Tobacco chemical-induced mouse lung adenocarcinoma cell lines pin the prolactin orthologue proliferin as a lung tumour promoter

Nikolaos I. Kanellakis[1,5], Anastasios D. Giannou[1,5], Mario A. Pepe[2,5], Theodora Agalioti[1], Dimitra E. Zazara[1], Ioanna Giopanou[1], Ioannis Psallidas[1], Magda Spella[1], Antonia Marazioti[1], Kristina A. M. Arendt[2], Anne-Sophie Lamort[2], Spyridon Champeris Tsaniras[3], Stavros Taraviras[3], Helen Papadaki[4], Ioannis Lilis[1,5], and Georgios T. Stathopoulos[1,2,5]

[1] Laboratory for Molecular Respiratory Carcinogenesis, Department of Physiology, Faculty of Medicine; University of Patras; Rio, Achaia, 26504; Greece.

[2] Comprehensive Pneumology Center (CPC) and Institute for Lung Biology and Disease (iLBD); Ludwig-Maximilians University and Helmholtz Center Munich, Member of the German Center for Lung Research (DZL); Munich, Bavaria, 81377; Germany.

[3] Stem Cell Biology Laboratory, Department of Physiology, Faculty of Medicine; University of Patras; Rio, Achaia, 26504; Greece.

[4] Department of Anatomy, Faculty of Medicine; University of Patras; Rio, Achaia, 26504; Greece.

[5] Equal co-authors

Corresponding author: Georgios T. Stathopoulos, MD PhD. Faculty of Medicine, University of Patras, Biomedical Sciences Research Building, 2nd floor, Room B40, 1 Asklepiou Str., University Campus, 26504 Rio, Greece. Email: gstathop@upatras.gr. Telephone: +30-2610-969154 Fax: +30-2610-969176

Abstract

Lung adenocarcinoma (LADC) is the leading cause of cancer death worldwide. Nevertheless, syngeneic mouse models of the disease are sparse, and cell lines suitable for transplantable and immunocompetent mouse models of LADC remain unmet needs. We established multiple mouse LADC cell lines by repeatedly exposing two mouse strains (*FVB, Balb/c*) to the tobacco carcinogens urethane or diethylnitrosamine and by culturing out the resulting lung tumors for prolonged periods of time. Characterization of the resulting cell lines ($n = 7$) showed that they were immortal and phenotypically stable *in vitro*, and oncogenic, metastatic, and lethal *in vivo*. The primary tumors that gave rise to the cell lines, as well as secondary tumors generated by transplantation of the cell lines, displayed typical LADC features, such as glandular architecture and mucin and thyroid transcription factor 1 expression. Moreover, these cells exhibited marked molecular similarity with human smokers' LADC, including carcinogen-specific *Kras* point mutations ($Kras^{Q61R}$ in urethane- and $Kras^{Q61H}$ in diethylnitrosamine-triggered cell lines) and *Trp53* deletions and displayed stemness features. Interestingly, all cell lines overexpressed proliferin, a murine prolactin orthologue, which functioned as a lung tumour promoter. Furthermore, prolactin was overexpressed and portended poor prognosis in human lung adenocarcinoma. In conclusion, we report the first LADC cell lines derived from mice exposed to tobacco carcinogens. These cells closely resemble human LADC and provide a valuable tool for the functional investigation of the pathobiology of the disease.

Keywords

Lung adenocarcinoma; tobacco chemical; carcinogen; KRAS; TRP53.

Summary

Multiple cell lines were established from lung adenocarcinomas that developed in mice exposed to tobacco carcinogens. Their characterization showed high similarity to human lung adenocarcinoma, indicating they can serve as faithful models of the disease.

Introduction

Lung cancer is the leading cause of cancer death worldwide accounting for 1.6 million deaths in 2012, including 270,000 in the European Union and 170,000 in the United States, with lung adenocarcinoma (LADC) accounting for half of the cases (1,2). Lung cancer is mainly caused by chemical carcinogens of tobacco smoke (3-5). Smoking-induced carcinomas including LADC bear thousands of mutations per genome, including gain-of-function point substitutions in critical codons of the *KRAS* proto-oncogene and deletion or loss-of-function point substitutions of the tumour suppressor *TRP53*, encountered respectively in 20-40% and 80-90% of LADC (6-9). There is an unmet need for relevant mouse models of smoking-induced carcinomas such as LADC, which cannot be fully recapitulated by genetic models (10). Such vehicles would aid in the identification of new lung cancer genes, in the distinction of true cancer drivers from passenger events, and in the development of new therapies.

To more comprehensively understand lung cancer initiation, evolution, and signaling, faithful mouse models of the disease are invaluable (11). Although elaborate genetic mouse models of lung cancer are available, they do not fully recapitulate smoke-induced carcinogenesis, since they are based on a single or a few transgenes that are turned on artificially (10,12,13) and tumours often regress after transgenes are turned-off (14). Importantly, evidence suggests that mouse models of tobacco carcinogen-induced LADC are closely related to the human disease (10). In specific, urethane promotes LADC development through induction of *KRAS* and other oncogene mutations that are also found in human LADC (10,15). However, existing tobacco carcinogen-triggered mouse lung tumour models are not really thought to be

malignant, despite that these lung tumours feature high similarities to human lung cancer (16,17). Therefore truly malignant and transplantable mouse tobacco-carcinogen-derived lung cancer cell lines do not exist. Such cells are invaluable as they could be used in immunocompetent mice to faithfully recapitulate human lung cancer on a background of full tumour-host interactions (13).

Proliferins (PRL), also known as mitogen regulated proteins, are four murine glycoprotein orthologues of human prolactin (18-22). Proliferins are highly expressed in the murine placenta during embryogenesis, as well as in highly proliferative adult mouse tissues such as skin hair follicles and small intestinal crypts, where they function to drive cellular proliferation, angiogenesis, and wound healing (18,23,24). Proliferin expression levels have been correlated with fibrosarcoma progression in mice (24), however their role in LADC remains unknown.

Here we report the establishment of a battery of murine cell lines derived from LADC of two inbred mouse strains following exposure to two different tobacco carcinogens and we show that they are malignant. These cell lines, faithful mouse models of human smoking-induced LADC, unveiled an unexpected LADC-promoting role for proliferin in mice and for its orthologue prolactin in humans.

Materials and methods

Ethics approval

Experiments were carefully designed and approved a priori by the Veterinary Administration of the Prefecture of Western Greece (approval protocol numbers 3741/16.11.2010, 60291/3035/19.03.2012, and 118018/578/ 30.04.2014), and were conducted according to Directive 2010/63/EU (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri= OJ:L:2010:276:0033:0079:EN:PDF).

Murine cell lines used and authentication method

The murine cancer cell lines used were Lewis lung carcinoma (LLC), B16F10 skin melanoma, and PANO2 pancreatic adenocarcinoma (all from the National Cancer Institute Tumour Depository; Frederick, MD), as well as MC38 colon adenocarcinoma cells (obtained from Dr. Timothy Blackwell, Vanderbilt University, Nashville, TN) and AE17 pleural mesothelioma cells (obtained from Dr. Timothy Blackwell, Vanderbilt University, Nashville, TN). All cell lines have been previously described in detail (25). NIH 3T3 cells were from the American Type Culture Collection (Manassas, VA). Cells were cultured at 37^{o}C in 5% CO_2-95% air using DMEM 10% FBS, 2 mM L-glutamine, 1 mM pyruvate, 100 U/ml penicillin, and 100 mg/ml streptomycin and were tested biannually for identity by short tandem repeats (STR) and *Mycoplasma Spp.* by PCR.

Derivation of mouse lung adenocarcinoma cell lines

Ten months post first carcinogen (urethane, ethyl carbamate, EC, $CH_3CH_2OCNH_2$, CAS #51-79-6; diethylnitrosamine, N,N-Diethylnitrous amide, DEN, $C_4H_{10}N_2O$, CAS # 55-18-5) exposure mice were sacrificed, lung tumours were dissected from surrounding healthy lung parenchyma under sterile conditions, were halved, one half was processed for histology, and the other half was chopped into 1 mm pieces and seeded to cell culture dishes. Cells were cultured under standard conditions outlined in the Supplement. When adenocarcinoma was diagnosed for a given tumour, its corresponding culture was passaged *in vitro* over a period of 18 months and 60 passages, whichever occurred first. All mouse LADC cell lines were deposited at the Laboratory for Molecular Respiratory Carcinogenesis cell line facility (http://www.lmrc.upatras.gr) and are available upon request (lmrc@upatras.gr).

Availability of data

Microarray data are publicly available at GEO Datasets (https://www.ncbi.nlm.nih.gov/gds/) using accession IDs GSE94981 (LADC, lungs, airway epithelial cells, mast cells, and macrophages), GSE82154 (ATII cells), GSE58188 (other cancer cells including LLC, MC38, AE17, B16F10, and PANO2 cells), and GSE43458 (BATTLE trial).

Results

Novel mouse lung adenocarcinoma cell lines generated by exposure of inbred mouse strains to tobacco carcinogens

To develop murine LADC cell lines, we repeatedly exposed *FVB* and *Balb/c* mice to the tobacco carcinogens urethane (ethyl carbamate, EC) and diethylnitrosamine (N-nitrosodiethylamine, DEN). For this, mice received repetitive intraperitoneal EC (1 g/Kg) or DEN (200 mg/Kg) injections and were observed for prolonged periods of time in order for true LADC to develop (Figure 1A). Indeed, mice developed large tumours that were harvested under sterile conditions and bisected, one half of the tumours was always used for histologic examination and the other half minced for long-term culture under standard conditions (Figure 1B). Histology revealed that some tumours were LADC showing abundant mitoses, invasion of adjacent lung structures, and necrosis (Figures 1C, 1D). Cells from these tumours were cultured for a period of over 18 months and/or 60 passages, such that only truly malignant cells survived. This simple method has been shown not to introduce new artificial mutations that are not present pre-culture (26). The resulting LADC cell lines (n = 7) were named XYLA# with X signifying the mouse strain (F, *FVB*; B, *Balb/c*), Y the carcinogen used (U, EC; E, DEN), LA lung adenocarcinoma, and # their serial number by derivation date. All cell lines were immortal, phenotypically stable, and indefinitely proliferative *in vitro* where they displayed spindle shapes and anoikis (Figures 1E, 1F). In addition, all cell lines exhibited nuclear atypia and stemness as they were able to form tumour-spheres *in vitro*, a capacity unique to stem and cancer cells (Figures 1G-1I). Remarkably, upon subcutaneous delivery of one million

cells/mouse to syngeneic mice, LADC cell lines were able to form primary solid tumours at the injection site, as well as spontaneous pulmonary metastases (Figures 1J-1L). Intravenous delivery of 250,000 LADC cells to syngeneic mice caused lung metastases as well, and intrapleural injection of 150,000 LADC cells to syngeneic mice triggered malignant pleural effusions (Figures 1M-1O). All LADC cell lines were uniformly lethal regardless of injection route, confirming their malignant nature (Figure 1P). The primary tumors that gave rise to the cell lines, as well as secondary tumors generated by transplantation of the cell lines, all displayed typical LADC features, such as glandular architecture and mucin and thyroid transcription factor 1 expression (Figure 2). These results firmly support that the cell lines derived from lung tumours of tobacco carcinogen-treated mice are true LADC cells that can recapitulate the metastatic patterns of human LADC.

Tobacco chemical-induced mouse lung adenocarcinoma cell lines harbor *Kras* mutations and *Trp53* loss

Activating *KRAS* mutations and loss or mutation of *TRP53* are common in human LADC of smokers (7,27,28). We hence sought to determine whether our LADC cell lines were similar to human LADC in terms of *Kras* and *Trp53* status (Figure 3). Reverse transcription PCR (RT-PCR) followed by direct cDNA sequencing of *Kras* (target) and *Nras* (control) transcripts revealed the presence of heterozygous $Kras^{Q61R}$ mutations in all EC-induced cell lines and heterozygous $Kras^{Q61H}$ mutations in all DEN-induced cell lines, but no *Nras* mutations (Figures 3A, 3B and supplementary figure SF1). Interestingly, all $Kras^{Q61R}$-mutant cell lines generated using EC also expressed a nonsense-mediated decay (NMD) transcript, in addition

to the mutant and wild-type (^{WT}) transcripts (Figure 3A and supplementary figure SF1), a mechanism thought to prevent the expression of mutant proteins (29). To assess the *Trp53* status of our carcinogen-induced LADC cell lines, they were cross-examined with mouse tracheal epithelial cells obtained over a time course post-EC exposure and with two non-carcinogen-derived mouse cancer cell lines with defined *Kras* and *Trp53* status: Lewis lung carcinoma (LLC) cells with mutant *Kras*^{G12C} and *Trp53*^{WT} and MC38 colon adenocarcinoma cells with mutant *Kras*^{G13R} and mutant *Trp53*^{R178P} (25,30). RT-PCR and quantitative real-time PCR (qPCR) showed different patterns of mono- or bi-allelic *Trp53* loss (Figures 3C and 3D), while Western immunoblots and immunocytochemistry did not detect *Trp53* mutations (Figure 3E and supplementary figure SF2). Diverse *EGFR* expression patterns were determined via qPCR and Western immunoblots that did not correlate with carcinogen or mouse strain employed (Figures 3D and 3E), while Sanger sequencing yielded *Egfr*^{WT} in all cell lines. Notably these results show that our murine tobacco carcinogen-triggered LADC cell lines bear *Kras*^{MUT} alleles and exhibit patterns of *Trp53* loss that resemble to the human LADC of smokers. Moreover and in accord with a comprehensive genomic screen of carcinogen-induced LADC (10), the data indicate that each tobacco carcinogen inflicts a defined *KRAS* point mutation (Single Nucleotide Variation; SNV): EC causes *Kras*^{Q61R} and DEN *Kras*^{Q61H} mutation. Finally, the data identify for the first time mutant *KRAS*-associated NMD in our EC-generated cell lines with *Kras*^{Q61R} mutations, which together with codon bias can explain the notorious absence of mutant *KRAS* reads in RNA sequencing studies (29,31).

Tobacco chemical-induced mouse lung adenocarcinoma cell lines overexpress stemness and cancer genes

Global gene expression analysis of our LADC cells in comparison with total lung RNA from naïve mice and various other cell types [Gene Expression Omnibus (GEO) Datasets accession IDs GSE94981 for LADC cell lines, lungs, airway epithelial cells, mast cells, and macrophages; GSE82154 for alveolar epithelial type II cells (ATII cells) and GSE58188 for other cancer cells; freely available at https://www.ncbi.nlm.nih.gov/gds/] identified a distinct transcriptomic pattern of LADC cells, including differential expression of a 43 gene-signature that comprised several cancer and stemness genes, such as *Itga2* (32) and PRL transcripts *Prl2c2/Prl2c3/Prl2c4.* Gene expression was validated by qPCR (Figures 4A-4C). Similar gene expression analyses comparing LADC cell lines only to mouse lungs revealed a broader transcriptomic signature of the LADC cells, which indicated significant perturbation of pathways significant for cancer cells, such as DNA replication, cell cycle, and purine metabolism pathways (Figures 4D, 4E). To examine the stemness of LADC cells, we determined RNA and protein levels of the lung and cancer stemness markers *Lgr6* and *Itgb3* (33,34). Remarkably, LADC cells displayed significant expression levels of both genes (Figures 5A-5C and Supplementary Figures SF3, SF4), indicating a prominent cancer stemness potency in line with their tumour sphere forming capacity.

The prolactin orthologue proliferin drives the *in vitro* and *in vivo* growth of tobacco chemical-induced mouse lung adenocarcinoma cell lines

Since microarray analyses identified PRL transcripts to be the most abundantly and specifically overexpressed by LADC cell lines compared with other samples (Figures 4B and Supplementary figure SF5) we validated the microarray (Figure 4C and Supplementary figure SF6) and sought to functionally investigate its role in LADC development and evolution. Interestingly, PRL was overexpressed in experimental murine hyperplastic lesions and LADC from both urethane and *KRAS*^{G12D} models (Figure 5D). Furthermore, PRL expression was found to be mutant *KRAS*-associated and -driven, with *KRAS*^{MUT} cells displaying exclusive nuclear PRL immunoreactivity, in stark contrast to the cytoplasmic signal of *KRAS*^{WT} cells (Figure 5E). Silencing of *Kras* expression in *KRAS*^{MUT} cells resulted in decreased PRL expression levels whereas plasmid-mediated overexpression of Δ*Kras*2B^{G12C} transcript in *KRAS*^{WT} cells induced the levels of PRL (Figures 5F, 5G). Importantly, shRNA-mediated silencing of the most abundant PRL transcript *Prl2c2* side-by-side with the important cancer stemness transcript *Itgb3* led to comparable and significant decreases in *in vitro* cell proliferation and tumour-sphere formation, as well as in *in vivo* subcutaneous tumour growth rates and spontaneous metastatic capacity to the lungs (Figure 5H), indicating that PRL is an important lung tumour promoter as was previously shown for integrin β3 (33).

Prolactin is overexpressed in human lung adenocarcinoma and portends poor survival

We next assessed a potential role for the human proliferin orthologue prolactin (also abbreviated PRL) in human LADC. In a sample set (GEO Datasets accession ID: GSE43458) of 30 normal lung tissues from never smokers, 40 LADC from never

smokers, and 40 LADC from smokers from the BATTLE trial (35), a 77 gene-set representing the PRL signaling pathway [Wiki Prolactin Signaling Pathway, Homo Sapiens; ttp://www.wikipathways.org/index.php/Pathway:WP2037; (36)] could accurately cluster normal samples from cancer tissues (Figures 6A, 6B). Gene expression analysis revealed that *PRL* mRNA was significantly overexpressed in LADC tissues compared with normal lung tissues from the BATTLE study (Figure 6C). Moreover PRL immunoreactivity was stronger in LADC tissues from our center (37) compared with surrounding non-cancerous lung tissues (Figure 6D). Interestingly, a 2,378-gene transcriptomic signature of our LADC cell lines identified above (Figure 4D) managed to accurately cluster normal from cancer samples of the BATTLE study (Figure 6E). Furthermore gene set enrichment analysis (GSEA; (38); http://software.broadinstitute.org/gsea/index.jsp) of the transcriptomic signature and the BATTLE study dataset samples revealed highly positive enrichment of the transcriptomic signature in LADC samples from smokers but negative enrichment in LADC samples from never smokers (Figure 6F).

Importantly, lung cancer patients from the Kaplan-Meier Plotter database (http://kmplot.com/analysis/index.php?p=service&default=true; (39) with high PRL expression displayed significantly shorter survival compared with patients with lower expression levels. Interestingly, the dismal survival effect of PRL was restricted to female patients with LADC (Figures 7A-7E). Remarkably, multivariate Cox regression analyses revealed that together with increasing tumour stage, high PRL expression is an independent negative prognosticator of overall survival in patients with lung cancer, independent of sex, smoking history, and histology (Figures 7F, 7G).

Discussion

Here we report the first-ever derivation of multiple true lung adenocarcinoma cell lines obtained from lung tumours generated in mice by exposure to tobacco carcinogens. We comprehensively characterized their properties *in vitro* and *in vivo* and clearly show that they present true adenocarcinoma cell lines that: i) display cancer stem cell properties; ii) grow and metastasize in the lungs and pleural space of syngeneic mice similar to the human disease; iii) carry expression and mutation profiles that resemble human LADC of smokers. Importantly, a universal signature of these cell lines across different tobacco carcinogens and mouse strains used to generate them is shown to be present in human LADC, rendering them relevant tools for research on this disease. This signature revealed that these tobacco carcinogen-inflicted LADC cell lines depend on proliferin for sustained growth and metastasis. Moreover, the human proliferin orthologue prolactin was overexpressed in human LADC and was linked with poor survival. Hence our murine LADC cell lines prove for the first time that tobacco chemical-induced lung tumours in mice are indeed malignant, address the unmet need for faithful mouse models of smoking-induced human LADC in syngeneic immunocompetent mice, and present exciting new tools for the discovery of novel drivers and treatments of the human disease in the future.

This is the first study designed and implemented to develop transplantable mouse models of human tobacco carcinogen-induced LADC. While lung cancer is the leading cause of cancer death worldwide (1,2), LADC is its most common histologic subtype with increasing incidence, and tobacco smoking is the main cause of the disease (3-5), tools for research are still sparse. Although cell lines and

transplantable models have spearheaded lung cancer research and discovery, only a handful of murine cell lines for syngeneic transplantable models exist, complemented by a multitude of human cell lines for xenograft models in immunocompromised mice that lack an adaptive immune system (16,40). This shortcoming has been overcome by the development of transgenic mouse models that recapitulate salient features of human LADC (41,42). However, genetic LADC models are not metastatic (43), display copy number alterations rather than the heavy load of single nucleotide variants found in human LADC of smokers (6) and in chemical-induced LADC of mice (10), and often display histologic appearances not reminiscent of the human disease (40). Although chemical LADC models were long discovered and widely used, they were neglected in the era of transgenic models, thought to present adenomas rather than carcinomas. However, the strengths of chemical models, including their high mutation load, the predominance of single nucleotide variants, and their interaction between carcinogen, exposure protocol, and host genetic background render them lucrative (10,13,16). To this end, we used tobacco chemicals to identify a cardinal role for nuclear factor-κB signaling in LADC (44,45) and another group discovered important mechanisms of genomic context- and organ-specific *KRAS*-driven carcinogenesis using chemical models (46,47). Our work provides for the first time mouse models of LADC that combine the strengths of transplantable and chemical models: our LADC cell lines are readily transplantable in syngeneic mice, metastasize like human LADC, and carry *Kras*/*Trp53* lesions and gene expression profiles that resemble the human disease.

In addition to new research tools, our findings also provide important conceptual advances. By applying Robert Koch's postulates, we prove for the first time beyond

doubt that at least some chemical-induced lung tumours in mice are malignant adenocarcinomas. In addition, the isolation of the true tumour-initiating cells from these tumours will likely lead to the identification of new disease drivers and mechanisms, such as the proliferin/prolactin pathway reported here. Future sequencing of these cells will hopefully yield yet unknown perturbed genes and pathways that go undetected by large scale molecular fingerprinting projects that examine heterotypic tumours (10,27). This can be appreciated by the average ~70-fold overexpression of proliferin by our LADC cell lines relative to naïve murine lungs, as compared with the ~1,3-1,5-fold overexpression of prolactin in human LADC relative to naïve human lungs (Figure 6C and Supplementary Figure SF5). We are currently fingerprinting our cell lines, aiming at the functional identification of the genomic imprints of different tobacco (and other) carcinogens on the murine DNA in the nucleotide, trinucleotide, gene, locus, and chromosome levels, aiming to expand and validate the important findings of Alexandrov, Behjati, and colleagues, who initiated the process of defining carcinogenic DNA imprints by clinical correlation across multiple human cancers (48,49). Taking into account that tumour initiating cells might be resistant to conventional chemotherapy regimens, LADC cells may also present optimal models to study drug response and test novel therapies (50).

The method to generate chemical-induced LADC cell lines reported here may also be useful for future research, since it can be applied to any genetically modified mouse strain, yielding a powerful tool to study gene function in cancer, expanding the methods currently available for this, such as CRISPR/Cas9. For example, we have derived LADC cells from urethane-treated mice carrying conditionally deleted (floxed) *Trp53* alleles and have performed CRE-mediated recombination (i.e., *Trp53*

deletion) *in vitro*, gaining important insights into the role of functional *Trp53* in osteopontin signaling (30). We are currently applying this technique to an array of reporter, knock-in/out, and conditional mice, garnering important insights into LADC biology.

LADC cell lines were derived from two mouse strains (*FVB* and *Balb/c*) exposed repeatedly to the cigarette carcinogens urethane (ethyl carbamate, EC) and diethylnitrosamine (N-nitrosodiethylamine, DEN). This procedure simulates tobacco smoking in humans (10,16). Importantly, the culture of these cells does not introduce changes other than those induced by the carcinogenic process (26), a fact pending validation. To this end, transcriptomic analyses revealed ubiquitous altered expression of a signature comprised of 2,716 genes in our LADC cell lines that included a 43-gene-set tightly linked with proliferin. Proliferin and its human counterpart prolactin were validated as potent LADC drivers using observational studies in murine and human LADC, as well as functional studies in mice. The discovery of the role of proliferin/prolactin signaling in LADC underscores the value of our LADC cell lines as research vehicles and warrants further investigation of proliferin/prolactin as candidate therapeutic targets.

In conclusion, the tobacco carcinogen-inflicted murine LADC cell lines reported and made available here are valuable tools for research and discovery and can be used in multifaceted ways for future identification of molecular signatures, driver genes and pathways, and drugs against lung adenocarcinoma. These cell lines made possible the identification of proliferin/prolactin signalling as lung tumour promoter.

Authors' contributions

N. I. K., A. D. G., M. A. P., and G. T. S. collaboratively conceived, designed and carried out most experiments, analyzed the data, provided critical intellectual input, and wrote portions of the paper draft; T. A. performed genomic characterization of LADC cell lines; D. E. Z., I. L., and M. S. did tissue and cellular immune labeling, bright-field and confocal microscopy, and quantification; I. P. performed time-course experiments of lung tumourigenesis; I. G. carried out *Trp53* immunohistochemistry; A. M. performed *in vivo* experiments; H.P. evaluated tissue sections and diagnosed the histologic types and subtypes of tumours; G. T. S. and I. L. supervised the study, wrote the final paper version submitted for publication, and are the guarantors of the study's integrity.

Conflict of interest

The authors declare no competing interests.

Funding

This work was supported by European Research Council 2010 Starting Independent Investigator and 2015 Proof of Concept Grants (grant numbers 260524 and 679345 respectively; to G.T.S.). I.G. is a recipient of a Greek State Scholarship Foundation (IKY) programme co-financed by the European Union (European Social Fund-ESF), by Greek national funds through an action entitled "Reinforcement of Postdoctoral Researchers" (NSRF 2014 - 2020).

References

1. Torre, L.A., *et al.* (2016) Global Cancer Incidence and Mortality Rates and Trends--An Update. *Cancer Epidemiol Biomarkers Prev*, **25**, 16-27.
2. Torre, L.A., *et al.* (2015) Global cancer statistics, 2012. *CA Cancer J Clin*, **65**, 87-108.
3. Alberg, A.J., *et al.* (2013) Epidemiology of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*, **143**, e1S-e29S.
4. Sun, S., *et al.* (2007) Lung cancer in never smokers--a different disease. *Nat Rev Cancer*, **7**, 778-90.
5. Hecht, S.S. (1999) Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst*, **91**, 1194-210.
6. Garraway, L.A., *et al.* (2013) Lessons from the cancer genome. *Cell*, **153**, 17-37.
7. Ding, L., *et al.* (2008) Somatic mutations affect key pathways in lung adenocarcinoma. *Nature*, **455**, 1069-75.
8. Graziano, S.L., *et al.* (1999) Prognostic significance of K-ras codon 12 mutations in patients with resected stage I and II non-small-cell lung cancer. *J Clin Oncol*, **17**, 668-75.
9. Nelson, M.A., *et al.* (1996) Detection of K-ras gene mutations in non-neoplastic lung tissue and lung cancers. *Cancer Lett*, **103**, 115-21.
10. Westcott, P.M., *et al.* (2015) The mutational landscapes of genetic and chemical models of Kras-driven lung cancer. *Nature*, **517**, 489-92.
11. Ollila, S., *et al.* (2011) The tumor suppressor kinase LKB1: lessons from mouse models. *J Mol Cell Biol*, **3**, 330-40.
12. Jackson, E.L., *et al.* (2001) Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev*, **15**, 3243-8.
13. de Seranno, S., *et al.* (2010) Progress and applications of mouse models for human lung cancer. *Eur Respir J*, **35**, 426-43.
14. Desai, T.J., *et al.* (2014) Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature*, **507**, 190-4.
15. You, M., *et al.* (1989) Activation of the Ki-ras protooncogene in spontaneously occurring and chemically induced lung tumors of the strain A mouse. *Proc Natl Acad Sci U S A*, **86**, 3070-4.
16. Meuwissen, R., *et al.* (2005) Mouse models for human lung cancer. *Genes Dev*, **19**, 643-64.

17. Tuveson, D.A., *et al.* (1999) Modeling human lung cancer in mice: similarities and shortcomings. *Oncogene*, **18**, 5318-24.

18. Fassett, J.T., *et al.* (2000) Mrp4, a new mitogen-regulated protein/proliferin gene; unique in this gene family for its expression in the adult mouse tail and ear. *Endocrinology*, **141**, 1863-71.

19. Nilsen-Hamilton, M., *et al.* (1987) Detection of proteins induced by growth regulators. *Methods Enzymol*, **147**, 427-44.

20. Nilsen-Hamilton, M., *et al.* (1987) Relationship between mitogen-regulated protein (MRP) and proliferin (PLF), a member of the prolactin/growth hormone family. *Gene*, **51**, 163-70.

21. Wilder, E.L., *et al.* (1986) Expression of multiple proliferin genes in mouse cells. *Mol Cell Biol*, **6**, 3283-6.

22. Fang, Y., *et al.* (1999) Signaling between the placenta and the uterus involving the mitogen-regulated protein/proliferins. *Endocrinology*, **140**, 5239-49.

23. Corbacho, A.M., *et al.* (2002) Roles of prolactin and related members of the prolactin/growth hormone/placental lactogen family in angiogenesis. *J Endocrinol*, **173**, 219-38.

24. Toft, D.J., *et al.* (2001) Reactivation of proliferin gene expression is associated with increased angiogenesis in a cell culture model of fibrosarcoma tumor progression. *Proc Natl Acad Sci U S A*, **98**, 13055-9.

25. Agalioti, T., *et al.* (2017) Mutant KRAS promotes malignant pleural effusion formation. *Nat Commun*, **8**, 15205.

26. Pauli, C., *et al.* (2017) Personalized In Vitro and In Vivo Cancer Models to Guide Precision Medicine. *Cancer Discov*, **7**, 462-477.

27. Cancer Genome Atlas Research, N. (2014) Comprehensive molecular profiling of lung adenocarcinoma. *Nature*, **511**, 543-50.

28. Wu, K., *et al.* (2015) Frequent alterations in cytoskeleton remodelling genes in primary and metastatic lung adenocarcinomas. *Nat Commun*, **6**, 10131.

29. Chang, Y.F., *et al.* (2007) The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem*, **76**, 51-74.

30. Giopanou, I., *et al.* (2017) Tumor-derived osteopontin isoforms cooperate with TRP53 and CCL2 to promote lung metastasis. *Oncoimmunology*, **6**, e1256528.

31. Lampson, B.L., *et al.* (2013) Rare codons regulate KRas oncogenesis. *Curr Biol*, **23**, 70-5.

32. Su, Y.J., *et al.* (2015) Polarized cell migration induces cancer type-specific CD133/integrin/Src/Akt/GSK3beta/beta-catenin signaling required for maintenance of cancer stem cell properties. *Oncotarget*, **6**, 38029-45.

33. Seguin, L., *et al.* (2014) An integrin beta(3)-KRAS-RalB complex drives tumour stemness and resistance to EGFR inhibition. *Nat Cell Biol*, **16**, 457-68.

34. Oeztuerk-Winder, F., *et al.* (2012) Regulation of human lung alveolar multipotent cells by a novel p38alpha MAPK/miR-17-92 axis. *EMBO J*, **31**, 3431-41.

35. Kabbout, M., *et al.* (2013) ETS2 mediated tumor suppressive function and MET oncogene inhibition in human non-small cell lung cancer. *Clin Cancer Res*, **19**, 3383-95.

36. Kelder, T., *et al.* (2012) WikiPathways: building research communities on biological pathways. *Nucleic Acids Res*, **40**, D1301-7.

37. Giopanou, I., *et al.* (2015) Comprehensive Evaluation of Nuclear Factor-kappaBeta Expression Patterns in Non-Small Cell Lung Cancer. *PLoS One*, **10**, e0132527.

38. Subramanian, A., *et al.* (2007) GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics*, **23**, 3251-3.

39. Gyorffy, B., *et al.* (2013) Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS One*, **8**, e82241.

40. Gazdar, A.F., *et al.* (2016) Correction: "From Mice to Men and Back: An Assessment of Preclinical Model Systems for the Study of Lung Cancers". *J Thorac Oncol*, **11**, e88-9.

41. Kim, C.F., *et al.* (2005) Mouse models of human non-small-cell lung cancer: raising the bar. *Cold Spring Harb Symp Quant Biol*, **70**, 241-50.

42. Kwak, I., *et al.* (2004) Genetically engineered mouse models for lung cancer. *Annu Rev Physiol*, **66**, 647-63.

43. Vanharanta, S., *et al.* (2013) Origins of metastatic traits. *Cancer Cell*, **24**, 410-21.

44. Stathopoulos, G.T., *et al.* (2007) Epithelial NF-kappaB activation promotes urethane-induced lung carcinogenesis. *Proc Natl Acad Sci U S A*, **104**, 18514-9.

45. Vreka, M., *et al.* (2018) IκB Kinase α Is Required for Development and Progression of *KRAS*-Mutant Lung Adenocarcinoma. Cancer Res, **78**, 2939-51.

46. To, M.D., *et al.* (2013) Interactions between wild-type and mutant Ras genes in lung and skin carcinogenesis. *Oncogene*, **32**, 4028-33.

47. To, M.D., *et al.* (2008) Kras regulatory elements and exon 4A determine mutation specificity in lung cancer. *Nat Genet*, **40**, 1240-4.

48. Alexandrov, L.B., *et al.* (2016) Mutational signatures associated with tobacco smoking in human cancer. *Science*, **354**, 618-622.

49. Behjati, S., *et al.* (2016) Mutational signatures of ionizing radiation in second malignancies. *Nat Commun*, **7**, 12605.

50. Zhou, B.B., *et al.* (2009) Tumour-initiating cells: challenges and opportunities for anticancer drug discovery. *Nat Rev Drug Discov*, **8**, 806-23.

Figure Legends

Figure 1. Cell lines derived from lung tumours of inbred mice treated with tobacco carcinogens are true lung adenocarcinomas.

(A) Schematic of the method used and naming convention of cell lines derived from lung tumours of urethane (ethyl carbamate, weekly intraperitoneal injections of 1 g/Kg)- and diethylnitrosamine (DEN, weekly intraperitoneal injections of 200 mg/Kg)-treated *FVB* and *Balb/c* mice. Tumour culture was done at 37^{o}C in 5% CO_2-95% air using DMEM 10% FBS, 2 mM L-glutamine, 1 mM pyruvate, 100 U/ml penicillin, and 100 mg/ml streptomycin for at least 18 months or 60 passages, whichever occurred first. Cell lines were named XYLA# with X signifying the mouse strain (F, *FVB*; B, *Balb/c*), Y the carcinogen used (U, EC; E, DEN), LA lung adenocarcinoma, and # their serial number by derivation date.

(B-D) Macroscopic image (B) and hematoxylin & eosin-stained sections (C, D) of lungs with primary lung adenocarcinomas (LADC, dashed outlines) from urethane-treated *FVB* mice displaying hallmarks of malignancy including necrosis (black arrows) and invasion/distortion of adjacent lung structures (red arrows).

(E) Phase contrast image of FULA 1 cells in culture showing cobblestone (white solid arrows) and spindle (black solid arrows) shapes and anchorage-independent growth (anoikis, empty black arrows).

(F) May-Grünwald-Giemsa-stained cytocentrifugal specimen of FULA1 cells shows nuclear atypia.

(G) *In vitro* MTT growth curves of chemical LADC cell lines compared with Lewis lung carcinoma (LLC) cells (*n* = 4/group/time-point). Data are presented as mean±SD. *** denotes *P* < 0.001 for LLC cells compared with LADC cells at 96 hours by two-way ANOVA with Bonferroni post-tests.

(H) Primary tumour-spheres formed by FULA1 cells *in vitro*.

(I) *In vitro* tumour sphere formation potential of chemical LADC cell lines compared with LLC cells (*n* = 3/group). Data are presented as mean±SD. * and *** denote *P* < 0.05 and *P* < 0.001 for the indicated comparisons by one-way ANOVA with Bonferroni post-tests.

(J) Bioluminescence image of constitutively luminescent *FVB*-Tg(CAG-luc,-GFP)L2G85Chco/J mouse injected with 1 mg D-luciferin at 4 weeks post-subcutaneous injection of one million FULA1 cells showing reduction of the bioluminescent signal by the non-luminescent tumour (dashed outline) compared to the opposite healthy site (arrow).

(K-P) Mice syngeneic to the respective cell line (LLC cells: *C57BL/6* mice; FULA cells, FVB mice; BULA and BELA cells, *Balb/c* mice) received one million tumour cells subcutaneously, 250,000 cells intravenously, or 150,000 cells intrapleurally (*n* = 10/cell line/route) and were followed till the first signs of sickness. (K) Representative images of subcutaneous primary tumours (top left, dashed outline) and spontaneous lung metastases (top right, black arrows) at 1 month post-subcutaneous delivery, of forced lung metastases at 3 weeks post intravenous injection (bottom left, black arrows), and of malignant pleural effusion (MPE) at 12 days post-intrapleural

injection (bottom right, dashed lines). (L, M) Data summary of primary tumour volume (L) and spontaneous lung metastasis number at 24-34 days (M) post-subcutaneous delivery. (N) Lung tumour number at 2-3 weeks post intravenous injection. (O) MPE volume at 12-28 days post-intrapleural injection. (P) Kaplan-Meier survival plots and overall log-rank test probability values (*P*). (L-O) Data are presented as mean±SD. ns, *, **, and *** denote $P > 0.05$, $P < 0.05$, $P < 0.01$, and $P < 0.001$, respectively, for the comparisons indicated or for LLC cells compared with LADC cells (L), by one-way (M-O) or two-way (L) ANOVA with Bonferroni post-tests.

Figure 2. Tobacco carcinogen-induced mouse lung carcinomas are classified as adenocarcinomas.

Primary urethane (A, F, K) and DEN (B, G, L)-induced LADC tumors that gave rise to the LADC cell lines, secondary subcutaneous tumors generated by transplantation of FULA cells into syngeneic mice (C, H, M), *KRAS*^{G12D}-driven LADC (Ref. 45; D, I, N), as well as subcutaneous tumours of pancreatic adenocarcinoma PANO2 cells (Ref. 25; E, J, O), were stained for hematoxylin & eosin (H&E), Periodic acid–Schiff–diastase (PAS-D), and thyroid transcription factor 1 (TTF1).

(A-E) H&E-stained representative tumor sections. Note the typical glandular-solenoid structure of LADC tumours (A-D) and the solid form of PANO2 tumours (E).

(F-J) PAS-D stain for visualization of mucin. Note the adenocarcinoma-distinctive positive mucin staining of all LADC (F-I) and the negative PAS-D staining of PANO2 tumours (J).

(K-O) Immunostaining for TTF1 (NKX2-1). Note the LADC-distinctive nuclear immunoreactivity of all LADC (K-N) and the negative results from PANO2 tumours (O).

Figure 3. Tobacco carcinogen-induced mouse lung adenocarcinoma cell lines bear codon 61 *Kras* mutations and exhibit loss of *Trp53*.

(A) *Kras* and *Nras* mRNA expression by reverse transcription (RT-PCR) of select chemical-induced lung adenocarcinoma cell lines and sequences of the amplicons cut and extracted from the gels together with their matching Ensembl annotations. Note the shorter nonsense-mediated decay (NMD) transcript.

(B) cDNA Sanger sequencing traces of splenocytes of wild-type (wt) *C57BL/6* mouse and of the chemical-induced lung adenocarcinoma cell lines reported here. Note the heterozygous $Kras^{Q61R}$ and $Kras^{Q61H}$ single nucleotide variants (arrows) in all urethane- and DEN-induced cell lines, respectively. Note also the superimposition of wild-type, mutant, and NMD *Kras* traces in urethane-induced cell lines. No *Egfr* and *Nras* mutations were detected.

(C) *Trp53* mRNA expression by reverse transcription (RT-PCR) of mouse tracheal epithelial cells (mTECs) cultured from the lungs of urethane-exposed mice at various time-points post-injection, of select chemical-induced lung adenocarcinoma cell lines, and of Lewis lung carcinoma (LLC; $Kras^{G12C}$, $Trp53^{WT}$) and MC38 colon adenocarcinoma ($Kras^{G13R}$, $Trp53^{R178P}$) cells.

(D) *Trp53* and *Egfr* mRNA expression by qPCR of LLC, MC38, and chemical-induced LADC cell lines relative to *Gusb*. Data are presented as mean±SD (n = 3/group). *, **, and *** denote $P < 0.05$, $P < 0.01$, and $P < 0.001$, respectively, for comparison with LLC cells by one-way ANOVA with Bonferroni post-tests.

(E) EGFR, KRAS, NRAS, TRP53, and ACTB protein expression of LLC, chemical-induced LADC, and CT26 (*Kras*^{WT}, *Trp53*^{WT}) and MC38 colon adenocarcinoma cells by Western immunoblot. Note the absence of detectable labile TRP53^{WT} expression in all but MC38 cells that bear mutant *Trp53*^{R178P} that results in abnormally stable but non-functional TRP53 protein.The immunoblot has been cropped.

LADC cell line naming convention XYLA# denotes X for mouse strain (F, *FVB*; B, *Balb/c*), Y for carcinogen used (U, EC; E, DEN), LA for lung adenocarcinoma, and # for serial number by derivation date.

Figure 4. Comparative transcriptome profiling of carcinogen-induced mouse lung adenocarcinoma cell lines identifies focal overexpression of proliferins.

(A) Unsupervised hierarchical clustering of global transcriptomes of chemical-induced lung adenocarcinoma (LADC) cell lines, total mouse lung RNA, alveolar type 2 cells, airway epithelial cells, bone marrow-derived mast cells and macrophages, and other cancer cell lines by microarray [GEO Datasets accession IDs GSE94981 for LADC cell lines, lungs, tracheal epithelial cells, mast cells, and macrophages; GSE82154 for alveolar epithelial type 2 cells (ATII cells) and GSE58188 for other cancer cell lines; freely available at https://www.ncbi.nlm.nih.gov/gds/]. Cut-off used was statistical significance by ANOVA (*P*) and false discovery rate (FDR) $q < 10^{-12}$.

(B) PRL transcript (*Pr2c2*, *Prl2c3* and *Prl2c4*) and *Itga2* expression of mouse LADC cells and naïve lungs relative to *Gusb* by qPCR. Data are presented as mean±SD. *** denotes $P < 0.001$ for all comparisons with naïve lungs by one-way ANOVA with Bonferroni post-tests.

(C) The 43 transcripts differentially expressed in LADC cells versus all other groups and clustering together with PRL transcripts (A, red box) comprising the LADC proliferin (PRL) signature. *Δ*GE, differential gene expression; ANOVA, analysis of variance; *P*, probability.

(D, E) Summary of the murine genes differentially expressed in LADC cells and naïve lungs using the cut-offs shown and of their human orthologues comprising the LADC cell line signature (D) and pathway analysis thereof (E).

Figure 5. Carcinogen-induced mouse lung adenocarcinoma cell lines overexpress lung and cancer stemness markers. Proliferin drives lung adenocarcinoma growth *in vitro* and *in vivo.*

(A) Representative dotplots and data summary of flow cytometry of lung cells and Lewis lung carcinoma (LLC) and chemical-induced lung adenocarcinoma (LADC) cell lines for the cancer stem cell marker CD44 and the lung stem cell marker LGR6 identified significant proportions of CD44+LGR6+ cells (arrows). Data are presented as mean±SD (n = 3/group). *** denotes $P < 0.001$ for comparison with lung cells by one-way ANOVA with Bonferroni post-tests.

(B) *Lgr6* mRNA expression by qPCR of lungs and LLC, MC38, and chemical-induced LADC cell lines relative to *Gusb*. Data are presented as mean±SD (n = 3/group). *** denotes $P < 0.001$ for comparison with lungs by one-way ANOVA with Bonferroni post-tests.

(C) ITGB3, PRL, and ACTB protein expression of mouse tracheal epithelial cells (mTECs) cultured from the lungs of urethane-exposed mice at various time-points post-injection, of select chemical-induced LADC cell lines, and of Lewis lung carcinoma, MC38 colon adenocarcinoma, and CT26 colon adenocarcinoma cells by Western immunoblot. Immunoblot has been cropped.

(D) Immunoreactivity of murine lungs from the urethane and $KRAS^{G12D}$ LADC models for proliferin (red color) before tumour initiation (top), at early stages of tumour progression (middle), and when harboring LADC (bottom). Blue color

indicates nuclear Hoechst33258 counterstaining. Note the increased PRL expression in LADCs.

(E) Immunoreactivity of benign mouse tracheal epithelial cells (mTECs) and NIH 3T3 fibroblasts, and of select chemical-induced LADC cells for proliferin (red color). Blue color indicates nuclear Hoechst33258 counterstaining. Note the increased nuclear PRL expression in LADCs.

(F) *Prl2c2* mRNA expression by qPCR of different mouse cancer cell lines with (FULA, LLC, MC38) and without (PANO2, B16F10) Kras mutations (25) relative to *Gusb*. Data are presented as mean±SD (n = 3/group). *** denotes $P < 0.001$ for comparison with B16F10 cells by one-way ANOVA with Bonferroni post-tests.

(G) *Prl2c2* gene expression of parental and *Kras*-modulated (red: sh*Kras*-expressing; green: p*ΔKras2B*-expressing) cancer cell lines (25) relative to *Gusb* by qPCR shows that Prl2c2 expression is *KRAS*-driven. Data are presented as mean±SD (n = 3/group). *, **, and *** denote $P < 0.05$, $P < 0.01$, and $P < 0.001$, respectively, for comparison with parental cells by Student's t-test.

(H) FULA1 and BULA2 cells were stably transfected with target-specific shRNA against *Prl2c2* or *Itgb3*, or random shRNA pools (shC). Shown are representative results of PRL, ITGB3, and ACTB protein expression by Western immunoblot, MTT assay and tumour sphere formation capacity *in vitro*, as well as primary tumour growth and lung metastasis formation *in vivo* upon injection of one million cells to syngeneic *FVB* or *Balb/c* mice. Data presented are mean±SD (n = 3-6/group) obtained from FULA2 cells, but identical results were obtained using BULA1 cells. *

and *** denote $P < 0.05$ and $P < 0.001$, respectively, for comparison of the color-coded silenced cells with control-transfected cells by one or two-way ANOVA with Bonferroni post-tests.

Figure 6. Prolactin is overexpressed in human lung adenocarcinoma.

(A) Genes that comprise the Wiki Prolactin Signaling Pathway (Homo Sapiens; http://www.wikipathways.org/index.php/Pathway:WP2037) used as input for analyses of the BATTLE study in (B) (36).

(B) Unsupervised hierarchical clustering of the BATTLE trial transcriptomic dataset including 30 normal lung tissues from never smokers, 40 lung adenocarcinomas (LADC) from never smokers, and 40 LADC from smokers (GEO Datasets accession ID: GSE43458; (35)) by the Wiki Prolactin Signaling Pathway from (A) significantly distinguishes normal samples from LADC tissues (χ^2 and hypergeometric test $P < 0.0001$).

(C) Prolactin (PRL) transcript *PRL* normalized to β-actin (ACTB) transcript *ACTB* expression of patients from the BATTLE trial shows increased PRL expression in LADC compared to normal lung tissues. Data are presented as median with Tukey's whiskers (boxes: interquartile range; bars: 50% extreme quartiles) and raw data points (dots) (n = 30-40/group). P denotes overall one-way ANOVA probability and * and *** denote $P < 0.05$ and $P < 0.001$, respectively, for comparisons indicated by Bonferroni post-tests.

(D) Immunoreactivity of representative LADC tissues from our center (37) for prolactin (red color). Blue color indicates nuclear Hoechst33258 counterstaining.

Note the increased PRL expression in LADCs (dashed lines) compared with adjacent tissues.

(E) Unsupervised hierarchical clustering of the BATTLE trial transcriptomic dataset by a 2,378-gene transcriptomic signature of our LADC cell lines identified in this study (Figure 4D) significantly distinguishes normal samples from LADC tissues (χ^2 and hypergeometric test $P < 0.0001$).

(F) Gene set enrichment analysis (GSEA; (38); http://software.broadinstitute.org/gsea/index.jsp) of a 2,378-gene transcriptomic signature of our LADC cell lines identified in this study (Figure 4D) in smokers' and never-smokers' LADC from the BATTLE study. Note that the signature of our tobacco carcinogen-induced LADC cell lines was significantly positively enriched in LADC from smokers from the BATTLE trial, but negatively enriched in LADC from never smokers. NES, normalized enrichment score; *P*, nominal probability: FDR *q*, false discovery rate probability; FWER *P*, family-wise error rate probability. Note that FDR *q* and FWER $P < 0.25$ are considered significant in GSEA.

Figure 7. Prolactin expression is associated specifically with poor survival of female patients with lung adenocarcinoma.

(A-E) Kaplan-Meier survival plots with univariate Cox regression hazard ratios (HR) and 95% confidence intervals (95% CI) of patients with lung cancer (A), squamous cell lung carcinoma (B), lung adenocarcinoma (C), and females (D) and males (E) with lung adenocarcinoma, stratified by prolactin mRNA expression as determined by microarray (probe ID: 205445_at). Optimal cut-offs were determined by dichotomizing patient data by all possible percentiles of prolactin expression. Data were from the Kaplan-Meier Plotter database (http://kmplot.com/analysis/index.php?p=service&default=true; (39). Note the sex- and histology-specific impact of prolactin expression on survival.

(F, G) Kaplan-Meier survival plot with multivariate Cox regression HR and 95% CI of patients with lung cancer stratified by prolactin mRNA expression (F). Analyses were done as above, this time using multivariate function and entering prolactin expression, histologic subtype, stage, sex, and smoking history as co-variables. Results of multivariate Cox regression analysis (G) shows that together with increasing stage, high prolactin expression is an independent dismal predictor of survival in lung cancer.

RMA, robust multi-array units; OS, overall survival; HR, hazard ratio of high versus low expressing patients; 95% CI, 95% confidence intervals; *P*, log-rank test or Cox regression probability values.

Figure 1

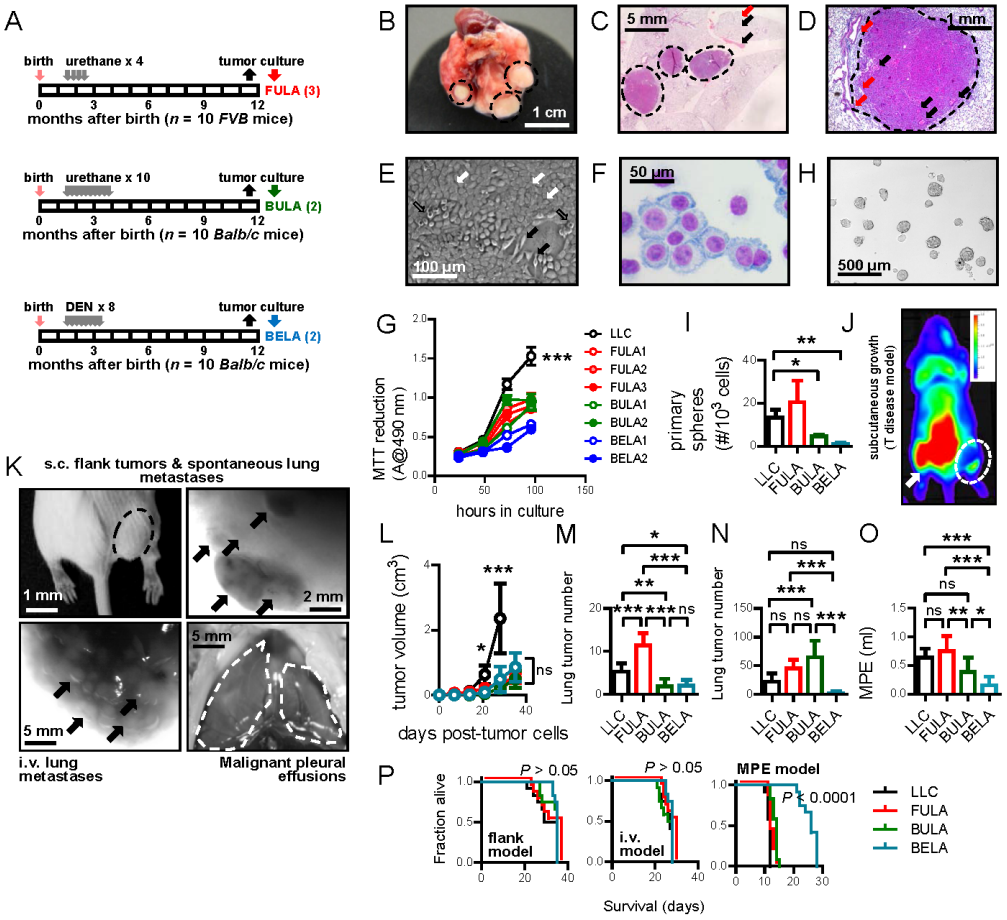


Figure 2

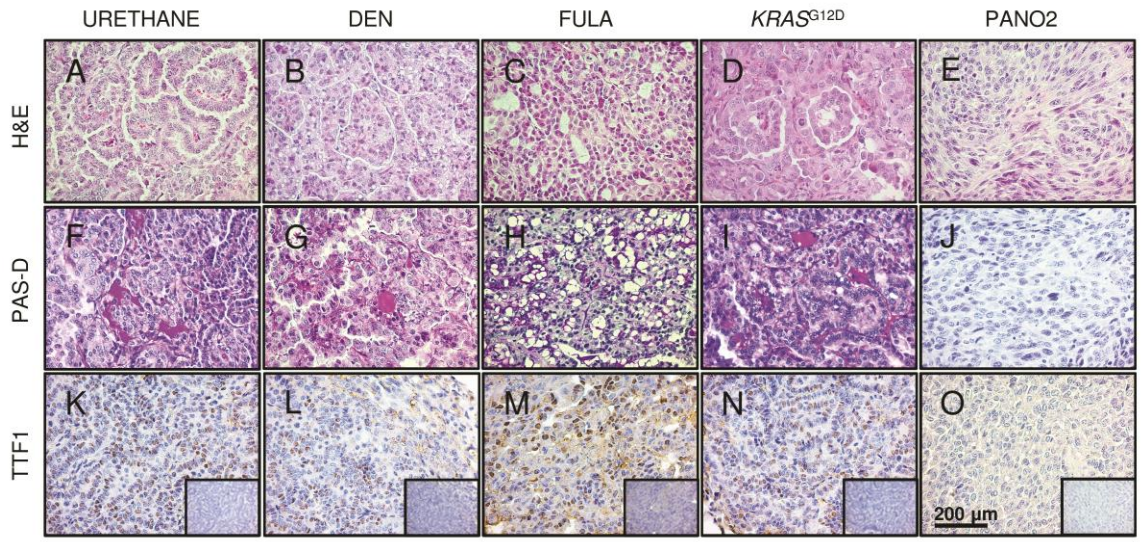


Figure 3

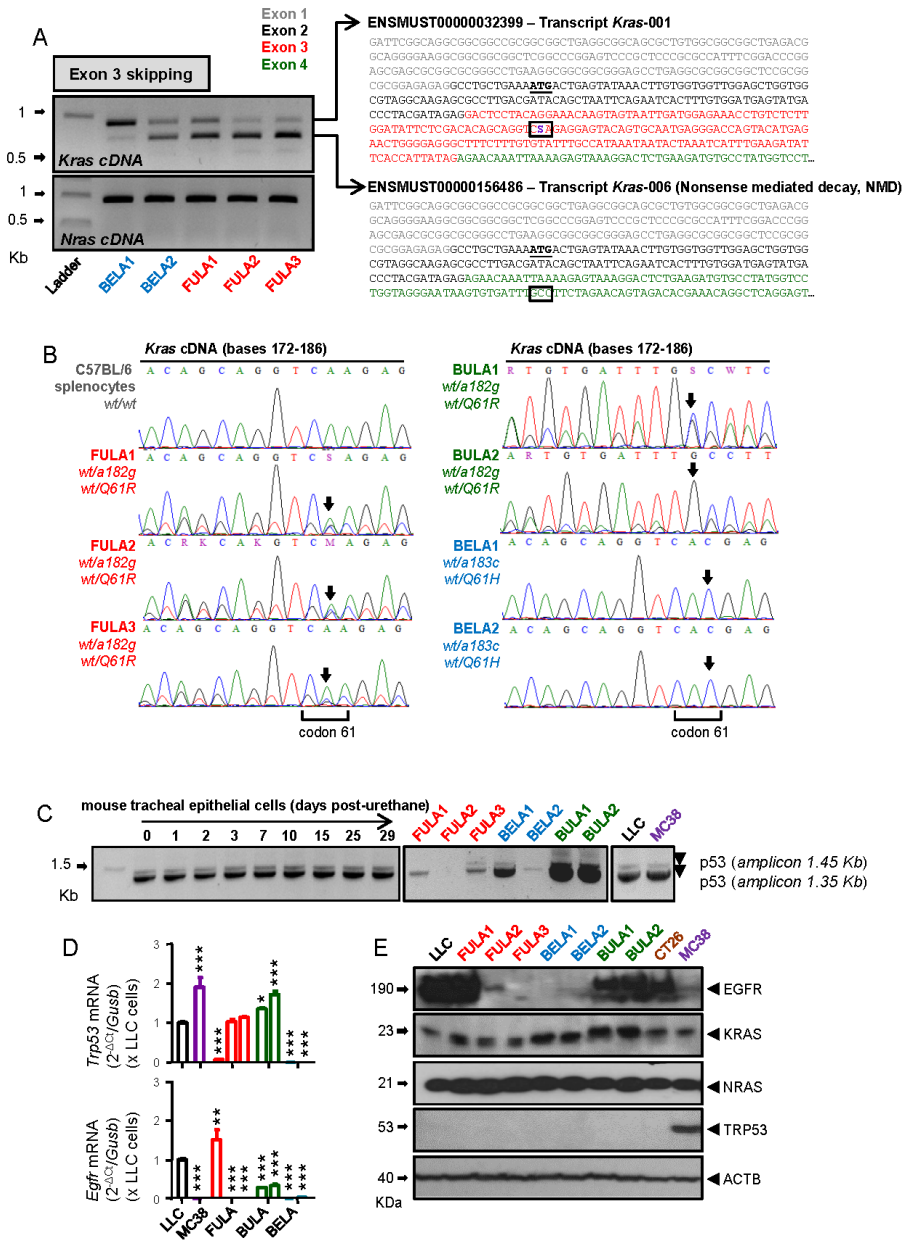
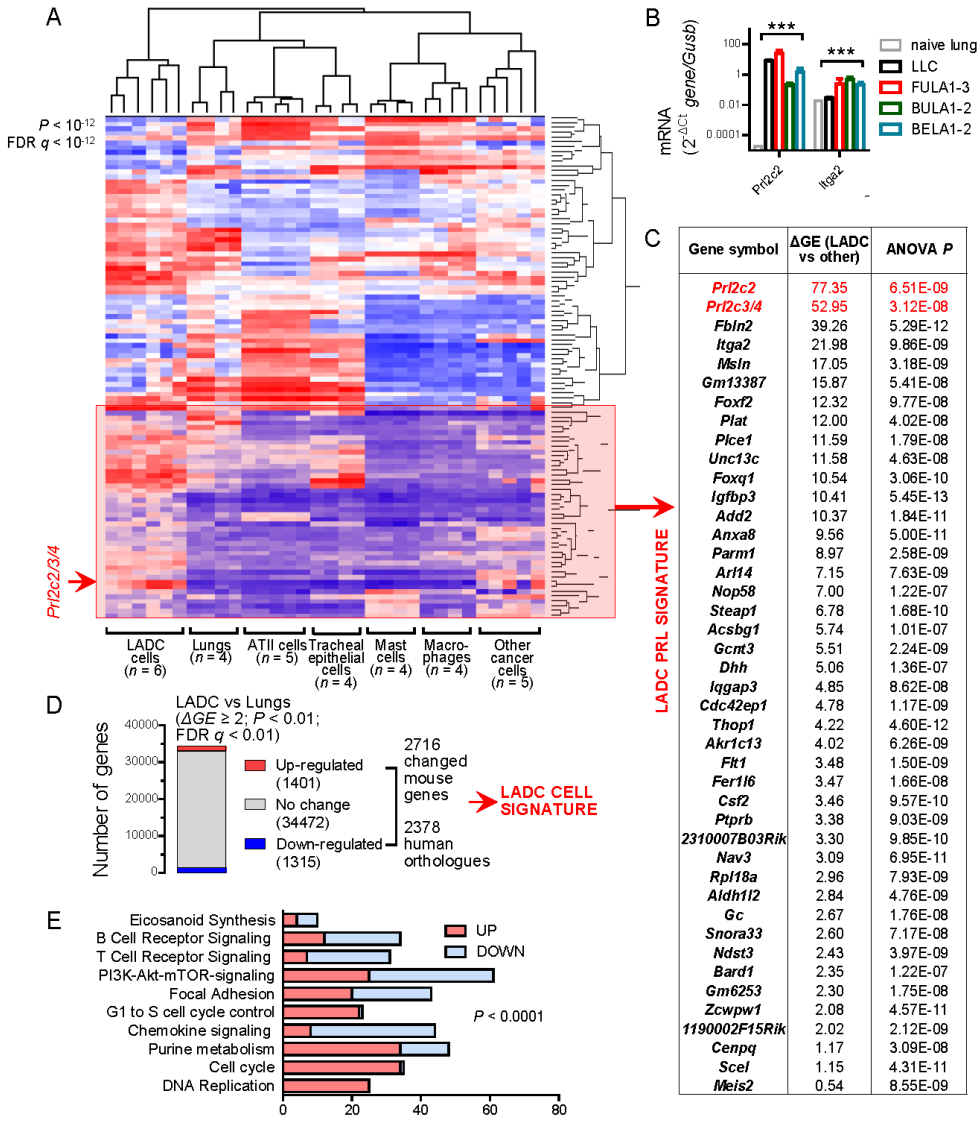


Figure 4



C

Gene symbol	ΔGE (LADC vs other)	ANOVA *P*
Prl2c2	77.35	6.51E-09
Prl2c3/4	52.95	3.12E-08
Fbln2	39.26	5.29E-12
Itga2	21.98	9.86E-09
Msln	17.05	3.18E-09
Gm13387	15.87	5.41E-08
Foxf2	12.32	9.77E-08
Plat	12.00	4.02E-08
Plce1	11.59	1.79E-08
Unc13c	11.58	4.63E-08
Foxq1	10.54	3.06E-10
Igfbp3	10.41	5.45E-13
Add2	10.37	1.84E-11
Anxa8	9.56	5.00E-11
Parm1	8.97	2.58E-09
Arl14	7.15	7.63E-09
Nop58	7.00	1.22E-07
Steap1	6.78	1.68E-10
Acsbg1	5.74	1.01E-07
Gcnt3	5.51	2.24E-09
Dhh	5.06	1.36E-07
Iqgap3	4.85	8.62E-08
Cdc42ep1	4.78	1.17E-09
Thop1	4.22	4.60E-12
Akr1c13	4.02	6.26E-09
Flt1	3.48	1.50E-09
Fer1l6	3.47	1.66E-08
Csf2	3.46	9.57E-10
Ptprb	3.38	9.03E-09
2310007B03Rik	3.30	9.85E-10
Nav3	3.09	6.95E-11
Rpl18a	2.96	7.93E-09
Aldh1l2	2.84	4.76E-09
Gc	2.67	1.76E-08
Snora33	2.60	7.17E-08
Ndst3	2.43	3.97E-09
Bard1	2.35	1.22E-07
Gm6253	2.30	1.75E-08
Zcwpw1	2.08	4.57E-11
1190002F15Rik	2.02	2.12E-09
Cenpq	1.17	3.09E-08
Scel	1.15	4.31E-11
Meis2	0.54	8.55E-09

Figure 5

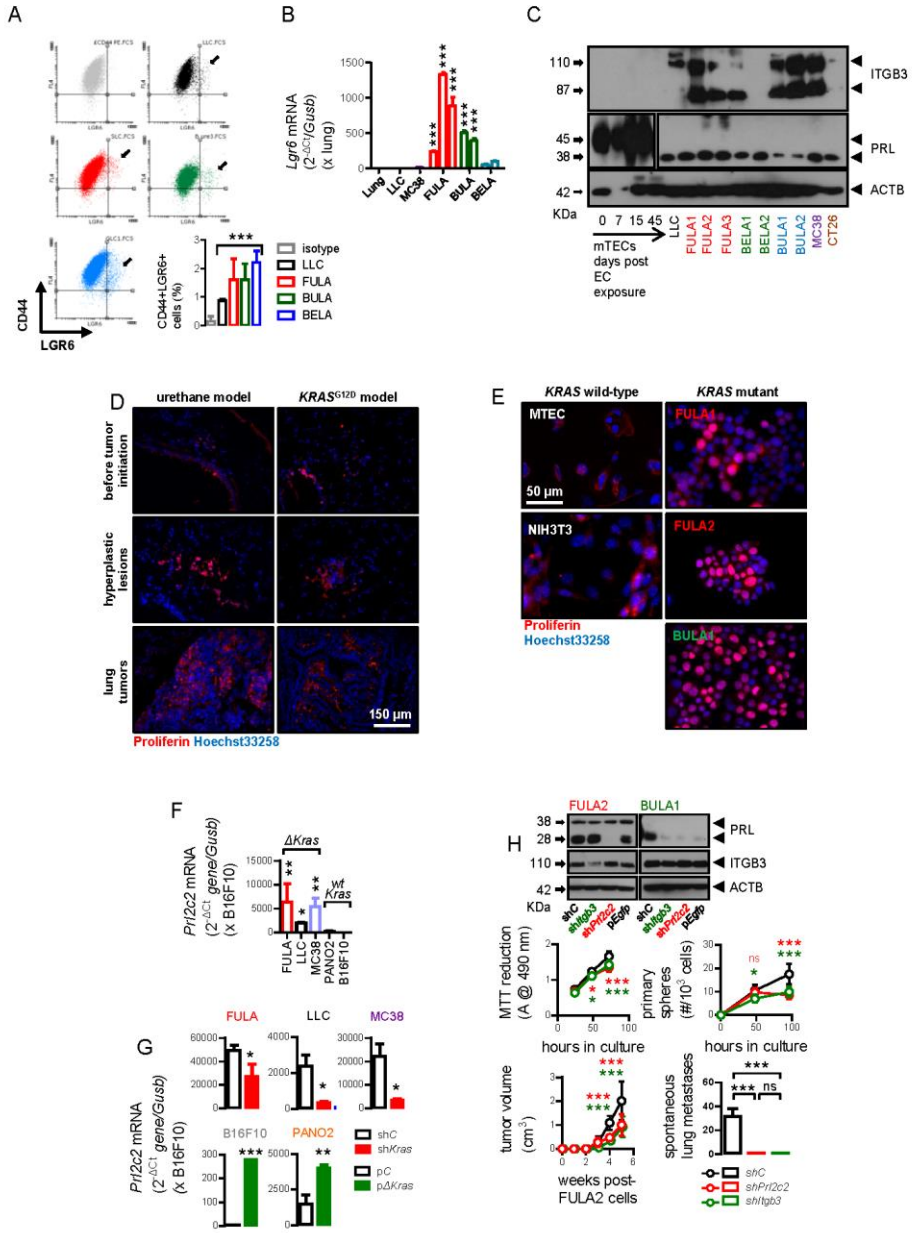


Figure 6

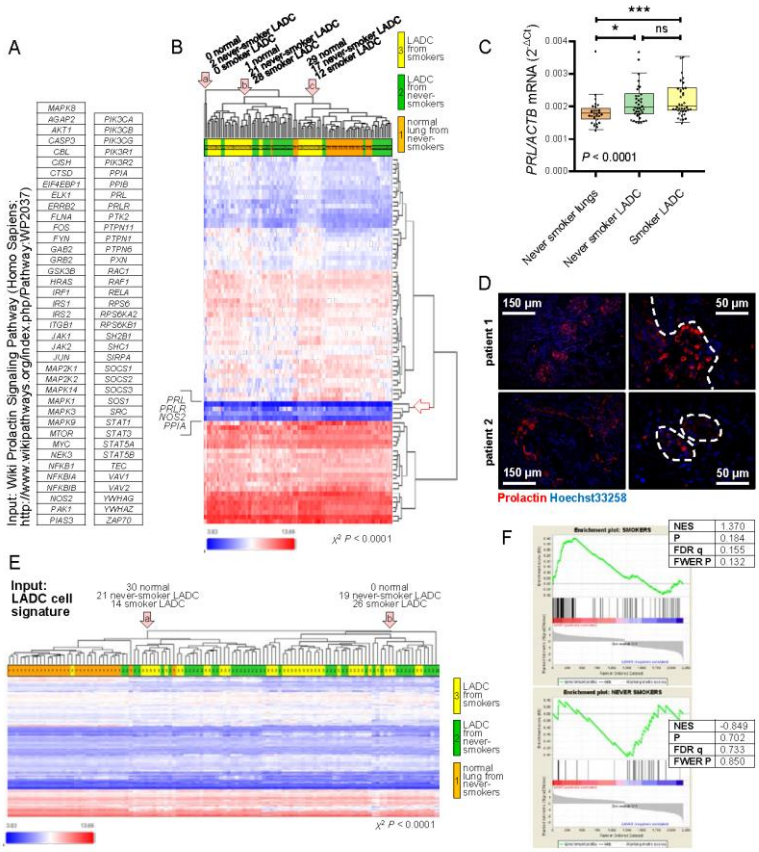
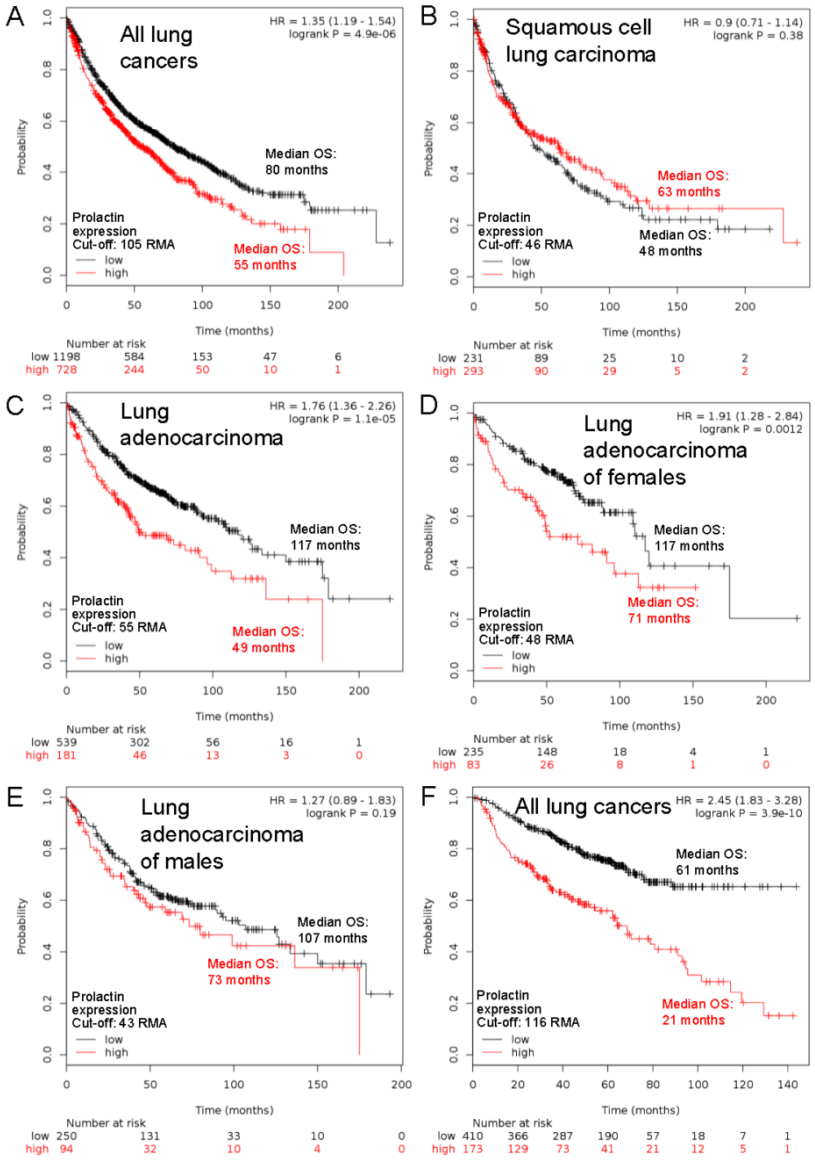


Figure 7



G

Factor	P	HR	95% CI
Histology	0.1995	1.26	0.89-1.78
Stage	< 0.0001	1.68	1.37-2.07
Sex	0.0717	1.34	0.97-1.84
Smoking	0.2196	0.74	0.46-1.20
Prolactin	0.0276	1.51	1.05-2.18