Science Translational Medicine
AAAS

www.sciencetranslationalmedicine.org/cgi/content/full/8/350/350ra104/DC1

Supplementary Materials for

Targeted BMI1 inhibition impairs tumor growth in lung adenocarcinomas with low CEBPα expression

Kol Jia Yong, Daniela S. Basseres, Robert S. Welner, Wen Cai Zhang, Henry Yang, Benedict Yan, Meritxell Alberich-Jorda, Junyan Zhang, Lorena Lobo de Figueiredo-Pontes, Chiara Battelli, Christopher J. Hetherington, Min Ye, Hong Zhang, Giorgia Maroni, Karen O'Brien, Maria Cristina Magli, Alain C. Borczuk, Lyuba Varticovski, Olivier Kocher, Pu Zhang, Young-Choon Moon, Nadiya Sydorenko, Liangxian Cao, Thomas W. Davis, Bhavin M. Thakkar, Ross A. Soo, Atsushi Iwama, Bing Lim, Balazs Halmos, Donna Neuberg, Daniel G. Tenen,* Elena Levantini*

*Corresponding author. Email: elevanti@bidmc.harvard.edu (E.L.); csidgt@nus.edu.sg (D.G.T.)

Published 3 August 2016, *Sci. Transl. Med.* **8**, 350ra104 (2016)
DOI: 10.1126/scitranslmed.aad6066

This PDF file includes:

Supplementary Materials and Methods:

Human tissue microarray (TMA) construction and immunohistochemistry: A total of 261 formalin-fixed, paraffin-embedded NSCLC samples were assembled onto the U.S. Biomax LC6161 TMA in duplicate. Tissue microarray sections were subsequently used for C/EBPα and BMI1 immunohistochemical analyses. Sections were mounted on pre-coated slides and then dewaxed. Antigen retrieval was performed with 10 mM citrate buffer (pH-6.0) in a MicroMED TT Microwave processor (Milestone) for 5 min at 120° C. Slides were then incubated with primary antibodies, CEBPα (Cell Signaling, cat # 2295, 1:50 dilution) for one hour at room temperature and BMI1 (Millipore, #05-637, clone# F6, 1:200 dilution) overnight at 4º C. Tissues were then counterstained with hematoxylin. Both CEBPα and BMI1 showed nuclear staining. Immunohistochemistry slides were separately examined by two pathologists who were blinded to the patients' clinicopathological data. If discrepancies were found, consensus was reached using a conference microscope. Intensity of nuclear staining was scored as follows: 0 to 3^+; 0, no staining; 1^+, mild staining; 2^+, moderate staining; 3^+, strong staining. For C/EBPα, 0 and 1^+ were considered as negative/low expression, and 2^+ and 3^+ were considered positive expression. For BMI1, 0 was considered negative expression, and 1^+-3^+ were considered positive expression. Detailed clinicopathologic information was available for most of the patients.

Human cell line culture conditions: Human adenocarcinoma cells were grown in RPMI 1640 supplemented with 10% fetal bovine serum.

Generation of the human H358 inducible cell lines: The human adenocarcinoma cell line H358 was transfected with a rapamycin-inducible C/EBPα construct (pL_6N_2-R_HS3H-ZF2-PL-mC/EBPα) or empty vector (pL_6N_2-R_HS3H-ZF2-PL), from ARIAD Pharmaceuticals, using LipofectAMINE PLUS transfection reagent (Invitrogen Life Technologies, Inc.) according to the manufacturer's instructions. Clones were selected on the basis of G418 resistance (G418 concentration of 500 μg/ml) and screened for inducibility of C/EBPα

expression after 36 hours of rapalog (100 nM, AP21967 from ARGENT) induction by Taqman q-RT-PCR, as described below.

Murine models: For lung-specific, doxycycline-induced recombination, *CCSP reverse tetracycline transactivator* (*CCSP-rtTA*)-transgenic mice and *tet(O)7CMV-Cre*-transgenic mice (fig. S3A) were crossed with *C/EBPα*$^{loxP/loxP}$ mice (fig. S3A). Resulting mice were of mixed background (C57/black; 129; FVB/N; balb/c). The targeted *C/EBPα* alleles, as well as the *CCSP-rtTA* and *tet-O-Cre* transgenes, were genotyped by PCR as previously described [14,17]. *C/EBPα* excision was evaluated by Southern blot, as described below [14]. Litters were kept on doxycycline-containing water (1 mg/ml) for four weeks at weaning in amber bottles. *C/EBPα*$^{loxP/loxP}$ mice, which were positive for both the *CCSP* and *Cre* alleles and underwent doxycycline treatment with subsequent deletion of *C/EBPα* in lung cells, are referred to as *C/EBPα*$^{Lung-\Delta}$. To obtain *C/EBPα*$^{Lung-\Delta}$ mice with decreased expression of BMI1 (fig. S4A), *C/EBPα*$^{loxP/loxP}$ *CCSP-rtTA*$^{+}$ *Cre*$^{+}$ mice were bred to *Bmi1*$^{WT/GFP}$ mice in which *GFP* replaces exon 2 of the *Bmi1* gene [21], and treated with doxycycline, as described above for the *C/EBPα*$^{Lung-\Delta}$ strain. *NOD-SCID IL2Rγ*(null) mice (NSG mice, Jackson Laboratories), which lack T, B, and natural killer cells, were subcutaneously injected with 3.5 x 10^5 tumor cells and 25 µl of Matrigel (BD Basement Membrane Matrix Phenol-red free #356237). Four weeks after injection, tumors were extracted and analyzed by IHC and q-RT-PCR. To study the in vivo effects of *Bmi1* knock-down on tumor cells, NSG mice were transplanted with cells transduced with two different GFP-coupled-shRNAs against murine *Bmi1* (sh1, and sh2)[44], with a control GFP-coupled shRNA (Luciferase-sh), or left untreated. GFP percentage within tumor cells and expression of BMI1 were assessed both before transplantation and at harvest 1 month later. To determine tumor volume by caliper measurement, the greatest longitudinal diameter (length) and the greatest transverse diameter (width) were determined. Tumor volume was calculated by the modified ellipsoidal formula: tumor volume= ½ (length x width^2) [45] . Tumors were first assessed when volume was at least 0.06 cm^3.

Mice were housed in a sterile-barrier facility, and all experiments were approved by the Institutional Animal Care and Use Committee at the Beth Israel Deaconess Medical Center.

Generation and culture conditions for the *C/EBPα* null pulmonary murine tumor cell line: Lung tumors from *C/EBPα*$^{\text{Lung-}\Delta}$ mice were dissected under aseptic conditions, chopped into small pieces, and incubated in 1 mg/mL collagenase/dispase (Roche Applied Science) and 0.001% DNAse (Sigma-Aldrich) for 3 hours in a 37ºC water bath. After incubation, the suspensions were passed through a 40 μm cell strainer (BD Falcon). The cells were centrifuged at 4° C for 20 minutes at 1,500 rpm and then suspended in red blood cell lysis buffer (eBioscience). After lysis, cells were suspended in DMEM/F12 medium and Matrigel (BD) (1:1). The cells were then subcutaneously transplanted into 4-6 week old NSG mice. Tumor xenografts were surgically removed when they reached 1 cm diameter. The xenografts were made into single cell suspension as described above. To derive the cell line, the cells were resuspended in complete culture medium: DMEM/F12 (Gibco #11330032) containing bovine serum albumin (BSA, 4 mg/mL, Sigma-Aldrich #A8806), ITS Universal Culture Supplement Premix (BD Biosciences #354350), epidermal growth factor (20 ng/mL, Invitrogen #13247051), basic fibroblast growth factor (10 ng/mL, Gibco #13256029), glutamine (2.5 mM), HEPES (15 mM), non-essential amino acids (1%), sodium pyruvate (1%, Gibco #11360), and antibiotic-antimycotic (1%, Invitrogen #15240). Cells were plated at 10,000 cells/cm^2. Fresh medium was replenished every 3 days until cells were 80% confluent. For passaging, cells were digested by accutase (Chemicon) into single cells and re-plated at a ratio of 1 to 3.

Histopathological analysis of murine samples: Mice were sacrificed by CO_2 euthanasia. Lungs were inflation fixed with 2.5 ml of 10% formalin. Xenograft tumors were extracted from under the skin of NSG mice and fixed, as described for lung tissues. Fixed specimens were embedded in paraffin and sectioned at 5 μm thickness. Tissue sections were stained with hematoxylin-eosin (HE) and the following antibodies: anti-C/EBPα (Cell Signaling #2295) diluted 1:50 and anti-BMI1 (Bethyl Laboratories Inc. #IHC-006-06) diluted 1:400.

Briefly, tissue sections were deparaffinized with Histoclear and hydrated in graded ethanols. Antigen retrieval was performed by boiling at 120°C in Citrate Buffer Target Retrieval Solution for 5 minutes in a pressure cooker. Non-specific signal was blocked by Peroxidase Blocking Solution for 10 minutes at room temperature (RT), followed by Protein Blocking Solution for 30 minutes at RT. Antibodies were incubated at RT for one hour in a humidified chamber, followed by 30 minutes of HRP-conjugated secondary antibody incubation at RT. Antibody binding was revealed by DAB, and the reaction was stopped by immersion of tissue sections in distilled water once brown signal appeared. Tissue sections were counterstained with hematoxylin, dehydrated in graded ethanols, and mounted. All reagents for immunohistochemistry were from Dako (Denmark A/S), unless otherwise specified.

Southern blot analysis: Fifteen micrograms of genomic DNA was digested overnight with BamHI (New England Biolabs), separated by 0.6% agarose gel electrophoresis, transferred to positively charged Biodyne nylon membranes (Pall Corp.) with 0.4 M NaOH overnight, and immobilized with a UV cross-linker (Stratagene). The nylon membranes were then hybridized to an 875-bp PstI-XbaI DNA probe located outside the targeted allele (Fig. 2C) [14].

RNA isolation and quantitative real-time PCR: RNA was extracted using the RNeasy Mini Kit (Qiagen). Real time analysis was performed on a Corbett Rotor Gene 6000. For quantitation, one-step reverse transcription PCR was performed using TaqMan Universal One-Step Mastermix (Applied Biosystems). Primers and probes were synthesized, and gene expression was compared with *18S* expression (Eukaryotic 18S rRNA endogenous control, Applied Biosystems). Human *C/EBPα* and *BMI1* sequences were as follows: human *C/EBPα* sense: 5'- TCGGTGGACAAGAACAG-3'; antisense 5'-GCAGGCGGTCATTG; probe 5'-FAM-TGGAGACGCAGCAGAAGGTG-BHQ1-3'; human *BMI1* sense: 5'-CCAGGGCTTTTCAAAAATGA -3'; antisense 5'-CCGATCCAATCTGTTCTGGT-3'; probe 5'-FAM-TGCTGATGCTGCCAATGGCTC- BHQ1-3' (set located across Exons 4-7). Murine *C/EBPα* sense: 5'-GACCATTAGCCTTGTGTGTACTGTATG-3';

antisense: 5'-TGGATCGATTGTGCTTCAAGTT-3'; probe: 5'-FAM-CGCCAGCCGCTGTTGCTGAA-BHQ-3' (set located in 3' UTR); murine *Bmi1* sense: 5'-CCAGCAAGTATTGTCCTATTTGTGA-3'; antisense: 5'-ATATCTTGAAGAGTTTTATCTGACCTT ATGTT-3'; probe: 5'-FAM-TCCAGGTTCACAAAACCAGACCACTCCT-BHQ-3' (set located in Exons 3-4).

Overexpression of C/EBPα in the *C/EBPα* null murine tumor cell line: A MSCV-C/EBPα-IRES-GFP retroviral construct and a control MSCV-IRES-GFP construct were used to transduce the *C/EBPα* null pulmonary tumor cell line, at MOI (multiplicity of infection) 20. GFP^+ cells were FACS-purified at 24 and 36 hours after infection and *C/EBPα* over-expression verified by SYBR Green q-RT-PCR with the following primers: sense 5'-ACGAGTTCCTGGCCGACCT-3'; antisense 5'-GGGCTCCCGGGTAGTCAAAG-3'.

Fluorescence activated cell sorting and analysis of pulmonary cells: Mice were sacrificed by CO_2 euthanasia. After tracheal injection of 2 ml dispase (BD Biosciences # 354235) with a 20G needle just until the gentle inflation of lungs, a second injection of 0.5-1 ml 1% low-melt agarose (made in water) was performed. Normal lung tissue or pulmonary tumors were then dissected out and chopped into small pieces. To obtain single cell suspensions, tissue pieces were incubated with 0.001% DNAse (Roche Applied Science # 10104159001), and 2 mg/ml collagenase/dispase (Roche Applied Science # 11097113001) in PBS 20% FBS, 37 C, 45-60 minutes, rotating. After filtering through 100 μm and 40 μm filters, cells were pelleted at 1200 rpm for 8 min and treated with ACK buffer to eliminate red blood cells. Cells were stained with different combinations of the following antibodies: anti-CD45.1 (1:100, APC-conjugated, Biolegend # 110714), anti-CD45.2 (1:100, APC-conjugated, Biolegend # 109814), anti-CD31 (1:100, APC-conjugated, Biolegend # 102410), and sorted according to standard procedures on MOFlo (MOFlo-MLS, Cytomation) or FACS Aria (BD Biosciences). Dead cells were excluded by propidium iodide staining. Data were analyzed with FlowJo software (Treestar, Inc.).

Drug treatment: The BMI1 inhibitor molecule (PTC-209) has been developed by PTC Therapeutics. The *C/EBPα* null cell line was treated for 48 hours with two PTC-209 concentrations (0.7 μM and 1.5 μM), and 0.5% DMSO as control. Mice carrying allograft or xenograft tumors were treated daily (fat pad injection) with 50 mg/kg of PTC-209 or vehicle control (DMSO/propanediol/polyethylene glycol 300) for 2-4 weeks. To determine tumor volume by caliper measurement, the greatest longitudinal diameter (length) and the greatest transverse diameter (width) were determined. Tumor volume was calculated by the modified ellipsoidal formula: tumor volume= ½ (length x width^2) [45] .

Western blot: Thirty micrograms of total protein were separated by SDS-PAGE, transferred to nitrocellulose, and blocked in TBS containing 5% non-fat dry milk. Proteins were immunoblotted overnight at 4C with a mouse monoclonal anti-BMI1 antibody (Millipore 05-637, clone #F6) at a dilution of 1:1000 or a rabbit anti-CEBPα antibody (Cell Signaling, cat # 2295) at a 1:1000 dilution. The EED antibody (R&D AF5827) was used at a 1:200 dilution in 5% BSA; the RING1 antibody (Millipore 05-1362, clone 4D6) was used at a 1:1000 dilution in 5% BSA; the EZH1 antibody (Millipore ABE281) was used at a 1:1000 dilution in 5% BSA; the EZH2 antibody (Cell Signaling 3147) at 1:500 dilution in 5% BSA. Membranes were then stripped and incubated for 1 h at RT with an anti-actin goat polyclonal antibody (Santa Cruz #1616) at a 1:1000 dilution to assess equal loading.

Cell cycle analysis with Hoechst 33342 and Pyronin Y: A combination of Hoechst 33342 and Pyronin Y was used for the differential staining of cellular DNA and RNA. Cells from the *C/EBPα* null line were permeabilized in phosphate-citrate buffer solution (pH 4.8), washed, and then resuspended in a solution of 2 μg/ml Hoechst 33342 (Molecular Probes Invitrogen) and 4 μg/ml pyronin Y (Polysciences). Cycle status was then evaluated by flow cytometry on a BD Aria II (BD Biosciences).

Lentivirus production and infection: Briefly, a lentivirus vector (CS-H1-shRNA-EF-1α-EGFP) expressing shRNA against mouse *Bmi1* (target sequence TAAAGGATTACTACACGCTAATG for sh-1; GAAGAG

GAACCTTTAAAGGATTA for sh-2)[44], and *Luciferase* was prepared. 293T cells were cotransfected using Lipofectamine 2000 with lentiviral packaging constructs (Gag-Pol and VSV-G Env). Virus was harvested and concentrated using a Centricon Plus-70 100000 MWCO column (Millipore). A single lentiviral transduction was performed in culture dishes (Falcon 1008; Becton Dickinson) in the presence of polybrene (8 μg/ml) (Sigma). The *C/EBPα* null pulmonary tumor cell line was infected with a MOI=20 for 18 hours, washed, and transplanted into NSG mice (0.35 x10^6 cells/injection). Because the efficiency of infection was extremely high and comparable among the three lentiviral constructs (GFP^+ cells ranged between 97.3 ±1.9% and 98.1 ±1.3%) (fig. S4E and F), cells did not need to be FACS-purified into GFP^+ cells before injection.

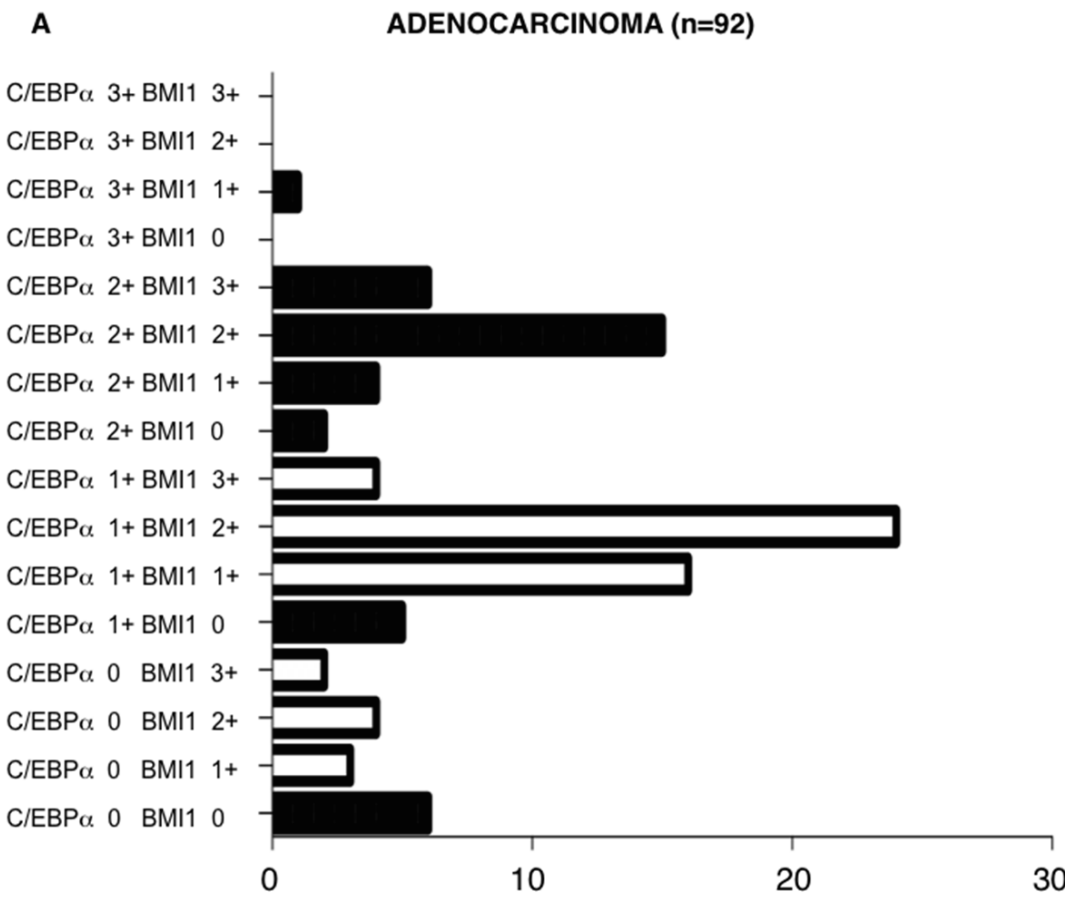
Microarray analysis: Total RNA was purified using a QIAGEN RNeasy Plus Micro kit, and cDNA was amplified by the Nugen Ovation Pico WTA System. The cDNA library was labeled and hybridized to Affymetrix Mouse 1.0 GeneChip with triplicates for both treated samples (PTC-209) and controls (Accession number: GSE56935). All raw data CEL files were analyzed together using the Robust Multichip Average method to obtain the gene expression intensities. Normalization was then performed across all samples based on the Cross Correlation method [46], and the normalized intensity data were further log2-tranformed and median-subtracted for heatmap profiling. Differentially expressed genes were then derived using a fold change cutoff of 2 and an FDR-adjusted p value cutoff of 0.05. Gene set enrichment analysis (GSEA) was performed based on the normalized mouse treatment data (DMSO versus PTC-209) using GSEA v2.0 tool (http://www.broad.mit.edu/gsea/) with the gene set database msigdb.v4.0 [23]. The normalized data set was split into two sample sets (DMSO and PTC-209) describing BMI1 up-regulated and down-regulated gene targets. Enriched gene sets between these two sample sets were identified by using the Kolmogorov-Smirnov statistics with a *p*-value cutoff of 0.05 and an FDR-derived *q*-value cutoff of 0.15.

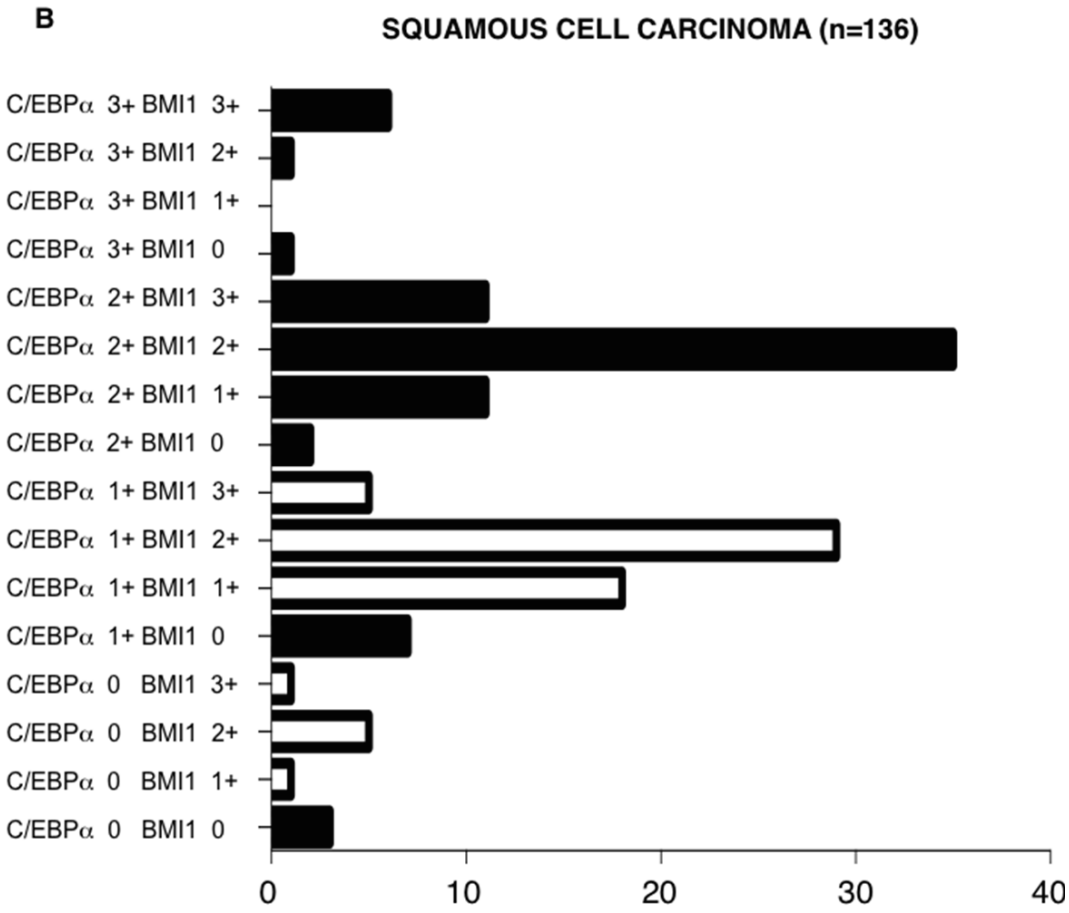
Survival data analysis: We downloaded raw gene expression data for three independent cohorts. After cross-correlation normalization for the cohorts (GSE14814, GSE50081, & GSE42127) generated from microarray platforms or total counts normalization of RNA-seq data for TCGA lung adenocarcinoma cohort, we used the

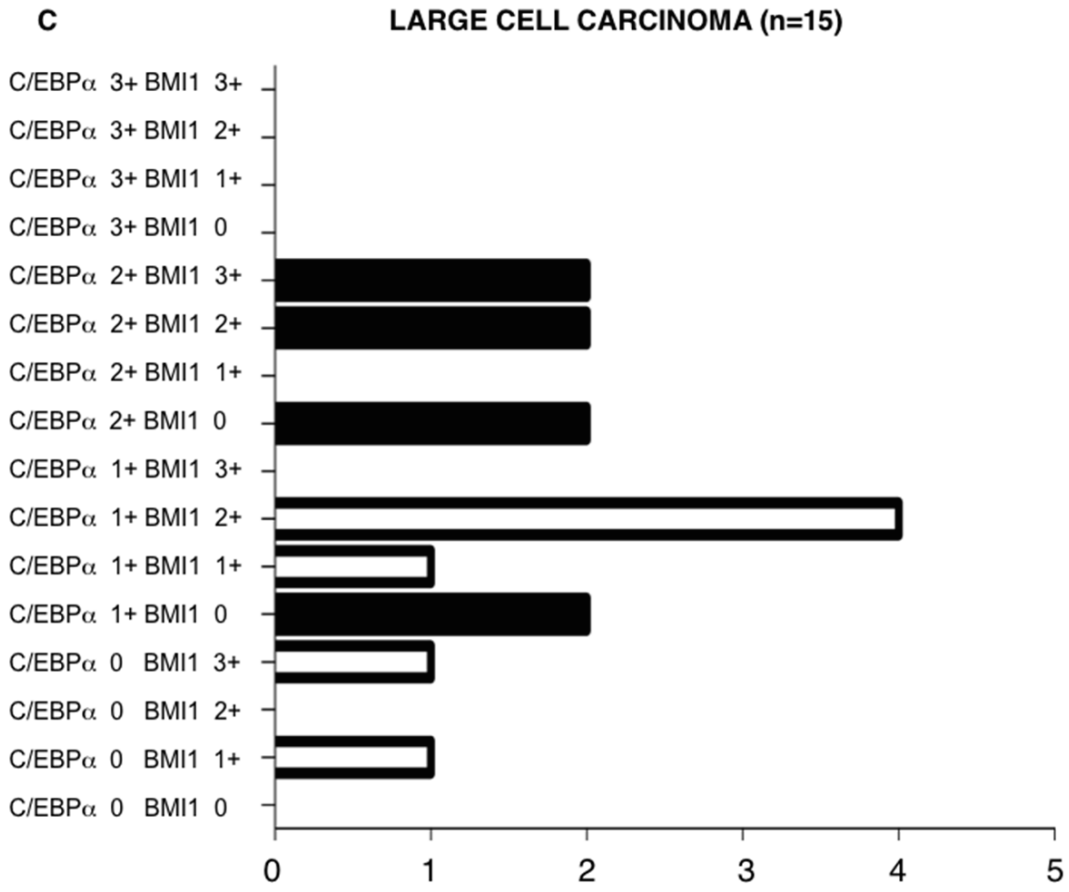
median *BMI1* expression to define *BMI1*^{low} and *BMI1*^{high} patient subgroups. Using the median to segregate C/EBPα high and low patients automatically assumes that equal numbers of patients have high and low C/EBPα expression. However, cancer gene expression in a cohort can be bimodal, consisting of two separate distributions with unequal masses or weightings [47,48]. Because the histograms of C/EBPα expression in all cohorts show a bimodal pattern, we used a bimodal distribution to model CEBPα expression for each cohort, and thus to segregate C/EBPα high and low patients. The Kaplan-Meier method was used for graphical displays of overall survival. The log rank test was used to assess differences in overall survival.

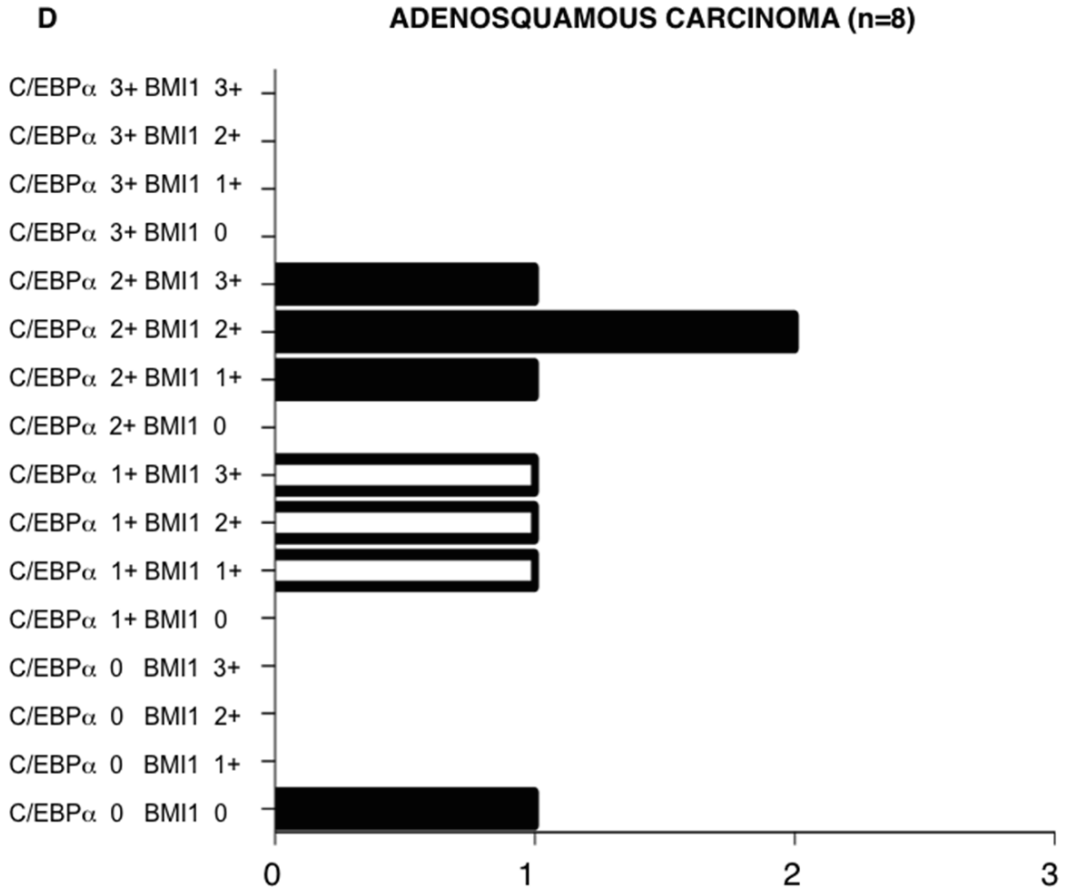
Statistical analysis: For comparison of continuous variables between groups, we used 2-sided Welch's t-test unless otherwise stated. Differences were considered statistically significant at $P < 0.05$ (indicated by asterisks). The association between categorical variables was investigated with Fisher's exact test for binary variables. For survival data analysis, we used cross-correlation normalization [46] and Kaplan-Meier method. Statistical analyses were performed in PASW Statistics 18 (SPSS Inc.) and R version 3.2.3 (The R Foundation for Statistical Computing) at 5% significance level. The specific details are provided in the text and in each figure legend.

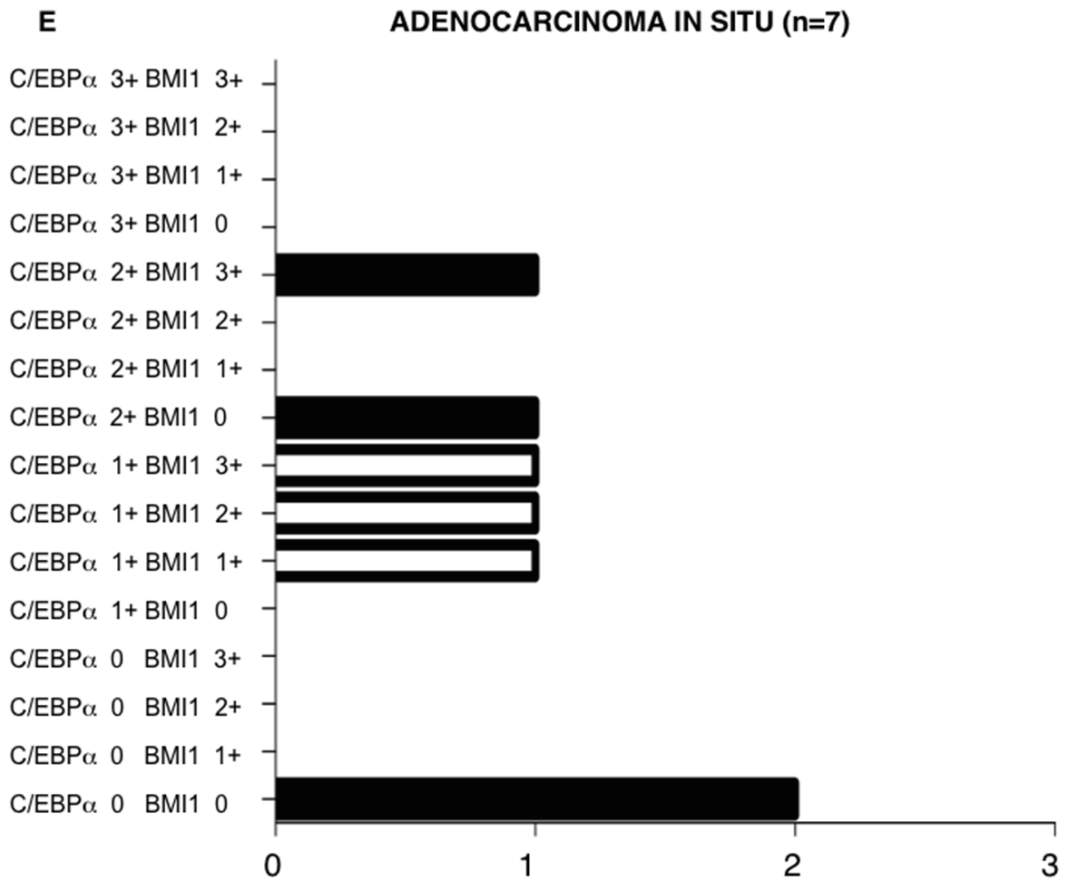
Supplementary Figure 1

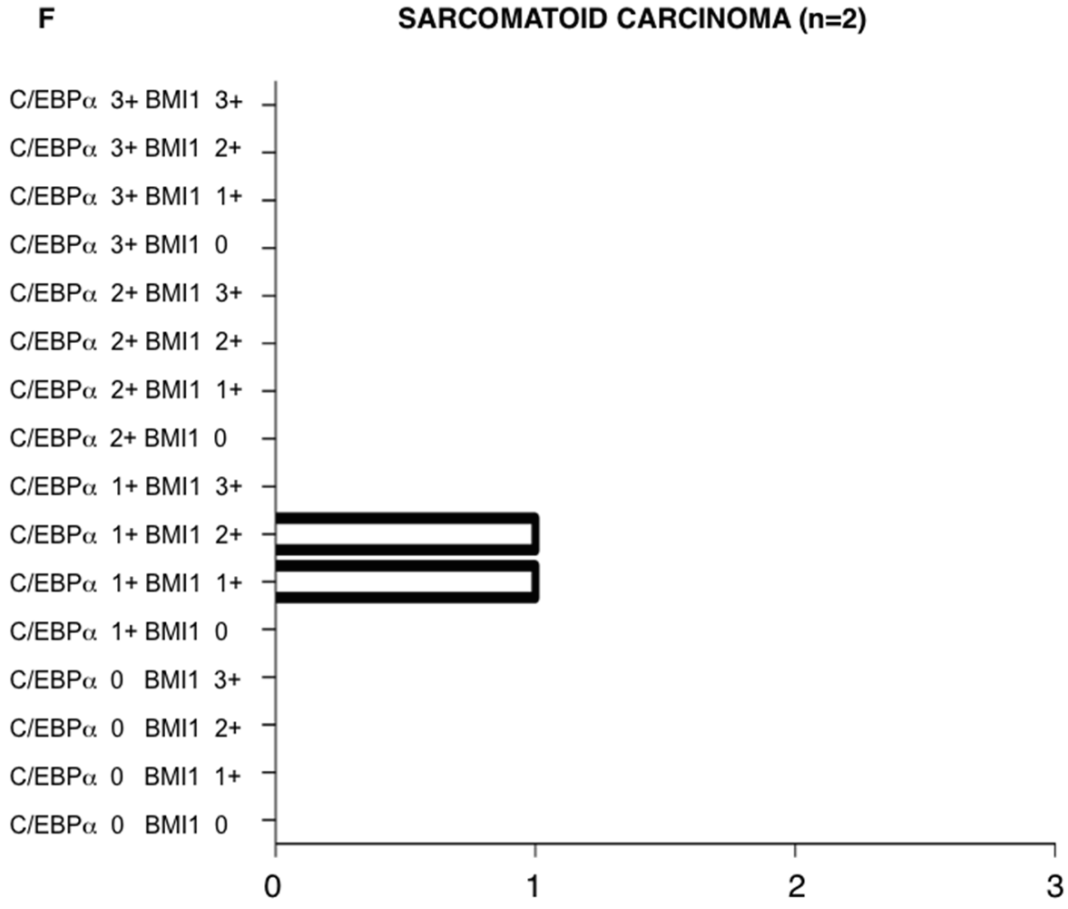


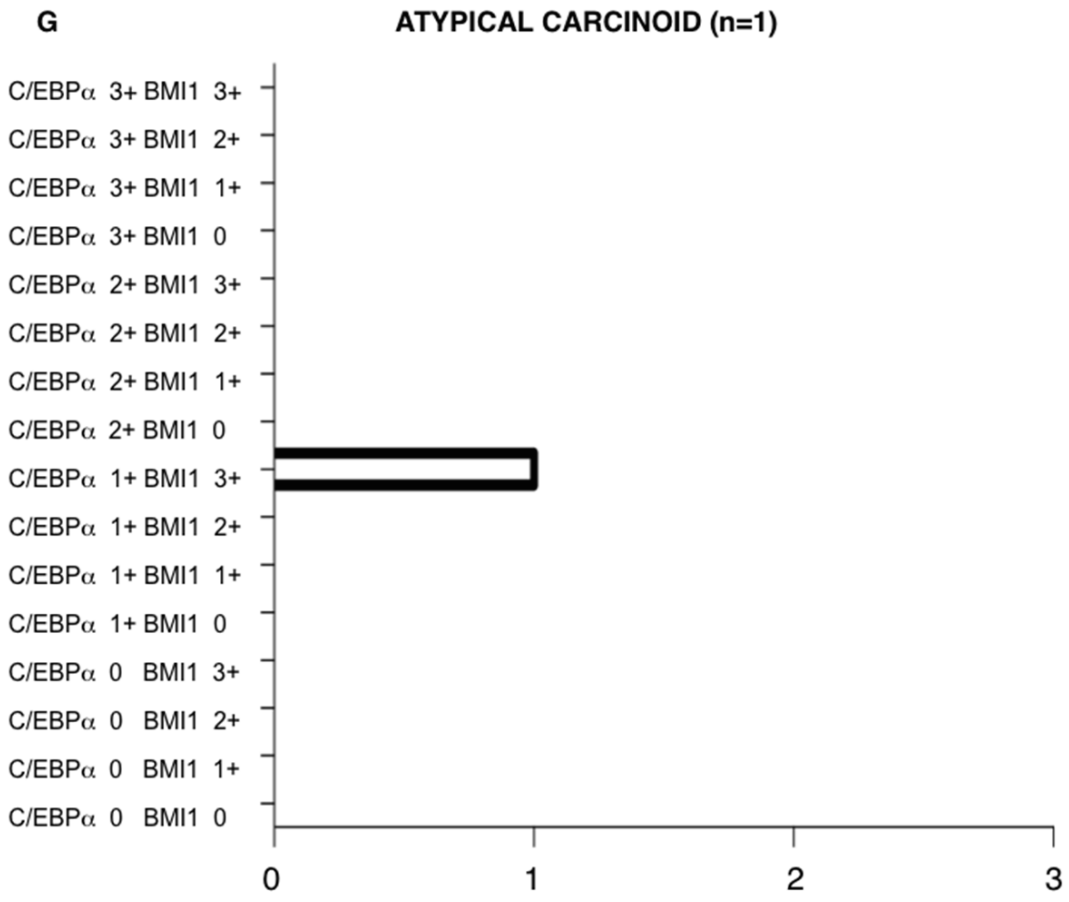



C
LARGE CELL CARCINOMA (n=15)
C/EBPα 3+ BMI1 3+
C/EBPα 3+ BMI1 2+
C/EBPα 3+ BMI1 1+
C/EBPα 3+ BMI1 0
C/EBPα 2+ BMI1 3+
C/EBPα 2+ BMI1 2+
C/EBPα 2+ BMI1 1+
C/EBPα 2+ BMI1 0
C/EBPα 1+ BMI1 3+
C/EBPα 1+ BMI1 2+
C/EBPα 1+ BMI1 1+
C/EBPα 1+ BMI1 0
C/EBPα 0 BMI1 3+
C/EBPα 0 BMI1 2+
C/EBPα 0 BMI1 1+
C/EBPα 0 BMI1 0
0
1
2
3
4
5


D
ADENOSQUAMOUS CARCINOMA (n=8)
C/EBPα 3+ BMI1 3+
C/EBPα 3+ BMI1 2+
C/EBPα 3+ BMI1 1+
C/EBPα 3+ BMI1 0
C/EBPα 2+ BMI1 3+
C/EBPα 2+ BMI1 2+
C/EBPα 2+ BMI1 1+
C/EBPα 2+ BMI1 0
C/EBPα 1+ BMI1 3+
C/EBPα 1+ BMI1 2+
C/EBPα 1+ BMI1 1+
C/EBPα 1+ BMI1 0
C/EBPα 0 BMI1 3+
C/EBPα 0 BMI1 2+
C/EBPα 0 BMI1 1+
C/EBPα 0 BMI1 0
0
1
2
3


E
ADENOCARCINOMA IN SITU (n=7)
C/EBPα 3+ BMI1 3+
C/EBPα 3+ BMI1 2+
C/EBPα 3+ BMI1 1+
C/EBPα 3+ BMI1 0
C/EBPα 2+ BMI1 3+
C/EBPα 2+ BMI1 2+
C/EBPα 2+ BMI1 1+
C/EBPα 2+ BMI1 0
C/EBPα 1+ BMI1 3+
C/EBPα 1+ BMI1 2+
C/EBPα 1+ BMI1 1+
C/EBPα 1+ BMI1 0
C/EBPα 0 BMI1 3+
C/EBPα 0 BMI1 2+
C/EBPα 0 BMI1 1+
C/EBPα 0 BMI1 0
0
1
2
3

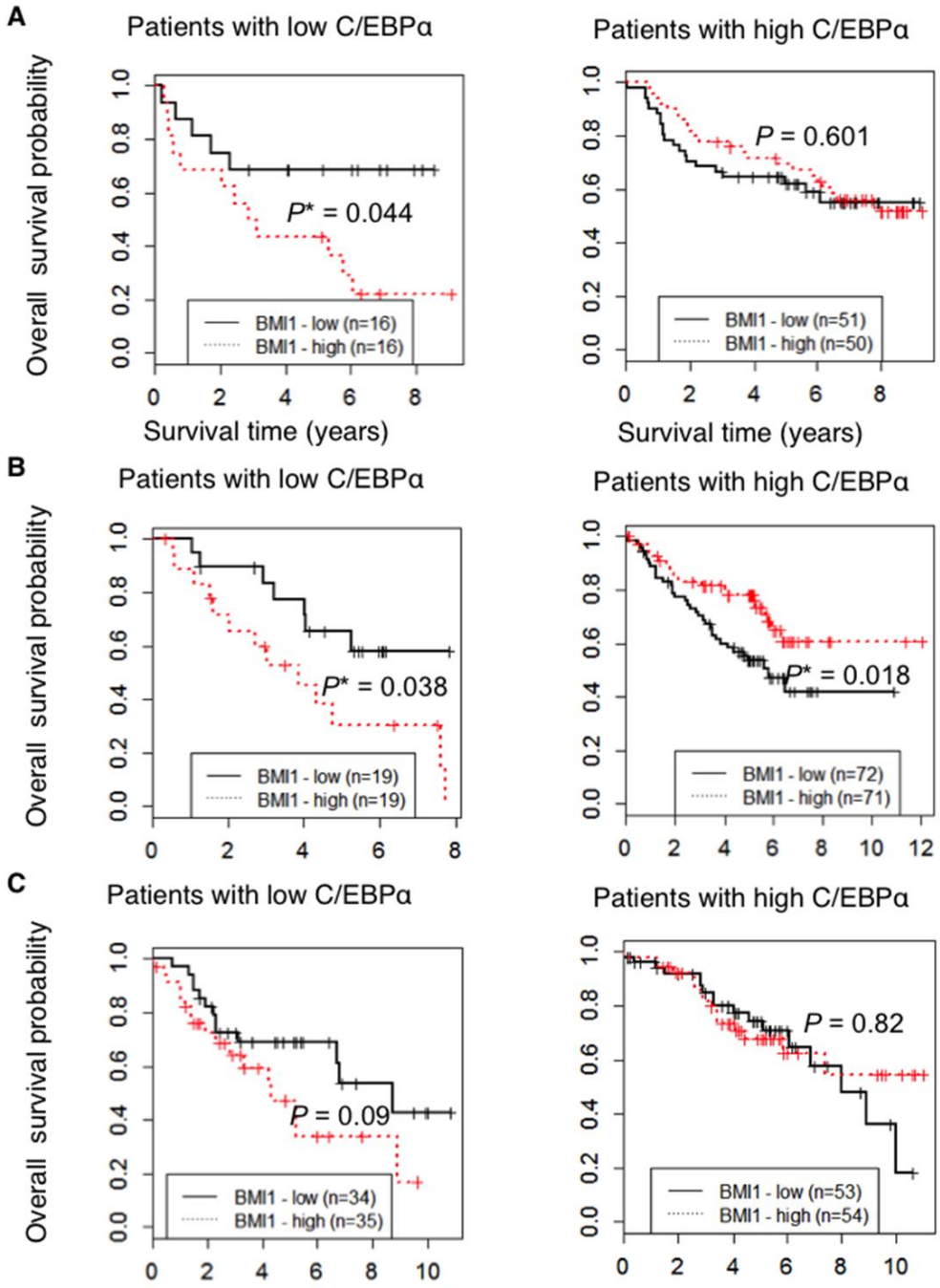

F
SARCOMATOID CARCINOMA (n=2)
C/EBPα 3+ BMI1 3+
C/EBPα 3+ BMI1 2+
C/EBPα 3+ BMI1 1+
C/EBPα 3+ BMI1 0
C/EBPα 2+ BMI1 3+
C/EBPα 2+ BMI1 2+
C/EBPα 2+ BMI1 1+
C/EBPα 2+ BMI1 0
C/EBPα 1+ BMI1 3+
C/EBPα 1+ BMI1 2+
C/EBPα 1+ BMI1 1+
C/EBPα 1+ BMI1 0
C/EBPα 0 BMI1 3+
C/EBPα 0 BMI1 2+
C/EBPα 0 BMI1 1+
C/EBPα 0 BMI1 0
0
1
2
3



Supplementary Figure 1: IHC data show BMI1 and C/EBPα expression in NSCLC patient subtypes.

A-G) Histograms showing distribution of both C/EBPα and BMI1 intensity of staining in each tumor analyzed, subdivided by NSCLC subtype. Adenocarcinoma cases are shown in **(A),** where it can be observed that 57.6% of them (white bars) are C/EBPα negative/low (0-1^+) and positive for BMI1 expression (1^+-3^+). The remaining samples (black bars) are either positive for C/EBPα expression (2^+-3^+) or low/negative for both C/EBPα (0-1^+) and BMI1 expression (0). Squamous cell carcinoma staining is reported in **(B)**; large cell carcinomas in **(C)**; a small number of cases of adenosquamous carcinoma and adenocarcinoma in situ are shown in **(D)** and **(E)**, respectively. **(F)** shows the staining for two cases of sarcomatoid carcinoma, and **(G)** depicts the staining of the only case of atypical carcinoid in our collection.

Supplementary Figure 2

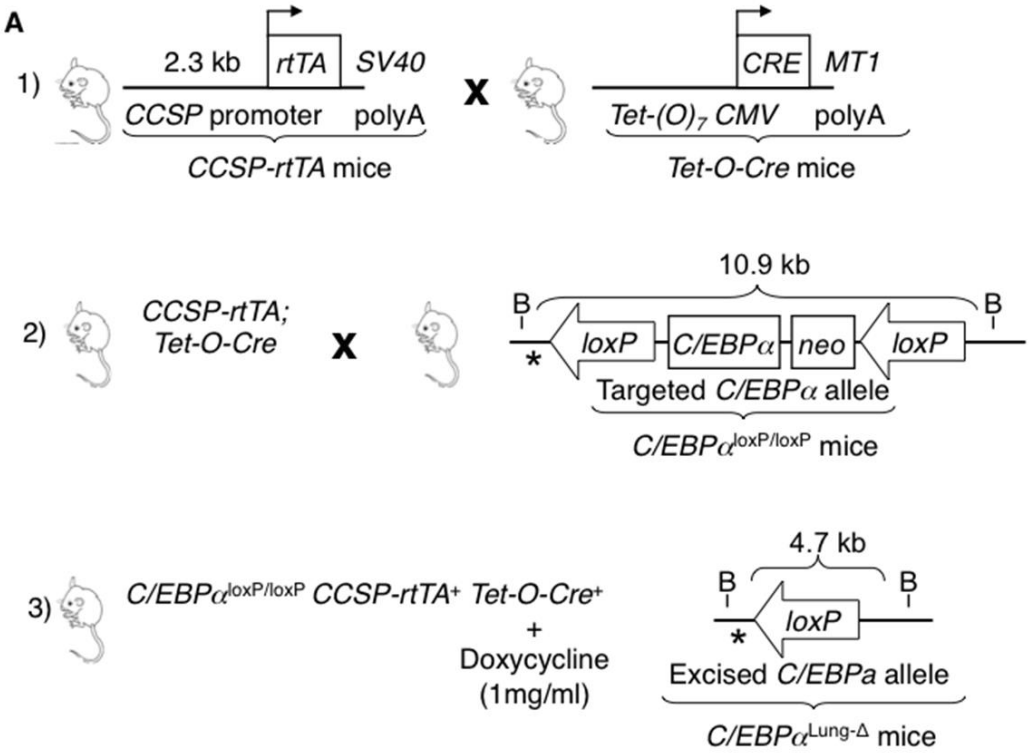


Supplementary Figure 2: C/EBPα determines BMI1-related overall survival probability in NSCLC patients.

Overall survival curves for **A)** 133 NSCLC patients from the Zhu et al dataset [18], **B)** the Der et al dataset (n=181) [19], and **C)** the Tang et al study (n=176) [20]. Patients are stratified according to low or high *C/EBPα* expression. The median BMI1 expression is used to define $BMI1^{low}$ and $BMI1^{high}$ patient subgroups. The p value (*P*) and the sample size (n) for each subgroup are indicated on each plot. In the Zhu et al. study, C/EBPα low

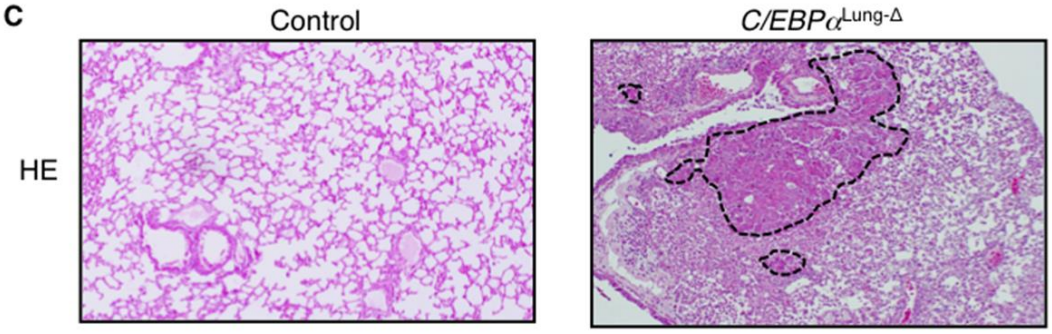
patient death events were as follows: 5 death events/16 patients who presented with low BMI1 expression and 12 death events/16 patients who presented with high BMI1. For the C/EBPα high patients, there were 21 death events/51 patients who presented with low BMI1 and 22 death events/50 patients who presented with high BMI1. In the Der et al. study, C/EBPα low patient death events were as follows: 7 death events/19 patients who presented with low BMI1 and 13 death events/19 patients who presented with high BMI1. For the C/EBPα high patients, there were 35 death events/72 patients who presented with low BMI1 and 20 death events/71 patients who presented with high BMI1. In the Tang et al. study, C/EBPα low patient death events were as follows: 13 death events/34 patients who presented with low BMI1 and 17 death events/35 patients who presented with high BMI1. For the C/EBPα high patients, there were 17 death events/53 patients who presented with low BMI1 and 17 death events/54 patients who presented with high BMI1.

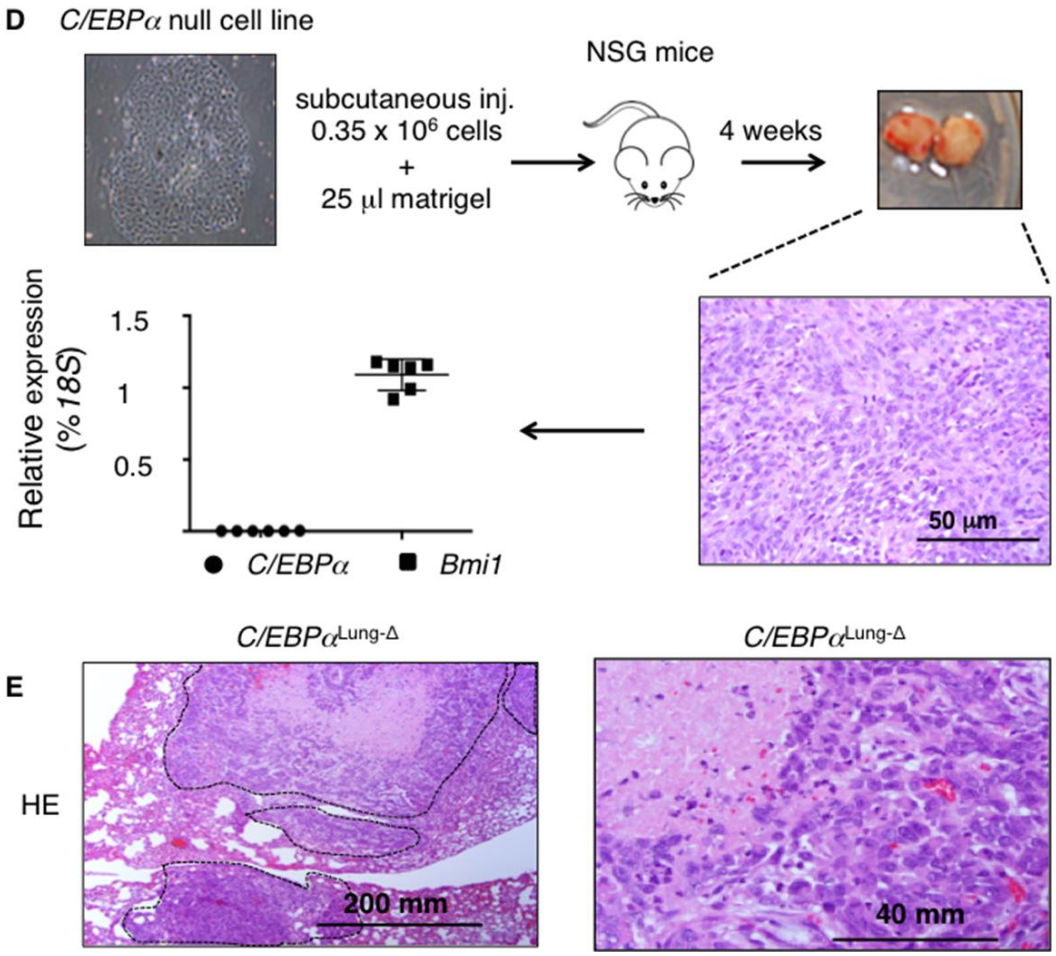
Supplementary Figure 3



B

	n=	mice with tumor	
C/EBPα Lung-Δ	102	34	
C/EBPα Lung Control	63	0	p<0.0001
C/EBPα Lung-Δ (het)	58	0	p<0.0001



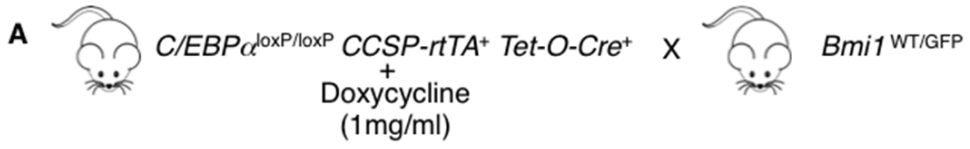


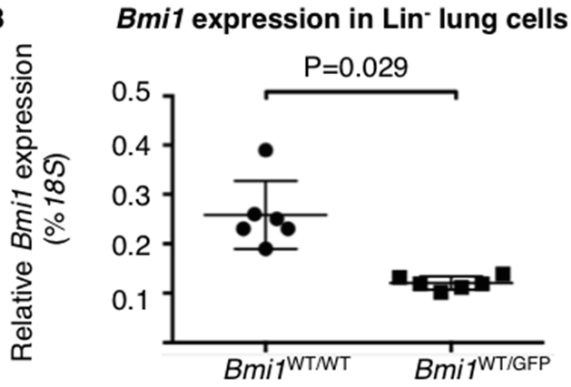
Supplementary Figure 3: C/EBPα$^{Lung-\Delta}$ mice develop pulmonary adenocarcinomas.

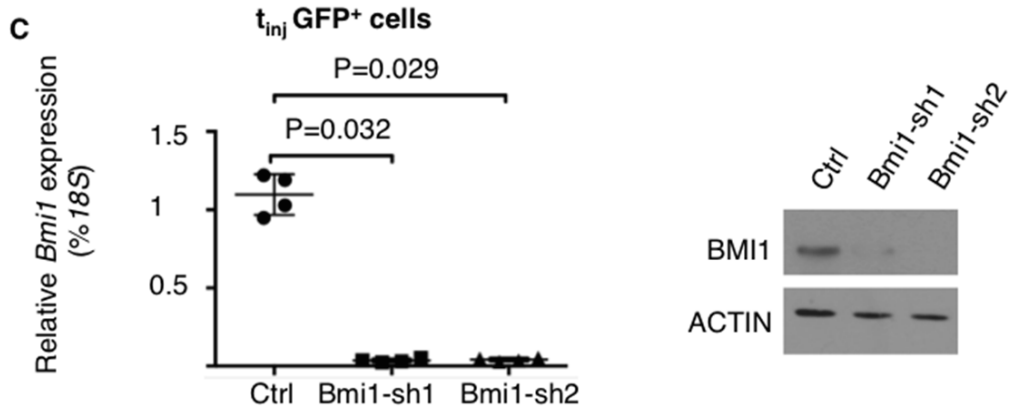
A) Constructs used to generate transgenic mice for doxycycline-regulatable *Cre* expression in the lung. **(1 left)** The *CCSP-rtTA* transgene consists of 2.3 kb of the rat *CCSP* promoter, the 1-kb *rTtA* coding sequence, and a 0.45-kb simian virus 40 *(SV40)* polyadenylation signal. **(1 right)** The *Tet(O)7-CMV-Cre* transgene consists of seven copies of the *tet* operator, a cytomegalovirus minimal promoter, the *Cre* recombinase coding sequence, and the *MT-1* polyadenylation sequence. **(2)** Schematic breeding strategy used to obtain *C/EBPα*$^{Lung-\Delta}$ mice (n=102), displaying conditional deletion of *C/EBPα* in CCSP^{+} pulmonary cells and their progeny, which included a mixed offspring of single and double transgenic littermates carrying one or two copies of the floxed *C/EBPα* allele. **(3)** Mice were treated with doxycycline in the drinking water (1 mg/ml) for 4 weeks after weaning. **B)** Assessment of the presence of tumorigenic foci in the lungs of *C/EBPα*$^{Lung-\Delta}$ mice (n=102, homozygous for *C/EBPα* deletion), *C/EBPα*$^{Lung(het)-\Delta}$ mice (n=58, heterozygous for *C/EBPα* deletion), as well

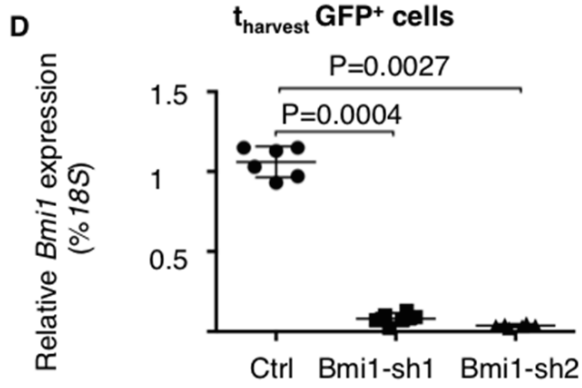
as control mice (n=63, littermates that do not have at least one of the transgenic alleles *CCSP-rtTA*, *Cre*, or *C/EBP*$\alpha^{loxP/loxP}$). The difference in tumorigenesis was estimated to be statistically significant by the 2-tailed Fisher's exact test ($p<0.0001$), demonstrating that only homozygous deletion of *C/EBP*α favors tumor formation. **C)** Representative lung sections of *C/EBP*$\alpha^{Lung-\Delta}$ mice that develop multiple foci per lung and control littermates stained with hematoxylin-eosin (HE). The dotted lines indicate where the tumors/foci are located. Scale bar is 500 μm. **D)** NSG mice (n=6) were subcutaneously injected with 3.5 x10^5 *C/EBP*α null tumor cells, as well as 25 μl of matrigel as three-dimensional support. After four weeks, mice were sacrificed and tumors growing under the skin were removed. HE staining was performed to assess their immunohistochemical features, and q-RT-PCR was performed to quantify *Bmi1* and *C/EBP*α expression in each tumor. The mean expression is presented as a percentage of the *18S* RNA amount. The p value for the paired comparison of *Bmi1/C/EBP*α ratio is 0.0008 (Welch's two-sided t-test). Scale bar is 50 μm. **E)** Representative lung sections of NSG mice injected with ~1000 primary *C/EBP*α null cells that develop multiple foci per lung, stained with HE. The dotted lines indicate where the tumors/foci are located. Scale bar is 200 μm (left panel) and 40 μm (right panel).

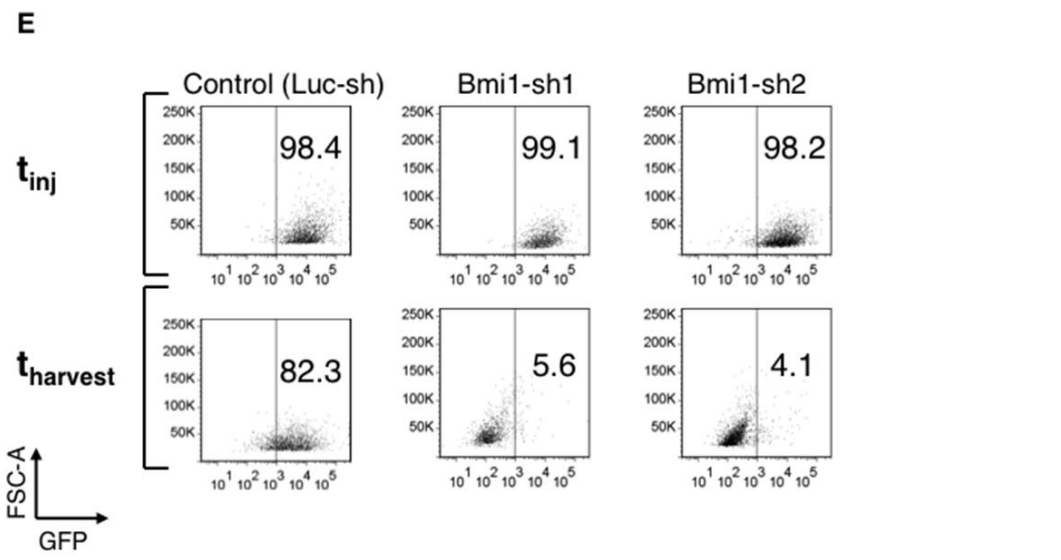
Supplementary Figure 4











F

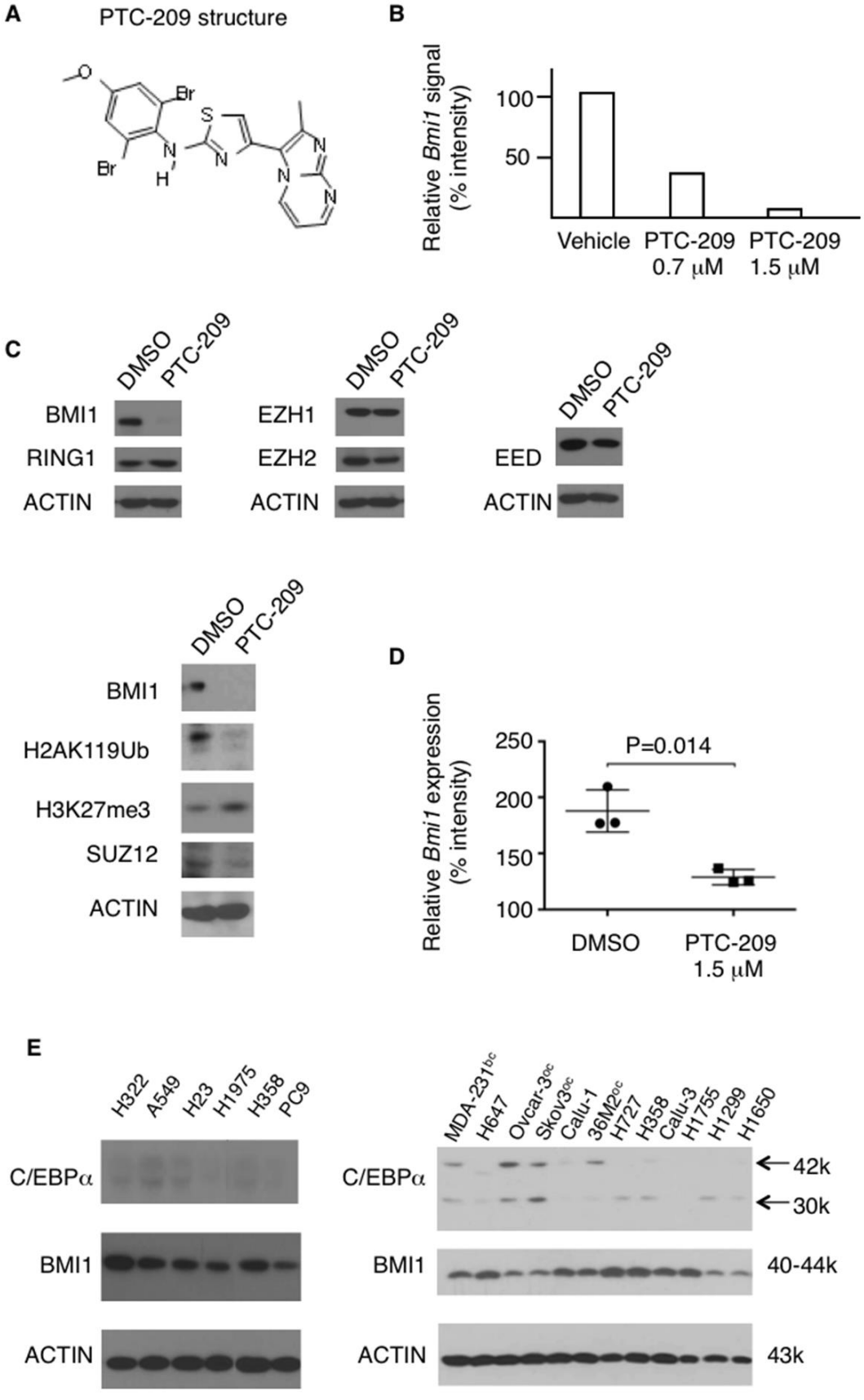
	GFP⁺ at t_{inj}	GFP⁺ at $t_{harvest}$	
Bmi1-sh1	97.3 +/- 1.9	6.6 +/- 1.2	p (at $t_{harvest}$)=0.00000057
Bmi1-sh2	98.5 +/- 0.4	3.5 +/- 0.7	p (at $t_{harvest}$)=0.00000072
Luc-sh	98.1 +/- 1.3	79.1 +/- 6.5	

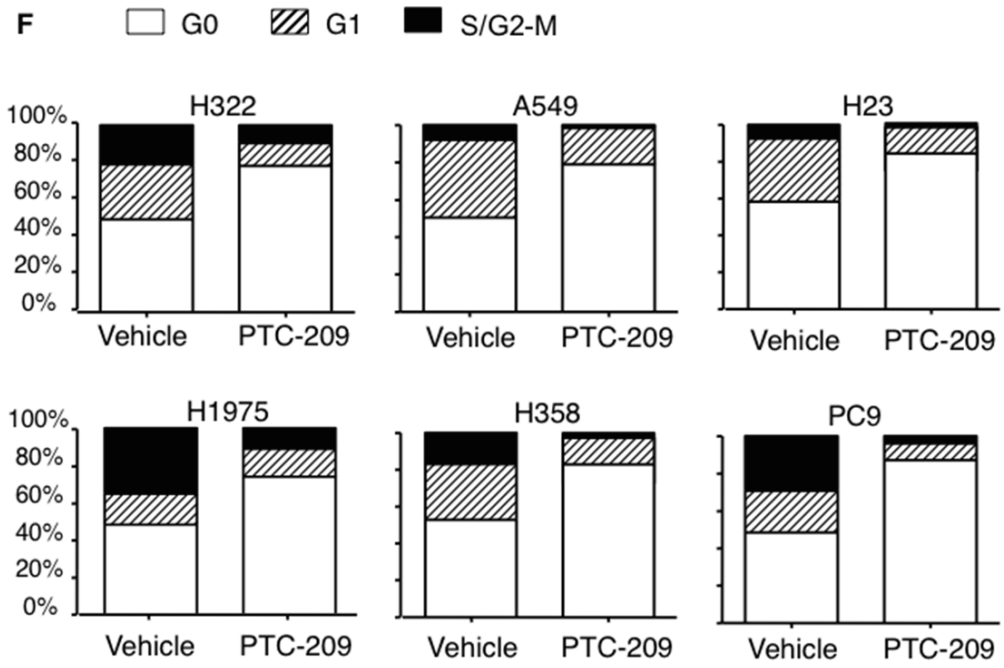
Supplementary Figure 4: Schematic representation shows generation of *Bmi1*-haploinsufficient *C/EBPα*$^{Lung-\Delta}$ mice and subcutaneous transplantation models.

A) Schematic representation of the breeding approach used to obtain *C/EBPα*$^{Lung-\Delta}$ mice with decreased BMI1 expression (*C/EBPα*$^{Lung-\Delta}$ x *Bmi1*$^{WT/GFP}$). *C/EBPα* conditional knockout mice (*C/EBPα*$^{loxP/loxP}$ *CCSP-rtTA*$^{+}$ *Cre*$^{+}$) mice were bred to *Bmi1*$^{WT/GFP}$ mice, in which one of the *Bmi1* alleles was replaced by *GFP*, to obtain *C/EBPα* conditional knockouts that constitutively express lower levels of BMI1 (*C/EBPα*$^{loxP/loxP}$ *CCSP-rtTA*$^{+}$ *Cre*$^{+}$ x *Bmi1*$^{WT/GFP}$) and control mice carrying normal copies of the *Bmi1* alleles (*C/EBPα*$^{loxP/loxP}$ *CCSP-rtTA*$^{+}$ *Cre*$^{+}$ x *Bmi1*$^{WT/WT}$), or lacking the *CCSP-rtTA* and/or the *tet-O-Cre* alleles (*C/EBPα*$^{Lung\ Control}$ x *Bmi1*$^{WT/GFP}$). Mice were treated with doxycycline in the drinking water (1 mg/ml) for 4 weeks after weaning. **B)** q-RT-PCR analysis for *Bmi1* expression was performed using lineage depleted (CD45.1^{-}, CD45.2^{-}, CD31^{-}) pulmonary cells (n=3 per group, in duplicate). Mice were sacrificed at 3 months of age. The mean expression (0.258 for WT/WT and 0.121 for WT/GFP) is presented as a percentage of *18S* RNA. Data were compared by Welch's one-sided t-

test. **C)** (**left panel**) q-RT-PCR in *C/EBP*α null cells transduced for 36 hours with Bmi1-sh1, Bmi1-sh2, or Luc-sh (control), after GFP^+ sorting. Y-axis indicates relative *Bmi1* expression presented as a percentage of *18S* at the time of injection into NSG mice, in two independent experiments (n=2 per group in duplicate). Data were analyzed by the one-sided Welch's t-test, and p values are indicated for each comparison. (**right panel**) Western blot analysis of *C/EBP*α null cells transduced for 36 hours with Bmi1-sh1, Bmi1-sh2, or Luc-sh (control), after GFP^+ sorting. Protein lysates were immunoblotted with an anti-BMI1 antibody, and loading was assessed with an anti-actin antibody. Expected size in kDa is indicated. **D)** q-RT-PCR analysis of *Bmi1* expression in GFP^+ cells at the time of tumor harvest. The mean expression is presented as a percentage of *18S* RNA. Results were obtained in duplicate in three tumors studied at harvest in each condition. The three Luc-sh-treated tumors were extracted from two different mice in which two tumors were taken from the left flanks, and one tumor was taken from the right flank; the three Bmi1-sh1-treated tumors were extracted from two different mice in which two tumors were taken from the left flanks and one from the right flank; the three Bmi1-sh2-treated tumors were extracted from two different mice in which one tumor was taken from the left flank and two tumors from the right flanks. **E)** Representative FACS analysis showing the efficiency of infection, scored as GFP^+ cells, at the time of injection (upper panels) and the percentage of GFP^+ cells at time of harvest (lower panels) 1 month after injection. **F)** Percentages of GFP^+ and GFP^- cells in Bmi1-sh1, Bmi1-sh2, and Luc-sh-control-treated cells at the time of injection and harvest. At the time of injection, GFP positivity was almost 100% in all sh-treated samples, whereas at the time of harvest, GFP positivity was markedly decreased in Bmi1-sh1 and Bmi1-sh2 treated cells, as compared to Luc-sh-control-treated cells.

Supplementary Figure 5





Supplementary Figure 5: The BMI1 inhibitor molecule PTC-209 inhibits BMI1 and induces cell cycle arrest in vitro.

A) PTC-209 is a substituted aminothiazole derivative with imidazolopyrymidine at the 4 position, identified from high-throughput screening with platform GEMSTM (gene expression modulation by small molecules)[49,50] (PTC Therapeutics). **B)** The histogram represents densitometry analysis of western blotting results of BMI1 staining normalized to actin staining. Murine *C/EBPα* null vehicle-treated cells are arbitrarily assigned a 100% value. After 48 hours of treatment with the BMI1 inhibitor, BMI1 translation was decreased to 35% (0.7 μM) and 5% (1.5 μM) of the control value. **C)** Western blot analysis was carried out in the murine *C/EBPα* null tumor cell line treated for 48 hours with the BMI1 inhibitor (1.5 μM) or 0.5% DMSO as vehicle control. Protein lysates (thirty micrograms per lane) were separated by SDS-PAGE and immunoblotted with antibodies recognizing C/EBPα, BMI1, RING1, EZH1, EZH2, EED, H2AK119Ub, H3K27me3, and SUZ12. Loading was assessed, after complete stripping of the membrane, with an anti-actin antibody. **D)** mRNA signal intensity was assessed in a microarray experiment comparing DMSO- versus PTC-209-treated cells (n=3). P-value was calculated using a two-sided Welch t-test. A second probe showed significant reduction with PTC209

treatment (p=0.003). **E)** Western blot analysis was carried out in the indicated NSCLC cell lines, as well as in breast cancer (MDA-231) and ovarian cancer cell lines (Ovcar-3, Skov3, and 36M2), designated as "bc" and "oc," respectively. Loading was assessed, after complete stripping of the membrane, with an anti-actin antibody. Expected size in kDa is indicated. **F)** The charts show the percentages of cells in G_0 (white columns), G1 (striped columns), and S/G2-M phase (black columns) for the indicated human adenocarcinoma cell lines, treated for 48 hours with PTC-209 or DMSO 5% as vehicle control. Cell cycle analysis was performed by Pyronin Y/Hoechst 33342 staining. The Welch's two-sided p-values for PTC-209-treated cells, as compared to vehicle-treated cells, are: H322 (p=0.0018), A549 (p=0.0098), H23 (p=0.0034), H1975 (p=0.0047), H358 (p=0.0010), and PC9 (p=0.0002).

Supplementary table 1: Patient demographic and clinicopathological parameters for all NSCLC cases.

	n	(%)
Age (years)		
Median	60	
Range	25-77	
Gender		
Male	197	75.5
Female	64	24.5
Pathology		
Adenocarcinoma	92	35.2
Squamous cell carcinoma	136	52.1
Adenosquamous carcinoma	8	3.1
Large cell carcinoma	15	5.7
Adenocarcinoma in situ	7	2.7
Atypical carcinoid	1	0.4
Sarcomatoid carcinoma	2	0.8
Cancer Stage		
I	101	38.7
II	78	29.9
III	82	31.4
Tumor Grade (n=217)		
1	28	12.9
2	122	56.2
3	67	30.9
Tumor Stage*		
T1	25	9.6
T2	179	68.6
T3	49	18.8
T4	8	3.0
Nodal stage*		
N0	122	46.7
N1	104	39.9
N2	31	11.9
N3	4	1.5

n=261, unless otherwise indicated in parentheses
*T and N were determined according to the TNM classification