

Syngeneic tobacco carcinogen-induced mouse lung adenocarcinoma model exhibits PD-L1 expression and high tumor mutational burden

Laura P. Stabile, ... , Timothy F. Burns, Christopher J. Bakkenist

JCI Insight. 2020. <https://doi.org/10.1172/jci.insight.145307>.

Technical Advance In-Press Preview Oncology

Human lung adenocarcinoma (LUAD) in current or former smokers exhibits a high tumor mutational burden (TMB) and distinct mutational signatures. Syngeneic mouse models of clinically relevant smoking-related LUAD are lacking. We established and characterized a tobacco-associated transplantable murine LUAD cell line, designated FVBW-17, from a LUAD induced by the tobacco carcinogen 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (NNK) in the FVB/N mouse strain. Whole exome sequencing of FVBW-17 cells identified tobacco-associated *Kras*^{G12D} and *Trp53* mutations and a similar mutation profile to that of classic alkylating agents with a TMB >500. FVBW-17 cells transplanted subcutaneously, via tail vein and orthotopically generated tumors in FVB/N mice that were histologically similar to human LUAD. FVBW-17 tumors expressed PD-L1, were infiltrated with CD8⁺ T cells, and responsive to anti-PD-L1 therapy. FVBW-17 cells were also engineered to express green fluorescent protein and luciferase to facilitate the detection and quantification of tumor growth. Distant metastases to lung, spleen, liver, and kidney were observed from subcutaneously transplanted tumors. This novel cell line is a robust representation of human smoking-related LUAD biology and provides a much needed pre-clinical model in which to test promising new agents and combinations including immune-based therapies.

Find the latest version:

<https://jci.me/145307/pdf>



Syngeneic tobacco carcinogen-induced mouse lung adenocarcinoma model exhibits PD-L1 expression and high tumor mutational burden

Laura P. Stabile^{1,2}, Vinod Kumar³, Autumn Gaither-Davis³, Eric Huang³, Frank P. Vendetti⁴, Princey Devadassan³, Sanja Dacic⁵, Riyue Bao^{2,3}, Richard A. Steinman^{1,2,3}, Timothy F. Burns^{1,2,3}, and Christopher J. Bakkenist^{1,2,4}

¹Department of Pharmacology & Chemical Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA

²UPMC Hillman Cancer Center, Pittsburgh, PA

³Department of Medicine, Division of Hematology/Oncology, University of Pittsburgh School of Medicine, Pittsburgh, PA

⁴Department of Radiation Oncology, University of Pittsburgh School of Medicine, Pittsburgh, PA

⁵Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, PA

Corresponding Author:

L.P.S.

UPMC Hillman Cancer Center

Research Pavilion, Suite 2.18

5117 Centre Avenue

Pittsburgh, PA 15213

Phone: 412-623-2015

stabilela@upmc.edu

Conflict of interest statement: The authors have declared that no conflict of interest exists.

Abbreviations: lung adenocarcinoma (LUAD); non-small cell lung cancer (NSCLC); immune checkpoint inhibitors (ICIs); programmed death 1 (PD-1); programmed death ligand 1 (PD-L1); tumor mutational burden (TMB); tumor microenvironment (TME); Lewis lung carcinoma (LLC); genetically engineered mouse models (GEMMs); 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (NNK); atypical adenomatous hyperplasia (AAH); variant allele frequency (VAF); whole exome sequencing (WES); green fluorescent protein (GFP); luciferase (Luc); fluorescence-activated cell sorting (FACS); 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL); conditioned media (CM); short tandem repeat (STR).

Abstract

Human lung adenocarcinoma (LUAD) in current or former smokers exhibits a high tumor mutational burden (TMB) and distinct mutational signatures. Syngeneic mouse models of clinically relevant smoking-related LUAD are lacking. We established and characterized a tobacco-associated transplantable murine LUAD cell line, designated FVBW-17, from a LUAD induced by the tobacco carcinogen 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (NNK) in the FVB/N mouse strain. Whole exome sequencing of FVBW-17 cells identified tobacco-associated *Kras*^{G12D} and *Trp53* mutations and a similar mutation profile to that of classic alkylating agents with a TMB >500. FVBW-17 cells transplanted subcutaneously, via tail vein and orthotopically generated tumors in FVB/N mice that were histologically similar to human LUAD. FVBW-17 tumors expressed PD-L1, were infiltrated with CD8+ T cells, and responsive to anti-PD-L1 therapy. FVBW-17 cells were also engineered to express green fluorescent protein and luciferase to facilitate detection and quantification of tumor growth. Distant metastases to lung, spleen, liver, and kidney were observed from subcutaneously transplanted tumors. This novel cell line is a robust representation of human smoking-related LUAD biology and provides a much needed pre-clinical model in which to test promising new agents and combinations including immune-based therapies.

25 word summary:

We established and characterized a clinically relevant tobacco carcinogen-induced LUAD syngeneic cell line that will be useful for preclinical evaluation of immune-oncology agents.

Keywords: Syngeneic mouse model, lung adenocarcinoma, tobacco carcinogen

Introduction

The therapeutic landscape for non-small cell lung cancer (NSCLC) has rapidly transformed based on histological subtype, molecular profile and immunophenotype. In the majority of patients, there are no targetable oncogenic drivers. Therefore the field of immuno-oncology has rapidly evolved for NSCLC and immune-based therapies, including immune checkpoint inhibitors (ICIs) such as antibodies to programmed death 1 (PD-1) and programmed death ligand 1 (PD-L1), have emerged as exciting therapeutic approaches (1). However, only approximately 20% of patients with advanced NSCLC respond clinically to current ICIs (2,3). While high tumor mutational burden (TMB), a characteristic of smoking-related cancers, and tumoral PD-L1 expression are the most well-validated predictive biomarkers for response to ICIs across multiple cancer types (4-6), preclinical autochthonous murine models of NSCLC are urgently needed to serve as a platform for novel immune-related biological insights and testing therapeutic strategies.

Our ability to understand changes in the lung tumor microenvironment (TME), especially the immune environment, is limited by the lack of appropriate preclinical models. The only robust and widely used syngeneic lung cancer model, the Lewis lung carcinoma (LLC) model (7), was derived from a spontaneous lung carcinoma in a C57BL/6 mouse. The LLC model does not recapitulate exposure to tobacco carcinogen which is the principal driver of lung cancer in patients as well as the primary mutagen that generates tumor neo-antigens that may be key to response to ICIs. The CMT64 model, also established from a spontaneous alveolar lung carcinoma in a C57Bl/6 mouse, and its metastatic sublines also have limited utility as they again do not represent exposure to tobacco carcinogen (8). LM2 cells, derived from a urethane induced lung tumor in A/J mice are not widely used and the molecular features and biological phenotypes are not well characterized (9,10). Similarly, while genetically engineered mouse models (GEMMs) of lung cancer are good for the evaluation of oncogenic drivers and targeted therapies, they do not represent clinical lung tumorigenesis and lack mutational diversity (11).

Among the tobacco carcinogenic compounds, nicotine-derived nitrosamines, such as 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (NNK), and its metabolites, are the most abundant and the strongest carcinogens (12,13). The FVB/N mouse does not develop spontaneous lung tumors, but it is susceptible to chemically-induced lung carcinomas (11), and a commonly used strain for generation of GEMMs (14-16). We previously characterized the development of NNK-induced lung adenocarcinoma (LUAD) in FVB/N mice, where pulmonary preneoplasias were found as early as 4 weeks post-NNK (17). These hyperplastic lesions, found in the small airways and the lung parenchyma, consist of tightly packed epithelial cells displaying a papillary growth pattern and decreased nuclear to cytoplasmic ratio (18). They are histologically similar to the human preneoplastic lesion and LUAD precursor, atypical adenomatous hyperplasia (AAH), often found in smokers (19), which progress to *KRAS* mutant LUAD over time (18,20). While gene expression signatures in NNK-induced lung tumors are highly comparable to those found in human LUAD (21), our knowledge of the murine mutational landscape, such as the smoking-associated mutational patterns characterized in human LUAD (22), is limited. The long latency time for full LUAD development in the NNK model, lack of metastatic potential and the labor intensive quantification of complete lung tumor burden that can only be performed postmortem restricts its usefulness for routine preclinical therapeutic testing.

To overcome these challenges, we developed and characterized a clinically relevant tobacco carcinogen (NNK)-induced LUAD syngeneic cell line to facilitate preclinical testing of potential therapeutic approaches in an immune-competent environment. FVBW-17 is the first tobacco carcinogen-induced murine lung tumor cell line that represents human LUAD disease. To our knowledge, this is the only syngeneic model derived from a lung tumor induced by a tobacco carcinogen that is easily transplantable in FVB/N mice, grows reproducibly and metastasizes to multiple sites. This tobacco carcinogen induced syngeneic model of LUAD with PD-L1 expression, high TMB and responsiveness to anti-PD-L1 therapy is a critical technical advance that will allow impactful preclinical experimental studies.

Results

Establishment and characterization of FVBW-17 cell line. We have previously described the use of the NNK-induced lung cancer model for both preventive and treatment therapies, and this model has been used extensively to study NSCLC initiation and progression (14,17,18,20,23). However, this model has a long latency period for adenocarcinoma and lacks metastatic potential. In order to improve the utility of the model for preclinical testing, we treated mice with NNK and isolated primary tumors to establish a syngeneic NNK-induced mouse lung cancer cell line as shown in the schema in **Figure 1A**. A parental line and 3 clonal lines were established. These cells exhibit typical NSCLC morphology under light microscopy and a rapid doubling time in culture of 35 hrs (**Figure 1B**). A DNA sample from the parental and clonal cell lines developed was submitted to ATCC for characterization and authentication using their Mouse Short Tandem Repeat (STR) Profiling Service (**Figure 1C**). In order to support the use of this line for studies evaluating antibodies targeting the PD-1/PD-L1 axis, in vitro PD-L1 expression on FVBW-17 cells was examined by flow cytometry. We observed robust basal expression of PD-L1, 6-fold higher than the control (**Figure 1D**), which was confirmed by Western analysis (not shown).

FVBW-17 cell lines display a prominent alkylating molecular signature and genetic features of human LUAD. We performed whole exome sequencing (WES) of FVBW-17 cell lines to evaluate the similarity to human LUAD. Signatures of mutational processes attributed to known mutagenic exposures have been previously reported (24). We analyzed the sequence-context specific mutation signature of FVBW-17 cells which revealed a striking similarity to human cancer Signature 11, with a low-high, low-high repetitive pattern within the C>T group (**Figure 2A and Supplemental Figure 1**), which does not preferentially occur at CpG sites from spontaneous deamination of methyl-cytosine leading to the age-related Signature 1 (24). Signature 11 has the mutational pattern reported in experimental studies of alkylating agents (25). NNK can be activated to a pyridyloxobutylating agent, and this alkylating agent creates O6-methylguanine

adducts which lead to G>A mutations (C>T on the opposite strand) (26). We next calculated the total TMB defined as the total number of nonsynonymous mutations from the whole exome in the four FVBW-17 cell lines and found a median TMB of 570 per exome (range 506-579), approximately 2-fold higher than the median TMB found in current and former smokers using the TCGA human LUAD cohort (**Figure 2B**). **Figure 2C** depicts the number of coding variants identified by type. Since *KRAS* and *TP53* are mutated at high frequencies in human LUAD, occur concurrently in many cases (27), and are typically found in tumors from patients who smoke (28,29), we specifically focused on identifying these mutations. We found a high variant allele frequency (VAF) of both *Trp53* (P275S; 0.957) and *Kras* (G12D; 0.60) (**Figure 2E**). While *Trp53* P275S (equivalent to human P278S) is not a common mutation, it is located in the DNA binding domain which is a mutational hot-spot. *Kras*^{G12D} was confirmed by Sanger sequencing and p53 protein expression (a surrogate which correlates with p53 hot-spot mutations) was confirmed by Western analysis (data not shown). No other oncogenic activating mutations were found.

FVBW-17 cells grow rapidly in syngeneic FVB/N mice and tumors maintain a similar histomorphology and immunophenotype to human LUAD. To evaluate the tumorigenic potential of FVBW-17 cells, we injected 1.25×10^5 – 7.5×10^5 cells subcutaneously (s.c.) in the hind flank and monitored tumor growth over 21 days. As shown in **Figure 3A**, FVBW-17 cells resulted in rapid tumor growth that was proportional to the number of cells injected. By 4 weeks, tumors were aggressive and exceeded 10% of the body weight in the 5×10^5 and 7.5×10^5 groups. All 3 clonal lines grew at a similar rate (data not shown). Tumors were removed, fixed, and paraffin embedded for histological evaluation. Hematoxylin & eosin (H&E) stained sections demonstrate malignant tumor with morphological features similar to LUAD in humans (**Figure 3B**). Similar to human LUAD, the tumor is composed of sheets and nests of large polygonal cells with vesicular nuclei, occasional prominent nucleoli and moderate amounts of cytoplasm (**Figure 3B**, red arrows). Tumor nests are surrounded by fibromyxoid desmoplastic stroma with scattered lymphocytes

(**Figure 3B**, black arrows). The tumor is poorly differentiated because no morphological features of glandular, squamous or neuroendocrine differentiation can be found on the routine H&E sections at least as a subcutaneous model.

To determine the extent of immune infiltration in s.c tumors, we performed immune profiling of CD45+ cells in a subset of mice at Day 21 (**Figure 3C**). CD8+ T cells, conventional CD4+ T cells (CD4+ Tconv) and regulatory T cells (Treg) were quantified as total number per 10⁴ cells (upper panels) as well as a percentage of CD45+ immune cells (lower panels). We found that T cell populations comprised less than 10% of the total CD45+ immune cells, and exhibited low CD8+ and CD4+ T cell/Treg ratios. In general, these results suggest that FVBW-17 tumors are weakly infiltrated with CD8+ T cells. However, these values are similar to what we have previously observed in the syngeneic CT26 colon cancer model (30). The relatively low T cell infiltrate that we observed may be due to the late time point of tumor extraction and the presence of necrosis in some tumors.

To determine if FVBW-17 cells could serve as a platform to understand lung metastasis and tumor growth within the lung environment, we evaluated whether FVBW-17 cells could colonize the lung following lateral tail vein injection or through direct implantation in the lungs. Ten mice (5 male/5 female) were injected intravenously (i.v.) with 1 x 10⁶ FVBW-17 cells. After 4-5 weeks, mice were sacrificed and lungs and other tissues were collected and fixed for histologic evaluation. We found metastatic lung nodules in all mice as shown in H&Es (**Figure 4A**) and in gross lungs (**Figure 4C**, top panel). In one mouse, we also observed bone metastases but no evidence of tumor growth in any other tissues at this time point (**Figure 4A**). To evaluate primary tumor growth within the lung microenvironment, we transplanted 1 x 10⁶ FVBW-17 cells orthotopically to the lungs of 22 FVB/N mice (12 female/10 male) which led to the formation of well-differentiated lung adenocarcinomas in 100% of the mice with a median primary tumor size of 1.31 mm³ (range: 0.49-2.72 mm³) as assessed by H&E stained sections (**Figure 4B**) and gross

lungs (**Figure 4C**, bottom panel, black arrow). By 2 weeks post-implantation, mice had difficulty breathing due to primary tumor burden. While we did not observe metastasis to distant tissues with this model, we did observe spread to the contralateral lobe of the lung in 5/22 mice (**Fig 4C**, bottom panel, gray arrows).

Establishment of cell lines that stably express green fluorescent protein and luciferase. FVBW-17 cells were stably transfected with a lentiviral construct conferring expression of green fluorescent protein (GFP) and firefly luciferase (GFP-Luc), allowing tumor growth to be monitored in real time using in vivo bioluminescence imaging. The GFP positive FVBW-17 cells were sorted into high and low GFP positive cells by fluorescence-activated cell sorting (FACS). The high GFP positive cells were expanded for use all in vivo experiments. **Figure 5A** shows high GFP expression by comparing the phase contrast image (left) to fluorescent image (right). We also detected a cell number dependent increase in luciferase expression on cells grown in vitro (**Figure 5C**). We found that FVBW-17 GFP-Luc cells implanted s.c. into FVB/N mice formed rapidly growing tumors that could be imaged and quantified over time, and showed a slightly higher luminescence signal as compared with caliper measurements due to the capability of 3D tomography in imaging (**Figure 5C**). The s.c. tumors were surgically excised and mice were monitored for distant metastases up to 2 additional months. In mice euthanized at 1 month or longer after primary tumor removal, we detected metastases to lung, spleen, kidney and liver as detected by *ex vivo* imaging (**Figure 5D**) and histological confirmation (**Figure 5E**). H&E sections of metastases show invasive adenocarcinoma comprised of malignant glands and solid sheets of epithelioid cells. **Supplemental Table 1** shows the frequency of metastases to these tissues. Central necrosis is also present in some cases. To confirm that the GFP-Luc cells could also grow within the lung and be imaged, we also injected these cells orthotopically. In vivo imaging detected tumor growth in the lung over time (**Figure 5F**).

Effect of PD-L1 inhibition on FVBW-17 tumor growth in vivo. To determine the therapeutic effect of PD-L1 targeting in this model, FVBW-17 cells were implanted s.c. into FVB/N mice. When tumors were palpable after 1 week, anti-PD-L1 or control treatments commenced with twice-weekly dosing for a total of four doses (2 weeks) by i.p. injection. Comparison of individual tumor growth in each treatment group demonstrated that a subset of tumors had suppressed growth of FVBW-17 compared with control treatment. Importantly, we found that 6/10 (60%) mice had evidence of lung metastases in the control group, compared to only 2/8 (25%) in the anti-PD-L1 treated group (**Figure 6A**). To determine the importance of the lung TME in mediating response to anti-PD-L1, we also tested a 3 week anti-PD-L1 treatment in the lung metastasis model and found a significant decrease in lung weight ($P=0.012$) and a lower median number of lung metastatic nodules in the anti-PD-L1 treated group compared to control treated mice (**Figure 6B**). These results suggest preliminary anti-PD-L1 activity, providing a new experimental model with robust growth to evaluate novel immunotherapeutic agents.

Discussion

A major barrier to identifying new treatment options and improving survival in NSCLC is the lack of preclinical models that mimic the human disease, including metastasis and immune competent models. We have generated and characterized a novel, clinically relevant tobacco carcinogen (NNK)-induced LUAD syngeneic cell line that will facilitate in vivo preclinical testing of targeted and immune therapies for NSCLC. Tobacco smoke is made up of thousands of chemicals and more than 60 carcinogens, including NNK (12). NNK is a heterocyclic nitrosamine with potent lung carcinogenic properties in both humans and mice (31-33). NNK metabolites have been detected in urine of smokers and those exposed to second-hand tobacco smoke (34), and elevated levels of NNK are associated with lung cancer risk (35). NNK induced lung carcinogenesis has been well characterized in mice and in other animal models including rats, hamsters, and ferrets (36). NNK initiates carcinogenesis in part through its metabolic conversion to the electrophilic metabolite 4-

(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) which binds to DNA, forming DNA adducts and in turn causing dysregulated gene replication (32,36). NNK-derived DNA damage occurs through metabolic activation of NNK and NNAL to DNA alkylating agents generating pyridylhydroxybutyl DNA adducts and methyl adducts, respectively (26). Recent studies have demonstrated NNK exerts lung tumor specific genotoxic effects by acting as an agonist for the $\alpha 7$ nicotinic acetylcholine receptor which regulates downstream activation of signaling networks involved in the promotion and progression of lung cancer including the PI3K/AKT and MAPK/ERK pathways (37,38). NNK has also been shown to activate the $\beta 2$ -adrenergic receptor which transactivates the EGFR and Raf-1/ERK1/2 signaling cascades, both of which are also well characterized for their roles in lung cancer progression and metastasis (39,40). We found that NNK exposure was not sufficient to mimic the known human tobacco-associated genetic signatures in lung cancer (i.e. Signatures 4 and 5) (22), which are likely dominated by benzopyrene, but was highly reflective of an alkylating signature consistent with the chemical attack of NNK on DNA (Signature 11) and induced the most frequent tobacco-associated alterations including *Kras* and *Trp53*. This is not surprising as the complex mixture found in tobacco smoke may initiate other mutational processes beyond that of NNK.

FVBW-17 cells grow reproducibly and aggressively in FVB/N mice, resulting in spontaneous metastasis to multiple sites. Although ICIs have revolutionized treatment of NSCLC patients, ~80% of patients do not respond. A key hurdle to advancement of immune based treatments is the availability of immune competent preclinical models that recapitulate human disease. While we found a relatively low immune infiltrate in the s.c. tumors at a late time point, immune infiltrate at earlier time points or in the lung following orthotopic implantation has not yet been evaluated. Further comprehensive immunophenotyping is also warranted as well as a direct head-to-head comparison with other syngeneic lines, including LLC. It is also possible that this model may have immune suppressive myeloid-derived suppressor cells or macrophages in the TME. Based on our preliminary finding of anti-PD-L1 activity, this model will be useful to interrogate whether blocking

the interaction of PD-1 and PD-L1 could cause tumor regression and block metastasis. Since only a subset of tumors responded to anti-PD-L1, tumor inhibition may require a multipronged approach of combinations of treatments to facilitate immune recruitment and activation of an anti-tumor response.

Overall, our findings show that the tobacco-carcinogen induced FVBW-17 LUAD murine cell line shares genomic and histological similarity to human LUAD, grows reproducibly in FVB/N mice and metastasizes to distant sites. This tobacco carcinogen induced syngeneic model of LUAD with PD-L1 expression and high TMB is a major technical advance that will allow impactful preclinical experimental studies.

Materials and Methods

Generation and characterization of the FVBW-17 cell line. Lung adenocarcinomas were induced by treating a 6-week old female FVB/N mouse with NNK (Toronto Research Chemicals; 120 mg/kg; 15 mg/ml solution in 0.9% NaCl) i.p. injections 3 times per week for 5 weeks starting at week 0. At week 25, the mouse was sacrificed, tumors were dissected from the lung and washed 2 times with DMEM complete medium (supplemented with 10% fetal bovine serum (FBS) and 1% antibiotic/antimycotic). Fresh lung tumor tissue was finely minced with scalpels and added to 24-well plates containing either DMEM complete medium or 50% DMEM complete medium supplemented with 5% sterile filtered conditioned medium (CM) from three bone marrow or robustly growing osteosarcoma derived cell lines, including MG-63 (ATCC CRL-1427), HS-5 (ATCC CRL-11882) and HS-27A (ATCC CRL-2496) cells (48-72hrs of exposure)/50% Ham's F12 + 1% FBS and 1% antibiotic/antimycotic. We had shown that the selected marrow lines, HS-5 and HS-27A (41) express mesenchymal stem cell (MSC) markers ((42)and data not shown) and CM from MSCs has been shown to augment growth of human lung cancer cell lines (43). We found that addition of CM from any of these sources to our DMEM culture medium augmented the growth of the FVBW-17 cells, with MG-63 CM increasing the growth most robustly. Once the cells

were expanded, we were able to decrease the CM percentage to 1% or use DMEM complete growth medium alone with maintenance of the growth rate.

Following tissue attachment, cells were monitored for colony formation followed by cell expansion. 2.5×10^5 cells in PBS were injected s.c into the right hind flank of a female FVB/N mouse for selecting transplantable cells. The established tumor was excised when it reached a size of approximately 250 mm^3 and cultured as for the primary culture. The parental line and three single cell clonal cell lines, B10, B11 and F8, were established by serial dilution in 96-well plates, followed by subculture and expansion in larger vessels. The morphology of cells was observed by phase contrast microscopy; all images were taken on a Leica DMI3000B inverted microscope. Mouse STR profiling was performed by ATCC on the parental and clonal lines at passage 3. Cell growth properties were assessed every 24 hrs over a 7-day period by counting the number of viable cells with a hemocytometer. Basal PD-L1 expression in FVBW-17 cells was assessed by flow cytometry after staining the cells with anti-mouse PD-L1 antibody (clone 10F.9G2) APC or viability dye without PD-L1 antibody (control). The fluorescence minus one technique was used to set gating controls. Cell lines were routinely tested to be mycoplasma-free.

Whole exome sequencing (WES). DNA from FVBW-17 parental and clonal cell lines was extracted using QIAamp DNA Mini kit (Qiagen). $1 \mu\text{g}$ genomic DNA per sample was used as input for sample preparation. Sequencing libraries were prepared using Agilent SureSelectXT Mouse All Exon kit (Agilent Technologies). Briefly, fragmentation was carried out using the hydrodynamic shearing system (Covaris) to generate 180-280bp fragments and remaining overhangs were converted into blunt ends. After adenylation of the 3' ends, adapter oligonucleotides were ligated and selectively enriched in a PCR reaction. The libraries were then hybridized with a biotin labeled probe, and magnetic beads with streptomycin were used to capture exons. Captured libraries were enriched in a PCR reaction to add index tags to prepare for hybridization. Products were

purified using the AMPure XP system (Beckman Coulter) and quantified using the Agilent Bioanalyzer 2100 System.

Somatic variant identification. The raw sequencing data were analyzed following protocols similar to our previous work (44) but customized to the mouse genome and the mouse strain in our study, FVB/N. The quality of raw reads was assessed by FastQC (45) (v0.11.7). Paired-end reads were mapped to mouse reference genome (GRCm38) using BWA-MEM (46) (v0.7.17) with a minimal requirement of mapping quality 30 and soft-clipping option activated by default. Post-alignment processing included duplicate removal using Sambamba (47) (v0.6.8), insertions/deletions realignment and base quality score recalibration (BQSR) using GATK4 (48) (v4.1.7.0). Since matched normal tissue was not available for sequencing, we detected putative somatic mutations by MuTect2 (49) (v4.1.7.0) using the tumor-only mode with vendor-provided target BED file. MuTect2 implements smarter handling of overlapping read pairs to account for PCR error and VAF calculation in this version, hence no prior read merging was required. Variants were further filtered by MuTect2 intrinsic filters, read orientation artifact filter, and strain-specific polymorphism filter. Variants that passed all filters were annotated using ANNOVAR (50) (release 11/27/2019). Our mouse strain is FVB/N, and we used C57BL/6J (GRCm38) as the reference for read alignment and variant calling, followed by proper measure to eliminate variants identified as strain-specific polymorphisms, with the following considerations: 1) the exome capture kit was designed based on C57BL/6J genome, with the target BED file in C57BL/6J genomic coordinates, and for our analysis, we lifted over the GRCm37 (mm9) coordinates provided in the BED file to GRCm38 (mm10) using UCSC chain files (<http://hgdownload.cse.ucsc.edu/goldenpath/mm9/liftOver/>); 2) The C57BL/6J gene annotation is the most comprehensive and well curated; 3) Most published studies are using C57BL/6J reference genome, and our variants reported in C57BL/6J genome coordinates will be more relevant in this context; and 4) We took steps to account for the strain-specific differences in

alignment refinement and variant filtering, including: (a) using FVB/N dbSNP (v142) (Sanger MGI) as the known sites of variation for post-alignment BQSR, and (b) removing all variants present in FVB/N dbSNP (v142) after the variant calling. The summary of variant filtering statistics is provided in **Table S2**.

Mutation signature analysis. The TMB was calculated by the total number of mutations that were predicted to cause a protein sequence change (a.k.a. non-synonymous somatic mutations, NSSMs), including non-synonymous, stopgain, and stoploss single nucleotide variants (SNVs), frameshift and non-frameshift insertions/deletions (indels), and variants that modify splicing sites. For mutation signature analysis, the trinucleotide sequence context of each variant was retrieved from the mouse genome using *genCountMatrixFromVcf* from Bioconductor package *signer* (v1.10.0) with object *BSgenome.Mmusculus.UCSC.mm10*. The mutation lollipop figures were generated using MutationMapper from cBioPortal (https://www.cbioportal.org/mutation_mapper) (accessed 9/22/2020). *Kras* mutation was validated by Sanger sequencing.

TCGA LUAD analysis. The level-2 open-access somatic mutation data and clinical tables were downloaded from Broad's Institute GDAC Firehose portal (<https://gdac.broadinstitute.org/>) (release 3/18/2016) for primary tumors from 512 LUAD patients in The Cancer Genome Atlas (TCGA). Mutation MAF file was imported using *maftools* (51) (v2.0.16), and TMB was calculated as the total number of NSSMs in tumor. The tobacco smoking history group of TCGA LUAD patients was retrieved from the clinical table and coded in 1 to 5. We were unable to locate documentation of the codes, therefore we attempted to decipher code 1 to 5 as follows: code 1 to 4 was deciphered by matching the column "tobacco_smoking_history" in clinical table from GDAC to "Smoking.status" column in Supplementary Table 7 from the TCGA LUAD study (52), and code 5 was deciphered by community knowledge (<https://groups.google.com/g/cbioportal/c/irEXZRj9Who>) with reference to NCI CDE Browser. As a result, code 1 to 5 were defined as: 1 = lifelong non-smoker, 2 = current smoker, 3 = current

reformed smoker for > 15 years, 4 = current reformed smoker for ≤ 15 years, and 5 = current reformed smoker, duration not specified. Patients with missing tobacco smoking history ($n=14$) or code 5 ($n=4$) were excluded, leaving a total of 494 patients for further analysis. Among those, 471 have exome data available, hence were included for TMB calculation.

Evaluation of in vivo tumor growth. FVBW-17 cells were grown to sub-confluent states, trypsinized, dispersed in 1x PBS, and injected s.c. (12.5×10^5 - 7.5×10^5 cells in 100 μ l PBS), i.v. (5×10^5 cells in 100 μ l PBS), or direct injection into the parenchyma of the left lung lobe through the rib cage using a 30G needle (1×10^6 cells in 40 μ l PBS containing matrigel at final concentration 1.35 mg/ml) into 6-8 week old FVB/N mice (both male and female mice). For the orthotopic injection, surgery was performed under inhaled isoflurane anesthesia and the incision was closed using VetBond™ adhesive. Mice were followed for up to 4 weeks. Flank tumors were measured twice weekly with digital calipers, and tumor volume calculated as length X width²/2. After 21 days, s.c tumors were removed and fixed in 10% formalin for histological evaluation. For metastasis studies, the s.c. tumors were surgically excised under isoflurane anesthesia and the excision wound was closed with VetBond™ tissue adhesive. Mice were monitored for distant metastases up to 2 additional months. For i.v. and orthotopic models, lungs were formalin-inflated via a cannula inserted through the trachea under 25 cm intra-alveolar pressure. Formalin fixed paraffin embedded (FFPE) tumor tissues from flank tumors, lungs or other tissues were prepared and sliced into 5 μ m sections for histologic assessment. Tumor morphology was evaluated by H&E staining and analyzed by a board-certified thoracic pathologist (S.D.). Bright field images were captured using a Leica DMI3000B scope and LASv4.7 software.

Tumor infiltrating immune profiling. In a separate group of mice, tumors (s.c.) and spleens were processed for flow cytometric analysis of T cell populations. Excised tumors were injected with 1.5 mL digestion solution (50 ug/mL Liberase DL research grade (Roche) and 10 U/mL DNase I (Sigma) in RPMI), incubated 3 minutes at room temperature, chopped into small pieces,

and incubated in a total volume of 5 mL digestion solution for 15 minutes at 37°C. Digested tumor pieces and spleens were mechanically dissociated between frosted glass slides and filtered through 70 µm cell strainers (Corning) to generate single cell suspensions. Red blood cells were lysed with 1 ml of 150 mM NH₄Cl, 10 mM NaHCO₃, and 0.1 mM EDTA at room temperature, and lysis was quenched with 4 ml RPMI after 30 seconds (spleen) or immediately (tumors). Cell suspensions were then counted, and 2 x 10⁶ cells were stained as follows: 15 minutes at 4°C in FCS buffer (2% FBS in PBS) with antibodies to surface antigens, 10 minutes at 4°C in PBS with eFlour780 viability dye, 15 minutes at room temperature in Fixation/Permeabilization reagent (eBioscience), and 45 minutes at room temperatures in 1x Permeabilization Buffer (eBioscience) with antibody to nuclear Foxp3. Cells were washed once in FCS buffer prior to surface staining, washed twice in FSC buffer before viability dye staining and fixation, washed twice in 1x Permeabilization Buffer before and after nuclear staining, and washed once in FSC buffer as a final wash before acquisition. Uncompensated data were collected using FACSDiva software and LSRFortessa cytometer (BD Biosciences). Compensation and analyses were performed using FlowJo v10 software. Single color stained, and unstained spleen cells were stained for single color compensation controls.

Generation of GFP-luciferase expressing cells. FVBW-17 cells were stably transfected with lentiviral vector encoding green fluorescent protein (GFP) and luciferase (LL310PA-1; System Biosciences) using polybrene according to the manufacturer's instructions. Briefly, 1 x 10⁶ FVBW-17 cells were seeded in a T25 flask. The next day, the transfection mixture of 4 mL of media, 4µL of polybrene (8 mg/mL) and 1 mL of virus was added to the cells and incubated overnight. After 24hr, the media was replaced with fresh media (without virus and polybrene) for an additional 48hr. After 48hr, 1 µg/mL puromycin was added to the media. The cells were trypsinized and cultured in T75 flasks for 10 days in growth media containing 1 µg/mL puromycin, with fresh media added every other day. After 10 days, puromycin was increased to 2.5 µg/mL and the cells were grown for an additional 10-15 days. Luciferase expression from stably transfected cells was

confirmed by IVIS imaging (IVIS Imaging System, Perkin Elmer, Inc.) of 10,000-80,000 cells seeded in a 12-well plate compared to cells without the GFP-Luc gene and luciferin alone. GFP expression was confirmed by fluorescence microscopy and flow cytometry. In vivo experiments were performed with stable luciferase and GFP-transfected cells cultured in media containing 0.5 µg/mL puromycin. For in vivo bioluminescence imaging, the mice were injected with D-luciferin potassium salt (150 mg/kg, i.p.), anesthetized with isoflurane (3% for induction; 1.5% for maintenance) and imaged. Tissues were removed and imaged *ex vivo* for evidence of metastatic spread. Analysis was done using the Living Imaging software (PerkinElmer, Inc.; version 4.3.1).

PD-L1 antibody blockade. Monoclonal antibody to mouse PD-L1 and mouse IgG1 control antibody were provided by AstraZeneca through a material transfer agreement. Tumor bearing mice were injected i.p. 2 times per week at a dose of 10mg/kg.

Availability of data and materials. Raw sequencing data were uploaded to the public Sequence Read Archive (SRA) database (<https://www.ncbi.nlm.nih.gov/sra>) in the National Center for Biotechnology Information (Accession code: PRJNA670224). We will send this cell line to anyone immediately upon completion of an MTA with the University of Pittsburgh.

Study approval. All procedures in mice were reviewed and approved by the University of Pittsburgh Institutional Animal Care and Use Committee and performed in accordance with relevant guidelines for tumor burden in mice.

Statistics. A two-tailed Student's t-test was used for the anti-PD-L1 experiments. A P value of less than 0.05 was considered significant. Prism 7.0 (GraphPad) was used for all analyses.

Author Contributions

L.P.S. conducted experiments, was responsible for the overall experimental design, data analysis and interpretation and wrote the manuscript; A.G-D. established the cell lines and coordinated WES and STR profiling; E.H. and P.D. performed in vivo experiments; V.K established the GFP-

Luc cells; F.P.V performed flow cytometry and data analysis; S.D provided all histological analysis; R.B. performed bioinformatics WES analysis and manuscript preparation; R.A.S. provided intellectual input and provided conditioned medium. T.F.B. oversaw the experimental design and data analysis and contributed to manuscript preparation; C.J.B provided intellectual input, oversaw the experimental design and data analysis and contributed to manuscript preparation.

Acknowledgements

This work was supported by NIH/NCI R01CA204173 to CJB and A Breath of Hope Foundation Award and American Lung Association Lung Cancer Discovery Award (LCD-257864) to LPS and TFB. LPS and RB acknowledge funding from the Hillman Fellows for Innovative Cancer Research Program. LPS acknowledges funding from Libby's Lungs. This study utilized the UPMC Hillman Cancer Center Animal Facility, Flow Cytometry Facility, In Vivo Imaging Facility and Cancer Bioinformatics Facility, supported by Cancer Center Support Grant NIH/NCI P30CA047904-32. We thank Adam Soloff for technical assistance with flow cytometry.

References

1. Somasundaram A, Burns TF. The next generation of immunotherapy: keeping lung cancer in check. *Journal of hematology & oncology*. 2017;10(1):87.
2. Topalian SL, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *The New England journal of medicine*. 2012;366(26):2443-2454.
3. Brahmer JR, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *The New England journal of medicine*. 2012;366(26):2455-2465.
4. Samstein RM, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nature genetics*. 2019;51(2):202-206.
5. Herbst RS, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature*. 2014;515(7528):563-567.
6. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nature reviews Cancer*. 2019;19(3):133-150.
7. Kellar A, Egan C, Morris D. Preclinical Murine Models for Lung Cancer: Clinical Trial Applications. *Biomed Res Int*. 2015;2015:621324.
8. Layton MG, Franks LM. Heterogeneity in a spontaneous mouse lung carcinoma: selection and characterisation of stable metastatic variants. *British journal of cancer*. 1984;49(4):415-421.

9. Patel MR, et al. Vesicular stomatitis virus expressing interferon-beta is oncolytic and promotes antitumor immune responses in a syngeneic murine model of non-small cell lung cancer. *Oncotarget*. 2015;6(32):33165-33177.
10. Malkinson AM, Dwyer-Nield LD, Rice PL, Dinsdale D. Mouse lung epithelial cell lines--tools for the study of differentiation and the neoplastic phenotype. *Toxicology*. 1997;123(1-2):53-100.
11. Rothenberger NJ, Stabile LP. Induction of Lung Tumors and Mutational Analysis in FVB/N Mice Treated with the Tobacco Carcinogen 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanone. *Methods in molecular biology (Clifton, NJ)*. 2020;2102:149-160.
12. Hecht SS. Tobacco smoke carcinogens and lung cancer. *Journal of the National Cancer Institute*. 1999;91(14):1194-1210.
13. Xue J, Yang S, Seng S. Mechanisms of Cancer Induction by Tobacco-Specific NNK and NNN. *Cancers*. 2014;6(2):1138-1156.
14. Stabile LP, Lyker JS, Land SR, Dacic S, Zamboni BA, Siegfried JM. Transgenic mice overexpressing hepatocyte growth factor in the airways show increased susceptibility to lung cancer. *Carcinogenesis*. 2006;27(8):1547-1555.
15. Panchenko AV, et al. Biomarkers of aging, life span and spontaneous carcinogenesis in the wild type and HER-2 transgenic FVB/N female mice. *Biogerontology*. 2016;17(2):317-324.
16. Tran PT, et al. Twist1 suppresses senescence programs and thereby accelerates and maintains mutant Kras-induced lung tumorigenesis. *PLoS genetics*. 2012;8(5):e1002650.
17. Siegfried JM, Farooqui M, Rothenberger NJ, Dacic S, Stabile LP. Interaction between the estrogen receptor and fibroblast growth factor receptor pathways in non-small cell lung cancer. *Oncotarget*. 2017;8(15):24063-24076.
18. Stabile LP, et al. Prevention of tobacco carcinogen-induced lung cancer in female mice using antiestrogens. *Carcinogenesis*. 2012;33(11):2181-2189.
19. Sivakumar S, et al. Genomic Landscape of Atypical Adenomatous Hyperplasia Reveals Divergent Modes to Lung Adenocarcinoma. *Cancer Res*. 2017;77(22):6119-6130.
20. Stabile LP, et al. Therapeutic targeting of human hepatocyte growth factor with a single neutralizing monoclonal antibody reduces lung tumorigenesis. *Molecular cancer therapeutics*. 2008;7(7):1913-1922.
21. Fujimoto J, Kadara H, Men T, van Pelt C, Lotan D, Lotan R. Comparative functional genomics analysis of NNK tobacco-carcinogen induced lung adenocarcinoma development in Gprc5a-knockout mice. *PloS one*. 2010;5(7):e11847.
22. Alexandrov LB, et al. Mutational signatures associated with tobacco smoking in human cancer. *Science (New York, NY)*. 2016;354(6312):618-622.
23. Stabile LP, Rothstein ME, Gubish CT, Cunningham DE, Lee N, Siegfried JM. Co-targeting c-Met and COX-2 leads to enhanced inhibition of lung tumorigenesis in a murine model with heightened airway HGF. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2014;9(9):1285-1293.
24. Alexandrov LB, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463):415-421.
25. Tomita-Mitchell A, Kat AG, Marcelino LA, Li-Sucholeiki XC, Goodluck-Griffith J, Thilly WG. Mismatch repair deficient human cells: spontaneous and MNG-inducible mutational spectra in the HPRT gene. *Mutation research*. 2000;450(1-2):125-138.
26. Mijal RS, Loktionova NA, Vu CC, Pegg AE, Peterson LA. O6-pyridyloxobutylguanine adducts contribute to the mutagenic properties of pyridyloxobutylating agents. *Chem Res Toxicol*. 2005;18(10):1619-1625.

27. Ding L, et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature*. 2008;455(7216):1069-1075.
28. El Osta B, et al. Characteristics and Outcomes of Patients With Metastatic KRAS-Mutant Lung Adenocarcinomas: The Lung Cancer Mutation Consortium Experience. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2019;14(5):876-889.
29. Halvorsen AR, et al. TP53 Mutation Spectrum in Smokers and Never Smoking Lung Cancer Patients. *Frontiers in genetics*. 2016;7:85.
30. Vendetti FP, et al. ATR kinase inhibitor AZD6738 potentiates CD8+ T cell-dependent antitumor activity following radiation. *The Journal of clinical investigation*. 2018;128(9):3926-3940.
31. Hecht SS, Hoffmann D. Tobacco-specific nitrosamines, an important group of carcinogens in tobacco and tobacco smoke. *Carcinogenesis*. 1988;9(6):875-884.
32. Hecht SS. Biochemistry, biology, and carcinogenicity of tobacco-specific N-nitrosamines. *Chem Res Toxicol*. 1998;11(6):559-603.
33. Humans IWGotEoCRt. Smokeless tobacco and some tobacco-specific N-nitrosamines. *IARC Monogr Eval Carcinog Risks Hum*. 2007;89:1-592.
34. Wei B, Blount BC, Xia B, Wang L. Assessing exposure to tobacco-specific carcinogen NNK using its urinary metabolite NNAL measured in US population: 2011-2012. *Journal of exposure science & environmental epidemiology*. 2016;26(3):249-256.
35. Yuan JM, et al. Urinary levels of cigarette smoke constituent metabolites are prospectively associated with lung cancer development in smokers. *Cancer Res*. 2011;71(21):6749-6757.
36. Ge GZ, Xu TR, Chen C. Tobacco carcinogen NNK-induced lung cancer animal models and associated carcinogenic mechanisms. *Acta Biochim Biophys Sin (Shanghai)*. 2015;47(7):477-487.
37. West KA, et al. Rapid Akt activation by nicotine and a tobacco carcinogen modulates the phenotype of normal human airway epithelial cells. *J Clin Invest*. 2003;111(1):81-90.
38. Jull BA, Plummer HK, 3rd, Schuller HM. Nicotinic receptor-mediated activation by the tobacco-specific nitrosamine NNK of a Raf-1/MAP kinase pathway, resulting in phosphorylation of c-myc in human small cell lung carcinoma cells and pulmonary neuroendocrine cells. *J Cancer Res Clin Oncol*. 2001;127(12):707-717.
39. Schuller HM, Cekanova M. NNK-induced hamster lung adenocarcinomas over-express beta2-adrenergic and EGFR signaling pathways. *Lung Cancer*. 2005;49(1):35-45.
40. Laag E, Majidi M, Cekanova M, Masi T, Takahashi T, Schuller HM. NNK activates ERK1/2 and CREB/ATF-1 via beta-1-AR and EGFR signaling in human lung adenocarcinoma and small airway epithelial cells. *Int J Cancer*. 2006;119(7):1547-1552.
41. Roecklein BA, Torok-Storb B. Functionally distinct human marrow stromal cell lines immortalized by transduction with the human papilloma virus E6/E7 genes. *Blood*. 1995;85(4):997-1005.
42. Huang J, et al. Downregulation of estrogen receptor and modulation of growth of breast cancer cell lines mediated by paracrine stromal cell signals. *Breast cancer research and treatment*. 2017;161(2):229-243.
43. Mussano F, et al. Cytokine, Chemokine, and Growth Factor Profile Characterization of Undifferentiated and Osteoinduced Human Adipose-Derived Stem Cells. *Stem cells international*. 2017;2017:6202783.
44. Godfrey J, et al. PD-L1 gene alterations identify a subset of diffuse large B-cell lymphoma harboring a T-cell-inflamed phenotype. *Blood*. 2019;133(21):2279-2290.
45. Andrews S. FastQC: A quality control application for high throughput sequence data. *Babraham Institute Project page*: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. 2016.

46. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:13033997v2 [q-bioGN]*. 2013.
47. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics*. 2015;31(12):2032-2034.
48. Van der Auwera GA, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013;43:1110-1133.
49. Cibulskis K, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature biotechnology*. 2013;31(3):213-219.
50. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research*. 2010;38(16):e164.
51. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747-1756.
52. Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511(7511):543-550.

Figures and Legends

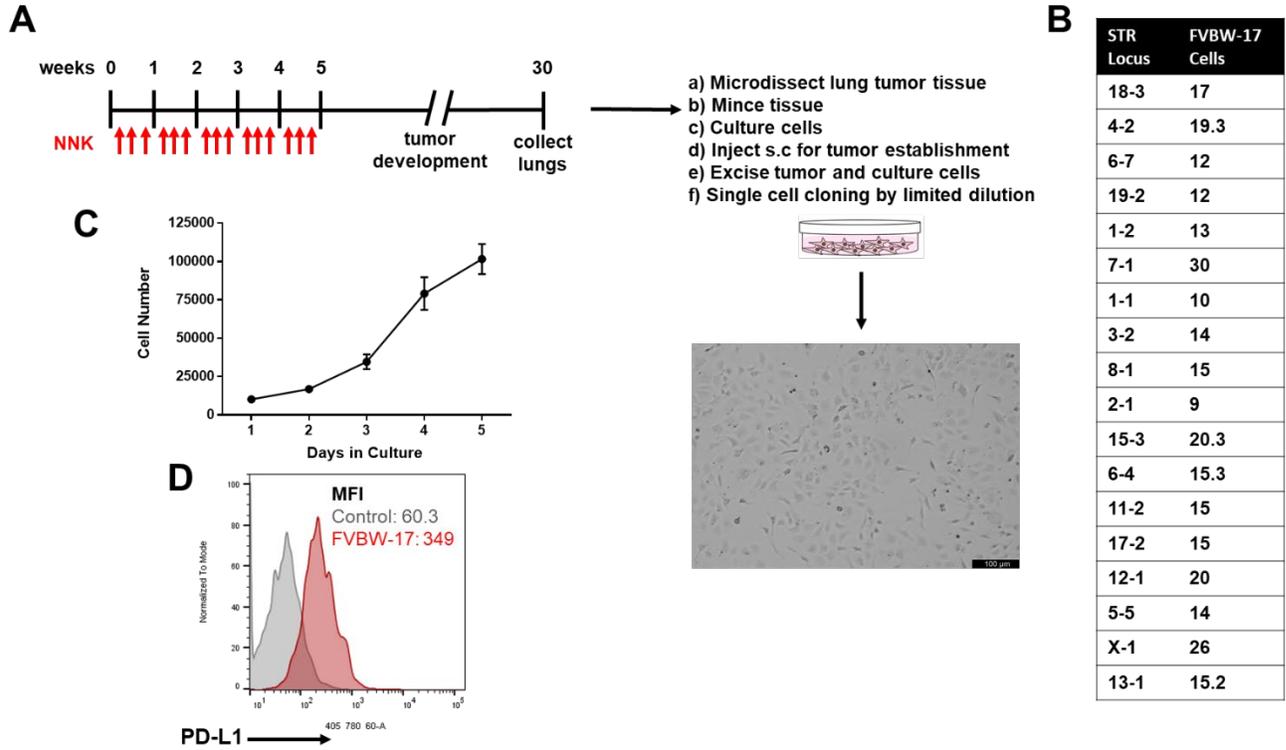


Figure 1. Generation and characterization of the murine FVBW-17 cell line. (A) Schema of the establishment of the parental FVBW-17 cell line and clonal lines and representative phase contrast microscopic image of FVBW-17 cells in culture. Scale bar, 100µm. **(B)** Cell growth properties of FVBW-17 cells in culture. Results are reported as the mean ± SD of 6 replicates. **(C)** STR profile of the FVBW-17 cell line. **(D)** Flow cytometry for PD-L1 expression of cultured cells. Results are expressed as mean fluorescence intensity (MFI) which is a measure of protein expression compared to the control (fluorescence minus one, which is viability dye without PD-L1 antibody).

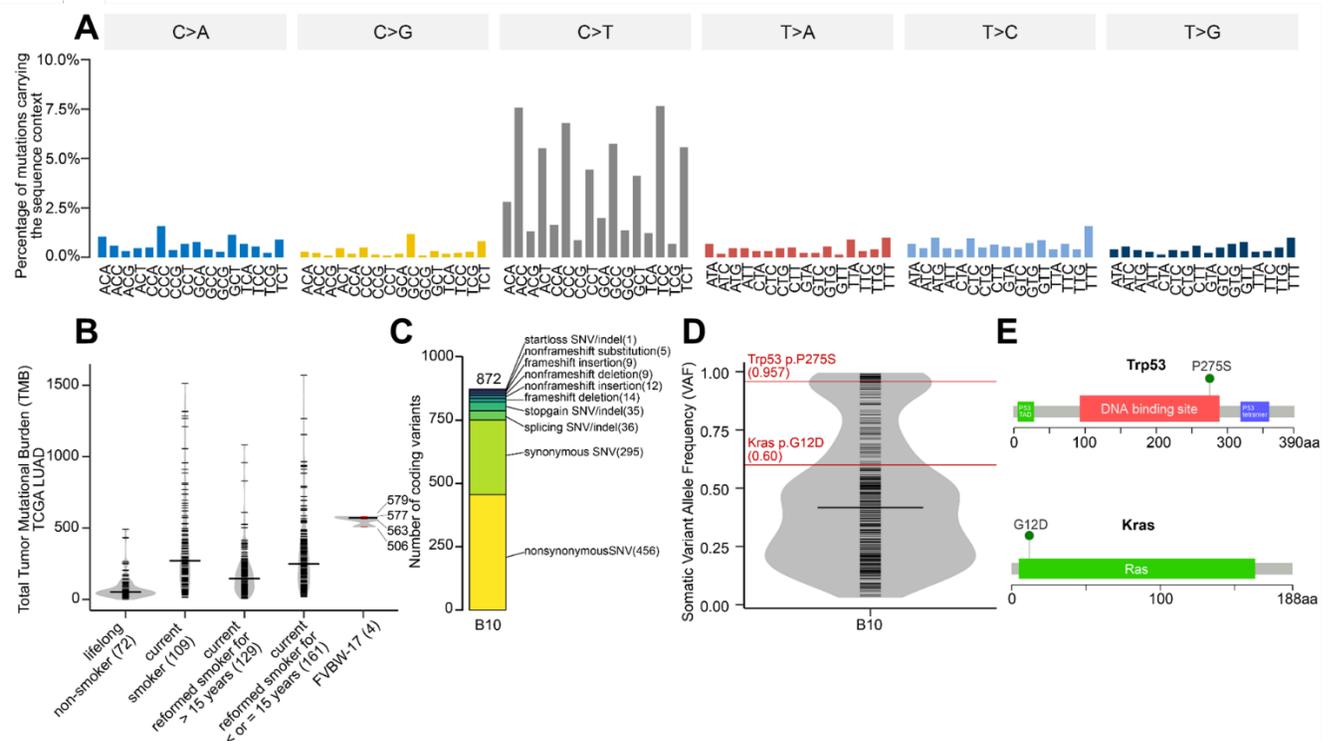


Figure 2. Somatic mutation profile of FVBW-17 represents a classic alkylating signature with validated *Kras* and *Trp53* mutations. (A) Pattern of the FVBW-17 mutation signature associated with NNK using a 96-substitution classification. Variants were grouped by the trinucleotide sequence context centered at the mutated base shown on the x-axis. The percentage of variants carrying the trinucleotide sequence context out of the whole mouse genome is shown on the y-axis. **(B)** Comparison of FVBW-17 TMB to that of primary tumors from TCGA LUAD patients grouped by tobacco smoking history. Each short red horizontal line indicates the TMB value of each FVBW-17 line (median= 570). The numbers on the right show the TMBs for each of the 4 lines. TMB was defined as the total number of somatic mutations predicted to affect protein sequences (NSSMs). $n=471$ samples shown. Each shorter black horizontal line represents one tumor. Longer black horizontal line represents median of TMB in each group. **(C)** Distribution of different variant types that affect the coding regions of the genome. Ten categories are shown based on exonic function annotation of each variant (synonymous

mutations included), with numbers from each category shown in between parenthesis. **(D)** Distribution of somatic VAF across coding variants from **c**. The VAFs of *Trp53* P275S and *Kras* G12D are shown as red horizontal lines. $n=872$ coding variants shown. Each shorter black horizontal line represents one variant. Longer black horizontal line represents median of VAF. **(E)** Lollipop plot showing the location of *Trp53* and *Kras* somatic mutations in the protein domains.

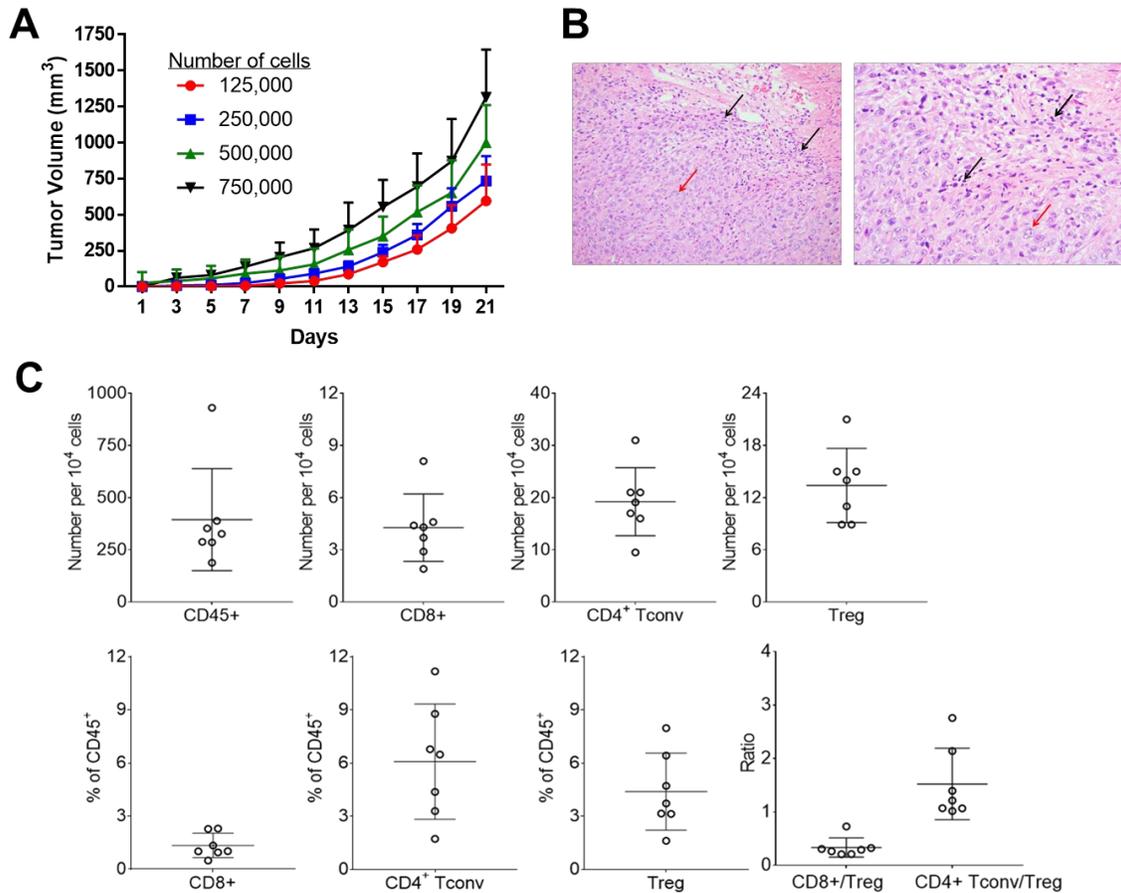


Figure 3. FVBW-17 cells grow as subcutaneous tumors in FVB/N mice, are histologically similar to human LUAD and contain tumor infiltrating lymphocytes. (A) 1.25×10^5 - 7.5×10^5 FVBW-17 cells were injected s.c. into hind flank in FVB/N mice (n=8-10 mice per cell density). Tumors were measured every 2 days for 21 days. Results represent the mean \pm SD. **(B)** Representative H&E images (4X, left panel; 20X; right panel) of s.c. FVBW-17 tumor showing malignant tumor with morphological features similar to human LUAD. Black arrows: representative TIL; Red arrow: representative tumor cell. Scale bar, 100 μ m. **(C)** Immunophenotyping of FVBW-17 cells grown as s.c. tumors collected at Day 21 (n= 7 mice). Each cell type is expressed as number per 10^4 cells (upper panel) or percentage of CD45+ cells (lower panel). The ratio of CD8+/Treg and CD4+/Treg is also shown. Results are expressed as the mean \pm SD.

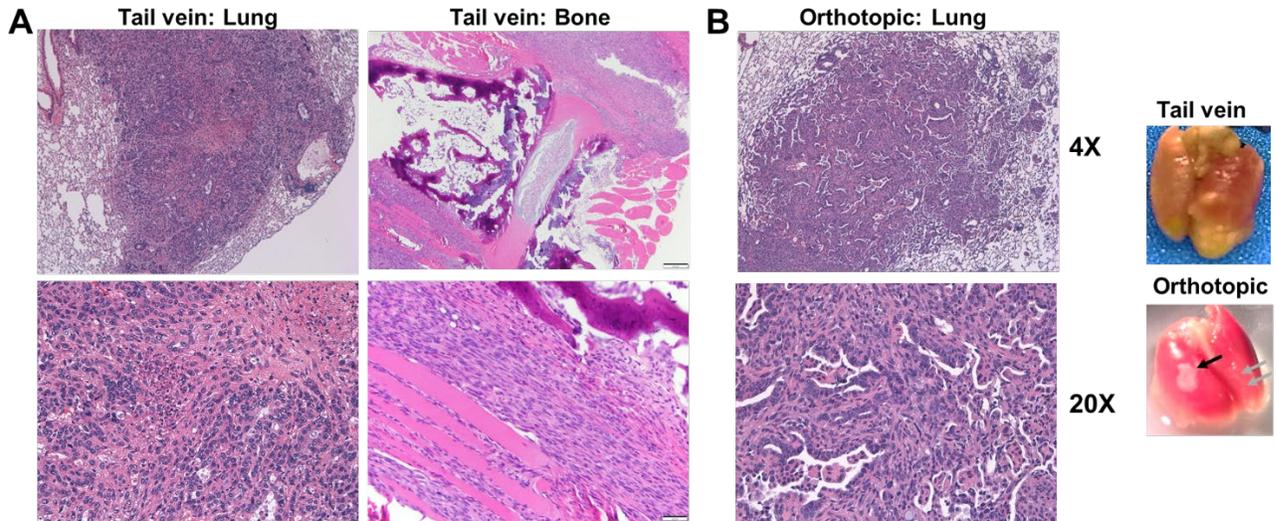


Figure 4. FVBW-17 cells seed the lung following intravenous injection and form lung nodules by direct orthotopic intrathoracic injection. (A) FVBW-17 cells (5×10^5) were injected into the lateral tail vein. Lungs were formalin-inflated 21 days post-injection for evidence of lung tumor burden by H&E, showing NSCLC with rare glandular structures and single cells invading the stroma and necrosis (left panels). Right panels show bone metastasis. Scale bar, 100 μm . **(B)** FVBW-17 tumor establishment in the lung 15 days after direct intrathoracic injection of 1×10^6 cells in PBS/1.35mg/ml matrigel. H&E shows evidence of invasive adenocarcinoma with malignant glandular structures. Scale bar, 100 μm . **(C)** Gross lung from tail vein model (top panel) or orthotopic model (bottom panel). Red arrow shows primary tumor, gray arrows show lung metastasis.

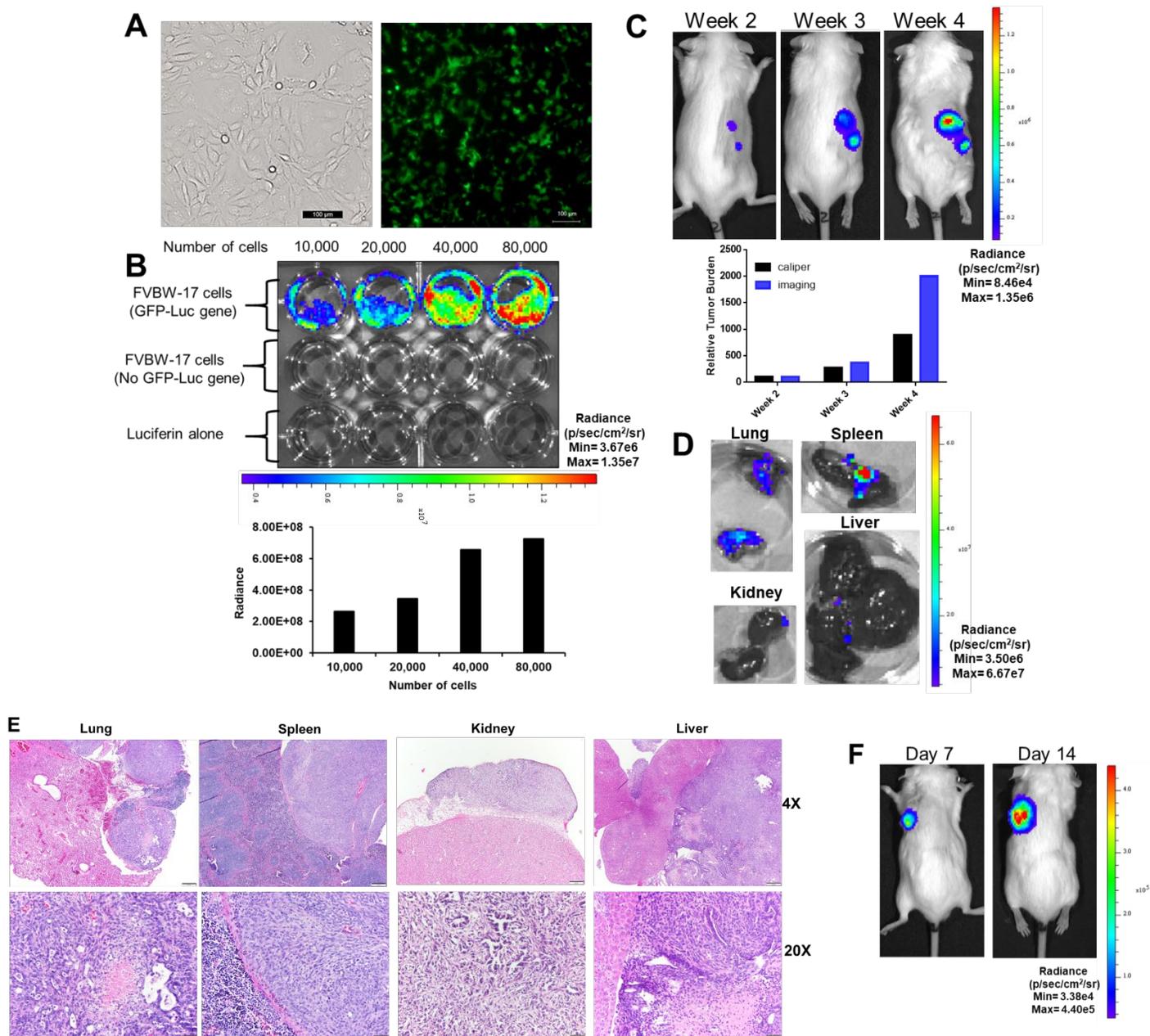


Figure 5. Bioluminescence imaging facilitates detection of subcutaneous tumor growth and metastasis. (A) FVBW-17 cells transduced with lentivirus for GFP and firefly Luc showed high transduction efficiency by comparing with phase contrast (left) and fluorescent microscopy (right). Scale bar, 100 μ m. (B) Luciferase assay of GFP-Luc cells and control cells with varying cell number and quantification of luciferase assay (radiance). (C) Subcutaneous injection of FVBW-17 GFP-Luc cells (2.5×10^5) and tumor growth over 4 weeks monitored in vivo by IVIS imaging. (D) Ex vivo bioluminescence imaging of various organs was used to identify metastatic

spread in lung, spleen, kidney and liver. **(E)** Sites of metastases found on IVIS imaging from subcutaneous FVBW-17 tumors confirmed by H&E histology (4X, top panel; 20X, bottom panel). Scale bar, 100 μ m. **(F)** Imaging of orthotopic injected cells.

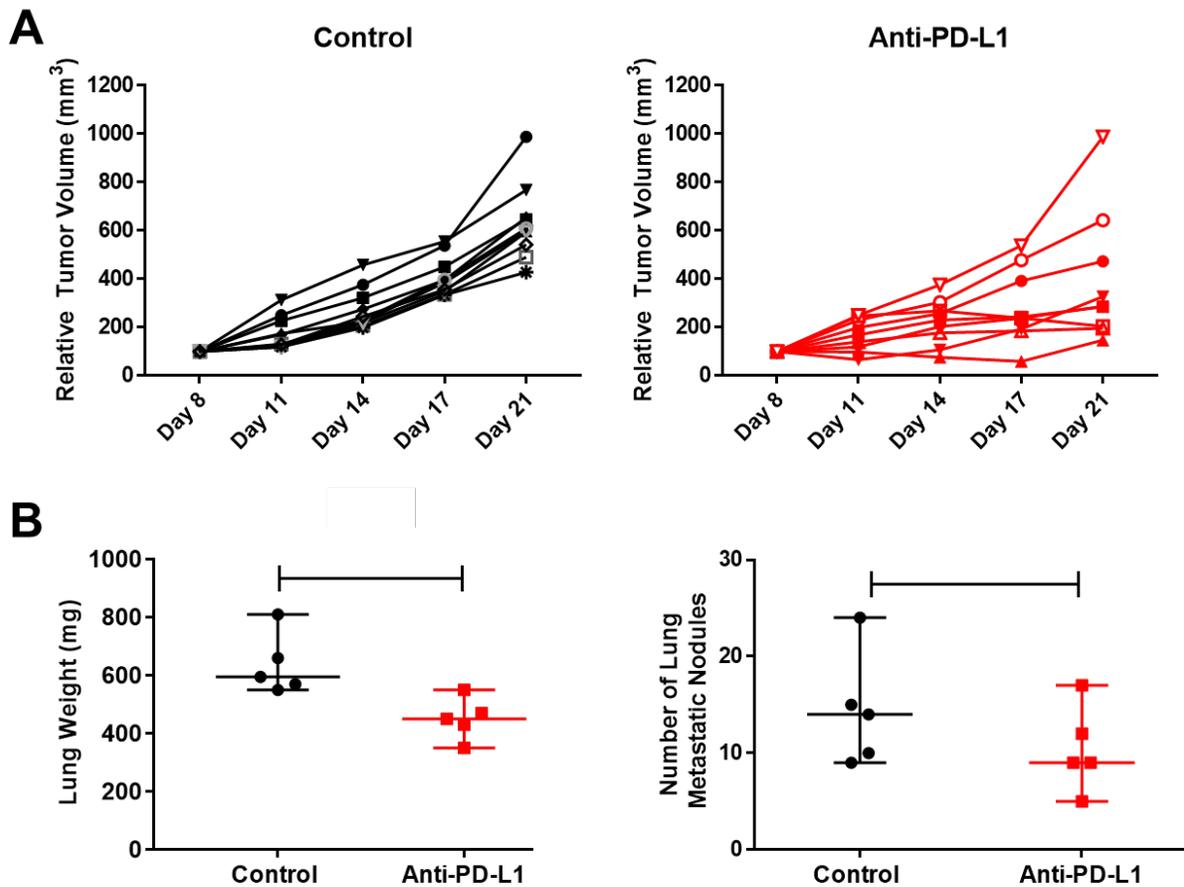


Figure 6. Growth kinetics of FVBW-17 cells treated with anti-PD-L1 in vivo. (A) 5×10^5 FVBW-17 cells were injected s.c. into FVBW-17 mice ($n=8-10$ per group) on Day 1. Palpable tumors were identified by Day 7. Anti-PD-L1 or isotype control antibody (10mg/kg) was administered 2 times per week for two weeks. Tumor measurements were recorded every 3 days and plotted as relative tumor volume for each individual tumor over time. Each line represents an individual tumor. (B) 1×10^6 FVBW-17 cells were injected IV into FVB/N mice ($n=5$ mice/group). Anti-PD-L1 or isotype control (10mg/kg) was administered 1 week following tumor injection and continued 2 times per week for 3 weeks. Both lung weight and number of lung metastatic nodules were evaluated. P-values calculated using Student's t-test.

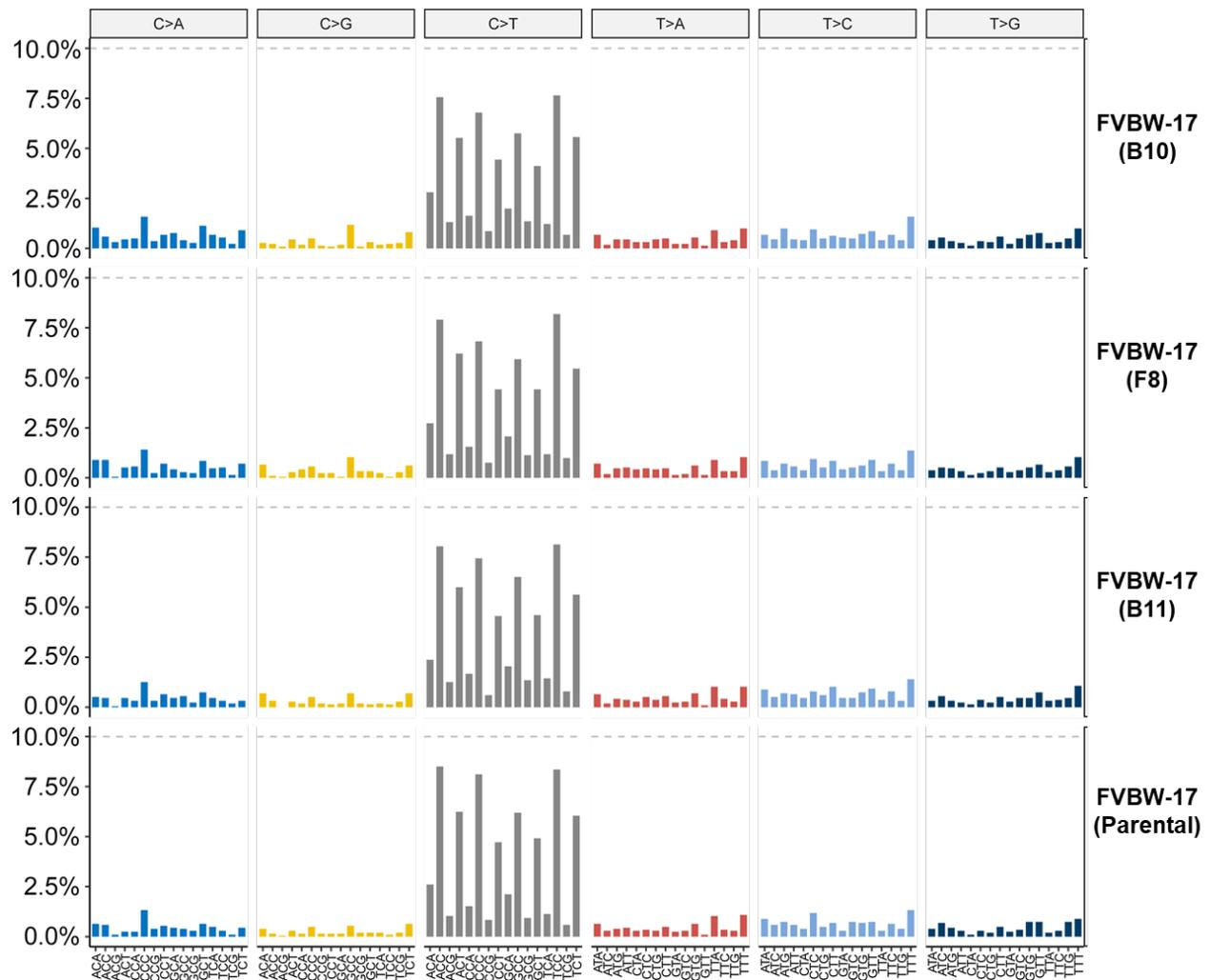


Figure S1. Mutation signature of all FVBW-17 cell lines associated with alkylating agents using a 96-substitution classification. FVBW-17 is the parental strain, and the other three are clonal cell lines.

Table S1. Number of mice and frequency of distant metastases of FVBW-17 cells from s.c. tumors.

	Lung	Spleen	Kidney	Liver
Male (n=15)	9 (60%)	3 (20%)	3 (20%)	2 (13%)
Female (n=15)	11 (73%)	4 (27%)	2 (13%)	2 (13%)

Table S2. Summary statistics of variant filtering steps.

Sample	Raw list of variants	After removing strain-specific polymorphisms (those present in FVB/NJ dbSNP v142)	Variants PASS (removing reads flagged by MuTect2 intrinsic filters in tumor-only mode or by read orientation filter)	Putative somatic variants in coding region (exonic or splicing)	Putative somatic variants predicted to affect protein sequences (a.k.a. non-synonymous somatic mutations, NSSMs)	NSSM single nucleotide variants (SNVs)	NSSM insertions/deletions (indels)
FVBW-17	176459	22316	3307	778	506	442	64
FVBW-17 B10	177689	23040	3519	873	577	504	73
FVBW-17 F8	177203	22745	3424	832	563	495	68
FVBW-17 B11	177149	22803	3414	874	579	506	73