 in inducing muscle cell atrophy is still unknown. In this study, we provide evidence that IL-8 is a biomarker of worse prognosis that has the potential to define the cachectic state in NSCLC patients.

The main strength of the present investigation is the identification of potential tumor-derived mediators of muscle wasting in patients with low muscularity, which have prognostic value in NSCLC. However, our study is based on the reuse of transcriptomic and clinical data, which results in limitations that can be pointed out. Firstly, the validation of the findings at protein levels in NSCLC patients with low muscularity would be a strategy to define the cachexia blood biomarkers useful for clinical routine. Secondly, our survival analyses were restricted to the validation set; the survival information was not available in our discovery dataset. Finally, since the *IL8* gene is not present in

the rodent genome, the atrophy phenotype observed in mice myotubes was likely induced by orthologue receptors to the human IL-8 [51]. In agreement with our study, Gerber et al., reported that IL-8 protein expression was significantly associated with tumor-free body weight and skeletal muscle weight in a human pancreatic cancer xenograft mouse model [52]. Further studies are needed to elucidate the mechanisms of action of IL-8 in human muscle cells.

4. Materials and Methods

4.1. Datasets

CT images and clinical data were downloaded from *The Cancer Imaging Archive* (TCIA, <http://cancerimagingarchive.net/>) database [53]. The dataset (NSCLC-Radiomics-Genomics collection) [27] contains information from 89 NSCLC adult patients treated at MAASTRO Clinic, The Netherlands, as previously published [26]. TCIA data are anonymized, and the institutional ethical review board approval is not needed [54]. CT images were taken on diagnosis, and the patients were treated with surgical procedure. Clinical data (age, gender, diagnosis, tumor stage), CT images and tumor microarrays data are available for all 89 patients. The NSCLC-Radiomics-Genomics microarrays data is available on Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>; microarrays dataset GSE58661) [26].

4.2. CT Imaging Analyses

The CT collection “NSCLC-Radiomics-Genomics” on TCIA database present CT images with radiomics features that can be used as noninvasive prognostic or predictive biomarkers [26]. This collection is also the most appropriate due to the homogeneity of the CT images. The pectoralis muscle was analyzed on a single axial slice of the image. This region was selected by a single trained physician (ENH) who identified the aortic arch and then selected the first image just above the arch. The cross-sectional area (cm²) of bilateral major and minor pectoralis muscles was measured by two independent examiners, using Slice-O-Matic software (v.5.0; Tomovision, Montreal, Quebec, Canada). Muscles were manually traced using the Region of Interest (ROI) tool by summing the appropriate pixels determined by CT Hounsfield unit (HU) for skeletal muscle (range −29 HU to 150 HU). The borders of the pectoralis muscles were corrected manually when necessary, as previously described [36,55,56]. The pectoralis muscle area (PMA) was calculated by adding up the four muscles area. To test the reproducibility of this analysis, an interobserver coefficient of variation was determined by comparing the results of the analyses conducted by the two observers. The mean of this coefficient of variation was 8.1%.

We also compared the PMA with muscle normalizations based on different radiomics features, as previously described [57–59], to test different approaches for screening muscularity in NSCLC patients. For this purpose, the pectoralis muscle area was also normalized by the following sternum measurements: 1) manubrium length; 2) sternum body length; 3) total manubrium and sternum body lengths; 4) distance between the beginning of manubrium and the end of sternum body measured in 90° (not considering the xiphoid process) (Supplementary figure 1a). Different T10 vertebrae measurements were also tested for muscle normalizations: 1) horizontal length of T10 body; 2) vertical length of T10 body; 3) distance between T10 body and spinous process; 4) distance between transverse processes; 5) distance between pedicles; and 6) T10 body area. We also analyzed the body cross-section anteroposterior diameter (APD) at the tenth thoracic vertebra (T10) level to normalize the muscle area (Supplementary Figure 1b). The bone images were selected in the cross-section where the bones appeared at a higher extent and dimension. The measurements were performed at the tenth thoracic vertebra (T10), which is a common region for all patients in this CT collection. Skeletal muscle index (or muscularity) was defined as the PMA divided by each bone or body measure (mentioned above) squared (cm²/cm²). The measurements generated were z-score normalized and submitted to a non-hierarchical k-means clustering analysis using Bioconductor Package Complex Heatmap (v 3.5) in RStudio software (<http://www.rstudio.org/>).

4.3. Gene expression analysis

Tumor gene expression analysis was performed by comparing low and high muscularity patients using the GEO2R tool (<http://www.ncbi.nlm.nih.gov/geo/geo2r/>) [60]. The adjusted p values (adj. p) were applied using Benjamini and Hochberg false discovery rate (FDR) method by default. The cut-off criteria to define differential expression were adj. $p < 0.05$ and $|\text{Fold Change (FC)}| > 1.5$.

4.4. Gene ontology enrichment analysis

Gene ontology (GO) functional enrichment analysis was performed to identify the overrepresented GO categories of differential expressed genes using Gene Ontology Consortium database (<http://geneontology.org/>) [[61]]. The GO categories with p-value and FDR < 0.05 were considered significant.

4.5. Protein-protein interactions (PPI) networks

PPI networks were generated using the STRING tool [62,63] ([HTTP:// string-db.org/](http://string-db.org/)). We considered experiments, database, co-expression, neighborhood, and co-occurrence as active interaction sources. The minimum required interaction score was 0.700 (high confidence), and the disconnected nodes in the network were hidden for display simplifications. The PPI enrichment p-value indicates the statistical significance provided by STRING.

4.6. In Silico Identification of Secreted Proteins

The over-expressed genes in the tumor of patients with low muscularity were filtered for genes encoding secreted proteins or proteins presented in microvesicles based on a pipeline of seven tools: SignalP 4.1 [64], SecretomeP 2.0 [65], ExoCarta [66], TargetP 1.1 [67], Human Cancer Secretome (HCS) [68], Vesiclepedia [69], and Evpedia [70]. Firstly, we accessed the UniProtKB database to obtain amino acid sequences of proteins in FASTA format [71]. These data were used in the prediction servers SignalP, TargetP, and SecretomeP at CBS portal (<http://www.cbs.dtu.dk/services/>). SignalP 4.1 server was used to identify classical secretory proteins (presenting signal peptide and D-value > 0.45). Proteins without signal peptide were evaluated in the SecretomeP 2.0 server to determine non-classical secreted proteins, using the cut-off for a neural network (NN) score > 0.6 . TargetP 1.1 server was used to selectively collect proteins involved in secretory pathways and exclude mitochondrial proteins [67]. These potentially secreted proteins were also investigated in lung cancer using the tools ExoCarta, HCS, Vesiclepedia, and Evpedia. Finally, the Plasma Proteome Database was consulted to identify human plasma proteins and their isoforms potentially encoded by the over-expressed genes from low muscularity patients [72]. The tumor over-expressed genes, detected by all eight prediction tools, were next used to assess their prognostic performance in predicting overall survival and time to recurrence in multiple NSCLC independent datasets (validation set).

4.7. Prognostic performance of secretory genes in predicting NSCLC outcome

SurvExpress [73] database (<http://bioinformatica.mty.itesm.mx/SurvExpress>) was used to assess the effect the differentially expressed genes on survival (datasets: Lung Meta-base, TCGA-LUAD and LUSC [74], GSE30219 [75], GSE31210 [76,77], and the Director's Challenge Consortium NCI [78]) and time to recurrence (dataset: GSE8894 [79]) of NSCLC patients. This tool allowed us to assess the expression of secretory genes and their association with the survival or time to recurrence by Cox Proportional Hazard regression according to the risk groups estimated by an optimization algorithm. The prognostic value of the secretory genes in predicting survival was further determined in 1053 NSCLC patients using Kaplan-Meier Plotter – KM plotter [35]. Here, gene expression was specifically associated with survival and time to recurrence (worse prognosis) due to the lack of other clinical

characteristics available in the databases. The datasets included in all survival analysis present other clinical variables, such as age, gender, histology, and stage, which were not discriminated here.

4.8. Functional assay using the C2C12 cell culture

C2C12 mouse myoblasts (ATCC® CRL-1772™) were cultured in Dulbecco's modified Eagle's medium (DMEM, Thermo Fisher Scientific, USA) with 1% Penicillin-Streptomycin (Thermo Fisher Scientific, USA) and 10% fetal bovine serum (FBS, Thermo Fisher Scientific, USA) at 37 °C and 5% CO₂ for growth and expansion. After reaching a confluence of 80-90%, the myoblasts were induced to differentiate in DMEM serum-free supplemented with 1% Penicillin-Streptomycin for five days. Human recombinant IL-8 (10, 100 or 1000 ng/ml; Abcam, USA) was added to a new differentiation medium for 24 hours. All experiments were conducted using three independent replicates per group. Control myotubes (Ctrl) received a sterile water solution containing bovine serum albumin 0.1%, the same solution used to dilute IL-8.

4.9. Immunofluorescence assay

C2C12 myotubes cultured in 6-well plates were fixed in 4% paraformaldehyde for 15 min, washed with phosphate buffered saline (PBS) and 0.1% TritonX-100 (Sigma, USA), and blocked with 3% bovine serum albumin (BSA), 1% glycine, 8% fetal bovine serum in PBS and 0.1% TritonX-100 for 1h at room temperature. Subsequently, the cells were incubated with primary (Myh2, 1:600 dilution, M7523, Sigma, USA) antibody overnight at 4°C. After washing, the cells were incubated with secondary antibody at the same concentration of the primary antibody (1h at room temperature), and counterstained with 4',6-diamidino-2-phenylindole - DAPI (ProLong Gold Antifade Mountant with DAPI, Thermo Fisher Scientific, USA). All images were acquired using scanning confocal microscope Fluoview FV10i (Olympus, Japan). The myotube diameter was measured by ImageJ software.

4.10. Statistical analysis

For the statistical analyses not previously described, we used the *GraphPad Prism®* (GraphPad Software, USA). Student's t-test or Mann-Whitney U-test was applied for independent samples with normal distributed or non-parametric data, respectively. The comparison of the effect of three different IL-8 concentrations on C2C12 myotubes with the respective controls was performed using one-way ANOVA followed by Tukey test. Data are expressed as mean \pm standard deviation (SD).

5. Conclusions

In conclusion, our study demonstrated that PMA is a clinical and practical method to distinguish NSCLC patients with different muscularity from routinely acquired CT images. Tumors from patients with low muscularity have a set of up-regulated genes coding for secreted proteins within the tumor microenvironment, including pro-inflammatory cytokines, which predict worse overall survival in NSCLC. Among the up-regulated genes, high *IL8* expression in tumor tissues is also associated with worse prognosis in NSCLC and recombinant IL-8 is capable of triggering atrophy in C2C12 myotubes.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: Different bone measurements used for Pectoralis muscle area normalization; Figure S2: Identification of clusters of patients according muscularity indexes and Pectoralis Muscle Area (PMA) cut-off values; Figure S3: Heatmaps showing gene expression findings of seven potential biomarkers in low- and high-risk groups in NSCLC validation sets; Figure S4: Highly expressed mRNAs in tumor tissue associated with poor prognosis in NSCLC; Figure S5: C2C12 myotubes treated with different concentrations of recombinant IL-8 (10, 100, 1000 ng/ml); Table S1: mRNA differentially expressed in tumors of 30 NSCLC patients with low muscularity compared with the gene expression of 59 NSCLC patients with high muscularity; Table S2: Potentially secreted proteins identified in secretome-related databases based on the over-expressed genes in NSCLC patients with low muscularity.

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used “conceptualization, S.S.C., S.R.R., and R.F.C; methodology, S.S.C., D.M., P.P.F., G.O., D.V.P.M., G.J.F., and E.N.H; validation, S.S.C. and D.M.; formal analysis, S.S.C., P.P.R., S.R.R., and R.F.C; investigation, S.S.C., D.M., P.P.F., G.O., G.J.F., E.N.H, and P.P.R.; resources, M.D.P.S., S.R.R., and R.F.C; data curation, M.D.P.S., P.P.R., S.R.R., and R.F.C; writing—original draft preparation, S.S.C., S.R.R., and R.F.C; writing—review and editing, all authors; supervision, S.R.R. and R.F.C; project administration, R.F.C; funding acquisition, M.D.P.S., S.R.R., and R.F.C.

Funding: This research was funded by São Paulo Research Foundation – FAPESP (grant #17/21223-9), Danish Government Scholarship under the Cultural Agreements Program, and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES).

Acknowledgments: The results shown here are in part based upon data generated by the TCGA Research Network: <http://cancergenome.nih.gov/>. We would like to thank Dr. Jean-Marie Delaissé and Jacob Bastholm Olesen, Vejle Hospital, University of Southern Denmark, for the contributions with the immunofluorescence assay.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D.M.; Forman, D.; Bray, F. Cancer incidence and mortality worldwide : Sources , methods and major patterns in GLOBOCAN 2012. *2015*, 386.
2. Herbst, R.S.; Morgensztern, D.; Boshoff, C. The biology and management of non-small cell lung cancer. *Nat. Publ. Gr.* **2018**, 553, 446–454.
3. Mitchell, G.; Mitchell, C. Lung cancer. *Aust Fam Physician* **2004**, 33, 321–325.
4. Baracos, V.E.; Reiman, T.; Mourtzakis, M.; Gioulbasanis, I.; Antoun, S. Body composition in patients with non-small cell lung cancer: a contemporary view of cancer cachexia with the use of computed tomography image analysis. *Am J Clin Nutr* **2010**, 91, 1133–1137.
5. Fearon, K.; Strasser, F.; Anker, S.D.; Bosaeus, I.; Bruera, E.; Fainsinger, R.L.; Jatoi, A.; Loprinzi, C.; MacDonald, N.; Mantovani, G.; et al. Definition and classification of cancer cachexia: An international consensus. *Lancet Oncol.* **2011**, 12, 489–495.
6. Mytelka, D.S.; Li, L.; Benoit, K. Post-diagnosis weight loss as a prognostic factor in non-small cell lung cancer. *J. Cachexia. Sarcopenia Muscle* **2018**, 9, 86–92.
7. Unit, L.; Hospital, R.M.; Road, D. Do patients with weight loss have a worse outcome when undergoing chemotherapy for lung cancers ? *Br J Cancer* **2004**, 90, 1905–1911.
8. Dahlberg, S.E.; Schiller, J.H.; Bonomi, P.B.; Sandler, A.B.; Brahmer, J.R.; Ramalingam, S.S.; Johnson, D.H. Body mass index and its association with clinical outcomes for advanced non-small-cell lung cancer patients enrolled on eastern cooperative oncology group clinical trials. *J. Thorac. Oncol.* **2013**, 8, 1121–1127.
9. Martin, L.; Birdsell, L.; MacDonald, N.; Reiman, T.; Clandinin, M.T.; McCargar, L.J.; Murphy, R.; Ghosh, S.; Sawyer, M.B.; Baracos, V.E. Cancer cachexia in the age of obesity: Skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. *J. Clin. Oncol.* **2013**, 31, 1539–1547.
10. Stene, G.B.; Helbostad, J.L.; Amundsen, T.; Sørhaug, S.; Hjelde, H.; Kaasa, S.; Grønberg, B.H. Changes in skeletal muscle mass during palliative chemotherapy in patients with advanced lung cancer. *Acta Oncol* **2015**, 54, 340–348.
11. Prado, C.M.M.; Lieff, J.R.; Mccargar, L.J.; Reiman, T.; Sawyer, M.B.; Martin, L.; Baracos, V.E. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts : a population-based study. 629–635.

- 505 12. Management, S.; Care, S.; Assessment, G. Nutritional Status , Body Surface , and Low Lean Body Mass
506 / Body Mass Index Are Related to Dose Reduction and Severe Gastrointestinal Toxicity Induced by
507 Afatinib in Patients With Non-Small Cell Lung Cancer. **2015**, 967–974.
- 508 13. Fløtten, Ø.; Hjermsstad, M.J.; Aass, N.; Jordhøy, M. Low muscle mass is associated with chemotherapy-
509 induced haematological toxicity in advanced non-small cell lung cancer. **2015**, 1–7.
- 510 14. Kinsey, C.M.; San Jose Estepar, R.; Van der Velden, J.; Cole, B.F.; Christiani, D.C.; Washko, G.R.; Cancer,
511 C.L.; Velden, J. Van Der; Cole, B.F.; Kinsey, C.M.; et al. Lower Pectoralis Muscle Area Is Associated with
512 a Worse Overall Survival in Non – Small. *Cancer Epidemiol. Biomarkers Prev.* **2017**, 26, 38–44.
- 513 15. Bye, A.; Sjøblom, B.; Wentzel-Larsen, T.; Grønberg, B.H.; Baracos, V.E.; Hjermsstad, M.J.; Aass, N.;
514 Bremnes, R.M.; Fløtten, Ø.; Jordhøy, M. Muscle mass and association to quality of life in non-small cell
515 lung cancer patients. *J. Cachexia. Sarcopenia Muscle* **2017**, 8, 759–767.
- 516 16. Kilgour, R.D.; Vigano, A.; Trutschnigg, B.; Hornby, L.; Lucar, E.; Bacon, S.L.; Morais, J.A. Cancer-related
517 fatigue: The impact of skeletal muscle mass and strength in patients with advanced cancer. *J. Cachexia.*
518 *Sarcopenia Muscle* **2010**, 1, 177–185.
- 519 17. Tsoli, M.; Robertson, G. Cancer cachexia : malignant inflammation, tumorkines, and metabolic mayhem.
520 *Trends Endocrinol Metab* **2013**, 24, 174–183.
- 521 18. Baracos, V.E.; Martin, L.; Korc, M.; Guttridge, D.C.; Fearon, K.C.H. Cancer-associated cachexia. *Nat. Publ.*
522 *Gr.* **2018**, 4, 1–18.
- 523 19. Twelkmeyer, B.; Tardif, N.; Rooyackers, O. Omics and cachexia. *Curr Opin Clin Nutr Metab Care* **2017**, 20,
524 181–185.
- 525 20. Hsiao, Y.; Chu, L.; Chen, J.; Yeh, T.; Yu, J. Proteomic profiling of the cancer cell secretome: informing
526 clinical research. *Expert Rev Proteomics* **2017**, 14, 737–756.
- 527 21. Schaaij-visser, T.B.M.; Wit, M. De; Lam, S.W.; Jiménez, C.R. The cancer secretome, current status and
528 opportunities in the lung , breast and colorectal cancer context. *BBA - Proteins Proteomics* **2013**, 1834,
529 2242–2258.
- 530 22. Pavlou, M.P.; Diamandis, E.P. The cancer cell secretome: A good source for discovering biomarkers? *J.*
531 *Proteomics* **2010**, 73, 1896–1906.
- 532 23. Fukawa, T.; Yan-Jiang, B.C.; Min-Wen, J.C.; Jun-Hao, E.T.; Huang, D.; Qian, C.-N.; Ong, P.; Li, Z.; Chen,
533 S.; Mak, S.Y.; et al. Excessive fatty acid oxidation induces muscle atrophy in cancer cachexia. *Nat. Med.*
534 **2016**, 22, 666–671.
- 535 24. Schäfer, M.; Oeing, C.U.; Rohm, M.; Baysal-Temel, E.; Lehmann, L.H.; Bauer, R.; Volz, H.C.; Boutros, M.;
536 Sohn, D.; Sticht, C.; et al. Ataxin-10 is part of a cachexokine cocktail triggering cardiac metabolic
537 dysfunction in cancer cachexia. *Mol. Metab.* **2016**, 5, 67–78.
- 538 25. McLean, J.B.; Moylan, J.S.; Horrell, E.M.W.; Andrade, F.H. Proteomic analysis of media from lung cancer
539 cells reveals role of 14-3-3 proteins in cachexia. *Front. Physiol.* **2015**, 6, 1–8.
- 540 26. Aerts, H.J.W.L.; Velazquez, E.R.; Leijenaar, R.T.H.; Parmar, C.; Grossmann, P.; Cavalho, S.; Bussink, J.;
541 Monshouwer, R.; Haibe-Kains, B.; Rietveld, D.; et al. Decoding tumour phenotype by noninvasive
542 imaging using a quantitative radiomics approach. *Nat. Commun.* **2014**, 5.
- 543 27. Aerts, Hugo J. W. L., Rios Velazquez, Emmanuel, Leijenaar, Ralph T. H., Parmar, Chintan, Grossmann,
544 Patrick, Carvalho, Sara, ... Lambin, P. Data From NSCLC-Radiomics-Genomics.The Cancer Imaging
545 Archive. **2015**.
- 546 28. Penafuerte, C.A.; Gagnon, B.; Sirois, J.; Murphy, J.; Macdonald, N.; Tremblay, M.L. Identification of
547 neutrophil-derived proteases and angiotensin II as biomarkers of cancer cachexia. *Br. J. Cancer* **2016**, 114,

- 548 680–687.
- 549 29. Kuroda, K.; Nakashima, J.; Kanao, K.; Kikuchi, E.; Miyajima, A.; Horiguchi, Y. Interleukin 6 is associated
550 with cachexia in patients with prostate cancer. *Urology* **2007**, *69*, 113–117.
- 551 30. Hou, Y.-C.; Wang, C.-J.; Chao, Y.-J.; Chen, H.-Y.; Wang, H.-C.; Tung, H.-L.; Lin, J.-T.; Shan, Y.-S. Elevated
552 Serum Interleukin-8 Level Correlates with Cancer-Related Cachexia and Sarcopenia: An Indicator for
553 Pancreatic Cancer Outcomes. *J. Clin. Med.* **2018**, *7*, 502.
- 554 31. Kandarian, S.C.; Nosacka, R.L.; Delitto, A.E.; Judge, A.R.; Judge, S.M.; Ganey, J.D.; Moreira, J.D.;
555 Jackman, R.W. Tumour-derived leukaemia inhibitory factor is a major driver of cancer cachexia and
556 morbidity in C26 tumour-bearing mice. *J. Cachexia. Sarcopenia Muscle* **2018**, *9*, 1109–1120.
- 557 32. Carson, J.A.; Baltgalvis, K.A. Interleukin 6 as a key regulator of muscle mass during cachexia. *Exerc.*
558 *Sport Sci. Rev.* **2010**, *38*, 168–76.
- 559 33. Silva, E.M.; Mariano, V.S.; Aguiar Pastrez, P.R.; Pinto, M.C.; Castro, A.G.; Syrjanen, K.J.; Longatto-Filho,
560 A. High systemic IL-6 is associated with worse prognosis in patients with non-small cell lung cancer.
561 *PLoS One* **2017**, *12*, 1–12.
- 562 34. Pettersen, K.; Andersen, S.; Degen, S.; Tadini, V.; Grosjean, J.; Hatakeyama, S.; Tesfahun, A.N.; Moestue,
563 S.; Kim, J.; Nonstad, U.; et al. Cancer cachexia associates with a systemic autophagy-inducing activity
564 mimicked by cancer cell-derived IL-6 trans-signaling. *Sci. Rep.* **2017**, *7*, 2046.
- 565 35. Surowiak, P.; Budczies, J. Online Survival Analysis Software to Assess the Prognostic Value of
566 Biomarkers Using Transcriptomic Data in Non-Small-Cell Lung Cancer. *PLoS One* **2013**, *8*, e82241.
- 567 36. McDonald, M.; Diaz, A.; Ross, J.; San Jose Estepar, R.; Zhou, L.; Washko, G. Quantitative computed
568 tomography measures of pectoralis muscle area and disease severity in chronic obstructive pulmonary
569 disease. A cross-sectional study. *Ann Am Thorac Soc* **2014**, *11*, 326–334.
- 570 37. Teigen, L.; John, R.; Kuchnia, A.; Nage, E.; Earthman, C.; Kealhofer, J.; Martin, C.; Cogswell, R.
571 Preoperative Pectoralis Muscle Quantity and Attenuation by Computed Tomography Are Novel and
572 Powerful Predictors of Mortality After Left Ventricular Assist Device Implantation. *Circ Hear. Fail* **2017**,
573 *10*(9), pii: e004069.
- 574 38. Fearon, K.C.H.; Glass, D.J.; Guttridge, D.C. Cancer cachexia: mediators, signaling, and metabolic
575 pathways. *Cell Metab.* **2012**, *16*, 153–66.
- 576 39. Jafri, R.; Previgliano, C.; Khandelwal, K.; Shi, R. Cachexia Index in Advanced Non-Small-Cell Lung
577 Cancer Patients. *Clin Med Insights Oncol* **2015**, *9*, 87–93.
- 578 40. Srdic, D.; Plestina, S.; Sverko-Peternac, A.; Nikolac, N.; Simundic, A.M.; Samarzija, M. Cancer cachexia,
579 sarcopenia and biochemical markers in patients with advanced non-small cell lung cancer—
580 chemotherapy toxicity and prognostic value. *Support. Care Cancer* **2016**, *24*, 4495–4502.
- 581 41. Pfizenmaier, J.; Vessella, R.; Higano, C.S.; Noteboom, J.L.; Wallace, D.; Corey, E. Elevation of cytokine
582 levels in cachectic patients with prostate carcinoma. *Cancer* **2003**, *97*, 1211–1216.
- 583 42. Tazaki, E.; Shimizu, N.; Tanaka, R.; Yoshizumi, M.; Kamma, H.; Imoto, S.; Goya, T.; Kozawa, K.; Nishina,
584 A.; Kimura, H. Serum cytokine profiles in patients with prostate carcinoma. *Exp. Ther. Med.* **2011**, *2*, 887–
585 891.
- 586 43. Alves, M.J.; Figuerêdo, R.G.; Azevedo, F.F.; Cavallaro, D.A.; Neto, N.I.P.; Lima, J.D.C.; Matos-Neto, E.;
587 Radloff, K.; Riccardi, D.M.; Camargo, R.G.; et al. Adipose tissue fibrosis in human cancer cachexia: The
588 role of TGF β pathway. *BMC Cancer* **2017**, *17*, 1–12.
- 589 44. Zhang, D.; Song, B.; Wang, S.; Zheng, H.; Wang, X. Association of interleukin-8 with cachexia from
590 patients with low-third gastric cancer. *Comp. Funct. Genomics* **2009**, *2009*, 1–6.

45. Lerner, L.; Hayes, T.G.; Tao, N.; Krieger, B.; Feng, B.; Wu, Z.; Nicoletti, R.; Isabel Chiu, M.; Gyuris, J.; Garcia, J.M. Plasma growth differentiation factor 15 is associated with weight loss and mortality in cancer patients. *J. Cachexia. Sarcopenia Muscle* **2015**, *6*, 317–324.
46. Kuroda, K.; Nakashima, J.; Kanao, K.; Kikuchi, E.; Miyajima, A.; Horiguchi, Y.; Nakagawa, K.; Oya, M.; Ohigashi, T.; Murai, M. Interleukin 6 Is Associated with Cachexia in Patients with Prostate Cancer. *Urology* **2007**, *69*, 113–117.
47. Richey, L.M.; George, J.R.; Couch, M.E.; Kanapkey, B.K.; Yin, X.; Cannon, T.; Stewart, P.W.; Weissler, M.C.; Shores, C.G. Defining cancer cachexia in head and neck squamous cell carcinoma. *Clin. Cancer Res.* **2007**, *13*, 6561–6567.
48. Krzystek-Korpacka, M.; Matusiewicz, M.; Diakowska, D.; Grabowski, K.; Blachut, K.; Kustrzeba-Wojcicka, I.; Banas, T. Impact of weight loss on circulating IL-1, IL-6, IL-8, TNF- α , VEGF-A, VEGF-C and midkine in gastroesophageal cancer patients. *Clin. Biochem.* **2007**, *40*, 1353–1360.
49. Dolan, R.D.; Almasaudi, A.S.; Dieu, L.B.; Horgan, P.G.; McSorley, S.T.; McMillan, D.C. The relationship between computed tomography-derived body composition, systemic inflammatory response, and survival in patients undergoing surgery for colorectal cancer. *J. Cachexia. Sarcopenia Muscle* **2018**.
50. van Dijk, D.P.J.; Krill, M.; Farshidfar, F.; Li, T.; Rensen, S.S.; Olde Damink, S.W.M.; Dixon, E.; Sutherland, F.R.; Ball, C.G.; Mazurak, V.C.; et al. Host phenotype is associated with reduced survival independent of tumour biology in patients with colorectal liver metastases. *J. Cachexia. Sarcopenia Muscle* **2018**.
51. Dunstan, C.A.N.; Salafranca, M.N.; Adhikari, S.; Xia, Y.; Feng, L.; Harrison, J.K. Identification of two rat genes orthologous to the human interleukin-8 receptors. *J. Biol. Chem.* **1996**, *271*, 32770–32776.
52. Gerber, M.H.; Underwood, P.W.; Judge, S.M.; Delitto, D.; Delitto, A.E.; Nosacka, R.L.; DiVita, B.B.; Thomas, R.M.; Permeth, J.B.; Hughes, S.J.; et al. Local and Systemic Cytokine Profiling for Pancreatic Ductal Adenocarcinoma to Study Cancer Cachexia in an Era of Precision Medicine. *Int. J. Mol. Sci.* **2018**, *19*, 3836.
53. Clark, K.; Vendt, B.; Smith, K.; Freymann, J.; Kirby, J.; Koppel, P.; Moore, S.; Phillips, S.; Maffitt, D.; Pringle, M.; et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *J. Digit Imaging* **2013**, *26*, 1045–1057.
54. Freymann, J.; Kirby, J.; Perry, J.; Clunie, D.; Jaffe, C. Image data sharing for biomedical research--meeting HIPAA requirements for De-identification. *J. Digit Imaging* **2012**, *25*, 14–24.
55. Kinsey, C.M.; Cole, B.F. Lower Pectoralis Muscle Area is Associated with a Worse Overall Survival in Non- Small Cell Lung Cancer. *Cancer Epidemiol Biomarkers Prev* **2017**, *26*, 38–43.
56. Kim, E.Y.; Kim, Y.S.; Park, I.; Ahn, H.K. Evaluation of sarcopenia in small-cell lung cancer patients by routine chest CT. *Support. Care Cancer* **2016**, *24*, 4721–4726.
57. Nilufer, G.; Ayse, Y. Estimation of stature and sex from sternal lengths : an autopsy study. **2015**, 89–96.
58. McCarthy, I.D.; McGregor, A.H.; Coombs, R.R.H. Geometrical dimensions of the lower lumbar vertebrae – analysis of data from digitised CT images. **2000**, 242–248.
59. Yaguchi, Y.; Kumata, Y.; Horikawa, M.; Kiyokawa, T. Clinical Significance of Area of Psoas Major Muscle on Computed Tomography after Gastrectomy in Gastric Cancer Patients. **2017**, 145–149.
60. Barrett, T.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.; Marshall, K.A.; Phillippy, K.H.; Sherman, P.M.; Holko, M.; et al. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res.* **2013**, *41*, D991–5.
61. Blake, J.; Christie, K.; Dolan, M.; Drabkin, H.; Hill, D.; Ni, L.; Sitnikov, D.; Westerfield, M. Gene Ontology Consortium: going forward. *Nucleic Acids Res* **2015**, *43*, 1049–1056.

- 634 62. Szklarczyk, D.; Morris, J.H.; Cook, H.; Kuhn, M.; Wyder, S.; Simonovic, M.; Santos, A.; Doncheva, N.T.;
635 Roth, A.; Bork, P.; et al. The STRING database in 2017: quality-controlled protein–protein association
636 networks, made broadly accessible. *Nucleic Acids Res.* **2017**, *45*, D362–D368.
- 637 63. Snel, B.; Lehmann, G.; Bork, P.; Huynen, M.A. STRING: a web-server to retrieve and display the
638 repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res.* **2000**, *28*, 3442–4.
- 639 64. Petersen, T.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: discriminating signal peptides from
640 transmembrane regions. *Nat Methods.* **2011**, *29*, 785–786.
- 641 65. Bendtsen, J.; Jensen, L.; Blom, N.; Von Heijne, G.; Brunak, S. Feature-based prediction of non-classical
642 and leaderless protein secretion. *Protein Eng Des Sel* **2004**, *17*, 349–356.
- 643 66. Mathivanan, S.; Fahner, C.; Reid, G.; Simpson, R. ExoCarta 2012: database of exosomal proteins, RNA
644 and lipids. *Nucleic Acids Res* **2012**, *40*, 1241–1244.
- 645 67. Emanuelsson, O.; Nielsen, H.; Brunak, S.; von Heijne, G. Predicting subcellular localization of proteins
646 based on their N-terminal amino acid sequence. *J Mol Biol.* **2000**, *300*, 1005–1016.
- 647 68. Feizi, A.; Banaei-esfahani, A.; Nielsen, J. Database tool HCSD : the human cancer secretome database.
648 *Database (Oxford)* **2015**, bav051.
- 649 69. Kalra, H.; Simpson, R.; Ji, H.; Aikawa, E.; Altevogt, P.; Askenase, P. Vesiclepedia: a compendium for
650 extracellular vesicles with continuous community annotation. *PLoS Biol* **2012**, *10*, e1001450.
- 651 70. Kim, D.; Lee, J.; Kim, S.; Choi, D.; Yoon, Y. EVpedia: a community web portal for extracellular vesicles
652 research. *Bioinformatics* **2015**, *31*, 933–939.
- 653 71. Boutet, E.; Lieberherr, D.; Tognolli, M.; Schneider, M.; Bairoch, A. UniProtKB/Swiss-Prot, the Manually
654 Annotated Section of the UniProt KnowledgeBase: How to Use the Entry View. *Methods Mol Biol* **2016**,
655 *1374*, 23–54.
- 656 72. Nanjappa, V.; Thomas, J.; Marimuthu, A.; Muthusamy, B. Plasma Proteome Database as a resource for
657 proteomics research: 2014 update. *Nucleic Acids Res* **2014**, *42*, 959–965.
- 658 73. Aguirre-gamboa, R.; Gomez-rueda, H.; Martu, E.; Chacolla-huaringa, R.; Rodriguez-barrientos, A.;
659 Tamez-Peña, J.; Treviño, V. SurvExpress : An Online Biomarker Validation Tool and Database for Cancer
660 Gene Expression Data Using Survival Analysis. *PLoS One* **2013**, *8*, e74250.
- 661 74. Tomczak, K.; Czerwińska, P.; Wiznerowicz, M. The Cancer Genome Atlas (TCGA): an immeasurable
662 source of knowledge. **2014**.
- 663 75. Rousseaux, S.; Debernardi, A.; Jacquiau, B.; Vitte, A. Ectopic activation of germline and placental genes
664 identifies aggressive metastasis-prone lung cancers. *Sci Transl Med* **2013**, *5*, 186ra66.
- 665 76. Yamauchi, M.; Yamaguchi, R.; Nakata, A.; Kohno, T. Epidermal growth factor receptor tyrosine kinase
666 defines critical prognostic genes of stage I lung adenocarcinoma. *PLoS One* **2012**, *7*, e43923.
- 667 77. Okayama, H.; Kohno, T.; Ishii, Y.; Shimada, Y. Identification of genes upregulated in ALK-positive and
668 EGFR/KRAS/ALK-negative lung adenocarcinomas. *Cancer Res* **2012**, *72*, 100–111.
- 669 78. Shedden, K.; Taylor, J.; Enkemann, S. Director's Challenge Consortium for the Molecular Classification
670 of Lung Adenocarcinoma. Gene expression-based survival prediction in lung adenocarcinoma: a multi-
671 site, blinded validation study. *Nat Med* **2008**, *14*, 822–7.
- 672 79. Lee, E.; Son, D.; Kim, S.; Lee, J. Prediction of recurrence-free survival in postoperative non-small cell
673 lung cancer patients by using an integrated model of clinical information and gene expression. *Clin*
674 *Cancer Res* **2008**, *14*, 7397–404.
- 675