



OPEN ACCESS

EDITED BY

Paola Ulivi,
Scientific Institute of Romagna for the Study
and Treatment of Tumors (IRCCS), Italy

REVIEWED BY

Mario Cioco,
Campus Bio-Medico University, Italy
Amir Feizi,
Novo Nordisk Research Centre Oxford
(NNRCO), United Kingdom

*CORRESPONDENCE

William W. Lockwood
✉ wlockwood@bccrc.ca

RECEIVED 31 August 2023

ACCEPTED 14 December 2023

PUBLISHED 08 January 2024

CITATION

Luu JK, Johnson FD, Jajarmi J, Sihota T,
Shi R, Lu D, Farnsworth D, Spencer SE,
Negri GL, Morin GB and Lockwood WW
(2024) Characterizing the secretome of EGFR
mutant lung adenocarcinoma.
Front. Oncol. 13:1286821.
doi: 10.3389/fonc.2023.1286821

COPYRIGHT

© 2024 Luu, Johnson, Jajarmi, Sihota, Shi, Lu,
Farnsworth, Spencer, Negri, Morin and
Lockwood. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Characterizing the secretome of EGFR mutant lung adenocarcinoma

Jennifer K. Luu^{1,2}, Fraser D. Johnson^{1,3}, Jana Jajarmi^{1,3},
Tianna Sihota^{1,2}, Rocky Shi^{1,3}, Daniel Lu^{1,3}, Dylan Farnsworth^{1,3},
Sandra E. Spencer⁴, Gian Luca Negri⁴, Gregg B. Morin^{4,5}
and William W. Lockwood^{1,2,3*}

¹Department of Integrative Oncology, British Columbia (BC), Cancer Research Institute, Vancouver, BC, Canada, ²Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada, ³Interdisciplinary Oncology Program, University of British Columbia (UBC), Vancouver, BC, Canada, ⁴Michael Smith Genome Sciences Centre, BC Cancer Research Institute, Vancouver, BC, Canada, ⁵Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada

Background: Lung cancer is the leading cause of cancer related death worldwide, mainly due to the late stage of disease at the time of diagnosis. Non-invasive biomarkers are needed to supplement existing screening methods to enable earlier detection and increased patient survival. This is critical to *EGFR*-driven lung adenocarcinoma as it commonly occurs in individuals who have never smoked and do not qualify for current screening protocols.

Methods: In this study, we performed mass spectrometry analysis of the secretome of cultured lung cells representing different stages of mutant *EGFR* driven transformation, from normal to fully malignant. Identified secreted proteins specific to the malignant state were validated using orthogonal methods and their clinical activity assessed in lung adenocarcinoma patient cohorts.

Results: We quantified 1020 secreted proteins, which were compared for differential expression between stages of transformation. We validated differentially expressed proteins at the transcriptional level in clinical tumor specimens, association with patient survival, and absolute concentration to yield three biomarker candidates: MDK, GDF15, and SPINT2. These candidates were validated using ELISA and increased levels were associated with poor patient survival specifically in EGFR mutant lung adenocarcinoma patients.

Conclusions: Our study provides insight into changes in secreted proteins during *EGFR* driven lung adenocarcinoma transformation that may play a role in the processes that promote tumor progression. The specific candidates identified can be harnessed for biomarker use to identify high risk individuals for early detection screening programs and disease management for this molecular subgroup of lung adenocarcinoma patients.

KEYWORDS

lung cancer, *EGFR*, secretome, transformation, early detection

Background

Lung cancer is the leading cause of cancer mortality in men and women worldwide, contributing to 1.8 million deaths in 2020 alone (1). Lung cancer consists of two subtypes, small cell lung cancer and non-small cell lung cancer (NSCLC), which comprise 15% and 85% of cases, respectively (2). The most common NSCLC subtype is lung adenocarcinoma (LUAD), which comprises ~60% of NSCLC cases (3). LUAD can be classified into oncogenic driver subgroups, where mutations in *KRAS* and *EGFR* are common (4). *KRAS* mutations are associated with smoking; in contrast, *EGFR* mutations are associated with never smokers, especially in women and in East Asia (5, 6). LUAD is thought to arise from a stepwise process of genetic and epigenetic changes, which begins with histologically normal epithelial cells and ends with invasive carcinoma (7, 8). The majority of *in vitro* studies aiming to investigate the genetic alterations required to enable transformation have centered on *KRAS*-driven LUAD, with limited investigation into *EGFR*-driven LUAD (9, 10).

Lung cancer, including LUAD, is typically diagnosed in late or metastatic stages; in these stages, long term patient survival is limited due to less effective treatment methods available such as chemotherapy and radiotherapy (2, 11). Prognosis and clinical stage are directly related, incentivizing earlier lung cancer detection for improvements to patient survival (12). Patients diagnosed at stage I have a five-year survival rate of 68.4%, in contrast with those diagnosed at stage IV, where the five-year survival rate is 5.8% (13, 14). Lose dose computed tomography (LDCT), a radiographic scanning technique used to image the lungs, is the standard for lung cancer screening (15). LDCT was initially demonstrated as an effective annual screening technique for high-risk individuals in the National Lung Screening Trial (NLST), where there was a 20% reduction in lung cancer mortality (16, 17). However, LDCT is limited by non-specificity, over diagnosis of benign pulmonary nodules, and potential harms of repeated radiation (16, 18). Furthermore, LDCT screening is not universally applicable to all populations susceptible to cancer; never smokers were not included in the NLST study and its effectiveness for this group is unclear (19, 20). One proposed strategy to supplement LDCT for better screening practices is the use of biomarkers of early cancer development (21). Biomarkers would complement LDCT by reducing screening costs through criteria refinement, supporting clinical decision making in unclear situations such as indeterminate pulmonary nodules, and personalized patient screening and treatment planning (21). Blood biomarkers are of particular interest, due to their capability for inexpensive and relatively non-invasive collection (22). However,

Abbreviations: BCCA, British Columbia Cancer Agency; DPE, Differential protein expression; DNA, Deoxyribonucleic acid; EGFR, Epidermal growth factor receptor; ELISA, Enzyme-linked immunosorbent assay; GFP, Green fluorescent protein; GSEA, Gene set enrichment analysis; GO, Gene ontology; HBEC, Human bronchial epithelial cells; KRAS, Kirsten rat sarcoma viral oncogene homolog; LFC, Log fold change; LDCT, Lose dose computed tomography; LUAD, Lung adenocarcinoma; MS, Mass spectrometry; NLST, National Lung Screening Trial; NSCLC, Non-small cell lung cancer; PCA, Principal component analysis; RNA, ribonucleic acid.

there are few confirmed protein biomarkers for LUAD and current candidates including NSE, proGRP, CEA, SCCA, and CYRFA 21-1 are non-specific for lung cancer and cannot be used to distinguish histology nor molecular subtype (23).

The secretome consists of proteins transported from a cell into the extracellular space and it is estimated to comprise 15% of all human proteins (24). Secretome proteins include cytokines, growth factors, extracellular matrix-degrading proteinases, and cell motility factors involved in local and systemic signaling (24, 25). A regulated secretome is important in maintaining homeostasis and changes in secretome protein abundance have been implicated in cancer (26). Tumor cells can release proteins that can affect functions such as angiogenesis, immunomodulation, basement membrane degradation, and extracellular matrix modeling (27). Secreted proteins enter bodily fluids such as blood and urine, which enables non-invasive collection and potential biomarker analysis (25, 28). A recent study of 32 types of primary tumors and normal-adjacent tissues found that proteins often found in the secretome are altered at the transcriptional level specifically in cancer, and found common expression decreases of proteins implicated in functions including adhesion and tumor suppression (29). These findings highlight the broad scope changes in the secretome during tumor progression and metastasis; in addition, this study showed the potential of the secretome as a reservoir of biomarker candidates such as matrix metalloprotease (MMP) family members, including MMP9 in breast and lung cancer (29–31). In NSCLC, recent secretome studies have identified proteins affecting erlotinib resistance, biomarkers for cisplatin response, and metastasis (32–34). However, secretome studies often profile immortalized cell lines, where cells are established from patient tumors and have already undergone malignant transformation (35, 36). Studying changes in secreted proteins that occur during the different steps of cancer progression from normal epithelium to invasive and metastatic cancer may therefore generate potential biomarker candidates to aid in early detection, diagnosis and prognosis. This is urgently needed in *EGFR*-driven LUAD, as the transformation process has not been fully elucidated and there are no concrete screening guidelines for the never smoker demographic where mutant *EGFR* LUAD cases are enriched (20, 37).

In this study, we investigated changes in the secretome during malignant transformation and identified potential biomarker candidates by performing proteomic analysis using an *in vitro* model of mutant *EGFR* driven transformation. We generated cell lines modeling the stepwise genetic alterations that occur during transformation, using non-transformed human bronchial epithelial cells (HBEC), and compared these against established LUAD cell lines to profile differences between stages of transformation (38). We initially identified 1020 secretome proteins and progressed through a series of groupwise and individual cell line comparisons to uncover 499 differentially expressed proteins between the untransformed and transformed states. Key selected proteins were validated with gene expression and patient survival data, to determine five biomarker candidates including MDK, GDF15, and SPINT2. This provides the first description of secretome changes during mutant *EGFR*-driven LUAD transformation and provides insight into the biological processes that can be applied for biomarker development.

Methods

Cell culture

All cell lines used were obtained from the American Type Culture Collection (ATCC) or gifted from Dr. Adi Gazdar (UT Southwestern Medical Center). PC-9, H1975 (NCI-H1975), HCC4006, HCC4011, H3255 (NCI-H3255) were cultured in RPMI-1640 (Thermo Fisher Scientific), supplemented with 10% fetal bovine serum (FBS; Thermo Fisher Scientific) and 1% Penicillin-Streptomycin (Thermo Fisher Scientific). HBEC (HBEC3-KT) cells were cultured in Keratinocyte serum-free medium (KSFM; Thermo Fisher Scientific), supplemented with accompanying bovine pituitary extract (BPE; Thermo Fisher Scientific), human recombinant epidermal growth factor (EGF; Thermo Fisher Scientific), and 1% Penicillin-Streptomycin. All cell lines were cultured at 37°C, in 5% CO₂.

Expression constructs and cell line generation

Lentiviral vector and overexpression plasmids used to construct overexpression constructs for *EGFR L858R* (Plasmids #82906, #17451) and *GFP* (Plasmid #17445) were obtained from Addgene. Retroviral *TP53* c-terminal fragment (CT) overexpression construct and pCX4 hisD vector control were gifted from Dr. Romel Somwar (Memorial Sloan Kettering Cancer Centre, NY). Lentivirus was produced using HEK 293TD cells (ATCC), psPAX2 (Plasmid #12260; Addgene) and pMD2.G (Plasmid #12259; Addgene). Retrovirus was produced using Phoenix-AMPHO cells (ATCC). HBEC cell lines expressing *GFP*, *EGFR L858R*, with *TP53* C-terminal (CT) domain dominant negative mutations were generated by lentiviral and retroviral infection, and selected with 5 µg/mL blasticidin (Thermo Fisher Scientific), and 2 mg/mL L-histidinol (Thermo Fisher Scientific). The HBEC cell line expressing *EGFR L858R* and *TP53* CT was additionally selected in 10 µM Nutlin-3a for 6 days (SelleckChem).

Western blot analysis

Protein from cell lysates were obtained by rinsing cells with cold Dulbecco's phosphate-buffered saline (DPBS) (Thermo Fisher Scientific) and lysed in RIPA buffer (VWR) with Halt protease and phosphatase inhibitor cocktail (Thermo Fisher Scientific). Samples were collected on ice, vortexed, and frozen at -80°C before being sonicated and centrifuge-separated at 15,000xg, 4°C for 10 minutes. Protein concentrations were detected using a Pierce BCA protein assay kit (Thermo Fisher Scientific), then samples were heated in 1x diluted NuPAGE LDS sample buffer (Thermo Fisher Scientific) containing 1:10 diluted 2-Mercaptoethanol (MilliporeSigma) at 75°C, for 10 minutes. 20-25 µg of samples were run on NuPAGE 4-12% Bis-Tris protein gels (Thermo Fisher Scientific) in NuPAGE MOPS SDS running buffer (Thermo Fisher Scientific) at 200V, for 50 minutes. Samples were

transferred from Bis-Tris gel to Immobilon-P PVDF (MilliporeSigma) either at 70V, 4°C, for 2 hours or 30V, 4°C overnight. Membranes were incubated in TBS-T (0.1% Tween-20) (TBS, Bio-Rad; Tween-20, Thermo Fisher Scientific) containing 5% BSA (MilliporeSigma) until primary antibody incubation.

Primary antibodies were prepared following manufacturer's instructions in TBS-T containing 5% BSA or 5% milk (MKP3); specific dilutions are noted. The following primary antibodies were used: p-ERK1/2(Thr202/Tyr204) (p-p44/42 (Thr202/Tyr204); Cell Signaling Technology, 9101); ERK1/2(p44/p42; Cell Signaling Technology, 4695); p-MEK1/2(Ser217/221) (Cell Signaling Technology, 9121); MEK1/2 (Cell Signaling Technology, 9122); p-EGFR(Tyr1068) (Cell Signaling Technology, 2234); EGFR L858R (Cell Signaling Technology, 3197); EGFR (Cell Signaling Technology, 2232); MKP3 (DUSP6) (Santa Cruz Biotechnology, sc-377070, 1:200); p53(Cell Signaling Technology, 2527); p53, to detect TP53 CT (MilliporeSigma, SAB4503011); GFP (Cell Signaling Technology, 2956); β-Actin (Cell Signaling Technology, 12620, 1:2000). Membranes were incubated in primary antibodies at 4°C overnight, then HRP-conjugated secondary antibodies according to manufacturer recommendations (Cell Signaling Technology). Proteins were detected after incubation with ECL, SuperSignal West Pico Plus Chemiluminescent Substrate (Thermo Fisher Scientific) or SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific) on a ChemiDoc MP imager (Bio-Rad).

Secretome sample collection

Cells were seeded at approximately 80% confluency in 6cm plates in triplicate overnight: 900,000 (HBEC); 1,000,000 (PC-9); 1,000,000 (H1975); 3,500,000 (HCC4006); 7,000,000 (HCC4011); 2,500,000 (H3255). Plates were rinsed twice with DPBS and media was changed to supplement-free KSFM containing 1% Penicillin-Streptomycin (HBEC) or serum-free RPMI-1640 containing 1% Penicillin-Streptomycin (PC-9, H1975, HCC4006, HCC4011, H3255). Plates containing only media were also prepared, and all plates were incubated for 24 hours at 37°C. Conditioned media was collected, centrifuged at 1000 RPM for 5 minutes, at 4°C, and filtered with a 0.45 µm filter (Sarstedt) to remove cell debris. The complete 4mL volume of filtered conditioned media was centrifuged in a Vivaspin Turbo 3kDa ultrafiltration unit at 3220xg, at 8°C until media was concentrated to approximately 150-200 µL. Concentrated media was buffer exchanged, where samples were centrifuged twice with 4mL 50mM HEPES buffer, pH 7.0, then once with 1mL HEPES at 3220xg, at 8°C to a final volume of 150-300 µL. Samples were stored at -80°C until mass spectrometry sample preparation.

Proteomic analysis

Samples were prepared for tandem mass spectrometry (MS/MS) analysis through a protocol of reduction, alkylation, and protein digestion. Samples were reduced by incubating with 16 µL

of 200 mM dithiothreitol (DTT, Bio-Rad) for 30 minutes at 55°C, then alkylated by incubating with 32 μ L of 400 mM iodoacetamide (IAA, Bio-Rad) for 30 minutes at room temperature. Samples were quenched with an additional 16 μ L 200 mM DTT. Trypsin/Lys-C mix was prepared for sample digestion, where 200 μ L of 200 mM HEPES pH 8.0 was added to 20 μ g Trypsin/Lys-C (Promega). Samples were digested by incubating with 16 μ L Trypsin/Lys-C mix on a ThermoMixer (Eppendorf) at 1000 RPM, overnight at 37°C. One tenth of each sample was pooled and prepared to confirm quality. Peptides were acidified by adding 10% (v/v) trifluoroacetic acid (TFA, Thermo Fisher Scientific) and diluted to a concentration of 1% TFA, then desalted following a Stop And Go Extraction (STAGE) tip protocol (39). Briefly, STAGE tips were packed with 3 punches of C18 resin which was washed (100 μ L 0.1% TFA in acetonitrile) and equilibrated (2x100 μ L 0.1% TFA in 18 M Ω water) then peptide was loaded. Salts were removed by rinsing (200 μ L 0.1% formic acid in HPLC water) then eluted in 100 μ L 0.1% formic acid in 60/40 acetonitrile/HPLC water. Desalted peptides were eluted and solvent evaporated by centrifuging samples in a SpeedVac Vacuum Concentrator (Thermo Fisher Scientific) until dry. Peptides were reconstituted in a 0.1% formic acid, 1% DMSO aqueous solution and assessed for quality on a LTQ Orbitrap VelosTM (Thermo Fisher Scientific).

The remaining digested peptides were tandem mass tag (TMT) labeled using a TMT 11-plex kit (Thermo Fisher Scientific) following manufacturer's instructions. Post-labeling, samples were pooled, dried by speed vacuum to evaporate excess solvent, and acidified with TFA as described above. Peptides were desalted following the STAGE tip protocol and excess solvent was reduced by vacuum centrifugation (39). Peptides were constituted in a 0.1% formic acid, 1% DMSO aqueous solution and run on an Orbitrap EclipseTM mass spectrometer (Thermo Fisher Scientific) set to MS2 mode. MS spectra were searched with Proteome Discoverer suite (v.2.4.0.305, Thermo Fisher Scientific) against Swissprot human reference database (20585 sequences, October 2020). Precursor and fragment ion tolerance were set to 20 ppm and 0.05 Da, respectively. Dynamic modifications included Oxidation (+15.995 Da, M), Acetylation (+42.011 Da, N-Term), and static modification included Carbamidomethyl (+57.021 Da, C) and TMT (+229.163 Da, K, N-Term). Peptide-spectrum matches (PSMs) were validated with Percolator, where only PSMs with false discovery rate (FDR) < 0.01 were retained in the analysis.

PSMs were filtered by removing PSMs with average signal-to-noise (SN) ratio lower than 10 and isolation interference higher than 50% and SN was summarized to the protein level for analysis. Protein level data was log₂-transformed and median normalized, where normalization was performed by taking the median total signal, calculating respective normalization factors for samples and media controls, and then missing values were imputed. Samples were compared against the appropriate media control (HBEC, K562; other cell lines, RPMI) on the log₂ scale, and enriched proteins were determined by analyzing the intersection between sample and media. Differences between technical replicates and cell lines were assessed with principal component analysis (PCA).

Proteins were filtered prior to statistical analysis, where only proteins seen in 2 or more technical replicates were retained. The average log₂ HBEC GFP;p53^{wt} signal intensity was subtracted from all samples to generate average log₂ fold changes. Sample log₂ fold changes were analyzed with the limma package (40) (version 3.50.0) with the moderated *t*-test in R, and adjusted *P*-values were calculated with the Benjamini-Hochberg procedure, with those < 0.05 considered significant (R version 4.0.5).

Protein annotation and gene ontology enrichment analysis

The high-throughput model of DeepLoc 2.0 was used for the prediction of subcellular localization for the identified proteins (41). Functioning as a multi-label predictor, it possesses the capability to anticipate one or more localizations for a given protein. Selected gene lists were analyzed for Gene Ontology (GO) terms using the clusterProfiler package in R for enrichment in biological processes, molecular function, and cellular compartments (42). *P*-values were adjusted using the Benjamini-Hochberg procedure and terms with *adjp* < 0.05 were retained (R version 4.0.5). Reactome enrichment analysis was performed using ShinyGO 0.77 with EnsemblIDs corresponding to the individual proteins where available (43, 44). The 'Curated.Reactome' database was assessed using default settings, consisting of FDR < 0.05, min pathway size *n*=2, and max pathway size *n*=2000 and the top 20 pathways were plotted, sorted by fold enrichment.

Microarray analysis

Z-score normalized Affymetrix gene expression data collected from 199 primary lung tumors was retrieved from Memorial Sloan Kettering Cancer Center (45). Statistical analysis was performed with the Wilcoxon rank-sum test (Mann-Whitney *U* test) and adjusted *P*-values were calculated using the Benjamini-Hochberg procedure. Analysis was performed using base R functions (version 4.0.5). Statistically significant microarray probes were mapped to their corresponding gene with the R package hgu133a.db (46) (version 3.2.3). Genes were filtered for optimal 1:1 probe:gene mapping with the R package jetset (47) (version 3.4.0) to yield a final gene list.

Survival analysis

Differentially expressed genes were analyzed for differences in patient survival with NCBI GEO gene expression dataset GSE31210 (48). Probes were mapped to the corresponding genes with the R package hgu133plus2.db (49) (version 3.2.3). Median overall survival was calculated by applying a median split in gene expression and the Logrank test in Graphpad Prism 6.

ELISA

Quantification of GDF-15 in the secretome was performed using the Human GDF-15 Quantikine ELISA Kit (R&D Systems, DGD150) according to the kit instructions. MDK and SPINT2 were quantified using the Human MDK ELISA Kit (Invitrogen, EH319RB) and SPINT2 (HAI-2) Human ELISA Kit (Invitrogen, EH319RB and EHSPINT2) according to kit instructions. Quantification was performed on conditioned media samples collected under secretome collection conditions, in parallel with secretome experiment collection. All samples were run in duplicate. Sample concentrations were determined by subtracting the media background control signal, then interpolating with a standard curve. Differences in concentrations were statistically computed with the unpaired Student's T-test in Graphpad Prism6.

Trypan Blue viability stain

Cells were seeded at approximately 80% confluency in 6-well plates overnight: 350,000 (HBEC); 340,000 (PC-9); 340,000 (H1975); 1,200,000 (HCC4006); 2,400,000 (HCC4011); 850,000 (H3255). Plates were rinsed twice with PBS and media was changed to supplement-free or serum-free conditions and incubated for 24 hours. HBECs were incubated with either supplement-free KSMF and 1% Penicillin-Streptomycin, or KSMF supplemented with BPE (50 µg/mL), EGF (5 ng/mL), and 1% Penicillin-Streptomycin. NSCLC cell lines were incubated with either serum-free RPMI-1640 and 1% Penicillin-Streptomycin, or RPMI-1640 supplemented with 10% FBS and 1% Penicillin-Streptomycin. Post-incubation, cells were trypsinized with 0.05% Trypsin-EDTA (Thermo Fisher Scientific; HBEC) or 0.25% Trypsin-EDTA (Thermo Fisher Scientific; NSCLC). Trypsinization was neutralized with either trypsin neutralizer (Thermo Fisher Scientific; HBEC) or RPMI-1640 containing 10% FBS (NSCLC), cells mixed in 0.4% Trypan Blue solution (Thermo Fisher Scientific) at a 1:1 ratio, and live cell population determined with a TC20 automated cell counter (Bio-Rad). The average percent live cell population was determined from the average of 3 wells with the unpaired Student's T-test in Graphpad Prism6 (2 counts per well).

Propidium Iodide viability stain

Cells were seeded and treated under supplement-free or serum-free conditions for 24 hours as described in the Trypan Blue viability analysis section. Post-treatment, cells were incubated with 1 µg/mL Hoescht 33342 (Thermo Fisher Scientific) for 30 minutes, and 1 µg/mL propidium iodide (Thermo Fisher Scientific) for 10 minutes, respectively. Stained cells were imaged with an EVOS FL fluorescence microscope (Thermo Fisher Scientific). Live cell population was determined by quantifying the average live cell population from 2 images per well using ImageJ software, and applying the unpaired Student's T-test in Graphpad Prism6. The formula $100\% - ((PI \text{ stained cell population} \div \text{Hoescht 33342 cell population}) \times 100)$ was used to determine the live population.

Results

Mass-spectrometry secretome profiling of mutant EGFR lung cell models

To study potential changes in secreted proteins during malignant mutant *EGFR*-driven LUAD transformation, we generated an *in vitro* model approximating mutant *EGFR* malignant transformation with HBEC stable cell lines and selected *EGFR* mutant NSCLC cell lines (Figure 1A). HBECs, a bronchial epithelial cell line immortalized with non-viral proteins hTERT and CDK4, was selected due to its ability to maintain a non-transformed phenotype post-immortalization *in vitro* and *in vivo* (38). HBEC cell lines stably expressing *EGFR*^{L858R} or GFP control, with or without dominant negative p53 C-terminal domain alterations (p53^{CT}) (GFP;p53^{wt}, GFP;p53^{CT}), *EGFR*^{L858R};p53^{wt}, *EGFR*^{L858R};p53^{CT} represent a profile of commonly mutated genes observed in *EGFR* mutant LUAD (4, 50). To confirm gene expression, HBEC cells were treated with the MDM2 inhibitor Nutlin-3a for 24 hours to assess p53 levels (51). HBEC GFP;p53^{CT} showed minor changes in p53 expression, consistent with a mutant p53 phenotype, comparable to the mutant p53 NSCLC cell line H1975 (52) (Figure 1B). Cell lines expressing p53^{wt} showed increased p53 expression, which is consistent with Nutlin treatment (53). Expression of *EGFR*^{L858R} was confirmed using mutant specific antibodies (Figure 1B).

For secretome collection and analysis, HBEC and *EGFR* mutant NSCLC cell lines were starved under supplement-free and serum-free conditions for 24 hours, with cell-free media serving as controls (Figure 1C). This was performed to improve the detection of low abundance proteins masked by FBS and minimize non-human contamination, which is common practice in secretome experiments (54). However, starvation conditions can negatively affect cell viability and increase cell cytolysis, potentially contaminating media with intracellular proteins (35, 54). To investigate whether supplement- or serum-free conditions would have an effect in this regard, we assessed cell viability prior to secretome collection (Supplementary Figures 1, 2). Cell lines treated with Alamar Blue or propidium iodide (PI) showed no major changes in cell viability, measured as a live cell population percentage (Supplementary Figures 2B, C).

Upon secretome collection, samples were analyzed by TMT-11 MS/MS. To improve detection of low abundance proteins, conditioned media was concentrated prior to mass spectrometry (Figure 1C). Data was checked for quality and potential sources of variation introduced during sample processing using principle component analysis (PCA), confirming that technical replicates were generally clustered per cell line tumor or tissue origin, separate from media control samples (Supplementary Figure 3). We identified 1020 proteins post-secretome collection. The initial output was filtered to identify proteins in conditioned media by removing the intersection between media control and cell line samples, yielding 852 candidate secreted proteins (see methods). We then assessed this subset for those predicted to contain a signal peptide for secretion and their predicted subcellular localization

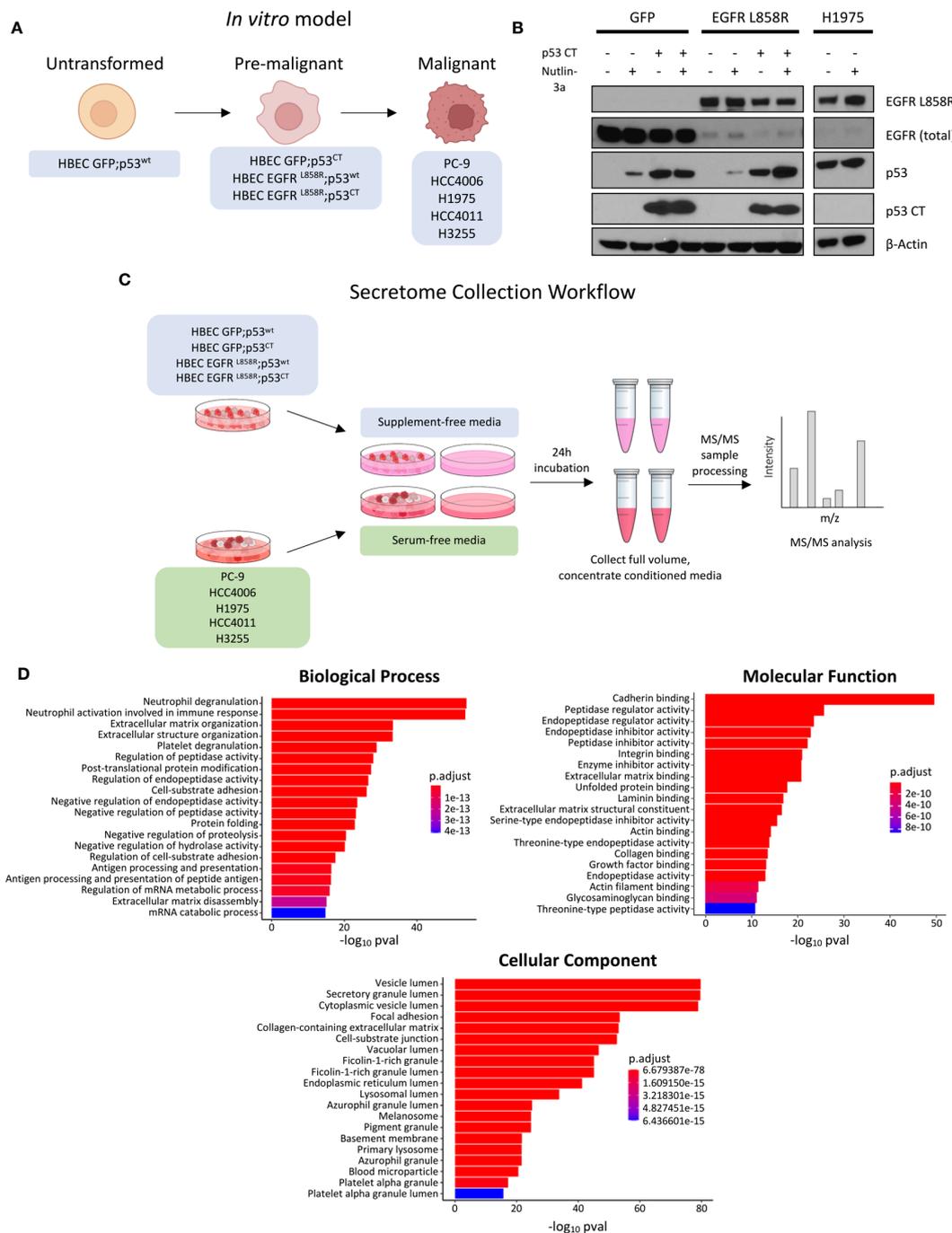


FIGURE 1

Experiment overview, model expression validation, and GO enrichment analysis of identified proteins. (A) Schematic overview of model, where each cell line represents a stage of malignant transformation. HBEC GFP;p53^{wt} represents the wild-type, untransformed state, HBEC cell lines expressing mutations in EGFR-driven transformation represent an intermediate pre-malignant stage, and EGFR mutant NSCLC cell lines PC-9, HCC4006, H1975, HCC4011, and H3255 represent the transformed, malignant state. (B) Western blot expression validation of HBEC stable cell lines expressing GFP, EGFR L858R, in combination with p53 CT. Cell lines were treated with 10μM Nutlin 3-a, an MDM2 inhibitor, for 24 hours to confirm mutant p53 expression. EGFR L858R basal expression was also confirmed. H1975 was used as a positive control. (C) Schematic overview of the secretome collection experiment. (D) Bar plots showing the top 20 GO enrichment terms sorted by adjusted p-value (p adj < 0.05, Benjamini-Hochberg adjustment) for biological processes, molecular functions, and cellular components, for proteins identified in 2 or more technical replicates (prior to statistical analysis).

using DeepLoc 2.0 (41, see methods). This predictive model can distinguish among 10 distinct localization and has the ability to forecast the presence of sorting signal peptides that influence the prediction of subcellular localization.

In total, image 359 (42%) of the candidate proteins are predicted to contain a signal peptide and 217 (26%) are predicted to have extracellular localization (Supplementary Table 1). This includes the proteins HSPG2, LAMA5, and AGRN that have previously been

demonstrated to be secreted (33). An additional 123 (14%) of the candidate secreted proteins were predicted to be localized to the cell membrane, including EGFR which is known to undergo shedding into the extracellular space (Supplementary Table 1) (55).

The enriched protein subset was examined for associated GO terms to further assess if proteins were secreted (Figure 1D). Cellular component GO terms were commonly associated with vesicular protein transport; the top enriched terms included “vesicle lumen”, “secretory granule lumen”, and “cytoplasmic vesicle lumen” (Figure 1D) (56). Other terms were associated with cellular compartments such as the lysosome and endoplasmic reticulum, which could suggest affiliation with either conventional or unconventional secretion (57, 58). Secretion-associated cellular component terms were complemented by biological process terms that are associated with extracellular proteins, with examples including “neutrophil degranulation”, “extracellular matrix organization”, and “platelet degranulation” (59–61). Of the molecular function GO terms identified, cadherin binding was the most significantly enriched; this may be attributed to cadherin and the associated catenin binding to facilitate cell adhesion (62). Proteolysis terms, such as peptidase and endopeptidase regulation, were also enriched (Figure 1D). These terms are consistent with protease functions, which range from cell proliferation to the immune response (63). To interrogate specific signaling pathways associated with the identified proteins we performed a separate enrichment analysis interrogating the Reactome database (see methods). This revealed the top enriched pathways to include post-translational protein phosphorylation, IGF signaling among others including platelet/neutrophil degranulation and non-integrin membrane-ECM interactions that closely resemble the results from the GO analyses (Supplementary Figure 4).

Identification of differentially expressed secreted proteins corresponding to stages of mutant EGFR mediated lung cell transformation

Differential protein expression (DPE) analysis was performed to investigate differences in secreted proteins between the pre-malignant and malignant stages of EGFR mutant LUAD transformation (Figure 1A). This was performed by comparing HBEC cell lines expressing mutant EGFR and/or p53 to EGFR mutant LUAD cell lines and assessing differences specific to each group (Figure 2A). 91 proteins were found to be differentially expressed, with 64 under-expressed and 27 over-expressed in the malignant vs non-malignant states, respectively (Figure 2A). Hierarchical clustering based on the differentially expressed proteins revealed distinct grouping between HBEC cell lines and EGFR mutant NSCLC cell lines (Figure 2B). This suggests that there may be distinct secretome profiles between pre-malignant and malignant stages of lung transformation. We assessed the top five over-expressed and under-expressed proteins in malignant vs non-

malignant states for their potential as biomarker candidates and found that two, NPC2 and MDK - both of which are predicted to have a signal peptide and extracellular localization (Supplementary Table 1) - have been found to be over-expressed in mouse LUAD plasma and NSCLC patient serum, respectively, confirming their secretion (64, 65). We queried the 91 differentially expressed proteins for GO terms associated with biological processes, molecular functions, and cellular components (Figure 2C). Four of the top five enriched cellular compartment GO terms were associated with secretory pathways, such as “secretory granule lumen” and “cytoplasmic vesicle lumen”, suggesting the presence of secreted proteins or proteins involved in secretion; this includes conventional secretion, but also unconventional secretion mediated by lysosomes, autophagosomes, and multivesicular bodies that become exosomes (58, 66). “Collagen-containing extracellular matrix” was the top enriched cellular component GO term, which could reflect the predominant role of collagen in extracellular matrix formation and integrity, as well as functions such as cell adhesion (67, 68). Similar to the observations made during initial secretome profiling, terms related to proteolysis represented the top 5 molecular function terms (Figures 1B, 2C). This is reflective of the broad functions of proteases which include extracellular matrix assembly and remodeling and aligns with the top GO cellular component terms (69). Likewise, protease-related functions were represented in biological process GO terms; interestingly, immune cell functions were also represented (Figure 1B). The presence of immune cell-related terms suggests changes in immune regulatory programs that occur during transformation. This aligns with previous observations where PD-L1, a key protein in immune homeostasis, was upregulated in EGFR mutant NSCLC cell lines and expression associated with clinical LUAD samples (70, 71). Reactome analysis also revealed enrichment in numerous immune related signaling pathways in addition to ATF6 and JAK-STAT signaling (Supplementary Figure 5).

To investigate the differences in secreted proteins during transformation, we analyzed differences in secreted proteins between untransformed and pre-malignant stages (Figure 1A, Supplementary Figure 6). This was done through comparison across the HBEC cell lines expressing different mutant proteins. DPE analysis was performed where HBEC GFP;p53^{CT}, HBEC EGFR^{L858R};p53^{wt}, and HBEC EGFR^{L858R};p53^{CT} were individually compared to HBEC GFP;p53^{wt} cells. In contrast to the malignant vs non-malignant comparison, only PFKP, FN1, SERPINA3, and SERPINB7 were found to be differentially expressed in the pre-malignant and non-transformed states (Supplementary Figures 6A–C). SERPINA3 and SERPINB7 were also differentially expressed in the malignant vs non-malignant comparison (Figure 2B). This suggests that their expression levels may change during multiple stages of transformation. As few differentially secreted proteins were identified through this analysis, it is possible that expression of cancer genes alone is insufficient to dramatically alter the secretome or that the pre-malignant stage of transformation may not be distinct from the histologically normal, untransformed stage in terms of secreted profiles.

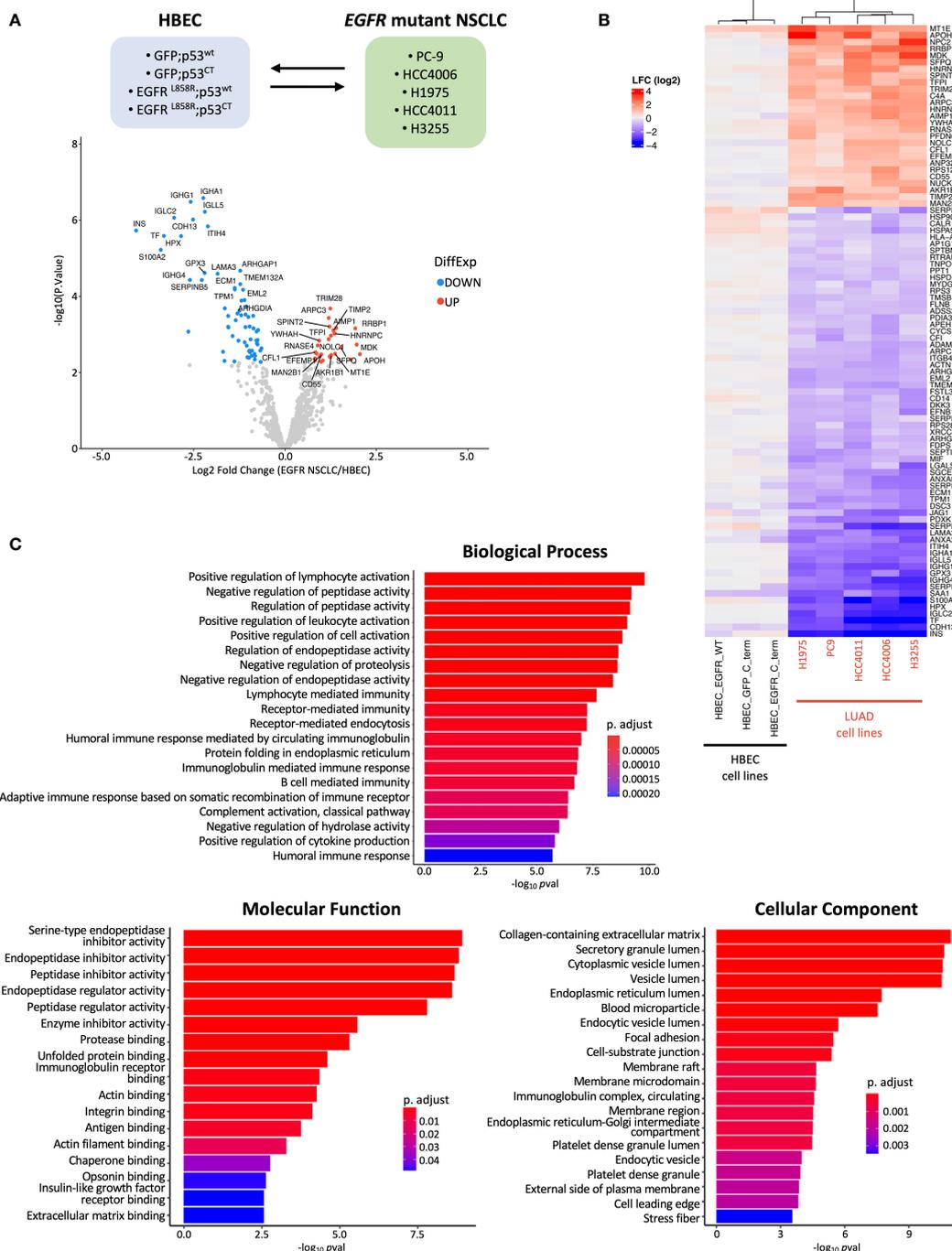


FIGURE 2 Groupwise comparison between HBEC cell lines and EGFR mutant NSCLC cell lines. **(A)** Schematic showing comparison performed during statistical analysis and volcano plot of log₂ fold change (LFC) differentially expressed proteins identified from MS/MS analysis. The top 20 significantly over- and under-expressed proteins ($p_{adj} < 0.05$ and absolute LFC > 0.6) are colored in red or blue, respectively, and labeled. **(B)** Heatmap of differentially expressed proteins identified from group-wise comparison in A., with hierarchical clustering ($n=91$) (cell line and protein, hierarchical clustering; cell line clustering distance, complete; protein clustering distance, average). **(C)** Bar plots showing the top 20 GO enrichment terms sorted by adjusted p-value ($p_{adj} < 0.05$, Benjamini-Hochberg adjustment) for biological processes, molecular functions, and cellular components for differentially expressed proteins found during MS/MS analysis.

Identification of secreted biomarker candidates specific to the malignant state

To identify protein candidates for further analysis and validation, we also performed DPE analysis of EGFR mutant

NSCLC cell lines individually against HBEC GFP;p53^{wt}, defined as the initial, untransformed stage (Figure 1A). This was done to capture cell line-specific differentially expressed proteins not observed in a group-wise comparison. Figure 3A outlines the filtering pipeline to identify potential protein biomarker

candidates using $p \text{ adj} < 0.05$ and $\text{LFC} > 0.6$. The number of differentially expressed proteins ranged from 119 to 316 per LUAD cell line, in contrast to 91 from the group-wise comparison (Supplementary Figure 7). Individual comparisons may highlight specific genetic alterations per cell line, as each cell line has varying *EGFR* and *TP53* mutations (72–74). Clustering based on differentially expressed proteins demonstrated that *EGFR* L858R driver mutation cell lines HCC4011 and H1975 grouped together, while *EGFR* exon19 deletion cell lines HCC4006 and PC-9 clustered together. The exception was H3255, an *EGFR* L858R mutant, which may be attributed to other differences in genetic alterations; an example is H1975 possessing *CDKN2A* and *PIK3CA* mutations that are not found in H3255 (74, 75). We then profiled the 499 differentially expressed proteins found across all LUAD cells compared to the non-transformed state for enrichment in GO terms associated with biological processes, molecular functions, and cellular compartments (Figure 3C). Similar to the group-wise comparison, GO cellular component terms were associated with the extracellular matrix, such as “collagen-containing extracellular matrix”, “laminin complex”, and “basement membrane” (60, 76). Also common were secretion associated terms, with “secretory granule lumen”, “cytoplasmic lumen”, and “vesicle lumen” comprising 3 of the top 5 cellular component terms (Figure 3D) (58, 66). Reactome analysis revealed enrichment in similar signaling pathways as the groupwise comparison, with the noted addition of *MET* related signaling as one of the most enriched pathways (Supplementary Figure 8).

Proteins identified from the individual cell line comparisons were filtered for further analysis (Figure 3A). Differentially expressed proteins were filtered for overlap in three or more *EGFR* mutant LUAD cell lines, and then for expression in the same direction. This resulted in 130 proteins for further investigation. As there was no relevant *EGFR* mutant LUAD proteomic dataset available, we aimed to assess whether the expression of the secreted proteins are specific to *EGFR* mutant LUADs using transcriptomic data. Differential gene expression analysis was performed on a cohort of 39 *EGFR* mutant and 154 *EGFR* wild-type tumors (45). This analysis revealed 16 genes differentially expressed between *EGFR* mutant and wild-type LUAD tumors with corresponding proteins that were differentially secreted in the *EGFR* mutant LUAD cell lines (Table 1, Figure 3D). *NPC2* demonstrated the highest level of expression, consistent with previous findings where *NPC2* expression was greater in LUAD compared to other lung tumor types (77, 78). *ENO1* and *RAC1* also showed high levels of expression in *EGFR* mutant LUADs (Figure 3D) aligning with previous studies demonstrating that *ENO1* expression is greater in LUAD tumor samples relative to non-cancerous tissue, and the *RAC1* splice variant *RAC1B* enhances LUAD tumor formation *in vivo* (79, 80). The analysis of our MS/MS results identified differentially expressed proteins from our secretome experiment with evidence that they may be useful candidate biomarkers in the clinical setting.

Secreted proteins with gene expression levels associated with poor outcome in *EGFR* mutant LUAD

The 16 protein candidates found to have *EGFR* specific expression levels in LUAD tumors were subsequently analyzed for survival difference between 125 *EGFR* mutant and 68 wild-type patient tumors in an independent dataset (Table 1) (48). Kaplan-Meier curves were plotted, where overall survival duration between patients with high and low expression of genes was compared based on *EGFR* mutation status (Figure 4). High expression of *ENO1* ($p < 0.001$), *PFKP* ($p < 0.05$), *RAC1* ($p < 0.001$), and *SPINT2* ($p < 0.05$) in patients with *EGFR* mutant tumors was associated with shorter overall survival than *EGFR* mutant tumors with low gene expression (Figures 4A–D). The association with poor survival was not seen in patients with *EGFR* wild-type tumors, suggesting that survival differences could be *EGFR* mutation-specific. High *MDK* expression was also associated with lower overall survival, in *EGFR* mutant and wild-type patient tumors ($p < 0.05$) (Figure 4E). This observation aligns with a previous study where NSCLC patients displayed increased protein expression of serum *MDK* compared to healthy individual controls, and expression was associated with lower overall survival (65).

Orthogonal validation of secreted proteins

To validate levels of secreted protein expression in the *EGFR* mutant LUAD cell lines, *MDK*, *GDF15*, and *SPINT2* were assessed by ELISA assays (Table 1, Figure 5). *MDK* and *SPINT2* were selected due to the observed differences in overall survival while *GDF15* levels have been associated with different stages of lung cancer in patients (81). Protein concentration was measured in conditioned media samples collected and concentrated in parallel with the MS/MS samples. As we were interested in comparing between malignant and untransformed, histologically normal states, we compared protein concentration between HBEC GFP;p53^{wt} and the selected LUAD *EGFR* mutant cell lines (Figure 5). The mean *MDK* concentration in LUAD cell lines ranged from 0.94 - 17.28 ng/mL, and when compared to HBEC GFP;p53^{wt}, concentrations were significantly different ($p < 0.01$; Figure 5A). *GDF15* mean protein concentration in NSCLC cell lines varied between 5.85 pg/mL - 6.44 ng/mL, compared to the mean HBEC GFP;p53^{wt} concentration of 2.41 pg/mL. This corresponded to an increase of secreted *GDF15* concentration in NSCLC cell lines up to 2000x that observed in HBEC GFP;p53^{wt} (PC-9, $p < 0.0001$; HCC4006, $p < 0.001$; H1975, $p < 0.001$; HCC4011, $p < 0.001$; H3255, $p < 0.0001$; Figure 5B). *SPINT2* mean protein concentration in NSCLC cell lines ranged between 98.30 - 580.97 pg/mL, relative to 98.58 pg/mL in HBEC GFP;p53^{wt}. With the exception of PC-9, protein concentrations were significantly greater than HBEC GFP;p53^{wt}, where concentrations were 1.7- 5.9 times greater (HCC4006, $p <$

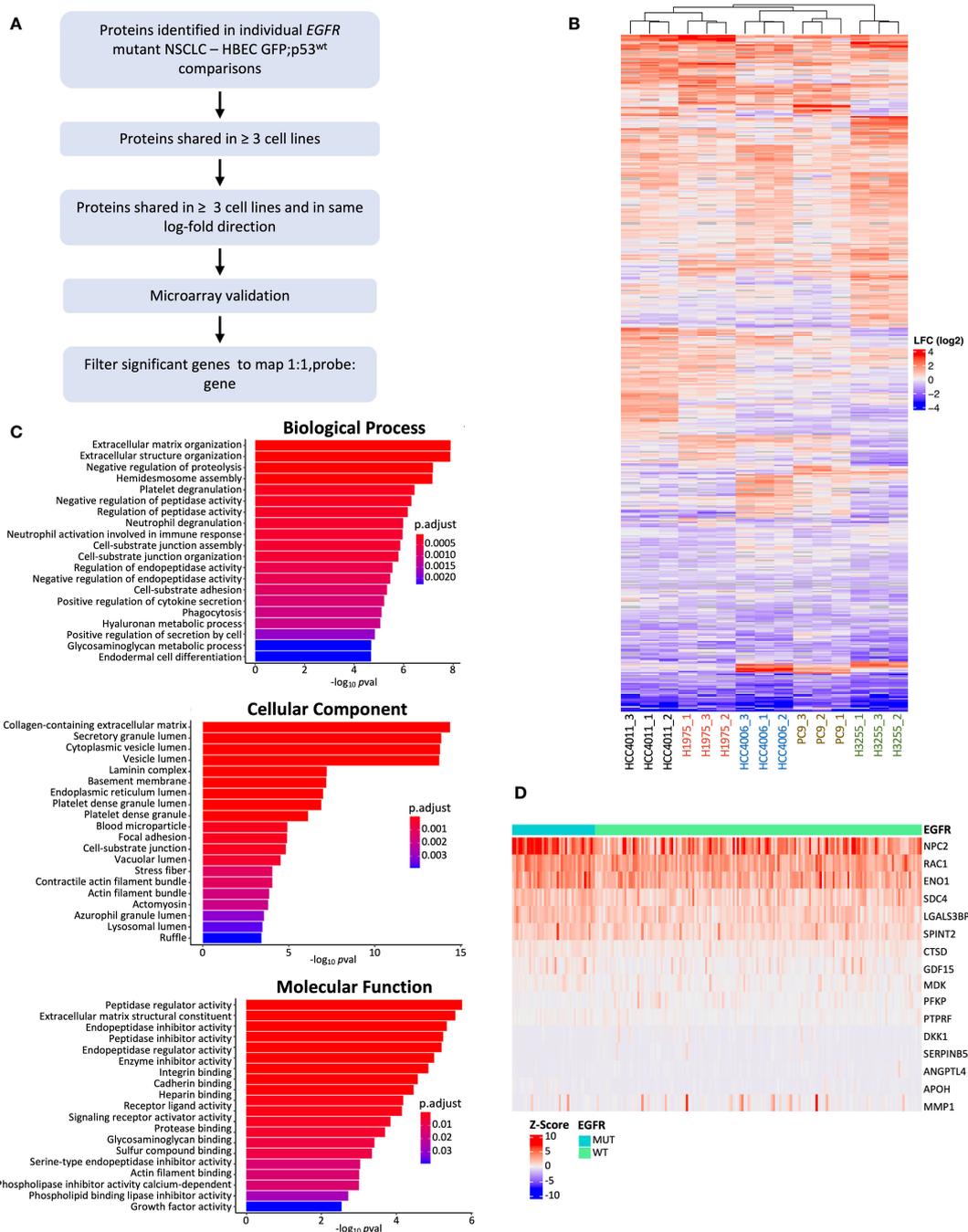


FIGURE 3

Individual *EGFR* mutant NSCLC cell line comparison against HBEC GFP;p53^{wt} and filtering pipeline. **(A)** Filtering pipeline used to identify protein candidates for biomarker validation. Differentially expressed proteins from cell line comparisons with HBEC GFP;p53^{wt} (p adj < 0.05, absolute LFC > 0.6) were determined, and filtered for microarray gene expression validation if found in 3 or more *EGFR* mutant cell line comparisons and LFC expression was in the same direction (LFC values were all positive or negative). Filtered differentially expressed proteins were analyzed for differential gene expression with Z-score normalized Affymetrix gene expression data from 199 primary LUAD tumors (45), and differentially expressed genes that were found to be significant (p adj < 0.05, Benjamini-Hochberg adjustment) were further filtered for optimal 1:1 probe to gene mapping for additional stringency (47). **(B)** Heatmap of differentially expressed genes found from individual comparison between *EGFR* mutant NSCLC cell lines and HBEC GFP;p53^{wt} (cell line and protein, hierarchical clustering; cell line clustering distance, complete; protein clustering distance, average). LFC ranges from high (red) to low (blue). **(C)** Bar plots showing the top 20 GO enrichment terms sorted by adjusted p-value (p adj < 0.05, Benjamini-Hochberg adjustment) for biological processes, molecular functions, and cellular components for differentially expressed proteins found during individual cell line comparisons between *EGFR* mutant NSCLC cell lines and HBEC GFP;p53^{wt}. **(D)** Heatmap of differentially expressed genes post-microarray analysis for further validation (p adj < 0.05). Samples are grouped by *EGFR* status (mutant, wild type) (gene clustering method, Euclidean; gene clustering distance, complete).

TABLE 1 Protein candidates identified from microarray gene expression analysis, listed by gene symbol, sorted by adjusted p value ($p_{adj} < 0.05$; Benjamini-Hochberg adjustment).

Accession	Gene	Peptides	Unique peptides	Quantified peptides	adj.P.Val
Q08380	LGALS3BP	23	23	20	9.5519E-05
P61916	NPC2	10	10	10	9.5519E-05
O43291	SPINT2	2	2	1	0.001714715
P06733	ENO1	30	28	27	0.003759408
P21741	MDK	12	12	11	0.003759408
P63000	RAC1	3	2	1	0.005590461
Q01813	PFKP	2	2	1	0.007683377
P31431	SDC4	7	7	5	0.012165019
Q99988	GDF15	6	6	4	0.020488018
P02749	APOH	2	2	2	0.02368962
P36952	SERPINB5	14	14	8	0.027201045
P10586	PTPRF	19	19	15	0.031704624
P07339	CTSD	20	20	20	0.034872449
O94907	DKK1	7	7	7	0.034872449
P03956	MMP1	2	2	2	0.035149978
Q9BY76	ANGPTL4	6	6	4	0.03699355

0.001; H1975, $p < 0.0001$; HCC4011, $p < 0.01$; H3255, $p < 0.01$; Figure 5C). Together, these assays confirm the MS results, validating the increased secretion in EGFR mutant LUAD.

Discussion

Earlier detection of LUAD is key to long-term patient survival, where LDCT screening could benefit from the inclusion of biomarkers to complement screening (12). This is especially important given that LDCT screening for never smokers, which have increased incidence of EGFR mutant LUAD, does not have concrete guidelines (20, 37). Compared to ever smokers, one study found that the rate of diagnosis with LDCT screening for a never smoker cohort was 0.45%, which was lower than the NLST ever smoker rate of 1.0% (18, 82). More recently, a LDCT screening study for primarily non-smoking Asian women, the demographic commonly associated with EGFR mutant LUAD, found an increase in cancer incidence in early-stage cancers (stages 0-I) yet no change in late-stage incidence (stages II-IV) (83). These findings suggested that additional cases identified by LDCT were attributed to over-diagnosis, and that LDCT would have limited use for populations affected by EGFR mutant LUAD, furthering the necessity of biomarker-based detection to inform clinical decisions (21, 83). Currently, NSCLC biomarker candidates are limited and are neither specific for lung cancer nor histology (23). The cancer secretome is a valuable resource to uncover prognostic and diagnostic biomarker candidates (25, 84).

To date, secretome studies focusing on NSCLC have primarily analyzed established cancer cell lines, identifying changes in

secreted proteins that provide further insight into processes such as tumor growth and metastasis (33, 85). However, cancer cell lines may not capture changes that occur during earlier stages of malignant transformation, resulting in missing potential biomarkers for earlier detection. Using an *in vitro* model representative of malignant transformation, this study analyzed the secretome of EGFR-driven LUAD malignant transformation *in vitro*. Our study began with generating a model and serum-free culture conditions compatible with mass spectrometry analysis to profile the secretome. Due to the lack of a transformation phenotype *in vitro* and *in vivo*, the HBEC cell line was selected to serve as an untransformed basal state to compare secretome changes during different transformation stages (38). By introducing common genetic alterations found in EGFR and p53 into HBECs, we could also profile secretome changes during an intermediate, pre-malignant stage of transformation (50, 86, 87). An additional benefit to the HBEC cell line was its ability to be cultured in non-serum conditions; serum proteins often obscure low abundance proteins and may introduce non-human contamination (33, 54).

Together with selected EGFR mutant NSCLC cell lines representing the transformed state, we generated a model that represents key genetic alterations occurring during EGFR-driven malignant transformation (Figures 1A, B) (9, 50). We initially identified 1020 proteins, where 852 proteins from secretome conditioned media were enriched. HSPG2, LAMA5, and AGRN were identified which is consistent with a previous NSCLC secretome study, suggesting that secreted proteins could be detected with our approach (33). We further validated the presence of secreted proteins by performing GO analyses, where

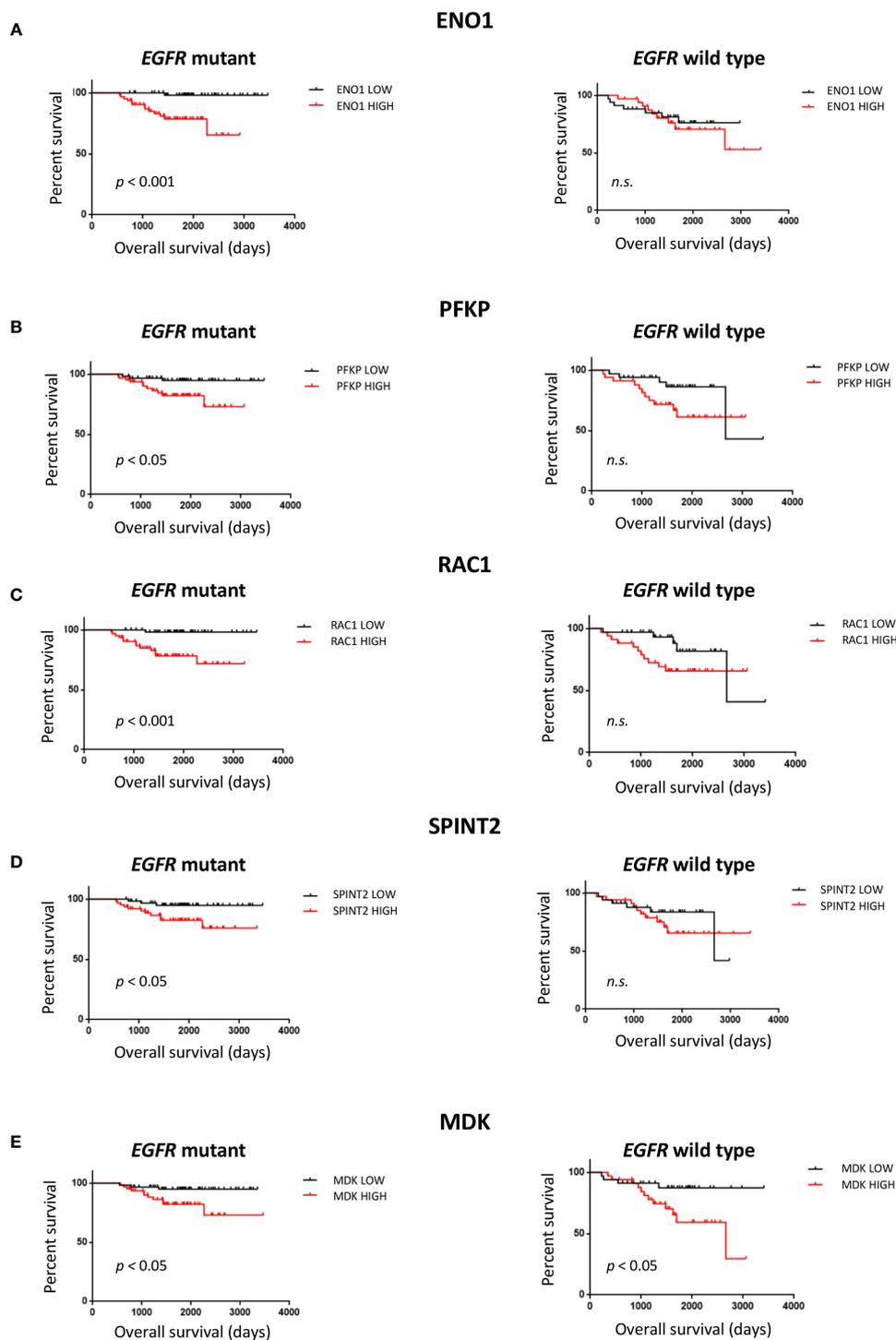
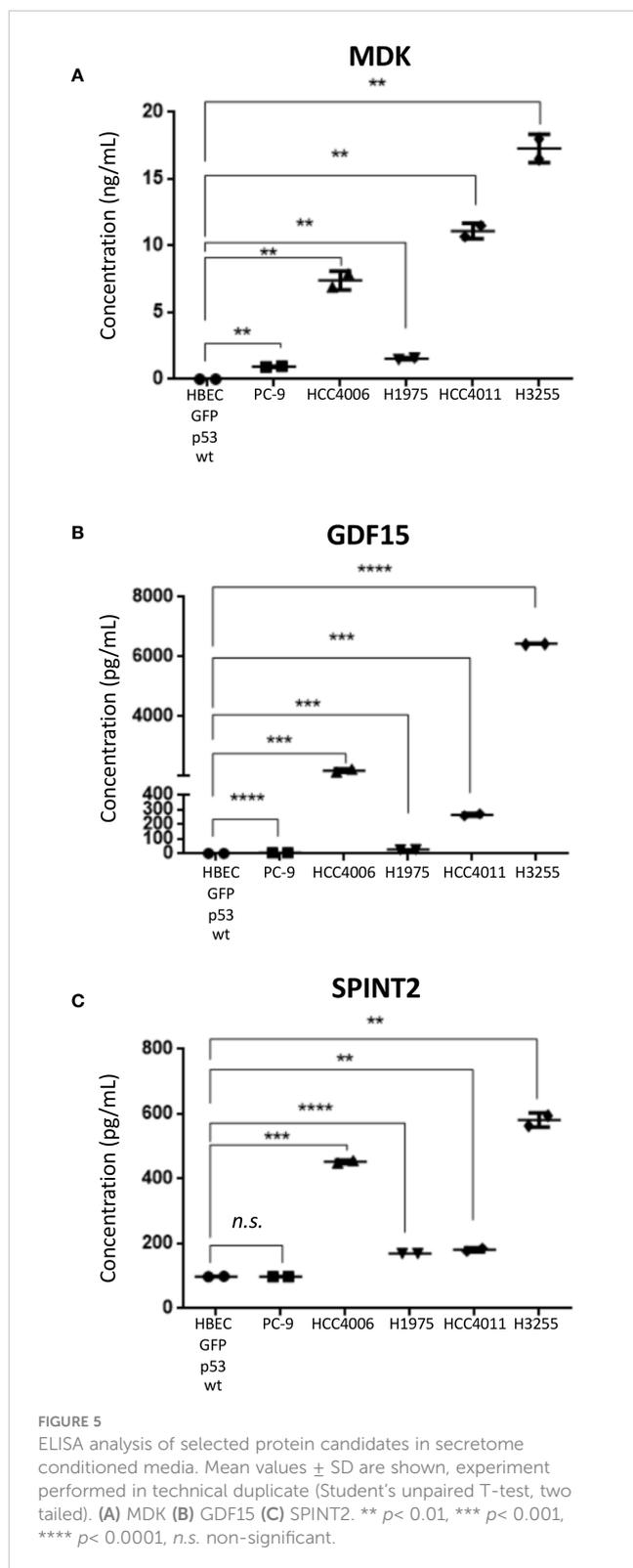


FIGURE 4

Survival analysis of protein candidates identified from filtering pipeline. Overall patient survival was analyzed (log rank test, median split) for *EGFR* mutant tumors, $n=125$ and *EGFR* wild type tumors, $n=68$ from the NCBI GEO GSE 31219 dataset. (A) ENO1. (B) PFKP (C) RAC1 (D) SPINT2 (E) MDK. *n.s.* non-significant.

cellular component terms referenced secretory pathways or locations associated with secretion (Figure 1D) (58, 66). We identified 91 differentially expressed proteins in *EGFR* mutant NSCLC cell lines relative to HBEC cell lines (Figures 2A, B) with secretome profiles of the transformed states including changes in

immune functions. This aligns with previous NSCLC studies where *EGFR* mutations were associated with immune changes such as increased expression of PD-1 and PD-L1, and decreased CD8+ T cell infiltration (Figure 2C) (88, 89). We were unable to identify notable differences when comparing the basal HBEC GFP;p53^{wt} cell



line to HBEC cell lines expressing *EGFR* and *p53* alterations (Supplementary Figure 5) suggesting the introduction of additional genetic alterations may be needed to establish a more advanced pre-malignant state (9).

To broaden the pool of potential candidates and identify the most notable differences between untransformed and transformed

states, we also compared *EGFR* mutant NSCLC cell lines individually to the basal HBEC cell line (Figure 3). 499 proteins across all comparisons were identified, with 130 proteins differentially expressed (Figures 3A, B). The secretome profile of the LUAD cell lines had a broad scope of biological functions (Figure 3C) which may be due to additional genetic changes specific to each NSCLC cell line, beyond *EGFR* and *p53*, such as *p16* that can affect the secretome profile (90, 91). Using transcriptome data, we found genes for 16 of these proteins which were expressed specifically in *EGFR* mutant LUAD, suggesting they may be biomarker candidates for this molecular subtype of lung cancer (Figure 3D) (65, 78). A subset of these were also associated with patient survival in *EGFR* mutant LUAD, and we validated MDK, GDF15, and SPINT2 by ELISA, confirming their secretion and association with the malignant state.

MDK is a growth factor that binds to heparin and is involved in promoting cell growth and survival *in vitro* and tumor growth *in vivo* in a model of LUAD (92). GDF15 is a member of the transforming growth factor- β superfamily and varying biological effects have been observed with expression changes (93). In one study, GDF15 overexpression suppressed cell proliferation *in vitro* and tumor formation *in vivo*, while in another study overexpression promoted tumor growth *in vivo*, and proliferation *in vitro* when stimulated with C5a (94, 95). SPINT2 is a serine protease inhibitor where decreased expression facilitated STYK1-mediated tumor progression (96). With the exception of the PC-9 cell line when measuring SPINT2 concentration, *EGFR* mutant NSCLC cell lines had significantly higher concentrations of the selected proteins than HBEC GFP;*p53*^{wt} (Figure 5). This suggests that there may be changes in MDK, GDF15, and SPINT2 expression during *EGFR*-driven malignant transformation that could be indicative of progression and studied for biomarker use (81, 97, 98).

While our study sought to identify changes in the secretome during LUAD transformation *in vitro*, there are limitations that should be considered for future studies. Firstly, the HBEC cell lines which represented the pre-malignant model stages were not validated for transformation capacity anchorage-independent growth *in vitro* or growth *in vivo* (99). As a result, this hampered the accuracy of the model when compared to clinical stepwise transformation and thus the accuracy of secretome changes occurring during the pre-malignant stage (50). Secondly, there may be additional genetic alterations that occur during *EGFR*-driven transformation. A previous study modeling transformation in HBECs found that *EGFR* and *TP53* mutations were unable to promote transformation *in vivo*, while another identified that alterations in *APC*, *RB1*, and *RBM10* promoted tumor growth *in vivo* (9, 100). Thirdly, limited incubation time under serum-free conditions can restrict the scope of secretome profiling, as secretome protein abundance has been observed to increase over time, despite minimizing cell death (35, 101, 102).

Conclusions

In summary, we have profiled the secretome of non-transformed and *EGFR* mutant transformed lung cells and

identified 3 protein candidates that were validated for differential expression in EGFR mutant patients. These proteins show promise as candidates for lung cancer biomarker applications, although further mechanistic and validation studies are needed. The data and findings shown provide an insight into secretome changes under a variety of conditions and will serve as a valuable resource to support future studies in LUAD biomarker discovery and molecular changes occurring during EGFR-driven malignant transformation.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: [ProteomeXchange/PXD045328].

Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

JL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. FJ: Investigation, Methodology, Writing – review & editing. JJ: Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – review & editing. TS: Investigation, Methodology, Writing – review & editing. RS: Investigation, Methodology, Writing – review & editing. DL: Investigation, Methodology, Writing – review & editing. DF: Investigation, Methodology, Writing – review & editing. SS: Investigation, Methodology, Project administration, Writing – review & editing. GL: Investigation, Methodology, Writing – review & editing. GM: Project administration, Supervision, Writing – review & editing. WL: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Canadian Institutes of Health Research (CIHR, MOP-142313, PJT – 169129 and PJT-148725) and BC Cancer Foundation to WL.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1286821/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Cell viability analysis of HBEC stable cell lines under secretome conditions. HBEC cell lines were seeded and incubated overnight, then media was changed to secretome media conditions (KSFM, 1% PenStrep) or standard culture conditions (KSFM, supplemented with BPE, EGF, and 1% PenStrep) for 24 hours, then cell viability was assessed with Trypan Blue and Propidium Iodide (PI) staining. **(A)** Quantification of HBEC live cell population, as stained with Trypan Blue. **(B)** Quantification of HBEC live cell population, as stained with PI. **(C)** Representative images of DAPI and PI channels used to quantify PI staining. Experiment was performed in biological triplicate. *n.s.* non-significant.

SUPPLEMENTARY FIGURE 2

Cell viability analysis of EGFR mutant NSCLC cell lines under secretome conditions. Cell lines were seeded and incubated overnight, then media was changed to secretome media conditions (RPMI, 1% PenStrep) or standard culture conditions (RPMI, supplemented with 10% FBS, and 1% PenStrep) for 24 hours, then cell viability was assessed with Trypan Blue and Propidium Iodide (PI) staining. **(A)** Quantification of live cell population, as stained with Trypan Blue. **(B)** Quantification of live cell population, as stained with PI. **(C)** Representative images of DAPI and PI channels used to quantify PI staining. Experiment was performed in biological triplicate. ** $p < 0.01$, *n.s.* non-significant.

SUPPLEMENTARY FIGURE 3

Secretome experiment PCA. PCA was performed on all proteins identified during MS/MS **(A)** PCA excluding media control samples. **(B)** PCA including media control samples.

SUPPLEMENTARY FIGURE 4

Reactome pathway analysis of candidate secreted proteins from all cell lines. Plot showing the top 20 pathways identified using the Curated.Reactome database (FDR < 0.05). Minimum pathway size was $n = 2$, maximum pathway size $n = 2000$.

SUPPLEMENTARY FIGURE 5

Groupwise Reactome analysis of secreted proteins between HBEC cell lines and EGFR mutant NSCLC cell lines. Plot showing the top 20 pathways identified using the Curated.Reactome database (FDR < 0.05). Minimum pathway size was $n = 2$, maximum pathway size $n = 2000$.

SUPPLEMENTARY FIGURE 6

Differential protein expression analysis of HBEC cell lines expressing EGFR L858R, with or without expression of p53 c-terminal, relative to HBEC GFP; p53^{wt} with different absolute LFC parameters to identify differentially expressed proteins. The top 20 significantly over- and under-expressed proteins (p adj < 0.05 and absolute LFC > 0.6) are colored in red or blue, respectively, and labeled. (A) HBEC GFP;p53^{CT} i) minimum absolute LFC < 0.6 ii) minimum absolute LFC < 0.3 (B) HBEC EGFR^{L858R};p53^{wt} i) minimum absolute LFC < 0.6 ii) minimum absolute LFC < 0.3 (C) HBEC EGFR^{L858R};p53^{CT} i) minimum absolute LFC < 0.6 ii) minimum absolute LFC < 0.3 (D) Venn diagram describing overlap in differentially expressed proteins among HBEC stable cell lines. "abs" = absolute.

SUPPLEMENTARY FIGURE 7

Differential protein expression analysis of EGFR mutant NSCLC cell lines, relative to HBEC GFP;p53^{wt}. The top 20 significantly over- and under-expressed proteins (p adj < 0.05 and absolute LFC > 0.6) are colored in red or blue, respectively, and labeled. (A) PC-9. (B) HCC4006. (C) H1975 (D) HCC4011 (E) H3255.

SUPPLEMENTARY FIGURE 8

Reactome pathway enrichment of proteins detected in individual EGFR mutant NSCLC cell lines vs HBEC GFP;p53^{wt}. Plot showing the top 20 pathways identified using the Curated.Reactome database (FDR < 0.05). Minimum pathway size was $n = 2$, maximum pathway size $n = 2000$.

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
- Herbst RS, Heymach JV, Lippman SM. Lung cancer. *New Engl J Med* (2008) 359(13):1367–80. doi: 10.1056/NEJMra0802714
- Arbour KC, Riely GJ. Systemic therapy for locally advanced and metastatic non-small cell lung cancer. *JAMA* (2019) 322(8):764. doi: 10.1001/jama.2019.11058
- Skoulidis F, Heymach JV. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. *Nat Rev Cancer* (2019) 19(9):495–509. doi: 10.1038/s41568-019-0179-8
- Takamochi K, Oh S, Suzuki K. Differences in EGFR and KRAS mutation spectra in lung adenocarcinoma of never and heavy smokers. *Oncol Lett* (2013) 6(5):1207–12. doi: 10.3892/ol.2013.1551
- Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature* (2018) 553(7689):446–54. doi: 10.1038/nature25183
- Larsen JE, Minna JD. Molecular biology of lung cancer: clinical implications. *Clinics Chest Med* (2011) 32(4):703–40. doi: 10.1016/j.ccm.2011.08.003
- Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, et al. The 2015 world health organization classification of lung tumors. *J Thorac Oncol* (2015) 10(9):1243–60. doi: 10.1097/JTO.0000000000000630
- Sato M, Vaughan MB, Girard L, Peyton M, Lee W, Shames DS, et al. Multiple Oncogenic Changes (K-RAS(V12), p53 Knockdown, Mutant EGFRs, p16 Bypass, Telomerase) Are Not Sufficient to Confer a Full Malignant Phenotype on Human Bronchial Epithelial Cells. *Cancer Res* (2006) 66(4):2116–28. doi: 10.1158/0008-5472.CAN-05-2521
- Sato M, Shay JW, Minna JD. Immortalized normal human lung epithelial cell models for studying lung cancer biology. *Respir Invest* (2020) 58(5):344–54. doi: 10.1016/j.resinv.2020.04.005
- Lemjabbar-Alaoui H, Hassan OU, Yang Y-W, Buchanan P. Lung cancer: Biology and treatment options. *Biochim Biophys Acta (BBA) - Rev Cancer* (2015) 1856(2):189–210. doi: 10.1016/j.bbcan.2015.08.002
- Mithoowani H, Febraro M. Non-small-cell lung cancer in 2022: A review for general practitioners in oncology. *Curr Oncol* (2022) 29(3):1828–39. doi: 10.3390/curoncol29030150
- Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WEE, et al. The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (Eighth) edition of the TNM classification for lung cancer. *J Thorac Oncol* (2016) 11(1):39–51. doi: 10.1016/j.jtho.2015.09.009
- Ganti AK, Klein AB, Cotarla I, Seal B, Chou E. Update of incidence, prevalence, survival, and initial treatment in patients with non-small cell lung cancer in the US. *JAMA Oncol* (2021) 7(12):1824. doi: 10.1001/jamaoncol.2021.4932
- Nooreldeen R, Bach H. Current and future development in lung cancer diagnosis. *Int J Mol Sci* (2021) 22(16):8661. doi: 10.3390/ijms22168661
- Tammemagi MC, Lam S. Screening for lung cancer using low dose computed tomography. *BMJ* (2014) 348(may27 7):g2253–3. doi: 10.1136/bmj.g2253
- Gierada DS, Black WC, Chiles C, Pinsky PF, Yankelevitz DF. Low-dose CT screening for lung cancer: evidence from 2 decades of study. *Radiology: Imaging Cancer* (2020) 2(2):e190058. doi: 10.1148/rycan.2020190058
- The National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomography screening. *New Engl J Med* (2011) 365(5):395–409. doi: 10.1056/NEJMoa1102873
- Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers — a different disease. *Nat Rev Cancer* (2007) 7(10):778–90. doi: 10.1038/nrc2190
- Lam S. Lung cancer screening in never-smokers. *J Thorac Oncol* (2019) 14(3):336–7. doi: 10.1016/j.jtho.2018.12.019
- Seijo LM, Peled N, Ajona D, Boeri M, Field JK, Sozzi G, et al. Biomarkers in lung cancer screening: achievements, promises, and challenges. *J Thorac Oncol* (2019) 14(3):343–57. doi: 10.1016/j.jtho.2018.11.023
- Sozzi G, Boeri M. Potential biomarkers for lung cancer screening. *Trans Lung Cancer Res* (2014) 3(3):139–48. doi: 10.3978/j.issn.2218-6751.2014.06.04
- Duffy MJ, O'Byrne K. Tissue and blood biomarkers in lung cancer: A review. *Adv Clin Chem* (2018) 86:1–21. doi: 10.1016/bs.acc.2018.05.001
- Uhlén M, Karlsson MJ, Hober A, Svensson A-S, Scheffel J, Kotol D, et al. The human secretome. *Sci Signaling* (2019) 12(609). doi: 10.1126/scisignal.aaz0274
- Xue H, Lu B, Lai M. The cancer secretome: a reservoir of biomarkers. *J Trans Med* (2008) 6(1):52. doi: 10.1186/1479-5876-6-52
- Madden EC, Gorman AM, Logue SE, Samali A. Tumour cell secretome in chemoresistance and tumour recurrence. *Trends Cancer* (2020) 6(6):489–505. doi: 10.1016/j.trecan.2020.02.020
- Ritchie S, Reed DA, Pereira BA, Timpson P. The cancer cell secretome drives cooperative manipulation of the tumour microenvironment to accelerate tumorigenesis. *Faculty Rev* (2021) 10. doi: 10.12703/r/10-4
- Hsiao Y-C, Chu LJ, Chen J-T, Yeh T-S, Yu J-S. Proteomic profiling of the cancer cell secretome: informing clinical research. *Expert Rev Proteomics* (2017) 14(9):737–56. doi: 10.1080/14789450.2017.1353913
- Robinson JL, Feizi A, Uhlén M, Nielsen J. A systematic investigation of the Malignant junctions and diagnostic potential of the cancer secretome. *Cell Rep* (2019) 26(10):2622–2635.e5. doi: 10.1016/j.celrep.2019.02.025
- Izbicka E, Streeper RT, Michalek JE, Loudon CL, Diaz A, Campos DR. Plasma biomarkers distinguish non-small cell lung cancer from asthma and differ in men and women. *Cancer Genomics Proteomics* (2012) 9(1):27–35.
- Yousef EM, Tahir MR, St-Pierre Y, Gaboury LA. MMP-9 expression varies according to molecular subtypes of breast cancer. *BMC Cancer* (2014) 14(1):609. doi: 10.1186/1471-2407-14-609
- Bosse K, Haneder S, Arlt C, Ihling CH, Seufferlein T, Sinz A. Mass spectrometry-based secretome analysis of non-small cell lung cancer cell lines. *PROTEOMICS* (2016) 16(21):2801–14. doi: 10.1002/pmic.201600297
- Hu R, Huffman KE, Chu M, Zhang Y, Minna JD, Yu Y. Quantitative secretomic analysis identifies extracellular protein factors that modulate the metastatic phenotype of non-small cell lung cancer. *J Proteome Res* (2016) 15(2):477–86. doi: 10.1021/acs.jproteome.5b00819
- Böttger F, Schaaïj-Visser TB, de Reus I, Piersma SR, Pham TV, Nagel R, et al. Proteome analysis of non-small cell lung cancer cell line secretomes and patient sputum reveals biofluid biomarker candidates for cisplatin response prediction. *J Proteomics* (2019) 196:106–19. doi: 10.1016/j.jprot.2019.01.018
- Brandi J, Manfredi M, Speziali G, Gosetti F, Marengo E, Ceconi D. Proteomic approaches to decipher cancer cell secretome. *Semin Cell Dev Biol* (2018) 78:93–101. doi: 10.1016/j.semcdb.2017.06.030
- Hynds RE, Frese KK, Pearce DR, Grönroos E, Dive C, Swanton C. Progress towards non-small-cell lung cancer models that represent clinical evolutionary trajectories. *Open Biol* (2021) 11(1). doi: 10.1098/rsob.200247
- Rudin CM, Avila-Tang E, Harris CC, Herman JG, Hirsch FR, Pao W, et al. Lung cancer in never smokers: molecular profiles and therapeutic implications. *Clin Cancer Res* (2009) 15(18):5646–61. doi: 10.1158/1078-0432.CCR-09-0377
- Ramirez RD, Sheridan S, Girard L, Sato M, Kim Y, Pollack J, et al. Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. *Cancer Res* (2004) 64(24):9027–34. doi: 10.1158/0008-5472.CAN-04-3703
- Rappsilber J, Ishihama Y, Mann M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Analytical Chem* (2003) 75(3):663–70. doi: 10.1021/ac026117i

40. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43(7):e47–7. doi: 10.1093/nar/gkv007
41. Thummuluri V, Almagro Armenteros JJ, Johansen AR, Nielsen H, Winther O. DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Res* (2022) 50(W1):W228–34. doi: 10.1093/nar/gkac278
42. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* (2021) 2(3):100141. doi: 10.1016/j.xinn.2021.100141
43. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* (2009) 4(8):1184–91. doi: 10.1038/nprot.2009.97
44. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* (2020) 36(8):2628–9. doi: 10.1093/bioinformatics/btz931
45. Chitale D, Gong Y, Taylor BS, Broderick S, Brennan C, Somwar R, et al. An integrated genomic analysis of lung cancer reveals loss of DUSP4 in EGFR-mutant tumors. *Oncogene* (2009) 28(31):2773–83. doi: 10.1038/onc.2009.135
46. Carlson M. *hgu133a.db*. Bioconductor (2017).
47. Li Q, Birkbak NJ, Györfy B, Szallasi Z, Eklund AC. Jset: selecting the optimal microarray probe set to represent a gene. *BMC Bioinf* (2011) 12(1):474. doi: 10.1186/1471-2105-12-474
48. Okayama H, Kohno T, Ishii Y, Shimada Y, Shiraishi K, Iwakawa R, et al. Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. *Cancer Res* (2012) 72(1):100–11. doi: 10.1158/0008-5472.CAN-11-1403
49. Carlson M. *hgu133plus2.db*. Bioconductor (2017).
50. Inamura K. Clinicopathological characteristics and mutations driving development of early lung adenocarcinoma: tumor initiation and progression. *Int J Mol Sci* (2018) 19(4):1259. doi: 10.3390/ijms19041259
51. Zhao Y, Aguilar A, Bernard D, Wang S. Small-molecule inhibitors of the MDM2–p53 protein–protein interaction (MDM2 inhibitors) in clinical trials for cancer treatment. *J Medicinal Chem* (2015) 58(3):1038–52. doi: 10.1021/jm501092z
52. Inoue Y, Nikolic A, Farnsworth D, Shi R, Johnson FD, Liu A, et al. Extracellular signal-regulated kinase mediates chromatin rewiring and lineage transformation in lung cancer. *ELife* (2021) 10. doi: 10.7554/eLife.66524
53. Kucab JE, Hollstein M, Arlt VM, Phillips DH. Nutlin-3a selects for cells harbouring TP53 mutations. *Int J Cancer* (2017) 140(4):877–87. doi: 10.1002/ijc.30504
54. Shin J, Rhim J, Kwon Y, Choi SY, Shin S, Ha C-W, et al. Comparative analysis of differentially secreted proteins in serum-free and serum-containing media by using BONCAT and pulsed SILAC. *Sci Rep* (2019) 9(1):3096. doi: 10.1038/s41598-019-39650-z
55. Gazdar AF, Minna JD. Deregulated EGFR signaling during lung cancer progression: mutations, amplicons, and autocrine loops. *Cancer Prev Res* (2008) 1(3):156–60. doi: 10.1158/1940-6207.CAPR-08-0080
56. Bonifacino JS, Glick BS. The mechanisms of vesicle budding and fusion. *Cell* (2004) 116(2):153–66. doi: 10.1016/S0092-8674(03)01079-1
57. Nickl W. Pathways of unconventional protein secretion. *Curr Opin Biotechnol* (2010) 21(5):621–6. doi: 10.1016/j.copbio.2010.06.004
58. Rabouille C. Pathways of unconventional protein secretion. *Trends Cell Biol* (2017) 27(3):230–40. doi: 10.1016/j.tcb.2016.11.007
59. Lacy P. Mechanisms of degranulation in neutrophils. *Allergy Asthma Clin Immunol* (2006) 2(3):98. doi: 10.1186/1710-1492-2-3-98
60. Yue B. Biology of the extracellular matrix. *J Glaucoma* (2014) 23:S20–3. doi: 10.1097/IJG.0000000000000108
61. Golebiewska EM, Poole AW. Platelet secretion: From haemostasis to wound healing and beyond. *Blood Rev* (2015) 29(3):153–62. doi: 10.1016/j.blre.2014.10.003
62. Kaszak I, Witkowska-Pilaszewicz O, Niewiadomska Z, Dworecka-Kaszak B, Ngosa Toka F, Jurka P. Role of cadherins in cancer—A review. *Int J Mol Sci* (2020) 21(20):7624. doi: 10.3390/ijms21207624
63. López-Otin C, Bond JS. Proteases: multifunctional enzymes in life and disease. *J Biol Chem* (2008) 283(45):30433–7. doi: 10.1074/jbc.R800035200
64. Taguchi A, Politi K, Pitteri SJ, Lockwood WW, Faça VM, Kelly-Spratt K, et al. Lung cancer signatures in plasma based on proteome profiling of mouse tumor models. *Cancer Cell* (2011) 20(3):289–99. doi: 10.1016/j.ccr.2011.08.007
65. Stern L, Mueller E, Bellon E, Reeh M, Grotelueschen R, Guengoer C, et al. Serum midkine as non-invasive biomarker for detection and prognosis of non-small cell lung cancer. *Sci Rep* (2021) 11(1):14616. doi: 10.1038/s41598-021-94272-8
66. Meldolesi J. Unconventional protein secretion dependent on two extracellular vesicles: exosomes and ectosomes. *Front Cell Dev Biol* (2022) 10:877344. doi: 10.3389/fcell.2022.877344
67. Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. *J Cell Sci* (2010) 123(24):4195–200. doi: 10.1242/jcs.023820
68. Rozario T, DeSimone DW. The extracellular matrix in development and morphogenesis: A dynamic view. *Dev Biol* (2010) 341(1):126–40. doi: 10.1016/j.ydbio.2009.10.026
69. Werb Z. ECM and cell surface proteolysis: regulating cellular ecology. *Cell* (1997) 91(4):439–42. doi: 10.1016/S0092-8674(00)80429-8
70. Azuma K, Ota K, Kawahara A, Hattori S, Iwama E, Harada T, et al. Association of PD-L1 overexpression with activating EGFR mutations in surgically resected nonsmall-cell lung cancer. *Ann Oncol* (2014) 25(10):1935–40. doi: 10.1093/annonc/mdl242
71. Chen N, Fang W, Zhan J, Hong S, Tang Y, Kang S, et al. Upregulation of PD-L1 by EGFR activation mediates the immune escape in EGFR-driven NSCLC: implication for optional immune targeted therapy for NSCLC patients with EGFR mutation. *J Thorac Oncol* (2015) 10(6):910–23. doi: 10.1097/JTO.0000000000000500
72. Kubo T, Yamamoto H, Lockwood WW, Valencia I, Soh J, Peyton M, et al. MET gene amplification or EGFR mutation activate MET in lung cancers untreated with EGFR tyrosine kinase inhibitors. *Int J Cancer* (2009) 124(8):1778–84. doi: 10.1002/ijc.24150
73. Pikor LA, Lockwood WW, Thu KL, Vucic EA, Chari R, Gazdar AF, et al. YEATS4 is a novel oncogene amplified in non-small cell lung cancer that regulates the p53 pathway. *Cancer Res* (2013) 73(24):7301–12. doi: 10.1158/0008-5472.CAN-13-1897
74. Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald ER, et al. Next-generation characterization of the cancer cell line encyclopedia. *Nature* (2019) 569(7757):503–8. doi: 10.1038/s41586-019-1186-3
75. Heavey S, Cuffe S, Finn S, Young V, Ryan R, Nicholson S, et al. In pursuit of synergy: An investigation of the PI3K/mTOR/MEK co-targeted inhibition strategy in NSCLC. *Oncotarget* (2016) 7(48):79526–43. doi: 10.18632/oncotarget.12755
76. Jayadev R, Sherwood DR. Basement membranes. *Curr Biol* (2017) 27(6):R207–11. doi: 10.1016/j.cub.2017.02.006
77. Permemalm M, De Petris L, Eriksson H, Brandén E, Koyi H, Kanter L, et al. Use of narrow-range peptide IEF to improve detection of lung adenocarcinoma markers in plasma and pleural effusion. *PROTEOMICS* (2009) 9(13):3414–24. doi: 10.1002/pmic.200800814
78. Shoshan-Barmatz V, Bishitz Y, Paul A, Krelin Y, Nakdimon I, Peled N, et al. A molecular signature of lung cancer: potential biomarkers for adenocarcinoma and squamous cell carcinoma. *Oncotarget* (2017) 8(62):105492–509. doi: 10.18632/oncotarget.22298
79. Zhou C, Licciulli S, Avila JL, Cho M, Troutman S, Jiang P, et al. The Rac1 splice form Rac1b promotes K-ras-induced lung tumorigenesis. *Oncogene* (2013) 32(7):903–9. doi: 10.1038/onc.2012.99
80. Fu Q-F, Liu Y, Fan Y, Hua S-N, Qu H-Y, Dong S-W, et al. Alpha-enolase promotes cell glycolysis, growth, migration, and invasion in non-small cell lung cancer through FAK-mediated PI3K/AKT pathway. *J Hematol Oncol* (2015) 8(1):22. doi: 10.1186/s13045-015-0117-5
81. Deng J, Zhang M, Zhang H, Lu C, Hou G, Feng Y, et al. Value of growth/differentiation factor 15 in diagnosis and the evaluation of chemotherapeutic response in lung cancer. *Clin Ther* (2021) 43(4):747–59. doi: 10.1016/j.clinthera.2021.02.001
82. Kang H-R, Cho JY, Lee SH, Lee YJ, Park JS, Cho Y-J, et al. Role of low-dose computerized tomography in lung cancer screening among never-smokers. *J Thorac Oncol* (2019) 14(3):436–44. doi: 10.1016/j.jtho.2018.11.002
83. Gao W, Wen CP, Wu A, Welch HG. Association of computed tomographic screening promotion with lung cancer overdiagnosis among Asian women. *JAMA Internal Med* (2022) 182(3):283. doi: 10.1001/jamainternmed.2021.7769
84. Makridakis M, Vlahou A. Secretome proteomics for discovery of cancer biomarkers. *J Proteomics* (2010) 73(12):2291–305. doi: 10.1016/j.jpro.2010.07.001
85. Hartwig T, Montinaro A, von Karstedt S, Sevko A, Surinova S, Chakravarthy A, et al. The TRAIL-induced cancer secretome promotes a tumor-supportive immune microenvironment via CCR2. *Mol Cell* (2017) 65(4):730–742.e5. doi: 10.1016/j.molcel.2017.01.021
86. Soh J, Toyooka S, Ichihara S, Asano H, Kobayashi N, Suehisa H, et al. Sequential molecular changes during multistage pathogenesis of small peripheral adenocarcinomas of the lung. *J Thorac Oncol* (2008) 3(4):340–7. doi: 10.1097/JTO.0b013e318168d20a
87. Siegelin MD, Borczuk AC. Epidermal growth factor receptor mutations in lung adenocarcinoma. *Lab Invest* (2014) 94(2):129–37. doi: 10.1038/labinvest.2013.147
88. Akbay EA, Koyama S, Carretero J, Altabel A, Tchaicha JH, Christensen CL, et al. Activation of the PD-1 pathway contributes to immune escape in EGFR-driven lung tumors. *Cancer Discovery* (2013) 3(12):1355–63. doi: 10.1158/2159-8290.CD-13-0310
89. Le X, Negrao MV, Reuben A, Federico L, Diao L, McGrail D, et al. Characterization of the immune landscape of EGFR-mutant NSCLC identifies CD73/adenosine pathway as a potential therapeutic target. *J Thorac Oncol* (2021) 16(4):583–600. doi: 10.1016/j.jtho.2020.12.010
90. Tam KW, Zhang W, Soh J, Stastny V, Chen M, Sun H, et al. CDKN2A/p16 inactivation mechanisms and their relationship to smoke exposure and molecular features in non-small-cell lung cancer. *J Thorac Oncol* (2013) 8(11):1378–88. doi: 10.1097/JTO.0b013e3182a46c0c
91. Buj R, Leon KE, Anguelov MA, Aird KM. Suppression of p16 alleviates the senescence-associated secretory phenotype. *Aging* (2021) 13(3):3290–312. doi: 10.18632/aging.202640
92. Hao H, Maeda Y, Fukazawa T, Yamatsuji T, Takaoka M, Bao X-H, et al. Inhibition of the growth factor MDK/midkine by a novel small molecule compound to

treat non-small cell lung cancer. *PLoS One* (2013) 8(8):e71093. doi: 10.1371/journal.pone.0071093

93. Wischhusen J, Melero I, Fridman WH. Growth/differentiation factor-15 (GDF-15): from biomarker to novel targetable immune checkpoint. *Front Immunol* (2020) 11:951. doi: 10.3389/fimmu.2020.00951

94. Lu X, He X, Su J, Wang J, Liu X, Xu K, et al. EZH2-mediated epigenetic suppression of GDF15 predicts a poor prognosis and regulates cell proliferation in non-small-cell lung cancer. *Mol Ther - Nucleic Acids* (2018) 12:309–18. doi: 10.1016/j.omtn.2018.05.016

95. Zhao C, Li Y, Qiu W, He F, Zhang W, Zhao D, et al. C5a induces A549 cell proliferation of non-small cell lung cancer via GDF15 gene activation mediated by GCN5-dependent KLF5 acetylation. *Oncogene* (2018) 37(35):4821–37. doi: 10.1038/s41388-018-0298-9

96. Ma Z, Liu D, Li W, Di S, Zhang Z, Zhang J, et al. STYK1 promotes tumor growth and metastasis by reducing SPINT2/HAI-2 expression in non-small cell lung cancer. *Cell Death Dis* (2019) 10(6):435. doi: 10.1038/s41419-019-1659-1

97. Pereira MSCRV. *Study of epigenetic silencing of the tumor suppressor gene SPINT2 in lung cancer*. Integrated Master's Degree in Medicine, University of Porto, Portugal (2019). Available at: <https://repositorio-aberto.up.pt/handle/10216/121546>.

98. Filippou PS, Karagiannis GS, Constantinidou A. Midkine (MDK) growth factor: a key player in cancer progression and a promising therapeutic target. *Oncogene* (2020) 39(10):2040–54. doi: 10.1038/s41388-019-1124-8

99. Ahuja D, Sáenz-Robles MT, Pipas JM. SV40 large T antigen targets multiple cellular pathways to elicit cellular transformation. *Oncogene* (2005) 24(52):7729–45. doi: 10.1038/sj.onc.1209046

100. Foggetti G, Li C, Cai H, Hellyer JA, Lin W-Y, Ayeni D, et al. Genetic determinants of EGFR-driven lung cancer growth and therapeutic response *in vivo*. *Cancer Discovery* (2021) 11(7):1736–53. doi: 10.1158/2159-8290.CD-20-1385

101. Hammond DE, Kumar JD, Raymond L, Simpson DM, Beynon RJ, Dockray GJ, et al. Stable isotope dynamic labeling of secretomes (SIDLS) identifies authentic secretory proteins released by cancer and stromal cells. *Mol Cell Proteomics* (2018) 17(9):1837–49. doi: 10.1074/mcp.TIR117.000516

102. Bajaj R, Rodriguez BL, Russell WK, Warner AN, Diao L, Wang J, et al. Impad1 and Syt11 work in an epistatic pathway that regulates EMT-mediated vesicular trafficking to drive lung cancer invasion and metastasis. *Cell Rep* (2022) 40(13):111429. doi: 10.1016/j.celrep.2022.111429