Torres-Martinez Susana (Orcid ID: 0000-0001-8366-8337)
Calabuig-Fariñas Silvia (Orcid ID: 0000-0001-7749-174X)
Moreno-Manuel Andrea (Orcid ID: 0000-0003-2996-2505)
Durendez-Saez Elena (Orcid ID: 0000-0002-7304-7796)
Jantus-Lewintre Eloisa (Orcid ID: 0000-0001-7395-4380)

Soluble galectin-3 as a microenvironment-relevant immunoregulator with prognostic and predictive value in lung adenocarcinoma

Susana Torres-Martínez[1,2,3], Silvia Calabuig-Fariñas[1,2,3,4,*], Andrea Moreno-Manuel[1], Giulia Bertolini[5], Alejandro Herreros-Pomares[3,6], Eva Escorihuela[1,2], Elena Duréndez-Saéz[1,2], Ricardo Guijarro[3,7,8], Ana Blasco[2,3,9], Luca Roz[5], Carlos Camps[1,2,3,9,10], Eloisa Jantus-Lewintre[1,2,3,6,11*].

[1] Molecular Oncology Laboratory, Fundación Investigación Hospital General Universitario de Valencia, 46014 Valencia, Spain;

[2] TRIAL Mixed Unit, Centro Investigación Príncipe Felipe—Fundación Investigación Hospital General Universitario de Valencia, 46014 Valencia, Spain;

[3] Centro de Investigación Biomédica en Red Cáncer, CIBERONC, 28029 Madrid, Spain;

[4] Department of Pathology, Universitat de València, 46010 Valencia, Spain

[5] Tumor Genomics Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy.

[6] Department of Biotechnology, Universitat Politècnica de València, 46022 Valencia, Spain

[7] Department of Surgery, Universitat de València, 46010 Valencia, Spain

[8] Department of Thoracic Surgery, Hospital General Universitario de Valencia, 46014 Valencia, Spain

[9] Department of Medical Oncology, Hospital General Universitario de Valencia, 46014 Valencia, Spain

[10] Department of Medicine, Universitat de València, 46010 Valencia, Spain

[11] Joint Unit: Nanomedicine, Centro Investigación Príncipe Felipe—Universitat Politècnica de Valencia, 46022 Valencia, Spain

*** Corresponding authors:**

Eloisa Jantus Lewintre (E.J.-L.); ejantus@btc.upv.es, (+34) 963 187 556

Silvia Calabuig Fariñas (S.C.-F.); calabuix_sil@gva.es, (+34) 963 187 556

Running title: sGalectin-3 as immunoregulator factor in LUAD

Keywords: sGAL-3, tumorspheres, T_{REG}, biomarker, lung adenocarcinoma.

Abbreviations

NSCLC, non-small cell lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous carcinoma; TME, tumor microenvironment; ctDNA, circulating tumor DNA; RTqPCR, reverse transcription–quantitative real-time PCR; GAL-3, galectin-3; 3D, three-dimensional; sGAL-3, soluble galectin-3; PRE, pretreatment; FR, first response assessment; NGS, next-generation

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/1878-0261.13505

sequencing; ATCC, American Type Culture Collection; SPH, tumorspheres; ADH, adherent; OS, overall survival; RFS, relapse-free survival; EVs, extracellular vesicles; PFS, progression-free survival; ORR, overall response rate; DCB, durable clinical benefit; T_{REGS}, regulatory T cells; AUC, area under curve; HR, hazard ratio; TCGA, The Cancer Genome Atlas; ICBs, immune checkpoint blockers; CSCs, cancer stem cells; TPS, tumor proportion scores; NGS, next-generation sequencing; ATCC, American Type Culture Collection; STR, short tandem repeat analysis; FBS, Fetal Bovine Serum; EGF, Epidermal Growth Factor; bFGF, Fibroblast Growth Factor; ITS, Insulin-Transferrin-Selenium; FBM, Fibroblast Basal Medium; CM, Conditioned media; PBMCs, peripheral blood mononuclear cells; PBS, phosphate-buffered saline; RNA, ribonucleic acid; RT, Reverse transcription; FFPE, formalin-fixed, paraffin-embedded; cDNA, complementary DNA; PVDF, polyvinylidine difluoride; PE, phycoerythrin; BSA, bovine serum albumin; EDTA, Ethylenediaminetetraacetic acid; 7AAD, 7-amino-actinomycin D; BV, Brilliant Violet; QC, Quality controls; CV, coefficients of variation; LLOQ, lower limit of quantification; RECIST 1.1, Response Evaluation Criteria in Solid Tumors version 1.1; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; NK, natural killer; TIICs, tumor-infiltrating immune cells; PCA, principal component analysis; IQR, interquartile range; ROC, Receiving Operating Curve; CI, confidence interval; SPSS, Statistical Package for the Social Sciences; LGALS3BP, Galectin-3 binding protein; ACTB, Beta-actin; IBs, Immunoblots; IF, immunofluorescence; TNM, tumor node metastasis; EMA, European Medicines Agency; ELISA, enzyme-linked immunosorbent assay.

Abstract

Despite the success of therapies in lung cancer, more studies of new biomarkers for patient selection are urgently needed. The present study aims to analyze the role of galectin-3 (GAL-3) in the lung tumor microenvironment (TME) using tumorspheres as a model and explore its potential role as a predictive and prognostic biomarker in non-small cell lung cancer (NSCLC) patients. For *in vitro* studies, lung adenocarcinoma (LUAD) and lung squamous carcinoma (LUSC) primary cultures from early-stage patients and commercial cell lines were cultured, using tumorsphere-forming assays and adherent conditions for the control counterparts. We analyzed the pattern of secretion and expression of GAL-3 using reverse transcription–quantitative real-time PCR (RTqPCR), immunoblot, immunofluorescence, flow cytometry and immunoassay analysis. Our results using three-dimensional (3D) models of lung tumor cells revealed that soluble GAL-3 (sGAL-3) is highly expressed and secreted. To more accurately mimic the TME, a co-culture of tumorspheres and fibroblasts was used, revealing that GAL-3 could be important as an immunomodulatory molecule expressed and secreted in the TME, modulating immunosuppression through regulatory T cells (T_{REGS}). In the translational phase, we confirmed that patients with high expression levels of GAL-3 had more T_{REGS}, which suggests that tumors may be recruiting this population through GAL-3. Next, we evaluated levels of sGAL-3 before surgery in LUAD and LUSC patients, hypothesizing that sGAL-3 could be used as an independent prognostic biomarker for overall survival and relapse-free survival in early-stage LUAD patients. Additionally, levels of sGAL-3 at pretreatment and first response assessment from plasma to predict clinical outcomes in advanced LUAD and LUSC patients treated with first-line pembrolizumab were evaluated, further supporting that sGAL-3 has a high efficiency in predicting durable clinical response to pembrolizumab with an area under curve (AUC) of 0.801 (p=0.011). Moreover, high levels might predict decreased progression-free survival and overall survival to anti-PD-1 therapy, with sGAL-3 being a prognosis-independent biomarker for advanced LUAD.

1. Introduction

Lung cancer is the second most diagnosed cancer in both men and women and the leading cause of cancer death worldwide [1] . Non-small cell lung cancer (NSCLC) is the most represented of lung cancer cases (eighty-five percent), including lung squamous cell carcinoma (LUSC; ~30%), adenocarcinoma (LUAD; ~50%) and others (~20%) [2,3]. On the other hand, in early stage, the first therapeutic option is the surgery, but the prognosis of NSCLC has gradually improved through advanced in therapeutic approaches like neoadjuvant chemotherapy and immunotherapy [4]. However, 10-60% of patients relapse within 5 years after radical resection and frequently cases are diagnosed at advanced stages, when surgery is not possible [5]. Therefore, the identification of useful biomarkers through a non-invasive approach to predict relapse is a priority. On the other hand, in advanced stages, the blockade of immune checkpoints has opened up a new standard of treatment for cancer patients, producing an effective antitumor response in tumor microenvironment (TME), concretely PD1/PDL1 axis inhibitors have been extensively studied and have drastically changed the therapeutic scenario for NSCLC with a plethora of clinical data demonstrating superior outcomes related to conventional therapies or molecular targeted therapy [6–9]. However, the efficacy of cancer immunotherapy is limited by multiples immunosuppressive mechanisms presents in tumor microenvironment (TME). Therefore, the better comprehension of the interactions in TME between the immune system and tumor cells are necessary to develop new immunotherapeutic strategies more effective in NSCLC. The expression level of PD-L1 on tumor cells or tumor infiltrating immune cells is considered the most available and implemented biomarker to select patients. However, significant percentage of PD-L1-positive NSCLCs cases do not respond to immune checkpoint blockers (ICBs), opposite a significant number of PD-1-negative tumors are sensitive to this therapy limiting its use in clinical practice [6,10,11]. Taking into consideration the abovementioned features, the identification of new reliable biomarkers, preferably tested in a minimal invasive manner, to guide patient selection and provide indications of efficacy and/or prognosis is a priority. In this line, exists intense interest in identifying predictive biomarkers derived from peripheral blood or minimal invasive samples. Some plasmatic biomarkers such as circulating tumor DNA (ctDNA) have been associated with clinical benefit and survival [12,13]. However, the prognostic and/or predictive value of soluble

plasma biomarkers in NSCLC have been sparsely validated in prospective studies and its role is not clearly understood.

Regarding TME, fibroblast, cancer stem cells (CSCs), tumor cells and immune cells can interact contributing to immunosuppression. One important protein that contribute to TME immunosuppression is the glycoprotein galectin-3 (GAL-3). GAL-3 is a carbohydrate-binding protein that might have a crucial role promoting tumor growth and helping tumors to escape immune surveillance through immunosuppression [14]. In human genome galectin-3 is coded by a single gene *LGALS3* which is suited on chromosome 14, locus q21–q2 [15]. Data have been shown that the intracellular Gal-3 promoted tumor growth, metastasis and survival and the extracellular GAL-3 may facilitate metastasis by promoting immune scape which has been poorly investigated [16,17].

To study the TME, multiple 3D model systems have been proposed as new approaches to examine it, ranging from the simple co-culture of cells in hydrogels, to complex multicomponent microfluidics, each with their own advantages and limitations [18]. Specifically, tumorspheres model provide an environment more similar to the tumor, with self-imposed nutrient, with better immuno-modulatory abilities and hypoxic gradients adding dimensions that not happened with monolayer 2D cell cultures [19].

Galectin-3 could be an immunosuppressive molecule involved in tumor scape from immune surveillance with the TME implicated so we proposed to study the expression and secretion of GAL-3 on 3D models of lung tumor cells analyzing its influence on $T_{REGS.}$ Moreover, as the clinical importance on recurrence of galectin-3 after surgery in NSCLC patients has not been elucidated fully, we aimed to evaluate the prognostic and recurrence predictive value of sGAL-3 on these patients. Finally, taking into account that there is a necessity of looking for new reliable biomarkers for ICBs, the objective of this study it is not only analyzed the role of galectin-3 in early patients but also in advanced patients to improve immune therapeutic strategies.

2. Materials and methods

2.1. Patients and plasma samples collection

This study included 137 individuals from the General University Hospital of Valencia divided in two different cohorts. Early cohort comprised 48 patients with early-stage lung adenocarcinoma (LUAD) and 42 patients with early-stage lung squamous cell carcinoma (LUSC) collected from July 2004 to September 2019. Plasma samples were obtained before

surgery and selected by following eligibility criteria: candidate for surgical resection, non-pretreated, over 18 years, non-pregnant, stage I–IIIA (according to the American Joint Committee on Cancer staging manual) and with a histological diagnosis of NSCLC. Cryopreserved tumor tissue samples from 19 patients were used in this study. Data of expression of FOXP3, CD4, and CD8 in both tumor and stromal areas (via immunohistochemistry and RT-qPCR) from these patients were collected from Usó M et al [20]. Advanced cohort included 47 patients treated with first-line pembrolizumab in monotherapy (200 mg every 21 days) (34 patients with advanced LUAD and 13 with advanced LUSC) (collected from February 2018 to July 2021) and fitted the following eligibility criteria: candidate for pembrolizumab treatment, non-pretreated, over 18 years, non-pregnant, irresectable stage IIIA-IV (according to the American Joint Committee on Cancer staging manual) and with a histological diagnosis of NSCLC. According to guidelines, PD-L1 expression ≥50% (assed by tumor proportion scores (TPS) and defined as the number of positive tumor cells divided by the total number of viable tumor cells multiplied by 100%) was present in tumor samples from all patients treated with pembrolizumab in monotherapy [21]. 34 plasma samples at pretreatment (PRE) were collected prior to the first administration of pembrolizumab and 25 plasma samples at first response assessment (FR) for LUAD advanced cohort and 13 samples at pre-treatment (PRE) were collected prior to the first administration of pembrolizumab and at first response assessment (FR) for LUSC advanced cohort. All patients were followed up until December 2022. All peripheral blood samples were collected in 10mL-EDTA tubes plasma, were isolated by centrifugation at 4ºC and then stored at -80ºC until the analysis.

This study was conducted in accordance with the Declaration of Helsinki, and along with the protocol, were approved by the ethical review board of the General University Hospital of Valencia (Nº 5/2015). All patients and healthy volunteers signed an informed consent for sample acquisition for research purposes before the beginning of this study.

2.2. Establishment of Primary Cell Cultures

Following the tumor dissociation protocol previously described by our group surgical tumor specimens from patients were established as monolayers and tumorspheres [22]. For this study, 3 primary patient-derived lung cancer long-term cultures (PC301, PC435 and PC471) were employed. Tumor profiling of each patient-derived culture was determined by next-generation sequencing (NGS) using Oncomine Focus Assay (Thermofisher Scientific, Waltham, MA, USA) and Ion GeneStudio S5 System (Thermofisher Scientific, Waltham, MA, USA) to get complete tumor profiling of each patient.

2.3. Commercial NSCLC and Fibroblast Cell Lines

Fifteen human NSCLC cell lines, LUAD cell lines (A549 (RRID:CVCL_0023), NCI-H1395 (RRID:CVCL_1467), NCI-H1650 (RRID:CVCL_1483), NCI-H1975 (RRID:CVCL_1511) NCI-H1993 (RRID:CVCL_1512) , NCI-H2228 (RRID:CVCL_1543), NCI-H23 (RRID:CVCL_1547), NCI-H358 (RRID:CVCL_1559), HCC827 (RRID:CVCL_2063), PC9 (RRID:CVCL_B260)) and LUSC cell lines ((SW900 (RRID:CVCL_1731), LUDLU-1 (RRID:CVCL_2582), NCI-H520 (RRID:CVCL_1566), NCI-H1703 (RRID:CVCL_1490) and SK-MES-1 (RRID:CVCL_0630)) were used for in vitro experiments. LUAD cell lines were obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA) and LUSC cell lines were kindly provided by Dr. J. Carretero (University of Valencia, Spain) unless SW900 that was purchased from ATCC. Immortalize primary fibroblast, CAF154-hTERT cells are originally from cancer-associated primary fibroblasts and were kindly provided by Dr. Luca Roz (Istituto Nazionale dei Tumori, Italy) The generation and the characteristics of them have been described previously [23]. All cell cultures (primary and commercial) were tested for mycoplasma before all the experiments. All human cell cultures were authenticated by short tandem repeat analysis (STR) with AmpFISTR™ Identifiler™ Plus PCR Amplification Kit (Thermofisher Scientific, Waltham, MA, USA).

2.4. Cell Culture Conditions for Tumor Cells and Fibroblast

Tumor cells were grown in RPMI-1640 (commercial cell lines) or DMEM-F12 (primary cultures) containing 10% Fetal Bovine Serum (FBS), 100 µg/mL penicillin/streptomycin, 0.001% non-essential amino acids (for RPMI-1640) and 2 mM L-glutamine (for DMEM-F12) (Gibco™, Grand Island, NY, USA). In order to obtain tumorspheres, when cells reached 80% confluence, they were trypsinized using 0.1% trypsin-EDTA (Corning, NY, USA). After that, cells were cultured at low density in ultra-low attachment flasks (Corning, NY, USA) with serum-free (RPMI-1640/DMEM-F12) medium supplemented with 0.4% Bovine Serum Albumin (BSA), 50 µg/ mL Epidermal Growth Factor (EGF), 20 µg/mL basic Fibroblast Growth Factor (bFGF), Insulin-Transferrin-Selenium (ITS) PREMIX, 100 µg/mL penicillin/streptomycin (P/S) and 2% B27 (Gibco™, Grand Island, NY, USA). The following experiments took place after 5 days when the cells started to grow and form floating aggregated. CAF154-hTERT cells were grown in) Fibroblast Basal Medium (FBM supplemented with the Kit-Low serum (ATCC, Manassas, VA, USA). All cells were maintained at 37ºC in humidified atmosphere of 5% CO2 and 95% air.

2.5. Co-cultures Conditions

For co-cultures, 3 x 10^5 CAF154-hTERT were cultured for 2 hours with the proper medium in 6-well plates. After 2 hours, 1x10^5 adherent or tumorspheres PC435 were cultured together with CAF154-hTERT in 50% of FBM and 50% SPH DMEM F12 for 48 hours. Conditioned

media were collected from different conditions (tumorspheres PC435 or co-culture tumorspheres PC435 + CAF154-hTERT). Conditioned media (CM) will be used in the following experiment to test the effect on regulatory T cells (also called T_{REG}).

2.6. PBMCs Cultures and CM Treatment

Human peripheral blood mononuclear cells (PBMCs) from 9 healthy volunteers were plated at $1x10^6$ cells/well in 6-well plates and incubated at 37°C for 4 h. After the incubation, non-adherent cells (T cells) were collected and used for the experiments. $1x10^7$ cells/well were treated with different CM collected from PC435 cultures and PC435+CAF154-hTERT co-cultures. At the same time, the galectin-3 monoclonal antibody (clone B2C10) (100ng/ml) (Thermofisher Scientific, Waltham, MA, USA) were added to the culture in order to blocked soluble Gal-3 in culture media to test its effect on the T_{REG} population.

2.7. Cellular pellets and supernatants collection

Both adherent cells and tumorspheres were seed at different densities for the following experiments (10.000 cell/ml and 100.000 cell/ml) in 24-well plates. Supernatant were collected at two time periods post-seeded (12h and 24h) and store at -80ºC until further analysis. Cell pellets were collected at the same points with TRIZol reagent (Invitrogen, Waltham, MA, USA) and frozen at -80ºC until the experiments for gene expression analysis.

2.8. Isolation of Extracellular Vesicles from Cell Cultures.

To isolate tumor-derived extracellular vesicles (EVs) from cultures, cells were grown in T175 cm^2 flasks until 70–80% confluence for 72 h in 30 mL of FBS-depleted media (in the case of tumorspheres cultures). After 72 h, detritus was eliminated by differential centrifugation at 500× *g* for 5 min, and then at 3000× *g* for 15 min. Subsequently, the supernatant was filtered through a 0.2 µm filter (Corning, NY, USA) and ultracentrifuged at 110,000× *g* for 90 min (CP-NX, P50AT2 Rotor; Hitachi, Japan). To wash the first pellet, second ultracentrifugation was performed; EVs were then resuspended in 30 mL of phosphate-buffered saline (PBS). All centrifugations were performed at 4 °C. At last, EVs were resuspended in a tiny volume (30–60 µL) of filtered PBS and stored at −80 °C until the corresponding analysis.

2.9. Gene expression analysis

The extraction of total cellular ribonucleic acid (RNA) from cell pellets and frozen tissue samples was performed using standard TRIZol method according to manufactures' instructions. Exosomal total RNA derived from cell cultures was isolated using the Total RNA

Purification Kit (Norgen Biotek, Thorold, ON, Canada). RNA concentrations were evaluated by Nanodrop (Thermofisher Scientific, Waltham, MA, USA). Reverse transcription–quantitative real time PCR (RTqPCR) was carried out to analyze the relative expression of *LGALS3* gene and reference genes on a Roche LightCycler®480 II system (Roche Ltd., Basel, Switzerland) (Table S1). Reverse transcription (RT) reactions were performed from 1.0 µg of total RNA (frozen tissue and FFPE samples) 0.5 µg of total RNA (cells samples) and 0.150 µg (EVs samples) using random hexanucleotides and a High-Capacity complementary DNA (cDNA) Reverse Transcription Kit (Applied Biosystems, USA) according to the manufacturer's instructions. The resulting cDNA was used for RTqPCR reaction and was carried out with assays based on hydrolysis probes using 1 µL of cDNA, TaqMan Gene Expression Master Mix, and a TaqMan Gene Expression Assay (Applied Biosystems, Waltham, MA, USA) in final reaction volume of 5 µl. We used random-primed qPCR Human Reference cDNA (Clontech, USA) for efficiency calculations. Using GeNorm software (https://genorm.cmgg.be/ accessed on 9 July 2015) [24], *ACTB*, *GUSB*, and *CDKN1B* were selected as endogenous controls for cells and frozen tissue, whereas *ACTB* and *GAPDH* were selected as endogenous controls for EVs samples selected as endogenous controls using GeNorm software. Relative gene expression levels of *LGALS3* and *LGALS3BP* were calculated as the ratio of target gene expression to the geometric mean of the endogenous gene expressions according to Pfaffl formula [25]. All samples were tested in triplicate.

2.10. Immunoblot analysis

Tumorspheres were washed with cold phosphate-buffered saline (PBS), whereas adherent cells were also scraped out of the dishes before lysis. Protein pellets were lysed using a lysis buffer composed of 100mM Tris pH8, 2% NP40, 1% Na deoxicholate, 0.2% SDS and 300mM NaCl, 1mM sodium orthovanadate, 25mM NaF and protease inhibitor cocktail (Roche, Basel, Switzerland). BCA Protein Assay (Thermofisher Scientific, Waltham, MA, USA) was employed to quantify the total protein concentration; 30 ug of total protein were separated on 12% SDS-polyacrylamide gel and electro-transferred to a 0.45 um polyvinylidine difluoride (PVDF) membrane (MilliporeSigma, Burlington, MA, USA). The membrane was then blocked with 5% skim milk for 1 h and immunoblotted overnight at 4 °C with the Anti-Galectin 3 antibody (Clone A3A12) (ab2785, Abcam, Cambridge, UK). Afterwards, membranes were incubated with anti-IgG (whole molecule)-Peroxidase secondary antibody (Thermo Fisher Scientific, Waltham, MA, USA) for 1 h at room temperature. Chemiluminescent detection with the high-sensitivity Amersham ECL Select™ detection reagent (GE Healthcare, Chicago, IL, USA) was employed (Table S2). All results were normalized over β-actin (Sigma-Aldrich, St. Louis, MO, USA).

2.11. Flow Cytometry Analysis

To analyze tumor cell surface markers, single cell solution was washed in staining buffer (PBS1×+ 0.5% BSA+ 2mM EDTA) and incubated for 30 min at 4C° with phycoerythrin (PE) anti-GAL-3 (clone M3/38) (Biolegend, San Diego, CA, USA) (Table S2). For these analyses, dead cells were excluded using 7AAD Viability Staining (Thermofisher Scientific, Waltham, MA, USA) (Table S2).

For analysis of Treg phenotype, T cells treated before with CM (tumorspheres or co-culture) with and without galectin-3 monoclonal antibody, were first incubated with surface antibodies in staining buffer for 30 min at 4°C: Brilliant Violet V510 (BV510) Mouse Anti-Human CD3 (Clone HIT3a), Brilliant Violet V421 (BV42) Anti-Human CD4 (Clone SK3), Allofhycocyanin (APC) Anti-Human CD25 (clone M-A251); then fixed and permeabilized with Transcription Factor Buffer Set (Thermo Fisher Scientific, Waltham, MA, USA), according to the datasheet instructions, and finally incubated with PE anti-Human FoxP3 (Clone 259D/C7) (all from BD Biosciences, Cambridge, UK) for 30 min at 4°C (Table S2). T_{REGS} were identified within live cell gate as CD3+CD4+Foxp3+CD25high. For these analyses, dead cells were excluded using Fixable Viability Stain 780 (BD Horizon, Franklin Lakes, NY, USA) (Table S2). Signal were acquired using a FC500 MPL Flow Cytometer and CytExpert v2.3 software (Beckman-Coulter, Inc., Brea, CA, USA).

2.12. Immunofluorescence Analysis

Cells were fixed in 4% paraformaldehyde in PBS at room temperature for 15 min, washed and permeabilized with 0.4% Triton X-100 in PBS for 10 min, and washed again with PBS. Permeabilized cells were blocked with PBS containing 1% BSA for 1 h, and subsequently incubated with GAL-3 anti-mouse [1:200] (ab2785, Abcam, Cambridge, UK) antibody in blocking buffer overnight at 4 °C (Table S2). Thereafter, cells were washed with PBS and incubated with Alexa-labelled IgG secondary antibodies containing blocking buffer for 1 h. Slides were incubated with 4′,6-diamidino-2-phenylindole for 3 min, mounted with Fluoromount Aqueous Mounting Medium (Sigma-Aldrich, St. Louis, MO, USA), and analyzed using a Leica confocal microscope (Leica Microsystems, Buffalo Grove, IL, USA).

2.13. In silico analysis

Firstly, an in-silico analysis was carried out using two lung cancer data sets from the The Cancer Genome Atlas (TCGA) consortium to study the expression of galectin-3 in early NSCLC patients [26,27]. RNA-sequencing (Ilumina Hi Seq platform) and clinical information was downloaded from the ICGC Data Portal, https://dcc.icgc.org/releaes/current/projects/LUAD-US and

https://dcc.icgc.org/releases/current/projects/LUSC-US [28]. The Limma package from Bioconductor was used to obtain normalized RNA-seq data. Linear fit model for LGALS3 was obtained before constructing the different contrast matrixes. Given the linear models, empirical Bayes statistics were computed for differential expression analysis.

2.14. Immunoassay based on Luminex xMAP

Supernatants of cell cultures or plasma samples were assayed through multiplex magnetic bead-based immunoassay technology based on flow cytometry using Human Circulating Cancer Biomarker Magnetic Bead Panel 3, 96 Well Plate Assay, Cat. # HCCBP3MAG-58K and Human Immuno-Oncology Checkpoint Protein Panel 2 - Immuno-Oncology Multiplex Assay, Cat. #HCKP2-11K (Merck Millipore, Billerica, MA) to quantify levels of galectin-3 produced by tumor cells in the culture medium and in plasma, respectively. Quality controls (QC1 and QC2), as well as a calibration curve based on 1:4 dilutions of the highest standard were used for quantification and as internal controls for intra- and inter-assay reproducibility. Briefly, 25 µl of culture medium or plasma samples (diluted 1:2) were used for each sample and mixed with proper regents and monoclonal antibody to human galectin-3, which are covalently bound to the surface of magnetic microspheres dyed with accurate amounts of red and infrared fluorophores in order to produce a single spectral signature which can be detected in the Luminex platform (Luminex Corp, Austin, Tx). Soluble galectin-3 quantification is determined by the fluorescently labeled secondary antibody whose signal intensity is proportional to the detected analyte concentration. Fluorescent signal of all samples was read on a Luminex 100/200™ instrument (Luminex Corp). Based on the measurements of 7 diluted standard concentrations provided by the manufacturer, a five-parameter standard curve was used to convert optical density values into concentrations (pg/mL). Data for minimum of 50 beads per cytokine were collected for each standard and sample. The final concentrations (expressed in pg/ml) were calculated using Belysa™ software (Merck Millipore, Billerica, MA). All inter-assay and intra-assay coefficients of variation (CV) were below 15%. The lower limit of quantification (LLOQ) of galectin-3 for HCCBP3MAG-58K was 4 pg/ml and for HCKP2-11K was 48.8 pg/ml.

2.15. Exploratory Endpoints Patients Evaluation

Patients' clinical and follow-up data were abstracted from medical records. Exploratory endpoints for early cohort were relapse-free survival (RFS) and overall survival (OS) according to plasma concentrations of galectin-3. Relapse-free survival (RFS) and overall survival (OS) were described as the interval before diagnostic to the endpoint (objective disease relapse and death, respectively) or last follow-up. Exploratory endpoints for advanced cohort were overall response rate (ORR; evaluated using the Response

Evaluation Criteria in Solid Tumors version 1.1 (RECIST 1.1) and defined as the proportion of patients achieving complete (CR) or partial response (PR), stable disease (SD), and progressive disease (PD); durable clinical benefit (DCB; CR, PR, or SD lasting 6 months or more after initiation of pembrolizumab treatment) and non-DCB (PD within 6 months after treatment start), progression-free survival (PFS) and overall survival (OS), according to plasma concentrations of galectin-3. Progression-free survival (PFS) and overall survival (OS) were described as the interval from the begging of pembrolizumab treatment to the endpoint (objective disease progression and death, respectively) or last follow-up.

2.16. Data acquisition and analysis of tumor infiltration immune cells by CIBERSORTx.

We acquired a LUAD data set from the TCGA consortium. Clinical and RNA-sequencing (Illumina HiSeq platform) information was directly downloaded from the ICGC Data Portal [29] (https://dcc.icgc.org/projects/LUAD-US), and only patients who fit the eligibility criteria (pathology-confirmed LUAD and stage I-IIIA) were included in further analysis.

We prepared and uploaded the mixture dataset according to the instructions of CIBERSORTx online analysis platform (https://cibersortx.stanford.edu/). To deconvolve immune cell subsets, we used the LM22 signature matrix, which is a validated leukocyte gene signature matrix that contains 547 genes distinguishing 22 human hematopoietic cell phenotypes, including seven T-cell types, naïve and memory B cells, plasma cells, natural killer (NK) cells and myeloid subsets [30]. We selected “B-mode” for batch correction and we set permutations to 500. Other parameters retained the default.

After running CIBERSORTx, we obtained the absolute proportions of subsets of tumor-infiltrating immune cells (TIICs) in each sample with *p*-values measuring the confidence of the results for the deconvolution. All samples were considered eligible for having $p < 0.05$. Dataset from CIBERSORTx of all samples is shown in Fig. S1. Heatmap of different cellular subtypes is represented on Fig. S2. Based on our previous analysis, only the proportions of T_{REGS}, T cells CD4 memory activated, T cells CD8, macrophages M1, and macrophages M2 were considered in the subsequent exploratory analyses. Exploratory analyses were performed in R (version 4.3.0) using *k*-means clustering and principal component analysis (PCA). In addition, we analyzed the RNA-seq data of counts for 356 LUAD patients obtained from TCGA. Patients were grouped into high and low groups by median of *LGALS3*.

2.17. Statistical analysis

For cell culture experiments, triplicate tests were carried out for each sample. Results are expressed as the median ± the interquartile range (IQR). Expression and secretion of paired adherent cells and tumorspheres were analyzed using non-parametric Wilcoxon's signed-rank test. The comparison of median galectin-3 levels between groups was performed using non-parametric Mann-Whitney U-test and Kruskall-Wallis to compare continuous variables. A Spearman rank test was used to test for correlations between continuous variables. The association between discrete variables were evaluated by the X2 tests. Graphs comparing metrics across groups show the median and the interquartile range (IQR), assuming non-normally distributed data. Receiving Operating Curve (ROC) method was used to determine a cut-off level of sGAL-3 for ORR and DCB. Other predictive parameters were also evaluated, including sensitivity, specificity, cut-off value, positive predictive value, negative predictive value, and area under the ROC curve (AUC) with 95% confidence interval (CI), to assess the discrimination power of sGAL-3. Survival analyses were performed using univariate Cox regression analysis and Kaplan–Meier (logrank) test method with dichotomized sGAL-3 levels and clinicopathological variables. To analyze the independent value of the galectin-3, a Cox proportional hazard model for multivariate analyses was used. All significant variables from the univariate were entered into the multivariate analyses in a forward stepwise Cox regression analysis. Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS, Chicago, IL, USA) version 23.0. Statistical significance was set at $p<0.05$ (*), $p<0.01$ (**), $p<0.001$(***).

3. RESULTS

3.1. Generation of lung tumorspheres from NSCLC patients and cell lines

In our laboratory, short-term patient derived cultures were successfully established in 40% of cases as described in Herreros-Pomares et al [22]. In this work we stablished three long-term patient derived cultures, PC301, PC435 and PC471 which were able to grow tumor cells as monolayer and tumorspheres. Clinicopathological features from PC301, PC435 and PC471 are summarized in Table 1. Long-term primary patient-derived lung cancer cell cultures were stablished for 1 months before they were split for the first passage. No significant association were found between the establishment of primary cultures and clinicopathological variables. The morphology of cells from patient-derived cultures and cell lines was examined presenting heterogeneity on the adherent-cultures cells between samples. Regarding tumorspheres, tight spheroids were formed by HCC827, H1395, H23, H1650, H358, H2228 PC435, PC471 and PC301 whereas H1993, A549, PC9, H520, SK-MES-1 and H1703 formed loose and irregularly shaped, and SW900, LUDLU-1 and H1975

showed a mixed behavior (Fig. S3). All these cell lines and primary cultures were included in further gene and protein expression analyses. Analysis will be done separating LUAD from LUSC cell cultures.

3.2. LUAD tumorspheres express high levels of LGALS3 related to immunoregulation

The expression at mRNA of *LGALS3* described as an immunoregulatory factor was analyzed in tumorspheres and adherent cells from LUAD and LUSC of 3 patient-derived cells and 15 cell lines using RTqPCR. No statistical difference between cell lines with *EGFR* and *KRAS* driver mutations and the expression of LGALS3 were found. LUAD tumorspheres showed significantly higher expression of *LGALS3* compared to adherent-cultures cells in both conditions at 12h and 24 hours post-seeded according to Wilcoxon's signed-rank test in all primary cultures and cell lines (p=0.004 and p=0.003, respectively (Fig. 1A, B). However, no significative differences in the expression of *LGALS3* between tumorspheres and adherent cells were shown in LUSC cell cultures (Fig. S4). Next, we analized the gene expression levels of Galectin-3 binding protein (LGALS3BP) and its correlation with gene expression levels of *LGALS3*. LUAD tumorspheres showed significantly higher expression of *LGALS3BP* compared to adherent-cultures cells in both conditions at 12h and 24 hours post-seeded according to Wilcoxon's signed-rank test in all primary cultures and cell lines (Fig. S5 A,B). Moreover, the expression of Galectin-3 binding protein was correlated with the expression of Galectin-3 in LUAD cell cultures in both conditions at 12h and 24 hours post-seeded both in adherent cells and tumorspheres (R=0.62, *p=0.0014* and R=0.64, *p=0.00095*, respectively) (Fig. S5 C,D).

Gene expression analyses were complemented protein expression levels analyses by means of different experiments. Galectin-3 was significantly higher in tumorspheres than in adherent cells in most of LUAD cells according to IB with only one cell line (H1395) exception (Fig. 2). Original and complete immunoblots are found in Fig. S6. Interestingly, at membrane level, LUAD tumorspheres were highly enriched in GAL-3+ cells (*p=0.021*) (Fig. 3A,B). Moreover, LUAD tumorspheres secreted significantly higher levels of soluble Galectina-3 (sGAL-3) than adherent cells at 12h and 24 hours post-seeded at low and high cell density (Fig. 3C,D). According with RTqPCR analysis, in terms of protein levels, H23 and A549 show the lowest expression levels of Gal-3 as well. We did not find significantly differences in LUSC cells (Fig. S7).

Interestingly, differential subcellular localization of GAL-3 (membranous, nuclear and cytoplasmatic) was observed without significant differences between lung tumorspheres and

adherent cells by IF (Fig. 4). No signal was detected in A549 and H23, according with low expression and low secretion levels detected previously (Fig. S8).

3.3. LUAD tumorspheres-derived extracellular vesicles express high levels of Galectin-3 in correlation with LUAD tumorspheres cell cultures.

The *LGALS3* expression was examined in a larger number of extracellular vesicles **(**EVs) samples from NSCLC cell cultures (adherent vs. tumorspheres conditions) using quantitative RT-PCR (RT-qPCR).

Employing this technique, in concordance with our previous study, it was confirmed that *LGALS3* presented significantly higher expression in LUAD secreted-EVs derived from tumorspheres than LUAD secreted-EVs derived from adherent cells (*p=0.001*) (N=11, Fig. 5 A,B), while there were no significant differences of *LGALS3* in the LUSC group (N=6). The expression of Galectin-3 in LUAD cell-derived EVs was correlated with the expression of Galectin-3 in LUAD cell cultures (R=0.54, p=0.011) and even more correlated when we analyze only the subgroup of spheres (R=0.74, p=0.013) (Fig. 5 C,D). Moreover, a strongly correlated with the secretion of sGAL-3 in LUAD cell cultures was observed (R=0.74, *p=0.00011*) (Figure 5E). No significant correlations were found for LUSC group.

3.4. Galectin-3 as an immunoregulatory factor responsible to increase regulatory T cells (T_{REGS})

To functionally test the relevance of effects on T_{REGS} induced by galectin-3, the ability of conditioned media (CM) collected from tumorspheres, and the co-culture (tumorspheres+fibroblasts) treated or not with the blocking galectin-3 monoclonal antibody were tested. So, the effects of CM from tumorspheres and co-cultures in modulating T cells having regulatory function (T_{REGS}: $CD4^+Foxp3^+CD25^+$) were assessed. Tumorspheres CM and co-culture CM were able to increase the percentage of T_{REGS} compared to control (1.9 and 1.7 fold-increased*, p=0.008* and *p=0.011* respectively). Remarkably, blockade of Gal-3 in co-culture CM was sufficient to prevent the increase of T_{REGS} population significantly (*p=0.028*) (Fig. 6).

3.5. Correlation between *LGALS3* expression in tumor with FOXP3, CD4 and CD8.

Next, we aimed to delve deeper into the relationship between galectin-3 and various T cell markers, including FOXP3 (the most specific Treg marker), in a more translational manner. To achieve this, we correlated the expression of galectin-3 in frozen tumor samples with the infiltration of FOXP3+, CD4+ and CD8+ lymphocytes as well as the expression of these markers in FPEE from tumor and tumor-near stroma compartment. First of all, the number of positive cells per HPF (high-powered field) in the stromal compartment ranged from 0 to 21 for FOXP3, from 0 to 37 for CD4 and from 9 to 55 for CD8. On the other hand, in the tumor compartment, the number ranged from 0 to 8 for FOXP3, from 0 to 12 for CD4 and from 1 to 24 for CD8. We have observed a positive correlation between those patients with high FOXP3+ infiltration in tumor and those with high expression of *LGALS3* in tumor (R=0.6, *p=0.019*) (Fig. 7A). No other correlations were found with the other T cell markers.

Then, we evaluated the correlation between expression of *LGALS3* in tumor and gene expression levels of *FOXP3*, *CD4* and *CD8* in tumor and stroma area samples that were microdissected from FFPE samples. Results of correlations with individual genes were not significant. Next, we try to combine these genes in order to find correlation with *LGALS3* expression. We decided to combine T cell markers such as *CD4* (a T helper cell marker), and *CD8* (a T cytotoxic cell marker) in combination with *FOXP3*. We calculated new variables based on the ratio of these markers. From the different combinations that were correlated with *LGALS3* expression in tumor, we found that the ratio between *FOXP3* expression assessed in the tumor compartment and the expression of *CD4* in the stroma and tumor compartment correlates positively and significantly with *LGALS3* expression in tumor (R=0.59, *p=0.012* and R=0.59, *p=0.0097*, respectively). In particular, those patients with high FOXP3 expression levels in the tumor compartment and low *CD4* levels in the tumor or in the stroma had higher levels of *LGALS3* in tumor (Fig. 7 B,C). No other significative correlations were found in the remaining combinations.

3.6. *LGALS3* expression and patient-clusters based on different immune cell infiltration

Next, to validate the relationship between *LGALS3* expression and different cellular subtypes, including T_{REGS}, which are of interest to us, we used the CIBERSORTx platform in a patient cohort from TCGA. This study was performed considering the proportion of T_{REGS}, T cells CD4 memory activated, T cells CD8, macrophages M1, and macrophages M2 in the tumors of 356 resecable LUAD patients. Based on these lymphocytes subset profiles, we identified 4 distinctive subgroups by using k-means clustering: Hot tumors, Cold tumors, M2 high tumors and T_{REGS} high tumors (Fig. 8A). A scatterplot of the four clusters conducted by PCA is displayed in Fig. 8B. We further explore the association of patient-clusters and

LGALS3 expression. As displayed in the Figure 8A, there is a trend showing that tumors with a high proportion of T_{REGS} have a higher percentage of patients with an upregulation of *LGALS3,* although not significant. Specifically, 65.45% of the patients in this cluster have upregulated galectin-3 (Fig. 8C).

3.7. Analysis of prognostic value of galectin-3 in early-stage NSCLC cohort

Data from TCGA for LUAD and LUSC patients were used to associate galectin-3 with prognosis. Characteristics of 338 patients from TCGA (in silico set) from LUAD cohort are presented in Table 2. Patients with post-surgical complications were excluded from the survival analysis, and only those patients who had more than 1 month of follow-up were included (*n*=338). In TCGA cohort, Cox regression and Kaplan–Meier analyses indicated that patients with high levels of LGALS3 presented worse RFS (23.74 months vs. 37.61 months, *p=0.021*) and OS (40.49 months vs. 103.9 months, *p=0.0004*) than those patients with low levels of *LGALS3* (Table 3 and Fig. 9). Other significant association between survival and clinicopathological variables were found (Table 3 and Fig. S9). Characteristics of 313 patients from TCGA (in silico set) from LUSC cohort are shown in Table S3. No significance results were found for LUSC cohort.

To evaluate the potential use of *LGALS3* as an independent prognostic biomarker, a multivariate Cox regression analysis was performed including all the clinicopathological variables (gender, age, tumor node metastasis (TNM) staging, smoking status and *LGALS3*). Results obtained from this multivariate analysis indicated that TNM staging and *LGALS3* were independently associated with survival (Table 4).

An independent cohort of plasma from patients with resected lung cancer from HGUV was used for validation of sGAL-3 prognosis. Clinicopathological characteristics of LUAD cohort are summarized in Table 2 (Validation set). In the same way, clinicopathological characteristics of LUSC cohort are summarized in Table S3. In LUAD patients, with a median duration of follow-up of 48 months (Interquartile range, IQC: 2,80 – 172.70 months), 21 patients were deceased at the time of cut-off due to relapse (43.8%). Those with high levels of sGAL-3 presented worse OS and in the same way, levels of sGAL-3 tended to be higher in patients with worse PFS with Cox regression and Kaplan-Meier (Fig. 9 and Table 5). Other significant association between survival and clinicopathological variables were found (see Table 5 and Fig. S10). No significance results were found for LUSC cohort.

Multivariate Cox regression analysis including all clinicopathological variables (gender, age, tumor node metastasis (TNM) staging, *KRAS* mutation status, *EGFR* mutation status, smoking status and *LGALS3*) on RFS and OS confirmed that sGAL-3 could be a prognosis independent biomarker with a hazard ratio (HR) at 2.862 (IC95% 1.057-7.753; *p=0.039*) and

3.580 (IC95% 1.185-10.81; *p=0.024*), respectively. Gender for OS and performance status for RFS were also confirmed as prognosis independent factors (Table 6).

3.8. Analysis of prognostic and predictive value of sGal-3 in NSCLC advanced-stage cohort

Following, we analyzed the possible predictive and prognostic value of sGal-3 in NSCLC advanced-stage cohort. Characteristic of the 34 LUAD patients are presented in Table 7. Patients were mostly male (79.4%), current or former smokers (94.1%) and with IV stage disease at diagnosis (82.4%). All patients were tested through Next Generation Sequencing panel Oncomine Precision Assay for genomic profiling. None of the patients harbored targetable drivers approved by European Medicines Agency (EMA). Pembrolizumab was given as first-line in 100% of cases with PDL-1≥50%, and patients had good PS (0-1) at pembrolizumab initiation in 85.5% of cases. The overall response rate (ORR) with pembrolizumab in the global population was 44.1% (*n*=15), 55.9% (*n*=19) had DCB (3CR, 10 PR and 6 SD under pembrolizumab whereas the remaining 44.11% (n=15) had non-DCB. With a median duration of follow-up of 20.01 months (Interquartile range, IQC: 6.15-31.83 months), 23 patients were deceased at the time of cut-off due to tumor progression (67.7%). The median pembrolizumab progression-free survival (PFS) was 6.30 (IQC: 2.59-18.67). At pretreatment and first response assessment (2 months of treatment), median sGAL-3 concentrations were 10150.88 pg/ml (IQR: 7985.53-13082.43) and 10126.5750 pg/ml (IQR: 8150.89-14089.95), respectively. Characteristics of the 13 LUSC patients are presented in Table S4.

3.8.1. ORR, clinical benefit and survival in advanced-stage LUAD

In LUAD patients, in terms of durable clinical benefit (DCB), at first response assessment (FR), sGAL-3 concentrations were significantly higher in patients without clinical benefit with a median value of 11972.50 pg/ml (IQR, 8040.25-23224.5975) compared to 8815.97 pg/ml (IQR, 7540.93-10126.5750) in patients with clinical benefit (*p=0.010*) (Fig. 10A). To determine sGAL-3 levels predictive of patients with DCB, we performed a ROC curve analysis, which determined a cut-off concentration of 10438.115 pg/ml associated with a sensitivity of 75%, a specificity of 84.6%, a PPV of 81.8% and NPV of 78.6% to predict durable clinical response to pembrolizumab at first response assessment with an AUC of 0.801 (*p=0.011*) (Fig. 10B). Using this cut-off, we determined that patients with high sGALS3 concentrations (n=11) had an DCB rate of 18.2%, whereas patients who had low sGAL-3 concentrations (n=14) had a DCB rate of 78.6% (*p=0.003*). However, at pretreatment (PRE), median sGAL-3 concentrations tended to be higher in patients with clinical benefit with a median value of 11208,02 pg/ml (IQR, 8014.89-14623.86) compared to 9185.27 pg/ml

(IQR,7485.67-11330.53) in patients with clinical benefit (*p=0.157*). The overall response rate (ORR) analysis elucidates no statistical difference in sGAL-3 concentrations measured at pretreatment and at first response assessment in patients who were responders compared to non-responders to pembrolizumab (Fig. S11). No significance results were found for LUSC advanced cohort (data not shown).

Patients with high sGAL-3 concentrations (≥ median) at FR were associated in cox regression analysis with worse PFS and OS in LUAD patients (HR: 3.215, 95%CI: 1.226–8.431, log-rank *p=0.018* and HR: 3.639, 95%CI: 1.317–10.056, log-rank *p*=0.013, respectively). Kaplan-Meier analysis also showed a significant association of sGAL-3 at FR with patient prognosis. Patients with high sGAL-3 levels (>median) had shorter progression-free survival (PFS) (3.20 vs. 18.6 months, *p=0.012)* and overall survival (OS) (11.53 vs. 35.1 months, *p=0.008)* (Fig. 11). In contrast, median sGAL-3 concentrations at PRE tent to be higher in patients with worse PFS but there was no statistical difference in OS (Fig. S12). Other significant associations between survival and clinicopathological variables were found in Fig. S13. No significance results were found for LUSC advanced cohort (data not shown).

Multivariate analysis including all clinicopathological variables (gender, age, tumor node metastasis (TNM) staging, smoking status and sGAL-3 on PFS and OS confirmed that sGAL-3 could be a prognosis independent biomarker with a hazard ratio (HR) at 3.215 (IC95% 1.226-8.431; *p=0.015*) and 3.639 (IC95% 1.317-10.056; *p=0.013*), respectively.

4. DISCUSSION

Despite the recent advanced in the treatment of NSCLC, the prognosis remains very poor due to the delay in the detection of the disease. In the last decade, ICBs have considerably improved the treatment of advanced NSCLC producing powerful antitumor effects however the immune therapy prediction remains poor or limited. In this context, tumor microenvironment (TME), a complex ecosystem which comprises interactions between cancer cells including CSCs, immune cells, stromal cells such as fibroblast and extracellular matrix elements, plays an important role in promoting immune evasion and suppression [31].

In the last years, preclinical studies have been focused on understanding the mechanisms involved in immune evasion and immunosuppression in tumor microenvironment (TME). Cancer cells achieve immunosuppression through several mechanisms: for instance, recruit

different cellular types such as cancer-associated fibroblast, tumor-associated macrophages or regulatory T cells (T_{REGS}); they are able to activate inhibitory pathways in immune cells, impair antigen presentation and tumor cells can also secrete immunosuppressive and pro-apoptotic cytokines and chemokines [32–34]. The evaluation of immune molecules' expression on tumor cells could provide the knowledge to comprehend better tumor immune evasion mechanisms. For this purpose, some studies have been focused on using tumorspheres, a 3D model system with outstanding applications for in vitro studies [35,36]. Recently, Bertolini et al reports that spheroid from cell lines are enriched in metastasis initiating cells with immunosuppressive potential [37]. In this work we proposed tumorspheres as a model to study the role of an immunoregulatory protein, glycoprotein galectin 3 (GAL-3) in lung cancer. What is more, we go one step further and in order to mimic more accurately the TME, we used a co-culture of tumorspheres and fibroblast, one of TME components, revealing the importance of galectin-3 as a molecule expressed and secreted in TME modulating immunosuppression through T_{REGS}. Our results confirm that lung tumorspheres express significantly more GALS3 than adherent cells, additionally more significant levels of soluble galectin-3 (sGAL-3) compering with monolayer cells. Ling-Yeng Chung et al. studied the expression of galectin-3 from NSCLC commercial cell lines (A549 and H1299) and revealed that spheroids express relatively high levels of this molecule over serial passages compared to monolayers cells acting as a cofactor by interacting with β-catenin to augment the transcriptional activities of stemness-related genes [38]. Notably, we have analyzed the expression of GAL-3 obtaining the same results not only on a large number of lung tumorspheres from cell lines moreover in primary patient-derived cell cultures from our hospital, which are a suitable and translational platform as described by some other authors [39–41]. Galectin-3 exerts different biological effects depending on its cellular localization through specific interaction with intra- and extracellular proteins affecting numerous biological processes such as neoplastic transformation and metastasis [42–44]. In concordance, our results revealed that GAL-3 in our NSCLC cells could be found in the cytoplasm, within the nucleus, on the cell surface and in the extracellular compartment depending on the cell line. Galectin-3-binding protein (LGALS3BP) is a hyperglucosylated protein that acts as a ligand for Galectin-3 that can induce the survival of cancel cells during the metastatic process [45]. Because of its relationship with Galectin-3, we decided to study its expression in cell cultures and its correlation with *LGALS3*. We have demonstrated that LUAD tumorspheres expressed higher levels of *LGALS3BP* than adherent cells and exist a positive correlation with expression of *LGALS3* in LUAD cell cultures. A previous study has reported that in the microenvironment of human neuroblastoma, GAL-3BP interacts with Galectin-3 in bone marrow mesenchymal stem cells and induces transcriptional upregulation

of IL-6, via the Gal-3BP/Gal-3/Ras/MEK/ERK signaling pathway [46,47]. In lung cancer, no previous studies have been reported about their correlation. Our results suggest that these two genes may cooperatively participate in the pathological process of cancer. Future studies should be performed in order to elucidated the mechanisms involved.

Extracellular vesicles (EVs) are a subset of small extracellular vesicles secreted by different cells. EVs are an important part of TME acting as effective signaling molecules between cancer cells and the surrounding cells [48]. We had previously performed an exhaustive characterization of NSCLC EVs revealing that EVs cargo can reflect the molecular signatures and their capacity to be used as a tool for diagnosis and prognosis [49]. In view of potential role of secreted Gal-3 as an immunomodulator molecule, we analyzed EVs-associated Gal-3 in our cohort of cell cultures. We found that *LGALS3* presented significantly higher expression in LUAD secreted EVs derived from tumorspheres than LUAD secreted-EVs derived from adherent cells. Moreover, the expression of Galectin-3 in LUAD cell-derived EVs was correlated with the expression and secretion of Galectin-3 in LUAD cell cultures. Previously, Galectin-3 has been found in EVs from bladder cancer and colon cancer but no reports were found in EVs from lung cancer [50,51]. Our results reveal that not only Galectin-3 from tumor cells but also a vesicular form of Gal3 could act as an external factor such as within EVs to help cells in the microenvironment communicate with each other. Further proteomics and plasma EVs studies should be performed to deep further into this research pathway.

Focusing on immune TME, some studies revealed that extracellular sGAL-3 secreted by tumor cells restricts TCR movement, induces T cell apoptosis and potentiate TCR downregulation [52–55]. However, the specific effect of sGAL-3 on T_{REGS} in TME has been poorly studied. We have used the CM from the co-culture between lung tumorspheres from PC435 and a fibroblast cell line to examine the effect on T_{REGS} and the role that sGAL-3 may be playing on it. CM from co-culture (PC435 and fibroblast cell line) increased the T_{REGS} population and the blocking of sGAL-3 through an antibody anti-Galectin-3 recues this phenotype. Overall, our study revealed that some components of tumor microenvironment in lung cancer such as tumor cells with stem-like properties and fibroblast could be favors an immunosuppressive microenvironment possibly recruiting T_{REGS} through sGAL-3.

Carrying on this path, we aimed to further explore the relationship between Galectin-3 and different lymphocyte populations, including T_{REGS}, and determined if there is a correlation between them to further support our prior findings. Firstly, in a cohort of early-stage LUAD patients from HGUV we found that those patients with high FOXP3+ infiltration in tumor had high expression of *LGALS3* in tumor. Moreover, we found also a positive correlation

between *FOXP3* and *LGALS3* at gene expression level. Secondly, CIBERSORTx tool with the TCGA database was used to validate the relationship between galectin-3 and different cellular subtypes, including T_{REGS}. We identified 4 clusters, where the one characterized by high levels of T_{REGS} also had the highest percentage of patients with high levels of galectin-3 expression. With these experiments we are observing that depending on the high or low expression of galectin-3, patients have more or fewer T_{REGS}. As galectin-3 regulates immune cell function to promote tumor-driven immunosuppression [56]based on our results, we can hypothesize that lung tumor cells may attract the population of T_{REGS} as a mechanism of tumor immune evasion by galectin-3.

The prognosis of NSCLC remains poor and heterogeneous and new biomarkers are needed. Our previous study described that the proportion of T helper and cytotoxic cells vs. T_{REGS} in different locations of the tumor microenvironment have opposite prognostic impacts in resected NSCLC [20]. Furthermore we have also revealed an immune-checkpoint score (PD1 and CTLA4) with relevant prognostic for a better characterization of early-stage NSCLC [57]. In accordance with our prior analyses, we would like to verify the possible prognosis role of GAL-3 on NSCLC patients, focusing on early-stage due to the tumor resection for these patients offers the best hope of cure, however, recurrences rates post-surgery remaining extremely increased [58]. Firstly, for this purpose, RNAseq data from a tumor tissue from a TCGA cohort of 331 early NSCLC patients was analyzed. Our results have confirmed that the expression of galectin-3 on LUAD patients from TCGA database is an independent prognostic biomarker for relapse-free survival and overall survival. Despite this, some limitations such as partial clinical outcome information which might lead to some uncertainties in the results. Nevertheless, TCGA database is public, provide massive information and allows carry out in silico analysis such performed previously in our laboratory [22,59].

Nowadays, studies have been focused on looking for new minimal invasive methodologies such as soluble immune mediators analysis on plasma samples. Many circulating proteins have been investigated as prognostic biomarkers in the early lung cancer management; one of the most investigated proteins have been CEA and CYFRA 21-1 [60]. However, their use in the routine clinical practice has been limited by the lack of both independent validation and reproducibility. Therefore, there is a necessity of new reliable biomarkers for early-stage NSCLC, we propose sGAL-3 as a new potential prognostic and predictive biomarker in lung cancer. Tumors cells are able to release sGAL-3 to the media confirmed in the in vitro experiments. Generally, soluble ligands and receptors can be produced by mRNA expression or by the cleavage of membrane-bound proteins. Specifically, Galectin-3 can be

cleaved by matrix metalloproteinases (MMP) and found free on plasma [61]. Blood levels of Gal-3 were found to be significantly higher in cancer patients than in controls [62]. In consequence, our results revealed that the secretion of sGAL-3 on resected LUAD patients plasma (in a validation set) is an independent prognostic biomarker for relapse-free survival and overall survival. In accordance with our results, previous studies in early NSCLC reported that galectin-3 expression on tumor cells has been reported to be associated with progression, poor prognosis and recurrence after radical resection on tissue samples [63]. Using non-invasive methodologies, Yoke et al. were value of sGAL-3 on 42 early NSCLC sera, but no prognostic role has been found [64]. Luminex® MAP technology instead of an enzyme-linked immunosorbent assay (ELISA) conventional, allow higher throughput, smaller sample volume and higher sensitivity [65]. Moreover, this technology facilitates the evaluation of simultaneous multiple mediators. As far as we know, this is the first study elucidating the prognostic value of sGAL-3 on early LUAD patients underwent surgery. One of the robustness of our study is that we employed a validation cohort from HGUV with a relatively long follow-up (median of 48 months, IQR, 2.8-172 months).

Despite the big efforts to look for new prognostic and predictive biomarker to immunotherapy in advanced NSCLC, data remain very poor and heterogeneous [66]. The expression level of PD-L1 on tumor immune cells has emerged as the first reliable predictive biomarker for sensitivity to immune checkpoint blockade in advanced NSCLC patients treated with immunotherapy [67]. However, PD-L1 expression in tissue as a predictive biomarker has limitations: range of different antibodies used in clinical trials, different positive thresholds, heterogeneity in PD-L1 staining in the tumor, insufficient tumor tissue, among others [68]. Plasmatic biomarkers have many advantages of being repeatable and easily accessible. There are some studies about new plasmatic biomarkers as putative prognostic and predictive biomarkers associated with immune checkpoint inhibitors (ICIs) efficacy in NSCLC. For instance, Okuma et al. revealed that baseline plasma sPD-L1 levels could represent a novel predictive biomarker of nivolumab therapy against NSCLC [69]. Moreover, other plasmatic biomarkers such as sGranB were associated with the response to nivolumab and also together with sPD-L1 were associated with the PFS and OS [70]. However, studies with plasmatic biomarkers about predict prognosis and tumor response to pembrolizumab remain currently sparse. In our study the prognosis and predictive value of sGAL-3 in a cohort of advanced LUAD patients treated with pembrolizumab was evaluated. Our results demonstrate that sGAL-3 levels were significantly higher in patients without clinical benefit and worse PFS and OS. These clinical results are supported by a strong biological basis in which galectin-3 have been shown to attenuate the effect of immune cells contributing to tumor cell evasion [44]. Our results are consistent with a recent study that proposed a

galectin-3 signature for the selection of candidates for immunotherapy analyzing 34 NSCLC patients [71]. In this study, those patients with high galectin-3 tumor expression before treatment showed an early and dramatic progression after three cycles of treatment, and patients with negative or low/intermediate expression of galectin-3 showed an early and durable objective responsiveness [71]. Conversely to Capalbo's study, we analyzed baseline as well as first response assessment samples, confirming the predictive and prognosis value of sGAL-3 in LUAD patients using a non-invasive methodology. Our results contribute to use a fast and high-sensitivity methodology that could be implemented for evaluating the secretion of sGAL-3 in plasma samples, predicting tumor response in patients treated with immunotherapy. In accordance with our results, Jung Sum Kim et al. also revealed that high blood Gal-3 levels at pretreatment (serum or plasma depending on the availability) may predict worse overall survival in patients with advanced NSCLC treated with immune checkpoint blockades (ICBs). In our study, in addition to employing pretreatment samples, we also evaluated first-response assessment samples demonstrating on them not only the prognostic but also the predictive impact of efficacy of pembrolizumab in LUAD patients. Moreover, contrary to these authors that used heterogeneous samples (different types of ICB, line of treatments and source) we used homogeneous samples (plasma samples from patients treated in first-line with pembrolizumab) [72].

Our study suggests that plasma sGAL-3 levels will help to select suitable patients for pembrolizumab treatment in advanced NSCLC, probably by excluding those with high plasma levels of sGAL-3. In contrast, the addition of a Gal-3 inhibitor in patients with high Gal-3 levels may be a suitable treatment to improve outcomes [73]. To date, several galectin-3 inhibitors are under clinical investigation both alone and in combination with check-point inhibitors in different cancer settings. Galectin-3 has not been reported as marker for treatment efficacy during immunotherapy in NSCLC or other cancers so far. However, a galectin-3 inhibitor (GR-MD-02), in combination with pembrolizumab or an anti-CTLA4 inhibitor, is being currently evaluated for the treatment of patients with metastatic NSCLC, melanoma and squamous cell head and neck cancer patients (NCT02575404) highlighting that galectin-3 could be part of a panel of biomarkers that predicts the outcome for immunotherapy in NSCLC [74]. Furthermore, more recently, a new clinical trial has been opened to test the safety and efficacy of other Gal-3 inhibitor (GB12211) in combination with atezolizumab in patients with advanced NSCLC (NTC05240131) remarking the relevance of including Gal-3 as predictive biomarker for ICBs.

Although our study supports that sGAL-3 could be used a prognostic and predictive biomarker for advanced LUAD patients, some limitations have should be considered. Our

study includes a small number of patients, and the results need to be confirmed in a large cohort of patients with a larger follow up. If these results will be confirmed, a better selection of responders' candidates for immunotherapy using sGAL-3 could be feasible, preventing ineffective treatments. As far as we know, this is the first report to address the independent prognostic role and predictive tumor response of sGAL-3 found on advanced LUAD patients' plasma treated with pembrolizumab in the first line with a non-invasive methodology.

5. Conclusions

In summary, we present an in vitro and translational robust study of galectin-3 in NSCLC. Our in vitro study demonstrate that NSCLC tumor cells express and secret galectin-3 acting as a regulator of immune microenvironment through T_{REG}. Focusing on the translational research studies, sGAL-3 might be applied as a novel independent biomarker to predict clinical outcomes for surgery in early LUAD patients. Furthermore, sGAL-3 is useful, not only to assess the prognosis as an independent biomarker in early stages, but also to predict the clinical outcomes for pembrolizumab in advanced LUAD patients. Prospective validation of this biomarker in a larger study should be performed to confirm these findings.

Acknowledgements

The authors E.D-S and A.M-M have a predoctoral fellowship (PRDVA18015MORE) from Scientific Foundation of Asociación Española Contra el Cáncer, Valencia (AECC Valencia). L.R is supported by the Italian Association for Cancer Research (AIRC) (IG21431). S.T-M is supported by the Generalitat Valenciana and Fondo Social Europeo fellowship (ACIF/2018/275). A.H.-P. is supported by Ayudas Margarita Salas from Ministerio de Universidades, Unión Europea-Next generation EU.

This research was founded by Centro de Investigación Biomédica en Red de Cáncer (CIBERONC) (grant number CB16-12-00350), Instituto de Salud Carlos III (PI18/00266, PI22/01221, PI22/01277 and IFEQ21/00194), Generalitat Valenciana (AICO/2021/333) and ERA-NET EURONANOMED III project METASTARG (Grant Number JTC 2018-045).

Conflict of interest

The authors declare no conflict of interest.

Authors Contributions

Conceptualization, E.J.-L., S.C.-F., C.C; methodology, S.T,-M., E.E, A.M.-M, E.D.-S., A.B; software, S.T.-M., A.M.-M., A.H-P; validation and formal analysis, S.T.-M., G.B., A.M.-M., A.H-P., R.G., A.B; investigation S.T.-M., A.M.-M, E.D.-S., G.B., L.R., S.C.-F., E.J.-L.; resources, R.G., A.B., L.R.; data curation, S.T.-M., A.M.-M, S.C.-F., G.B., A.H-P., R.G; writing-original draft preparation, S.T,-M., S.C.-F.; writing-review and editing, S.T.-M., S.C.-F., E.J.-L., A.B., R.G; visualization and supervision, E.J.-L., S.C.-F., C.C.; project administration and funding acquisition, E.J.-L., S.C.-F., C.C. All authors have read and agree to the published version of the manuscript.

Data Accessibility

Data that supports these findings are available from the corresponding authors upon reasonable request.

References

1 Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin*. 2022;**72**:7–33. doi: 10.3322/caac.21708.

2 Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature*. 2018;**553**:446–454. doi: 10.1038/nature25183.

3 Chan BA, Hughes BGM. Targeted therapy for non-small cell lung cancer: Current standards and the promise of the future. *Transl Lung Cancer Res*. 2015;**4**:36–54. doi: 10.3978/j.issn.2218-6751.2014.05.01.

4 Allehebi A, Kattan K Al, Rujaib M Al, Dayel F Al, Black E, Mahrous M, et al. Management of Early-Stage Resected Non-Small Cell Lung Cancer: Consensus Statement of the Lung cancer Consortium. *Cancer Treat Res Commun*. 2022;**31**:100538. doi: 10.1016/j.ctarc.2022.100538.

5 Schneider BJ, Naidoo J, Santomasso BD, Lacchetti C, Adkins S, Anadkat M, et al. Management of Immune-Related Adverse Events in Patients Treated with Immune Checkpoint Inhibitor Therapy: ASCO Guideline Update. *Journal of Clinical Oncology*. 2021;**39**:4073–4126. doi: 10.1200/JCO.21.01440.

6 Brahmer J, Reckamp KL, Baas P, Crinò L, Eberhardt WEE, Poddubskaya E, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non–Small-Cell Lung Cancer. *New England Journal of Medicine*. 2015;**373**:123–135. doi: 10.1056/NEJMOA1504627/SUPPL_FILE/NEJMOA1504627_DISCLOSURES.PDF.

7 Gandhi L, Rodríguez-Abreu D, Gadgeel S, Esteban E, Felip E, De Angelis F, et al. Pembrolizumab plus Chemotherapy in Metastatic Non–Small-Cell Lung Cancer. *New England Journal of Medicine*. 2018;**378**:2078–2092. doi: 10.1056/nejmoa1801005.

8 Herbst RS, Baas P, Kim DW, Felip E, Pérez-Gracia JL, Han JY, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet*. 2016;**387**:1540–1550. doi: 10.1016/S0140-6736(15)01281-7.

9 Socinski MA, Nishio M, Jotte RM, Cappuzzo F, Orlandi F, Stroyakovskiy D, et al. IMpower150 Final Overall Survival Analyses for Atezolizumab Plus Bevacizumab and Chemotherapy in First-Line Metastatic Nonsquamous NSCLC. *Journal of Thoracic Oncology*. 2021;**16**:1909–1924. doi: 10.1016/j.jtho.2021.07.009.

10 Aguiar PN, Santoro IL, Tadokoro H, De Lima Lopes G, Filardi BA, Oliveira P, et al. The role of PD-L1 expression as a predictive biomarker in advanced non-small-cell lung cancer: a network meta-analysis. *Immunotherapy*. 2016;**8**:479–488. doi: 10.2217/IMT-2015-0002.

11 Gettinger SN, Horn L, Gandhi L, Spigel DR, Antonia SJ, Rizvi NA, et al. Overall survival and long-term safety of nivolumab (anti-programmed death 1 antibody, BMS-936558, ONO-4538) in patients with previously treated advanced non-small-cell lung cancer. *Journal of Clinical Oncology*. 2015;**33**:2004–2012. doi: 10.1200/JCO.2014.58.3708.

12 Giroux Leprieur E, Herbretau G, Dumenil C, Julie C, Giraud V, Labrune S, et al. Circulating tumor DNA evaluated by Next-Generation Sequencing is predictive of tumor response and prolonged clinical benefit with nivolumab in advanced non-small cell lung cancer. *Oncoimmunology*. 2018;**7**. doi: 10.1080/2162402X.2018.1424675.

13 Okuma Y, Hosomi Y, Nakahara Y, Watanabe K, Sagawa Y, Homma S. High plasma levels of soluble programmed cell death ligand 1 are prognostic for reduced survival in advanced lung cancer. *Lung Cancer*. 2017;**104**:1–6. doi: 10.1016/j.lungcan.2016.11.023.

14 Cardoso ACF, Andrade LN de S, Bustos SO, Chammas R. Galectin-3 determines tumor cell adaptive strategies in stressed tumor microenvironments. *Front Oncol*. 2016;**6**:127. doi: 10.3389/FONC.2016.00127/BIBTEX.

15 Raimond J, Zimonjic DB, Mignon C, Mattei MG, Popescu NC, Monsigny M, et al. Mapping of the galectin-3 gene (LGALS3) to human chromosome 14 at region 14q21-22. *Mammalian Genome*. 1997;**8**:706–707. doi: 10.1007/S003359900548/METRICS.

16 Farhad M, Rolig AS, Redmond WL. The role of Galectin-3 in modulating tumor growth and immunosuppression within the tumor microenvironment. *Oncoimmunology*. 2018;**7**. doi: 10.1080/2162402X.2018.1434467.

17 Fortuna-Costa A, Gomes AM, Kozlowski EO, Stelling MP, Pavão MSG. Extracellular galectin-3 in tumor progression and metastasis. *Front Oncol*. 2014;**4**:138. doi: 10.3389/FONC.2014.00138/BIBTEX.

18 Carter EP, Roozitalab R, Gibson S V., Grose RP. Tumour microenvironment 3D-modelling: simplicity to complexity and back again. *Trends Cancer*. 2021;**7**:1033–1046. doi: 10.1016/j.trecan.2021.06.009.

19 Ramgolam K, Lauriol J, Lalou C, Lauden L, Michel L, de la Grange P, et al. Melanoma Spheroids Grown Under Neural Crest Cell Conditions Are Highly Plastic Migratory/Invasive Tumor Cells Endowed with Immunomodulator Function. *PLoS One*. 2011;**6**:e18784. doi: 10.1371/JOURNAL.PONE.0018784.

20 Usó M, Jantus-Lewintre E, Bremnes RM, Calabuig S, Blasco A, Pastor E, et al. Analysis of the immune microenvironment in resected non-small cell lung cancer: the prognostic value of different T lymphocyte markers. *Oncotarget*. 2016;**7**:52849–52861. doi: 10.18632/oncotarget.10811.

21 Marcus L, Lemery SJ, Keegan P, Pazdur R. FDA approval summary: Pembrolizumab for the treatment of microsatellite instability-high solid tumors. *Clinical Cancer Research*. 2019;**25**:3753–3758. doi: 10.1158/1078-0432.CCR-18-4070.

22 Herreros-Pomares A, de-Maya-Girones JD, Calabuig-Fariñas S, Lucas R, Martínez A, Pardo-Sánchez JM, et al. Lung tumorspheres reveal cancer stem cell-like properties and a score with prognostic impact in resected non-small-cell lung cancer. *Cell Death Dis*. 2019;**10**. doi: 10.1038/S41419-019-1898-1.

23 Andriani F, Majorini MT, Mano M, Landoni E, Miceli R, Facchinetti F, et al. MiR-16 regulates the pro-tumorigenic potential of lung fibroblasts through the inhibition of HGF production in an FGFR-1- and MEK1-dependent manner. *J Hematol Oncol*. 2018;**11**:1–17. doi: 10.1186/s13045-018-0594-4.

24 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002;**3**:1–12. doi: 10.1186/GB-2002-3-7-RESEARCH0034/COMMENTS.

25 Pfaffl MW. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res*. 2001;**29**:e45–e45. doi: 10.1093/NAR/29.9.E45.

26 Collisson EA, Campbell JD, Brooks AN, Berger AH, Lee W, Chmielecki J, et al. Comprehensive molecular profiling of lung adenocarcinoma: The cancer genome atlas research network. *Nature*. 2014;**511**:543–550. doi: 10.1038/nature13385.

27 Hammerman PS, Voet D, Lawrence MS, Voet D, Jing R, Cibulskis K, et al. Comprehensive genomic characterization of squamous cell lung cancers. *Nature*. 2012;**489**:519–525. doi: 10.1038/nature11404.

28 Zhang J, Baran J, Cros A, Guberman JM, Haider S, Hsu J, et al. International Cancer Genome Consortium Data Portal—a one-stop shop for cancer genomics data. *Database (Oxford)*. 2011;**2011**. doi: 10.1093/DATABASE/BAR026.

29 Zhang J, Baran J, Cros A, Guberman JM, Haider S, Hsu J, et al. International cancer genome consortium data portal-a one-stop shop for cancer genomics data. *Database*. 2011;**2011**. doi: 10.1093/database/bar026.

30 Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;**12**:453–457. doi: 10.1038/nmeth.3337.

31 Altorki NK, Markowitz GJ, Gao D, Port JL, Saxena A, Stiles B, et al. The lung microenvironment: an important regulator of tumour growth and metastasis. *Nat Rev Cancer*. 2019;**19**:9–31. doi: 10.1038/s41568-018-0081-9.

32 Tuccitto A, Shahaj E, Vergani E, Ferro S, Huber V, Rodolfo M, et al. Immunosuppressive circuits in tumor microenvironment and their influence on cancer treatment efficacy. *Virchows Archiv*. 2018;**474**:407–420. doi: 10.1007/S00428-018-2477-Z.

33 Zou W. Regulatory T cells, tumour immunity and immunotherapy. *Nat Rev Immunol*. 2006;**6**:295–307. doi: 10.1038/nri1806.

34 Kitamura T, Qian BZ, Pollard JW. Immune cell promotion of metastasis. *Nat Rev Immunol*. 2015;**15**:73. doi: 10.1038/NRI3789.

35 García-Rocha R, Monroy-García A, Carrera-Martínez M, Hernández-Montes J, Don-López CA, Weiss-Steider B, et al. Evidence that cervical cancer cells cultured as tumorspheres maintain high CD73 expression and increase their protumor characteristics through TGF-β production. *Cell Biochem Funct*. 2022;**40**:760–772. doi: 10.1002/CBF.3742.

36 Darvin P, Sasidharan Nair V, Elkord E. PD-L1 Expression in Human Breast Cancer Stem Cells Is Epigenetically Regulated through Posttranslational Histone Modifications. *J Oncol*. 2019;**2019**. doi: 10.1155/2019/3958908.

37 Fortunato O, Belisario DC, Compagno M, Giovinazzo F, Bracci C, Pastorino U, et al. CXCR4 Inhibition Counteracts Immunosuppressive Properties of Metastatic NSCLC Stem Cells. *Front Immunol*. 2020;**11**:1–13. doi: 10.3389/fimmu.2020.02168.

38 Chung LY, Tang SJ, Wu YC, Sun GH, Liu HY, Sun KH. Galectin-3 augments tumor initiating property and tumorigenicity of lung cancer through interaction with β-catenin. *Oncotarget*. 2015;**6**:4936. doi: 10.18632/oncotarget.3210.

39 Kodack DP, Farago AF, Dastur A, Held MA, Dardaei L, Friboulet L, et al. Primary Patient-Derived Cancer Cells and Their Potential for Personalized Cancer Patient Care. *Cell Rep*. 2017;**21**:3298. doi: 10.1016/J.CELREP.2017.11.051.

40 Kim SY, Lee JY, Kim DH, Joo HS, Yun MR, Jung D, et al. Patient-Derived Cells to Guide Targeted Therapy for Advanced Lung Adenocarcinoma. *Sci Rep*. 2019;**9**:1–12. doi: 10.1038/s41598-019-56356-4.

41 Zhang Z, Wang H, Ding Q, Xing Y, Xu Z, Lu C, et al. Establishment of patient-derived tumor spheroids for non-small cell lung cancer. *PLoS One*. 2018;**13**:e0194016. doi: 10.1371/JOURNAL.PONE.0194016.

42 Dumic J, Dabelic S, Flögel M. Galectin-3: An open-ended story. *Biochimica et Biophysica Acta (BBA) - General Subjects*. 2006;**1760**:616–635. doi: 10.1016/J.BBAGEN.2005.12.020.

43 Vladoiu MC, Labrie M, St-pierre Y. Intracellular galectins in cancer cells : Potential new targets for therapy (Review). *Int J Oncol*. 2014;**44**:1001–1014. doi: 10.3892/ijo.2014.2267.

44 Ruvolo PP. Galectin 3 as a guardian of the tumor microenvironment. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*. 2016;**1863**:427–437. doi: 10.1016/J.BBAMCR.2015.08.008.

45 Capone E, Iacobelli S, Sala G. Role of galectin 3 binding protein in cancer progression: a potential novel therapeutic target. Journal of Translational Medicine. 2021;**19**. doi: 10.1186/s12967-021-03085-w.

46 He X, Zhang S, Chen J, Li D. Increased LGALS3 expression independently predicts shorter overall survival in patients with the proneural subtype of glioblastoma. *Cancer Med*. 2019;**8**:2031–2040. doi: 10.1002/cam4.2075.

47 Fukaya Y, Shimada H, Wang LC, Zandi E, DeClerck YA. Identification of galectin-3-binding protein as a factor secreted by tumor cells that stimulates interleukin-6 expression in the bone marrow stroma. *Journal of Biological Chemistry*. 2008;**283**:18573–18581. doi: 10.1074/jbc.M803115200.

48 Dang C Van, Zhao H, Yang L, Baddour J, Achreja A, Bernard V, et al. Tumor microenvironment derived exosomes pleiotropically modulate cancer cell metabolism. *Elife*. 2016;**5**. doi: 10.7554/eLife.10250.001.

49 Duréndez-Sáez E, Calabuig-Fariñas S, Torres-Martínez S, Moreno-Manuel A, Herreros-Pomares A, Escorihuela E, et al. Analysis of Exosomal Cargo Provides Accurate Clinical, Histologic and Mutational Information in Non-Small Cell Lung Cancer. *Cancers (Basel)*. 2022;**14**. doi: 10.3390/cancers14133216.

50 Liang B, Peng P, Chen S, Li L, Zhang M, Cao D, et al. Characterization and proteomic analysis of ovarian cancer-derived exosomes. *J Proteomics*. 2013;**80**:171–182. doi: 10.1016/j.jprot.2012.12.029.

51 Welton JL, Khanna S, Giles PJ, Brennan P, Brewis IA, Staffurth J, et al. Proteomics analysis of bladder cancer exosomes. *Molecular and Cellular Proteomics*. 2010;**9**:1324–1338. doi: 10.1074/mcp.M000063-MCP201.

52 Guha P, Kaptan E, Bandyopadhyaya G, Kaczanowska S, Davila E, Thompson K, et al. Cod glycopeptide with picomolar affinity to galectin-3 suppresses T-cell apoptosis and prostate cancer metastasis. *Proc Natl Acad Sci U S A*. 2013;**110**:5052–5057. doi: 10.1073/PNAS.1202653110/SUPPL_FILE/PNAS.201202653SI.PDF.

53 Chen HY, Fermin A, Vardhana S, Weng IC, Lo KFR, Chang EY, et al. Galectin-3 negatively regulates TCR-mediated CD4+ T-cell activation at the immunological synapse. *Proc Natl Acad Sci U S A*. 2009;**106**:14496–14501. doi: 10.1073/pnas.0903497106.

54 Kouo T, Huang L, Pucsek AB, Cao M, Solt S, Armstrong T, et al. Galectin-3 shapes antitumor immune responses by suppressing CD8+ T cells via LAG-3 and inhibiting expansion of plasmacytoid dendritic cells. *Cancer Immunol Res*. 2015;**3**:412. doi: 10.1158/2326-6066.CIR-14-0150.

55 Fukumori T, Oka N, Takenaka Y, Nangia-Makker P, Elsamman E, Kasai T, et al. Galectin-3 regulates mitochondrial stability and antiapoptotic function in response to anticancer drug in prostate cancer. *Cancer Res*. 2006;**66**:3114–3119. doi: 10.1158/0008-5472.CAN-05-3750.

56 GUO Y, SHEN R, YU L, ZHENG X, CUI R, SONG Y, et al. Roles of galectin-3 in the tumor microenvironment and tumor metabolism (Review). *Oncol Rep*. 2020;**44**:1799–1809. doi: 10.3892/OR.2020.7777/HTML.

57 Usó M, Jantus-Lewintre E, Calabuig-Fariñas S, Blasco A, Del Olmo E, Guijarro R, et al. Analysis of the prognostic role of an immune checkpoint score in resected non-small cell lung cancer patients. *Oncoimmunology*. 2017;**6**:1–9. doi: 10.1080/2162402X.2016.1260214.

58 Bowes K, Jovanoski N, Brown AE, Di Maio D, Belleli R, Chadda S, et al. Treatment patterns and survival of patients with locoregional recurrence in early-stage NSCLC: a literature review of real-world evidence. *Med Oncol*. 2023;**40**:4. doi: 10.1007/S12032-022-01790-0.

59 Duréndez-Sáez E, Calabuig-Fariñas S, Torres-Martínez S, Moreno-Manuel A, Herreros-Pomares A, Escorihuela E, et al. Analysis of Exosomal Cargo Provides Accurate Clinical, Histologic and Mutational Information in Non-Small Cell Lung Cancer. *Cancers (Basel)*. 2022;**14**:1–23. doi: 10.3390/cancers14133216.

60 Crosbie PAJ, Shah R, Summers Y, Dive C, Blackhall F. Prognostic and predictive biomarkers in early stage NSCLC : CTCs and serum / plasma markers. *Traslational Lung Cancer Research*. 2013;**2**:382–397. doi: 10.3978/j.issn.2218-6751.2013.09.02.

61 Ochieng J, Leite-Browning ML, Warfield P. Regulation of cellular adhesion to extracellular matrix proteins by galectin-3. *Biochem Biophys Res Commun*. 1998;**246**:788–791. doi: 10.1006/BBRC.1998.8708.

62 Cancer L, Blair BB, Funkhouser AT, Goodwin JL, Strigenz AM, Chaballout BH, et al. Increased Circulating Levels of Galectin Proteins in Patients. 2021;1–13.

63 Kusuhara S, Igawa S, Ichinoe M, Nagashio R, Kuchitsu Y, Hiyoshi Y, et al. Prognostic significance of galectin-3 expression in patients with resected NSCLC treated with platinum-based adjuvant chemotherapy. *Thorac Cancer*. 2021;**12**:1570–1578. doi: 10.1111/1759-7714.13945.

64 Kataoka Y, Igarashi T, Ohshio Y, Fujita T, Hanaoka J. Predictive importance of galectin-3 for recurrence of non-small cell lung cancer. *Gen Thorac Cardiovasc Surg*. 2019;**67**:704–711. doi: 10.1007/s11748-019-01074-x.

65 DuPont NC, Wang K, Wadhwa PD, Culhane JF, Nelson EL. Validation and comparison of luminex multiplex cytokine analysis kits with ELISA: Determinations of a panel of nine cytokines in clinical sample culture supernatants. *J Reprod Immunol*. 2005;**66**:175. doi: 10.1016/J.JRI.2005.03.005.

66 Augustus E, Zwaenepoel K, Siozopoulou V, Raskin J, Jordaens S, Baggerman G, et al. Prognostic and predictive biomarkers in non-small cell lung cancer patients on immunotherapy—the role of liquid biopsy in unraveling the puzzle. *Cancers (Basel)*. 2021;**13**. doi: 10.3390/cancers13071675.

67 Patel SP, Kurzrock R. PD-L1 expression as a predictive biomarker in cancer immunotherapy. *Mol Cancer Ther*. 2015;**14**:847–856. doi: 10.1158/1535-7163.MCT-14-0983/86482/AM/PD-L1-EXPRESSION-AS-A-PREDICTIVE-BIOMARKER-IN.

68 Davis AA, Patel VG. The role of PD-L1 expression as a predictive biomarker: an analysis of all US Food and Drug Administration (FDA) approvals of immune checkpoint inhibitors. *J Immunother Cancer*. 2019;**7**. doi: 10.1186/S40425-019-0768-9.

69 Okuma Y, Wakui H, Utsumi H, Sagawa Y, Hosomi Y, Kuwano K, et al. Soluble Programmed Cell Death Ligand 1 as a Novel Biomarker for Nivolumab Therapy for Non–Small-cell Lung Cancer. *Clin Lung Cancer*. 2018;**19**:410-417.e1. doi: 10.1016/j.cllc.2018.04.014.

70 Costantini A, Julie C, Dumenil C, Hélias-Rodzewicz Z, Tisserand J, Dumoulin J, et al. Predictive role of plasmatic biomarkers in advanced non-small cell lung cancer treated by nivolumab. *Oncoimmunology*. 2018;**7**. doi: 10.1080/2162402X.2018.1452581.

71 Capalbo C, Scafetta G, Filetti M, Marchetti P, Bartolazzi A. Predictive Biomarkers for Checkpoint Inhibitor-Based Immunotherapy: The Galectin-3 Signature in NSCLCs. *Int J Mol Sci*. 2019;**20**. doi: 10.3390/IJMS20071607.

72 Sun J, Soyeon K, Koh J, Kim M, Keam B, Kim TM, et al. Predictive role of galectin - 3 for immune checkpoint blockades (ICBs) in advanced or metastatic non - small cell lung cancer : a potential new marker for ICB resistance. *J Cancer Res Clin Oncol*. 2022;**1**. doi: 10.1007/s00432-022-04275-9.

73 Vuong L, Kouverianou E, Rooney CM, McHugh BJ, Howie SEM, Gregory CD, et al. An orally active galectin-3 antagonist inhibits lung adenocarcinoma growth and augments response to PD-L1 blockade. *Cancer Res*. 2019;**79**:1480–1492. doi: 10.1158/0008-5472.CAN-18-2244.

74 Curti BD, Koguchi Y, Leidner RS, Rolig AS, Sturgill ER, Sun Z, et al. Enhancing clinical and immunological effects of anti-PD-1 with belapectin, a galectin-3 inhibitor. *J Immunother Cancer*. 2021;**9**. doi: 10.1136/jitc-2021-002371.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Dataset from CIBERSORTx.

Fig. S2. Heatmap of different cellular subtypes representing absolute cell fraction of different cellular subtypes.

Fig. S3. Representative images of the primary patient-derived cancer cells and cell lines grown under adherent conditions and suspension conditions.

Fig. S4. Transcription levels of *LGALS3* in tumorspheres vs. adherent-culture in different LUSC primary cultures and cell lines.

Fig. S5. Transcription levels of *LGALS3BP* in tumorspheres vs. adherent-culture in different LUAD primary cultures and cell lines.

Fig. S6. Original and complete immunoblots for B-actin and Galectin-3.

Fig. S7. Expression of *LGALS3* as protein level in LUSC cells.

Fig. S8. Representative immunofluorescence images of Gal-3 in tumorspheres and adherent-cultured cells from ADC patients.

Fig. S9. Kaplan-Meier survival curves according to clinicopathological variables from TCGA in silico set.

Fig. S10. Kaplan-Meier survival curves according to clinicopathological variables in validation set.

Fig. S11. Analysis of predictive value in terms of Overall Response Rate (ORR) of sGal-3 in LUAD advanced cohort.

Fig. S12. Kaplan-Meier survival curves according to sGAL-3 concentrations at pretreatment (PRE).

Fig. S13. Kaplan-Meier survival curves according to clinicopathological variables in LUAD advanced cohort.

Table S1. TaqMan® Gene Expression Assay used in gene expression analyses.

Table S2. List of antibodies used for immunoblot (IB), immunofluorescence (IF) and flow cytometry (FC) analysis.

Table S3. Clinicopathological characteristics of the LUSC early patients included in the study.

Table S4. Patient's characteristics of advanced-stage LUSC cohort.

Figure legend

Fig. 1. Transcription levels of *LGALS3* in tumorspheres vs. adherent LUAD primary cultures and cell lines analyzed by RTqPCR at 12 and 24 hours after cell seeding. (A) The results shown the relative fold-change gene expression of LGALS3 to reference genes *ACTB*, *CDKN1B*, and *GUSB*. Errors bars represent standard deviation (SD) of three different experiments. (B) The results shown are the median of relative fold-change gene expression of *LGALS3* to reference genes *ACTB*, *CDKN1B*, and *GUSB*. Statistical analysis was carried out with the Wilcoxon test. Errors bars represent interquartile range (IQR) of all samples (n=12). Significance values were ** $p \leq 0.01$. ADH, adherent; SPH, tumorspheres; n, sample size.

Fig. 2. Expression of GAL-3 as protein level. (A) Immunoblots (IBs) showing the level of galectin-3 in adherent cells and tumorspheres. Beta-actin (ACTB) was used as loading control. The experiment was repeated three times and representative western blot results from one experiment were shown. (B) ImageJ analysis of IBs of panel a. Bar chart represents the relative expression of each protein according to immunoblots. Three grey values relative to the loading controls were measured in every case and averaged. (C) Values relative to the loading controls were measured in every cell line and averaged. Statistical analysis was carried out with the Wilcoxon test. Errors bars represent interquartile range (IQR) of all cell lines and primary cultures median (n=12). Significance values were **p<0.01. ADH, adherent; SPH, tumorspheres; *n*, sample size; IB, immunoblot, IQR, interquartile range.

Fig. 3. Flow cytometry and Immunoassay analysis of Galectin-3 in LUAD cells. (A,B) Flow cytometry analysis of surface GAL3 in LUAD adherent cells and tumorspheres. (A) The results shown are individual results for each cell line and primary culture. Errors bars represent SD of three different experiments. (B) The results shown are the median of all cells lines and primary cultures. Statistical analysis was carried out with the Wilcoxon test. Errors bars represent IQR of the median. (C,D) Immunoassay of sGAL-3 in LUAD adherent cells and tumorspheres analyzed by Luminex® Technology at 12 and 24 hours after cell seeding. (C) Median levels of sGAL-3 of all cell lines and primary cultures at 12 and 24 hours after 10.000 cells/ml seeding (low density). (D) Median levels of

sGAL-3 of all cell lines and primary cultures at 12 and 24 hours after 100.000 cells/ml seeding (high density). Statistical analysis was carried out with the Wilcoxon test. Errors bars represent interquartile range (IQR) of the median of all cell lines and primary cultures (n=12). Significance values were * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. ADH, adherent; SPH, tumorspheres; *n*, sample size; IQR, interquartile range; LUAD, lung adenocarcinoma.

Fig. 4. Representative immunofluorescence images of galectin-3 in adherent-cultured cells and tumorspheres from (A) PC435 and (B) H1650. Immunofluorescence green channel shows the indicated GAL-3 staining, blue channel shows DAPI staining, and merge shows all channels merged. The experiment was repeated three times and representative immunofluorescence image from one experiment were shown. Scale bar represents 50 µm. ADH, adherent; SPH, tumorspheres.

Fig. 5. *LGALS3* expression in LUAD tumor-derived extracellular vesicles (EVs) from tumorspheres and adherent cells and correlation with expression of *LGALS3* and secretion of Galectin-3 in culture cells. (A) The results shown the relative fold-change gene expression of *LGALS3* in LUAD tumor-derived EVs to reference genes *ACTB*, *CDKN1B*, and *GUSB*. Experiments were performed in duplicate. (B) The results shown are the median of relative fold-change gene expression of LGALS3 in LUAD tumor derived-EVs to reference genes *ACTB*, *CDKN1B*, and *GUSB*. Statistical analysis was carried out with the Wilcoxon test. Errors bars represent interquartile range (IQR) of all samples (n=11). (C) Correlation between *LGALS3* expression levels in LUAD tumor derived-EVs and *LGALS3* expression levels in LUAD tumor cell cultures (n=11). (D) Correlation between *LGALS3* expression levels in LUAD tumor derived-EVs from spheroids and *LGALS3* expression levels in LUAD tumorspheres cell cultures (n=22). E) Correlation between *LGALS3* expression levels in LUAD tumor derived-EVs and sGAL3 levels secreted by LUAD tumor cell cultures (n=11). Statistical analysis was carried out with the Spearman Correlation Coefficient. R represents the Spearman correlation coefficient. Significance values were ** $p \leq 0.01$. ADH, adherent; SPH, tumorspheres; n, sample size.

Fig. 6. Conditioned media (CM) from spheroids induces T_{REGS} that can be prevented by Gal-3 blockade. Flow cytometry analysis for T_{REG} population within T lymphocytes (T_{REG}: CD4+Foxp3+CD25+), from n=9 healthy volunteers. T lymphocytes were incubated for 72 h with CM from spheroids or co-culture, untreated or treated with anti- Gal-3 antibody. Data are the median value in % T_{REG} population $FOXP3^+/CD25^+$ relative to CD4+. We used as a control (CNT) (50% of fibroblast medium (FBM) and 50% tumorspheres DMEM F12). Statistical analysis was carried out with the Wilcoxon test. Bars represent minimum and maximum points. Significance values were * $p \leq 0.05$, ** $p \leq 0.01$. *n*, sample size; *CNT*, control medium; *CM*, conditioned medium; *SPH*, tumorspheres.

Fig. 7. Correlations between T cell markers in tumor or stroma compartments from FPEE samples and *LGALS3* expression levels in frozen tumor tissue. (A) Correlation between *LGALS3* expression levels in tumor and FOXP3+ infiltration in tumor (n=15) (B) Correlation between *LGALS3* expression levels in tumor and *FOXP3* tumor/CD4 tumor ratio (n=19) (C) Correlation between LGALS3

expression levels in tumor and FOXP3 tumor/CD4 stroma ratio (n=19). Statistical analysis was carried out with the Spearman Correlation Coefficient. R represents the Spearman correlation coefficient. P-value was statistical significative $p<0.05$.

Fig. 8. Results of immune cell infiltration clustering and expression of *LGALS3*. (A) K-means heatmap. Four distinctive clusters of patients (n=356) were identified by using hierarchical clustering algorithm with ComplexHeatmap package based on different immune cell infiltration. Clusters are distinguished by hot tumors (Hot), cold tumors (Cold), M2-enriched tumors (M2 high), and regulatory T cell-enriched tumors (Treg high). More red color designates higher expression for a given sample while blue designates lower expression. *LGALS3* expression is shown on top. Red color represents overexpression and green represents underexpression. (B) The scatterplot performed by PCA to show the four distinct clusters. C) Bar charts representing the percentage of patients with upregulated *LGALS3* and downregulated *LGALS3* in the 4 clusters.

Fig. 9. Kaplan-Meier survival curves according to *LGALS3* from TCGA in silico set. (A,B) and sGAL-3 concentrations before surgery in validation set (C,D). (A) Relapse-free survival (RFS) stratified in high (n=169) vs. low LGALS3 concentrations (n=169). B) Overall survival (OS) stratified in high (n=169) vs. low (n=169) LGALS3 concentrations. The groups were divided as low and high according to its median. Green lines represent patients with high levels of expression, whereas blue lines represent patients with low levels of expression. C) Relapse-free survival (RFS) stratified in high (n=24) vs. low sGAL-3 levels (n=24). D) Overall survival (OS) stratified in high (n=24) vs. low (n=24) sGAL-3 levels. The groups were divided as low and high according to its median. Orange lines represent patients with high levels of sGAL-3 (>9125.73 pg/ml), whereas purple lines represent patients with low levels of sGAL-3 (≤ 9125.73 pg/ml). P-values were obtained using the log-rank test. *RFS*, relapse-free survival; *OS*, overall survival.

Fig. 10. Analysis of predictive value in terms of durable clinical benefit (DBC) of sGal-3 in LUAD advanced cohort. (A) sGAL-3 concentrations at first response evaluation in patients with durable clinical benefit (DCB) response (*n*=13) and patients without DCB response (n=12). Data are the median values and bars represent minimum and maximum values. P-values were obtained using the Mann-Whitney test. (B) Receiving Operating Characteristics (ROC) curve of sGAL-3 for discriminating between patients with DCB and patients without DCB representing the area under the ROC curve (AUC). Statistical analysis was carried out with the Receiver operating characteristic analysis. Significance values were ** $p \leq 0.01$. ° Outliers; *RFS*, relapse-free survival; *OS*, overall survival; *AUC*, area under the curve; *ROC*, receiver operating characteristic.

Fig. 11. Kaplan-Meier survival curves according to sGAL-3 concentrations at first response assessment (FR) in LUAD advanced cohort A) Progression-free survival (PFS) stratified in high (n= 13) vs. low sGAL-3 levels (n=13). B) Overall survival (OS) stratified in high (n= 13) vs. low sGAL-3 levels (n=13). The groups were divided as low and high according to median value. Red lines represent patients with high levels of expression, whereas blue lines represent patients with low levels

of expression. P-values were obtained using the log-rank test. *FR;* first response assessment; *PFS*, progression-free survival; *OS*, overall survival.

TABLES

Table 1. Clinicopathological characteristics of the patients included in the study. DFS, disease-free survival; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.

Patient Code	Gender	Age	TNM Stage	Histology	Smoking Status	Progression /Exitus	DFS (months)	Mutational Status
435	male	73	IIB	LUAD	Former	NO	24	KRAS p.G12C, PIK3CA p.H1047R
471	female	85	IIA	LUAD	Never	NO	27	PIK3CA p.D538N
301	Male	71	IIB	LUSC	Former	NO	75.50	PIK3CA p.G118D TP53 p.S261V*fs84,

Table 2. Clinicopathological characteristics of the LUAD patients included in the study. NS, non-specified; n, sample size: IQR; interquartile range.

	In silico cohort		**Plasma Validation Set**	
	***n*= 338**	**%**	**n=48**	**%**
Age at surgery (median, range)	67 [IQR 38–88]		65.5 [IQR 42–84]	
Gender				
Male	161	47.6	28	58.3
Female	177	52.4	20	41.7
Stage				
I	195	57.7	23	47.9
II	86	25.4	15	31.3
IIIA	57	16.9	10	20.8
Performance Status				
0	NS	NS	39	81.3
1			9	18.8
Smoking status				
Current	81	24	21	43.8
Former	175	51.8	16	33.3
Never	82	24.3	11	22.9
EGFR				
Mutated	NS	NS	8	16.3
WildType			39	79.6
NS			2	4.1
KRAS				
Mutated	NS	NS	11	22.4
WildType			28	57.1
NS			10	20.4
Relapse				
No	196	58.0	26	54.2
Yes	121	35.8	22	45.8

NS	21	6.2		
Exitus				
No	226	66.9	27	56.3
Yes	112	33.1	21	43.8

Table 3. Results from the univariate Cox regression model for OS and RFS on LUAD *in silico* set. RFS, relapse-free survival; OS, overall survival; HR, hazard ratio; CI, confidence interval; LN, lymph node; TNM, tumor node metastasis

	***in silico* set (n=338)**					
		RFS			**OS**	
	HR	**95%CI**	**p-value**	**HR**	**95%CI**	***p-value***
LGALS3 (High vs Low)	1.551	1.136-2.117	0.003*	1.968	1.341-2.888	0.0001*
Gender Male vs Female	0.879	0.644-1.191	0.397	0.901	0.621-1.306	0.582
Age >65 vs. ≤65	1.291	0.933-1.786	0.123	1.308	0.881-1.941	0.183
TNM staging III vs. II vs. I	1.465	1.213-1.771	<0.0001*	1.560	1.243-1.958	<0.0001*
Tumor size T3/T4 vs. T2 vs. T1	1.207	1.097-1.328	<0.0001*	1.172	1.041-1.320	0.009*
LN involvement Yes vs. No	1.722	1.260-2.354	0.001	2.116	1.455-3.079	<0.0001*
Smoking Status Former/Current vs.Never	0.831	0.590-1.172	0.291	0.754	0.501-1.133	0.174

Table 4. Results from the multivariate Cox regression model for OS and RFS on LUAD *in silico* set. RFS relapse-free survival, OS overall survival, HR hazard ratio, CI confidence interval

	***in silico* set (n=338)**					
		RFS			**OS**	
	HR	**95%CI**	**p-value**	**HR**	**95%CI**	***p-value***
LGALS3 (High vs Low)	1.908	1.294-2.814	0.001	1.513	1.092-2.096	0.013
Tumor size T3/T4 vs. T2 vs. T1	1.568	1.249-1.968	<0.0001	1.451	1.193-1.763	<0.0001

Table 5. Results from the univariate Cox regression model for OS and RFS of LUAD validation Set. *RFS*, relapse-free survival; OS, overall survival; *HR*, hazard ratio; CI, confidence interval.

	Validation set (N=48)	
	RFS	**OS**

	HR	95%CI	p-value	HR	95%CI	*p-value*
sGAL-3 (High vs Low)	2.269	0.985-5.230	0.054	2.844	1.127-7.176	0.027*
Gender Male vs Female	2.802	1.117-7.031	0.028*	2.870	1.049-7.848	0.040*
Age >65 vs. ≤65	0.738	0.327-1.662	0.163	1.071	0.453-2.529	0.879
TNM staging III vs. II vs. I	1.762	1.086-2.857	0.022*	1.653	0.977-2.797	0.061
Tumor size T3/T4 vs. T2 vs. T1	1.792	0.976-3.293	0.060	1.506	0.805-2.815	0.200
Performance Status (PS) 0 vs 1	3.354	1.352-8.321	0.009*	2.803	1.072-7.331	0.036*
LN involvement Yes vs. No	2.023	0.878-4.661	0.098	1.556	0.626-3.866	0.341
Smoking Status Former/Current vs.Never	3.311	0.981-11.17	0.054	1.803	0.599-5.427	0.294

Table 6. Results from the multivariate Cox regression model for RFS and OS on LUAD validation set. *RFS*, relapse-free survival; *OS*, overall survival; *HR*, hazard ratio; *CI*, confidence interval.

	validation set (N=48)					
	RFS			OS		
	HR	95%CI	p-value	HR	95%CI	*p-value*
LGALS3 (High vs Low)	2.862	1.057-7.753	0.039	3.580	1.185-10.81	0.024
Gender Male vs Female	-	-	-	3.238	1.043-10.05	0.042
(PS) 0 vs 1	3.139	1.116-8.829	0.030	-	-	-

Table 7**.** Patient's characteristics of advanced-stage LUAD cohort. *n*, sample size. [1]PD-L1 expression was assessed by tumor proportion scores (TPS).

Patient characteristics	LUAD advanced cohort	
	***n* = 34**	**%**
Age at surgery (median, range)	67 [IQR 52-89]	
Gender		
Male	27	79.4
Female	7	20.6

Stage		
III	6	17.6
IVA	11	32.4
IVB	17	50
Performance Status		
0-1	29	85.3
2	4	11.8
Smoking status		
Current	25	73.5
Former	7	20.6
Never	2	5.9
PD-L1 TPS [1]		
100%	2	5.9
95%	3	8.8
90%	8	23.5
80%	5	14.7
70%	8	23.5
60%	8	23.5
Progression		
Yes	24	70.6
No	10	29.4
Exitus		
Yes	23	67.6
No	11	32.4

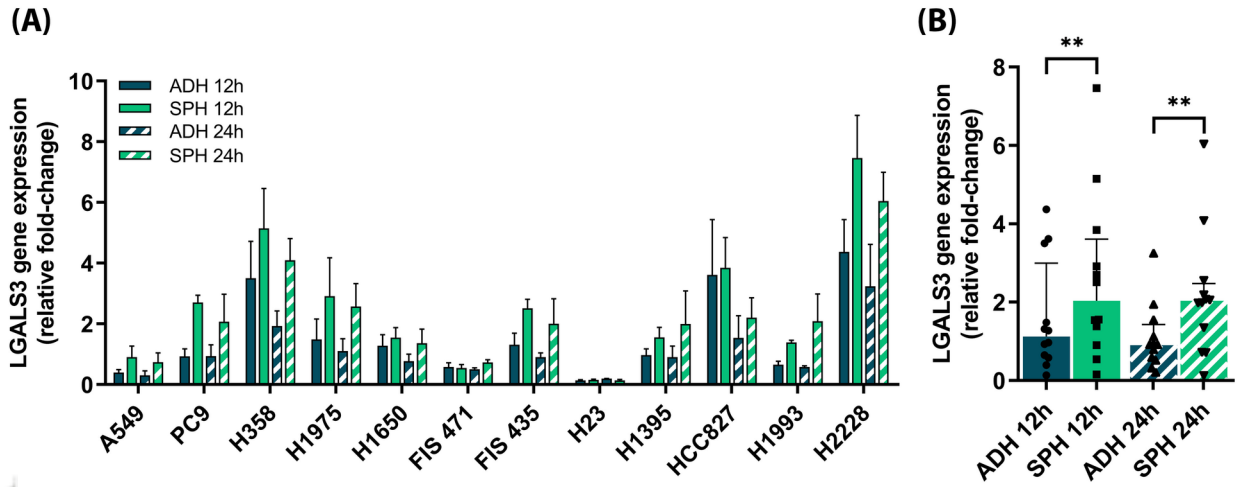

Figure_1.tif

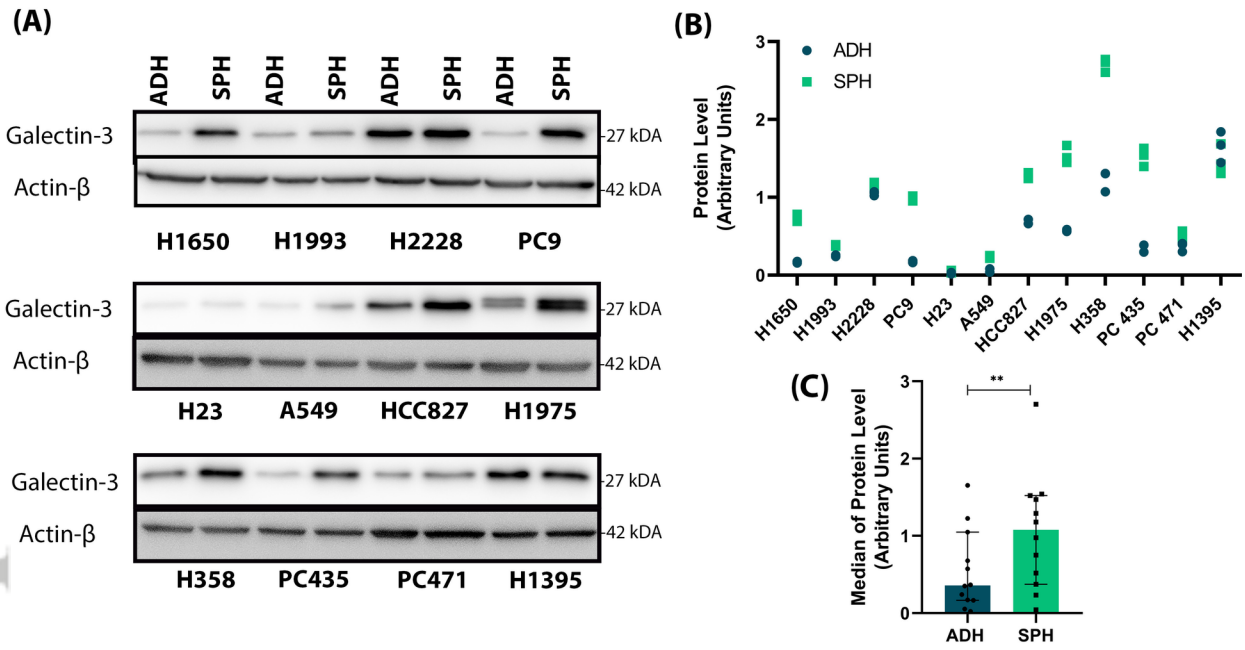

Figure_2.tif

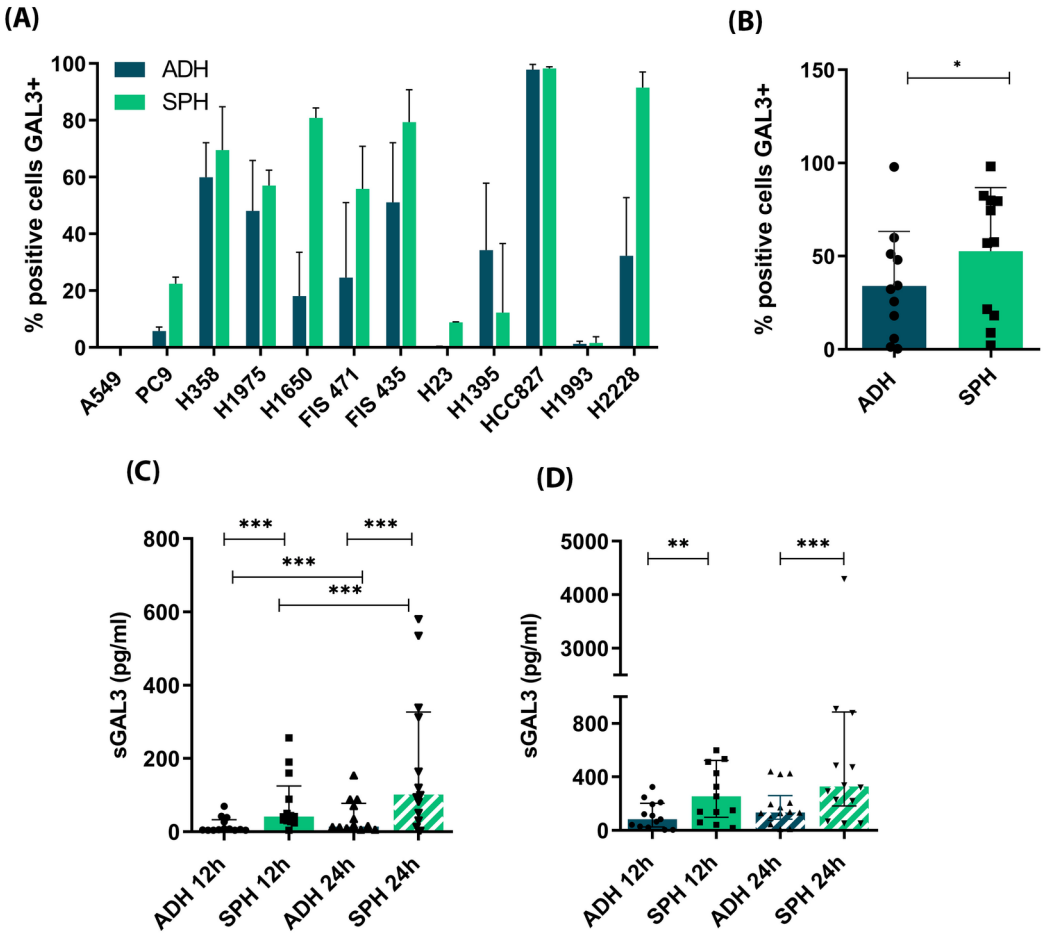

Figure_3.tif

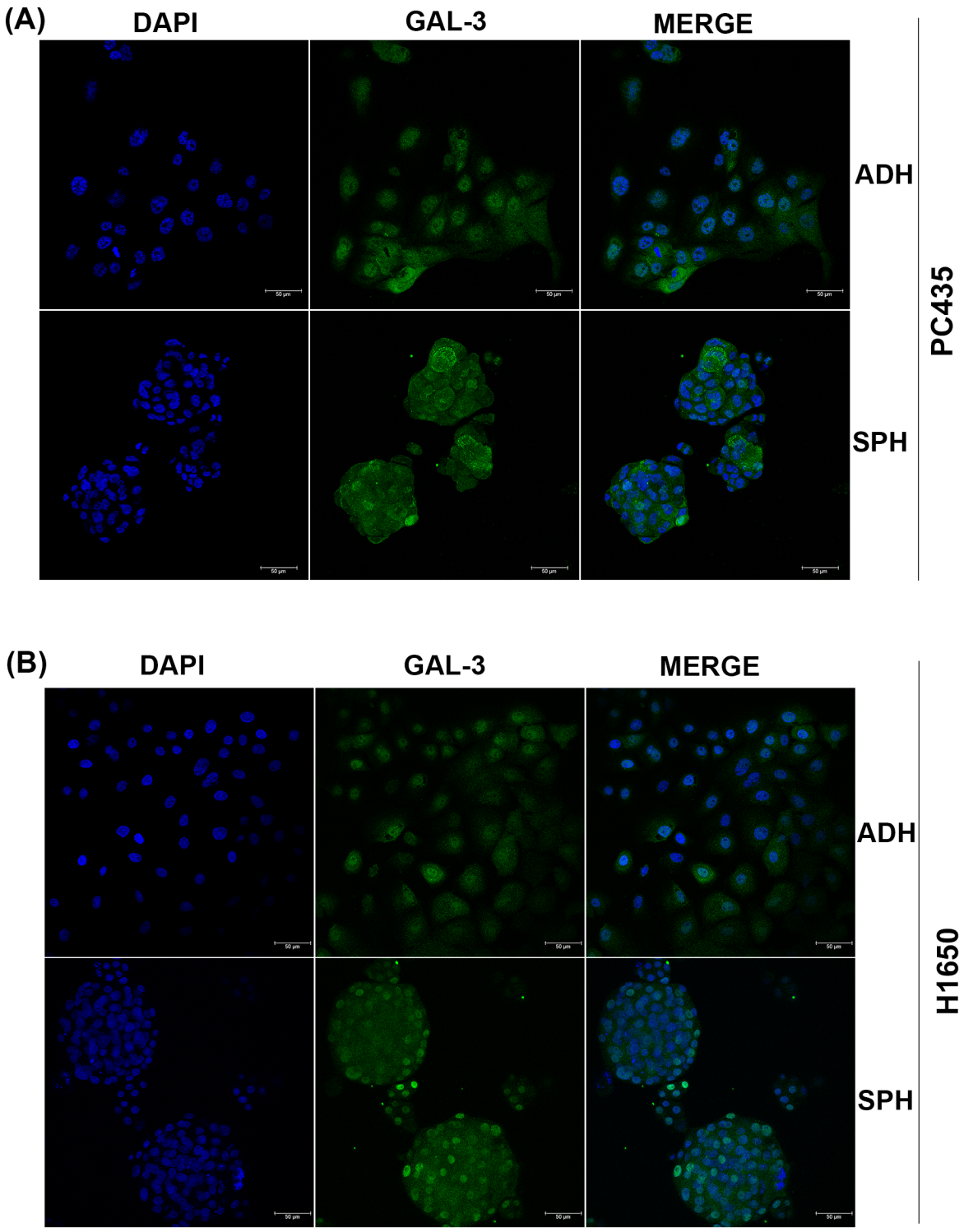

Figure_4.tif

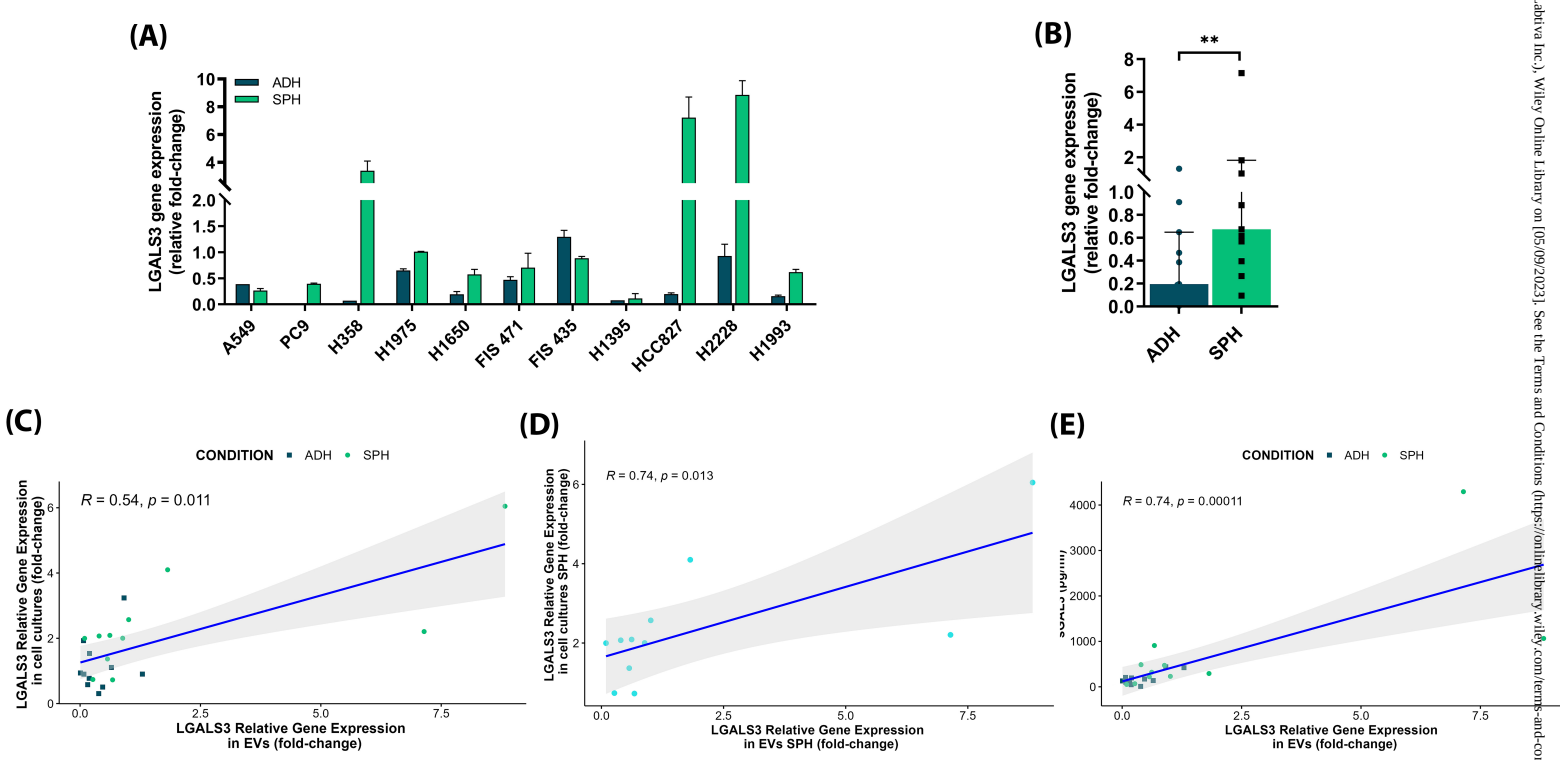

Figure_5.tif

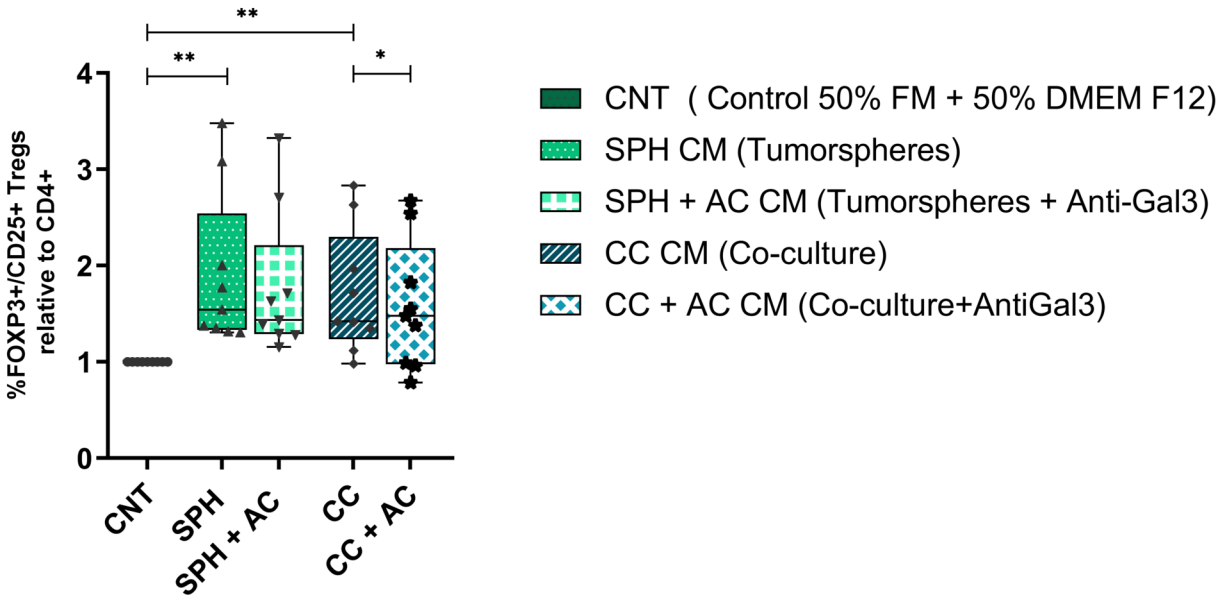

Figure_6.tif

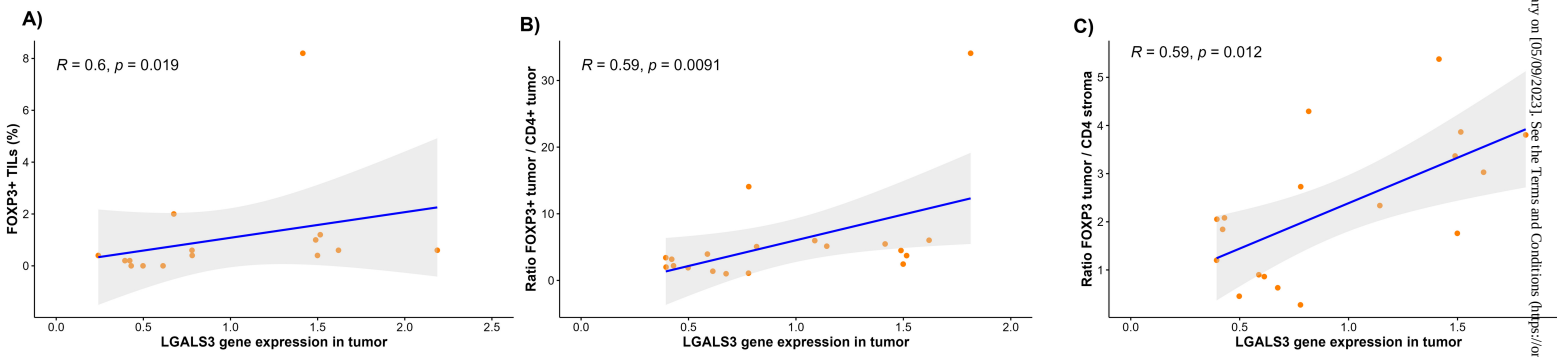

Figure_7.tif

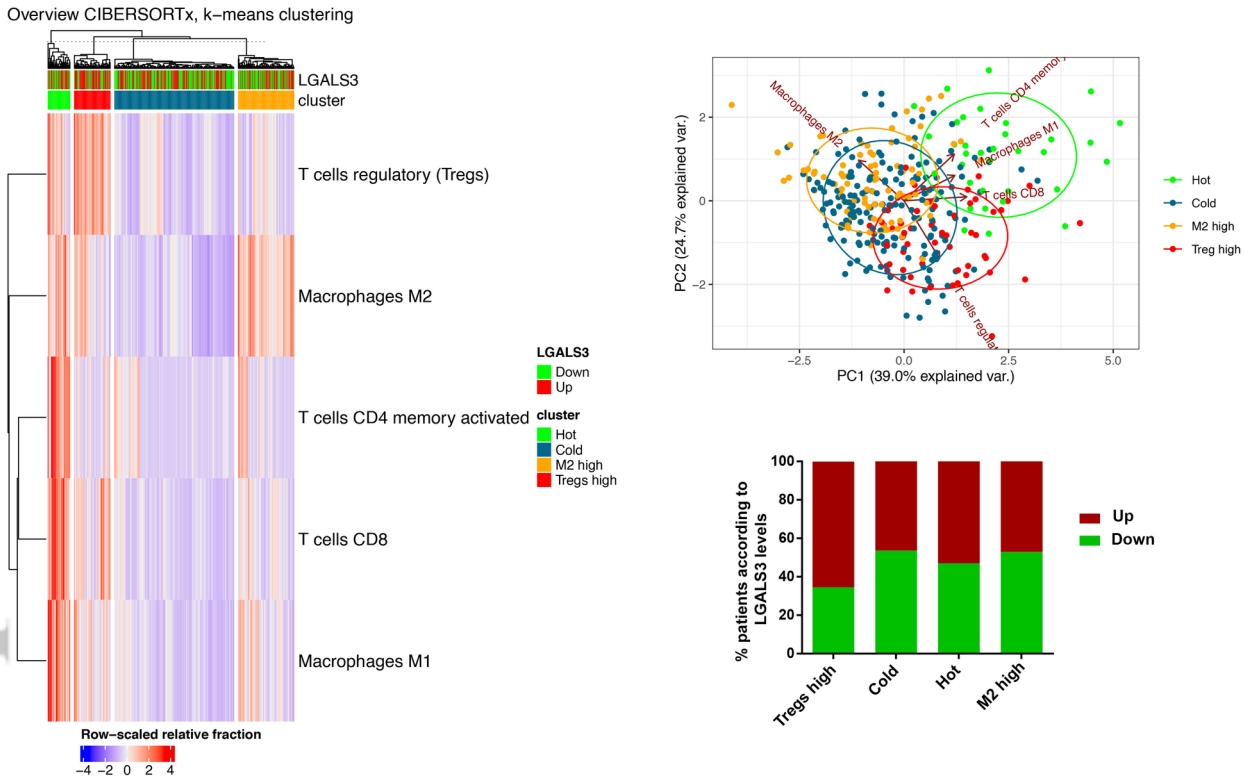

Figure_8.tif

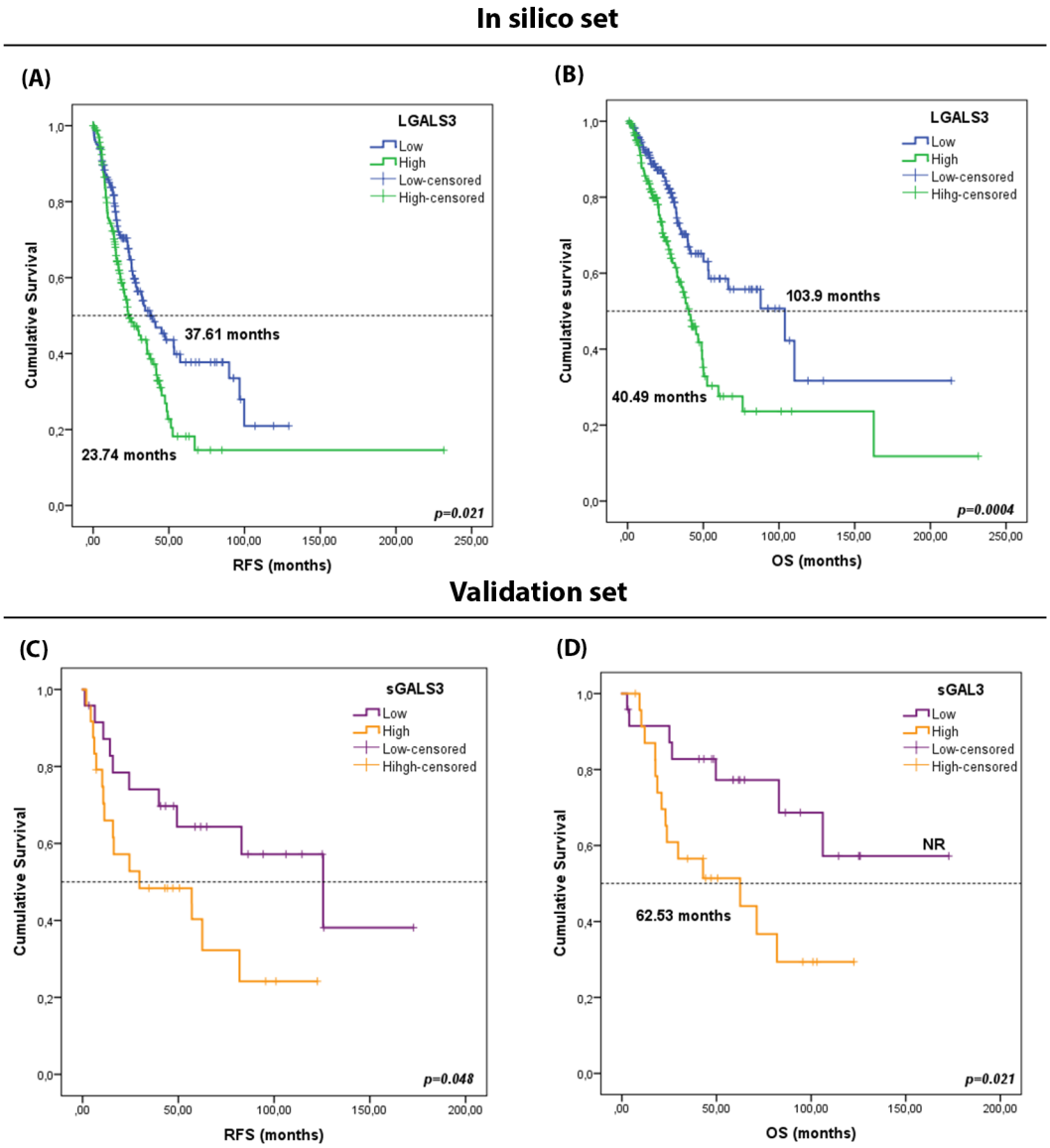

Figure_9.tif

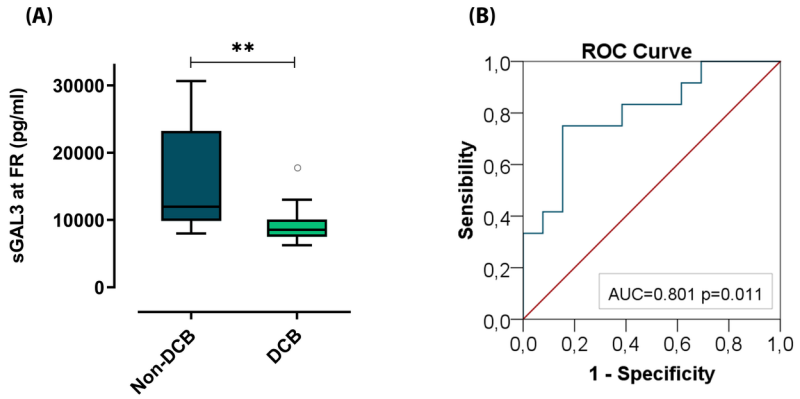

Figure_10.tif

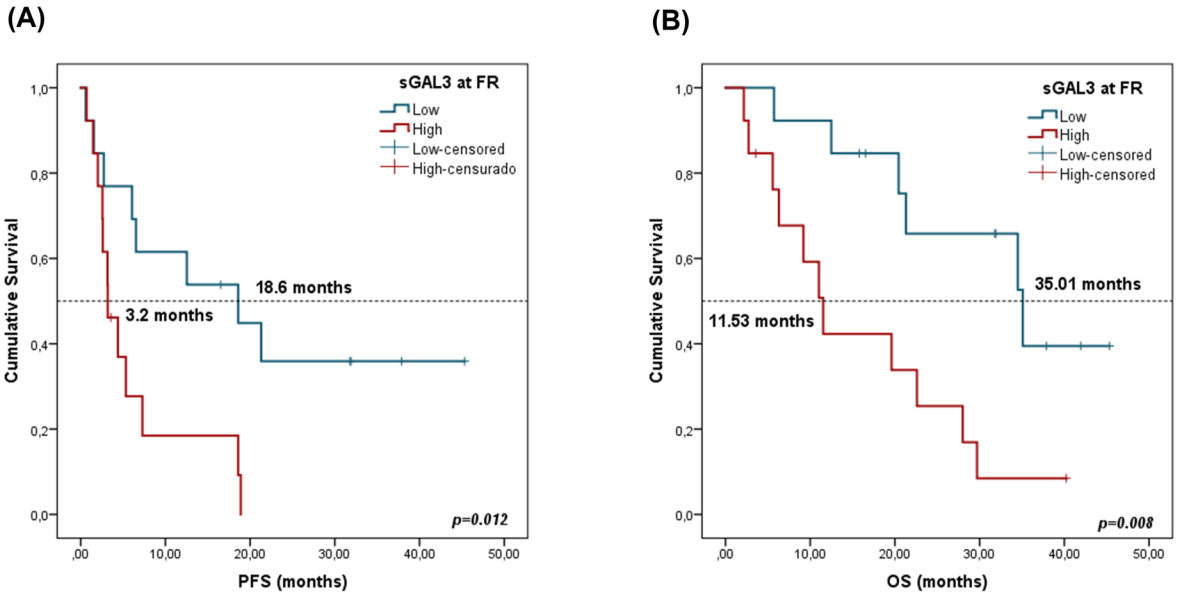

Figure_11.tif