## Supplementary Information for:

# *Keap1* loss promotes Kras-driven lung cancer and results in a dependence on glutaminolysis

Rodrigo Romero[1-2,#], Volkan I. Sayin[3,#], Shawn M. Davidson[1-2], Matthew R. Bauer[1], Simranjit X. Singh[3], Sarah E. LeBoeuf[3], Triantafyllia R. Karakousi[3], Donald C. Ellis[1-2], Arjun Bhutkar[1], Francisco J. Sanchez-Rivera[1-2], Lakshmipriya Subbaraj[1-2], Britney Martinez[3], Roderick T. Bronson[6,7], Justin R. Prigge[4], Edward E. Schmidt[4], Craig J. Thomas[8], Chandra Goparaju[9], Angela Davies[10], Igor Dolgalev[11], Adriana Heguy[11], Viola Allaj[12,13], John T. Poirier[12,13], Andre L. Moreira[3], Charles M. Rudin[12,13], Harvey I. Pass[9], Matthew G. Vander Heiden[1-2], Tyler Jacks[1-2,5*], and Thales Papagiannakopoulos[3,14*]

[1]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
[3]Department of Pathology, New York University School of Medicine, 550 First Avenue, New York, NY 10016, USA
[4]Department of Immunology and Infectious Diseases, Montana State University, Bozeman, MT 59717, USA
[5]Howard Hughes Medical Institute, NA, Chevy Chase, MD
[6]Tufts University, Boston, Massachusetts 02115, USA.
[7]Harvard Medical School, Boston, Massachusetts 02115, USA
[8]NIH Chemical Genomics Center, Division of Preclinical Innovation, National Center for Advancing Translational Sciences, National Institutes of Health, Bethesda, MD 20892, United States.
[9]Department of Cardiothoracic Surgery, New York University Langone Medical Center, New York, NY 10016.
[10]Champions Oncology, Hackensack, NJ, USA.
[11]Genome Technology Center, NYU School of Medicine, New York, USA[12]Molecular Pharmacology Program and Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA.
[13]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA.
[14]Perlmutter Cancer Center, New York University School of Medicine, New York, NY 10016, USA

[#]These authors contributed equally to this work
*Corresponding authors. Communication can be sent to: tjacks@mit.edu, papagt01@nyumc.org

**Supplemental Figure Legends**

**Supplementary Figure 1: Mutational analysis of sgKeap1 tumors**

**a)** Schematic representation of *Kras*$^{LSL\text{-}G12D/+}$*; p53*$^{fl/fl}$ (KP) mice intratracheally infected with pSECC lentiviruses containing sgKeap1 or control sgTom. Mouse tumor burden was tracked by micro-computed tomography (micro-CT) at 4 and 5 months post infection. Mouse lungs were harvested 21 weeks post infection. Whole lungs were subjected to immunohistochemistry (IHC) and tumors were micro-dissected for sequencing, IHC, and generation of tumor derived cell lines. **b)** Distribution of tumour grades in KP animals 21 weeks after infection with pSECC lentiviruses expressing: control (sgTom, KP; $n = 6$), sgKeap1.4 (KP; $n = 5$). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ obtained from two-sided Student's *t*-test. All error bars denote s.e.m. **c)** Fraction of mutant and wild-type reads within individual sgKeap1.2 tumors ($n = 11$) **d)** Total mutant reads summarized by mutation-type within tumors obtained from KP mice 21 weeks after infection with pSECC lentiviruses expressing sgKeap1.2 and **e)** sgKeap1.4. Figure legend corresponds to both d) and e). **f)** Representative alleles obtained from the KP tumor, 22T3 or **g)** the related metastasis 22LN, 21 weeks post infection with pSECC lentiviruses expressing sgKeap1.2. Left: pie chart contains the fraction of allele-specific reads over total mutant reads from tumor 22T3 or 22LN. All mutant reads falling below 1% sequencing reads are marked as other. Right panel: Summary of mutational analyses of locus-specific deep sequencing datasets (using the Illumina MiSEQ platform) showing the *Keap1*-WT locus containing the sgKeap1.2 binding site (black) and protospacer adjacent motif (PAM) sequence (green) along with representative mutant alleles. Dashes indicate deletion events and red arrows indicate insertion event. Sequences were obtained using MiSEQ. Note that alleles A and B from e) are found enriched in the metastatic tumor 22LN, obtained from the mediastinal lymph node indicating our ability to **h)** track genetic bottlenecks in the metastatic cascade using CRISPR and next generation sequencing.

**Supplementary Figure 2: Generation and validation of *Keap1* and *Nrf2*-mutant KP cells**

**a)** Schematic representation of the generation of the indicated cells (*n* = 2/genotype). Parental KP cells were electroporated with pX458 containing sgTom, sgNrf2.3, or sgKeap1.4 sgRNAs and sorted for GFP as single cells. **b**) **Left panel:** Keap1 locus containing the sgKeap1.4 or **right panel**: Nrf2 locus containing the Nrf2.3 binding site (black) and PAM sequence (green), Red arrows indicate insertion event. KPK1 and KPK2 alleles screened by next generation sequencing. **c)** Western blot analysis of nuclear and cytoplasmic fractions of indicated cells. HH3 and Hsp90 were used as a loading controls for the nuclear fraction cytoplasmic fractions respectively. Note, accumulation of nuclear Nrf2 and increases in cytoplasmic Gclc occurs only in KPK clones. **d)** Real-time quantitative PCR of Nrf2 target genes, *Nqo1*, *Hmox1*, and *Gclc* in the indicated cell lines. Y-axis depicts the fold change relative to KP1 for each respective target gene. Errors bars depict standard deviation (*n* = 3 per cell line per gene). **p* < 0.05, ***p* < 0.01, ****p* < 0.001, *****p* < 0.0001. Obtained from two-sided Student's *t*-test. Statistics are derived via the comparison of the KP1 sample for each target gene. **e)** Western blot analysis of CRISPR targeted KP clones with or without the Nrf2 activator, Sulforaphane (SFN; 10uM for 6 hrs). Gapdh was used as a loading control. Note, accumulation of Nrf2 occurs only in the KP samples, but is not further stabilized in either KPK samples. **f-h)** Real-time quantitative PCR of Nrf2 target genes, *Nqo1*, *Hmox1*, and *Gclc* in KP cell lines treated with the Nrf2 activator, SFN (10uM for 6 hrs). Y-axis depicts the fold change relative to the KP1-SFN treated sample. Errors bars depict standard deviation (*n* = 3 per cell line per gene). **p* < 0.05, ***p* < 0.01, ****p* < 0.001, *****p* < 0.0001. Obtained from two-sided Student's *t*-test. All statistics are derived via the comparison of the KP1-SFN sample for each target gene. **i)** Box-plot highlighting the *Keap1*-WT untreated cell lines (KP; *n* = 2) versus *Keap1*-WT SFN treated cell lines (*n* = 2) and *Keap1*-mutant untreated (KPK; *n* = 2) gene expression signature detected within the KP GEMM-derived isogenic pairs. The Y-axis shows a significantly different signature profile between the two sets of isogenic pairs (*p* = 0.009, Mann-Whitney-Wilcoxon test).

**Supplementary Figure 3: Differential response to oxidative stress in *Keap1* and *Nrf2*-mutant cells.**

**a)** Heat-map representing the standardized $IC_{50}$ values for each indicated cell line treated with the indicated compounds. Grey squares depict samples in which $IC_{50}$ values could not be determined. Z-score color scale depicted below the heatmap. **b-d)** Dose response curves of L-Buthionine sulfoximine (BSO), Auranofin (AUR), Erastin (ERA), and **e)** dimethyl fumarate (DMF, Nrf2 activator). Relative viability was obtained using cell-titer glo assay after 72 hrs of drug treatment. Relative viability for the NRF2 activator, dimethyl-fumarate (DMF) treated samples were obtained using Hoechst staining to count stained nuclei after 72 hrs of drug treatment. All values were normalized to their respective vehicle treated control. X-axis depicts respective drug concentration (uM) in $\log_{10}$ scale. Error bars depict s.e.m. ($n$ = 3 per cell line per treatment). **f)** Basal reduced glutathione (GSH): oxidized glutathione (GSSG) ratios of KP (grey shades), KPN (red shades), and KPK (blue shades) cell lines ($n$ = 3 per cell line). **$p$ < 0.01, ****$p$ < 0.0001. Obtained from two-sided Student's $t$-test. **g)** Top panel: Longitudinal GSH concentrations (μM) following treatment with 100μM BSO (0 hr). GSH concentrations were normalized to the 0 hr BSO samples of each individual cell line. Bottom panel: Cell viability after addition of 100μM BSO (0 hr). All samples were normalized to their respective vehicle treated control (relative luminescent units). All error bars depict s.e.m. ($n$ = 3 treatments per time point/cell line). **h)** Antioxidant (Trolox and N-Acetyl cysteine (NAC)) rescue of oxidative stressed (BSO, AUR, ERA or Control) treated KPN cell lines ($n$ = 3 per cell line per antioxidant per oxidative stress condition). ****$p$ < 0.0001, obtained from 1-way Anova with Tukeys post hoc test **i)** Representative schematic of the lentivirus containing the *HA*-tagged GOF-*Nrf2* cDNA driven by a *TRE*-promoter. **j)** Western blot depicting two independent KPN cell lines expressing a doxycycline inducible *HA*-tagged GOF-*Nrf2* cDNA (KPN-ix). The GOF-*Nrf2* cDNA was induced for 72 hrs using doxycycline. Induced cells were treated with SFN (10uM) for 6 hrs before harvesting. First panel blotting for HA. Second panel depicts a Nrf2 antibody blot. Note * depicts the expected wild-type Nrf2 size (not seen in KPN clones) and ** depicts the lower molecular weight GOF mutant lacking exon 2. Third

panel depicts Gclc blotting. Final panel: Gapdh loading control. Note: GOF-*Nrf2* cDNA is only expressed following doxycycline administration, which leads to increases in cytoplasmic Gclc. **k,l)** Dose response curves of the indicated KPN-ix cell lines treated with BSO and 72 hrs of dox induction. KPN-ix - dox in dark red, KPN-ix + dox in light red. Relative viability was obtained using cell-titer glo assay after 72 hrs of drug treatment. All values were normalized to their respective vehicle treated control. X-axis depicts respective BSO concentrations in log scale. Error bars depict s.e.m. ($n$ = 3 per cell line per treatment). Note: induction of the GOF-*Nrf2* cDNA results in an increase $IC_{50}$. **m-o)** Real-time quantitative PCR of Nrf2 target genes, *Nqo1*, *Gclc*, and *Slc7a11* in KPN-ix (- dox dark red; + dox light red). The GOF-*Nrf2* cDNA was induced for 72 hrs using doxycycline. Induced cells were treated with SFN (10uM) for 6 hrs before harvesting. Y-axis depicts the fold change relative to untreated KPN1-ix for each respective target gene. Errors bars depict standard deviation ($n$ = 3 per cell line per gene). *$p$ < 0.05, **$p$ < 0.01, ***$p$ < 0.001, ****$p$ < 0.0001. Obtained from two-sided Student's $t$-test. **p)** The amount of ROS in CRISPR targeted cells as judged by fluorescence-activated cell sorting (FACS) analyses of CM-DCF fluorescence ($n$ = 3 per cell line) ****$p$ < 0.0001 obtained from 1-way Anova with Tukeys post hoc test. **q)** The amounts of ROS in human lung cancer cell lines as judged by FACS analyses of CM-DCF fluorescence ($n$ = 3 per cell line). ****$p$ < 0.0001. Obtained from 1-way Anova with Tukeys post hoc test.

**Supplementary Figure 4: *Keap1*-mutant cells display a selective growth advantage *in vivo***

**a)** Subcutaneous tumor volumes of KP and KPK cells injected into nude mice ($mm^3$; $(a^2*b)*(\pi/6)$ where a is the smaller dimension and b is the larger dimension) measured over time for 22 days. Related to Supplementary Fig 4b. **b)** Final subcutaneous tumor masses ($n$ = 6 tumors/cell line) ***$p$ <0.001, obtained from two-sided Student's $t$-test. **c)** Quantification of Ki67 in tumors from mice orthotopically transplanted with KP and KPK cells ($n$ = 16 and 11; $p$ < 0.0001). Obtained from two-sided Student's $t$-test. **d)** Representative images from Ki67 stained KP and KPK orthotopic lung tumors. **e)** Orthotopic growth measurements of KP and

KPK cells (*n* = 4). Quantitation of luminescence (photon flux) in mice orthotopically transplanted with KP or KPK cells transduced with a vector expressing *Luciferase*. Relative photon flux calculated by normalizing all time points per animal to initial measurements at 14-days post transplantation. ***$p < 0.001$ obtained from 2-way Anova. **f)** Cumulative population doublings *in vitro* of indicated cell lines (*n* = 4). Scale bars are 100um.

**Supplementary Figure 5: *Keap1/KEAP1* loss accelerates *in vivo* tumor growth independent of *p53/TP53* status.**

**a)** p53 western blots of two independent *Kras*-mutant; *p53*-WT murine cells LKR cells lines[18] treated with the DNA intercalating agent, Doxorubicin (DOXO) at 0.2 ug/mL for 6 hours to induce p53 stabilization and the DNA damage response. **b)** NRF2 Western blot of the human *KRAS*-mutant; *P53*-WT lung cancer cell line, SW1573, transduced with five independent sgRNAs targeting *KEAP1* results in NRF2 stabilization. **c)** Keap1 western blots of isolated LKR clones transduced with either sgTom or sgKeap1.4 as described in Supplementary Fig 2a. LKR10 T2 (sgTom), LKR10 K7 (sgKeap1.4), LKR13 T1 (sgTom), and LKR13 K17 (sgKeap1.4) were selected for follow up growth assays. **d-g)** NRF2 target gene expression of the *KRAS*-mutant; *P53*-WT; *KEAP1*-mutant human lung adenocarcinoma cell line, A549, transduced with lentiviruses containing *PGK*-control or *PGK-KEAP1* cDNAs. All error bars denote s.e.m. Obtained from two-sided Student's *t*-test. *$p < 0.05$, ****$p < 0.0001$. **h-j)** Subcutaneous tumor volumes of the above outlined human cell lines (A549, transduced with *PGK-control* or *PGK-KEAP1* cDNAs, and SW1573 transduced with sg*Tom* or sg*KEAP1.3* or sg*KEAP1.5*) and **k-l)** *p53*-WT mouse cell line LKR10 and LKR13 transfected with plasmids containing sgTom or sgKeap1.4. **m)** Subcutaneous tumor masses of LKR13 *Keap1*-mutant (*n* = 6) or –WT(*n* = 9) cell lines, related to Supplementary Fig 5l. Depicted statistics obtained from 2-way Anova. **n)** $Kras^{G12D/+}$; $p53^{+/+}$ (K-only) autochthonous tumors grades derived from mice initiated with pSECC-sgTom (*n* = 70 tumors) or pSECC-sgKeap1.2 (*n* = 158 tumors). Depicted statistics obtained from Fisher's exact test. **o)** IHC of Ki67 positive nuclei per $mm^2$

of tumors derived from K-only mice initiated with pSECC-sgTom or pSECC-sgKeap1.2 (*n*=16/group). **p)** example Ki67 IHC related to Supplementary Fig 5o. Scale bars are 100um.

**Supplementary Figure 6: A human derived *KEAP1*-mutant and *NRF2* target gene signature predicts survival of LUAD patient survival.**

**a)** Box-plot highlighting the *KEAP1*-WT (*n* = 380) versus *KEAP1*-mutant (*n* = 79) gene expression signature detected within the TCGA LUAD cohort. The Y-axis shows a significantly different signature profile between the two sets of tumors ($p < 2.22e-16$, Mann-Whitney-Wilcoxon test). **b)** GSEA enrichment plot showing that the NRF2 core target signature is highly enriched in the *KEAP1*-mutant signature derived from the TCGA LUAD cohort (FDR = 0.0). **c)** GSEA enrichment plot of the published NFE2L2.V2 signature[18] exhibiting enrichment in the TCGA *KEAP1*-mutant signature (FDR = 0.0). **d)** Kaplan-Meier (KM) survival curves comparing *KRAS*-mutant TCGA LUAD patients stratified by their correlation with the *KEAP1*-mutant signature derived from TCGA patient expression profiles. The top 20% correlated patients (*n* = 24) exhibit decreased survival compared to the rest (*n* = 99) of the TCGA LUAD cohort ($p = 0.00013$, log-rank test). **e)** Empirical cumulative distribution function (CDF) plot showing correlation of individual tumors with the *KEAP1*-mutant signature across various tumor grades from the TCGA LUAD cohort. Each curve represents a unique tumor grade as depicted in the figure legend. Grade III/IV tumors (*n* = 60) exhibit significantly higher correlation with the *KEAP1*-mutant signature compared to grade I tumors (*n* = 146; $p = 0.02$, Kolmogorov-Smirnov test; $p_g$ = Kruskal-Wallis test across all grades). **f)** GSEA enrichment plot of the murine-derived *Keap1*-mutant signature within the human *KEAP1*-mutant signature derived from the TCGA LUAD cohort (FDR = 0.0).

**Supplementary Figure 7: CRISPR screen reveals that *Keap1*-mutant cells are sensitive to reduced glutamine levels**

**a)** Full representation of pooled sgRNA library screen related to Figure 3a. All bars denote individual sgRNA score. **b)** Crystal violet stain of KP and KPK cells after transduction with

sgTom, sgSlc1a5.1 or sgSlc1a5.2 cultured for 72h **c)** Relative viability assayed with cell-titer glo (relative luminescent units) on KP and KPK cells after treatment with GPNA for 72 hrs (*n* = 4 technical replicates/data point). **d)** Crystal violet stain of KP and KPK cells cultured with 2mM, 1mM or 0.5mM Glutamine cultured for 72h. **e)** Glucose consumption (left Y-axis) and lactate excretion (right Y-axis) in KP and KPK cells measured (*n* = 3 technical replicates/cell line). All samples were normalized to their respective vehicle treated control. $^{**}p < 0.01$ obtained from 1-way Anova with Tukey's post hoc test. All error bars depict s.e.m.

**Supplementary Figure 8: *Keap1*-mutant cells are more glycolytic**

**a)** Trypan blue exclusion viability counts of KP and KPK cells cultured in the presence or absence of 5mM 2DG for 72h (*n* = 4 technical replicates/cell line). $^{****}p < 0.0001$. Obtained from two-sided Student's *t*-test. **b)** Crystal violet stain of KP and KPK cells cultured with or without 5mM 2DG for 72h. All error bars depict s.e.m. **c)** diagram of **d)** flux of glucose derived carbon through pyruvate dehydrogenase (M+2 citrate) or pyruvate carboxylase (M+3 citrate) into the TCA cycle. Cells were cultured for 24 hours in RPMI with [U-13C]-L-glucose. Graph represents the fractional enrichment of the indicated isotopomer in the total citrate pool (*n* = 3). **e)** Total contribution of glucose carbon to TCA cycle intermediates. Cells were cultured for 24 hours in RPMI with [U-13C]-L-glucose. Percentage mole enrichment of $^{13}C$ carbon derived from glucose calculated for each metabolite indicated. (*n* = 3). All error bars depict s.e.m.

**Supplementary Figure 9: Glutaminase inhibition specifically inhibits growth of *Keap1/KEAP1*-mutant cells**

**a)** Crystal violet stain of KP and KPK cells receiving CB-839 (250nM), BPTES (5uM) or control for 72 hrs. **b)** Viability counts of human lung cancer cell lines that are *KEAP1*-WT, *KEAP1*-mutant or *NRF2*-mutant, treated with 500nM CB-839 plotted as % of control (Individual lines in Fig 4d for 72 hrs. $^{**}p < 0.01$ obtained from Mann-Whitney test. **c-f)** Relative growth of KP and KPK cells treated with and without CB-839 and with or without the

indicated antioxidant or metabolite NAC, trolox, cell permeable alpha-ketoglutarate (DMG), pyruvate, or glutamate.

**Supplementary Figure 10: Nrf2 pathway activity is required for sensitivity to glutaminase inhibition**

**a)** Western blot depicting two independent KP cell lines expressing a doxycycline inducible *HA*-tagged GOF-*Nrf2* cDNA (KP-ix). The GOF-*Nrf2* cDNA was induced for 72 hrs using doxycycline. Induced cells were treated with SFN (10uM) for 6 hrs before harvesting. First panel blotting for HA. Second panel depicts a Nrf2 antibody blot. Note * depicts the expected wild-type Nrf2 size and ** depicts the lower molecular weight GOF-mutant lacking exon 2. Third panel depicts Gclc blotting. Final panel: Gapdh loading control. **b)** Real-time quantitative PCR of Nrf2 target genes, *Nqo1*, *Gclc*, and *Slc7a11* in KP-ix. The GOF-*Nrf2* cDNA was induced for 72 hrs using doxycycline. Induced cells were treated with SFN (10uM) for 6 hrs before harvesting. Y-axis depicts the fold change relative to untreated KP1-ix for each respective target gene (*n* = 3 technical replicates/cell line/target gene). **c)** Viability counts of KP-ix cells containing a doxycycline inducible Nrf2 Δexon2 (GOF-*Nrf2* cDNA) treated with control, doxycycline, CB-839 (250nM), or both doxycycline and CB-839 for 72 hrs (*n* = 3 technical replicates/data point). **d)** Nrf2 and Nrf2 target gene expression by western blot of KPK cells transduced with *Keap1* or control cDNAs. **e)** Nrf2 target gene expression by real-time quantitative PCR in KPK transduced with control or *Keap1* cDNAs (*n* = 3 technical replicates/cell line/target gene. **f)** Subcutaneous tumor volumes of KPK cells transduced with control or *Keap1* cDNAs injected into the flank of nude mice. **g)** Relative cell growth of KP and KPK cells transduced with control or *Keap1* cDNAs in the presence or absence of 500nM CB-839 for 72 hrs. All samples relative to each genotype's vehicle treated sample. (*n* = 4 technical replicates/cell line/treatment).

**Supplementary Figure 11: *Keap1/KEAP1*-mutant lung adenocarcinoma cells and patient derived xenografts are sensitive to glutaminase inhibition *in vivo***

**a)** Schematic of subcutaneous or orthotopic transplants of *Keap1/KEAP1*-WT or *Keap1/KEAP1*-mutant cells into immunocompromised mice. **b)** Subcutaneous tumor volumes of KP and KPK cells treated with Vehicle or CB-839 starting from day 13 (arrow indicating treatment start) measured over time for 25 days (*n* = 6 tumors/group)**.** Related to Fig 4f. **c)** Orthotopic growth measurements of KP and KPK cells treated with Vehicle or CB-839 starting from day 13 (arrow indicating treatment start, *n* = 4 mice/group). Quantitation of luminescence (photon flux) in mice orthotopically transplanted with KP or KPK cells transduced with a vector expressing Luciferase. Relative photon flux calculated by normalizing all time points per animal to initial measurements at 10 days post transplantation. Related to Figure 4g. ***$p < 0.001$ obtained from 2-way Anova. All error bars depict s.e.m. **d)** Subcutaneous tumor volumes of KP cells transduced with inducible GOF-*Nrf2* cDNAs (KP-ix) treated with vehicle or CB-839 in the presence or absence of doxycycline (DOX). Related to Figure 4h (*n* = 6 tumors/group). **e)** Subcutaneous tumor volumes of the *KEAP1*-mutant human LUAD cell line, H2122 and **f)** *KEAP1*-WT human LUAD cell line, H2009, treated with CB-839 or vehicle. **g)** Subcutaneous tumor volumes of three independent *KEAP1*-mutant patient-derived xenografts (PDX) and **h)** *KEAP1*-WT PDX treated with CB-839 or vehicle. Mouse number and *p* values depicted in the respective figures. Related to Figure 4i. **i)** IHC of the NRF2 target gene, NQO1, on respective PDX model. Note weak CB-839 responder (MSK-LX337) has low NQO1 staining indicating low NRF2 pathway activation. Scale bars are 100um.

**Supplementary Figure 12:** Uncropped western blot images

**Supplementary Table 1**: Targeted exome capture of 88 LUAD tumors from the NYU Center for Biospecimen Research and Development

**Supplementary Table 2**: GEMM derived *Keap1*-mutant gene expression signature (z-score table) and top enriched gene sets from MSigDB curated and Oncosig collections. Genes with increasingly positive scores are over-expressed in *Keap1*-mutant samples whereas those

with negative scores exhibit relatively lower expression in *Keap1*-mutant samples (magnitude denotes strength of a gene's expression correlation with the signature).

**Supplementary Table 3**: Nrf2 core target gene set derived from the union of three published datasets and Nrf2-induced targets from individual datasets.

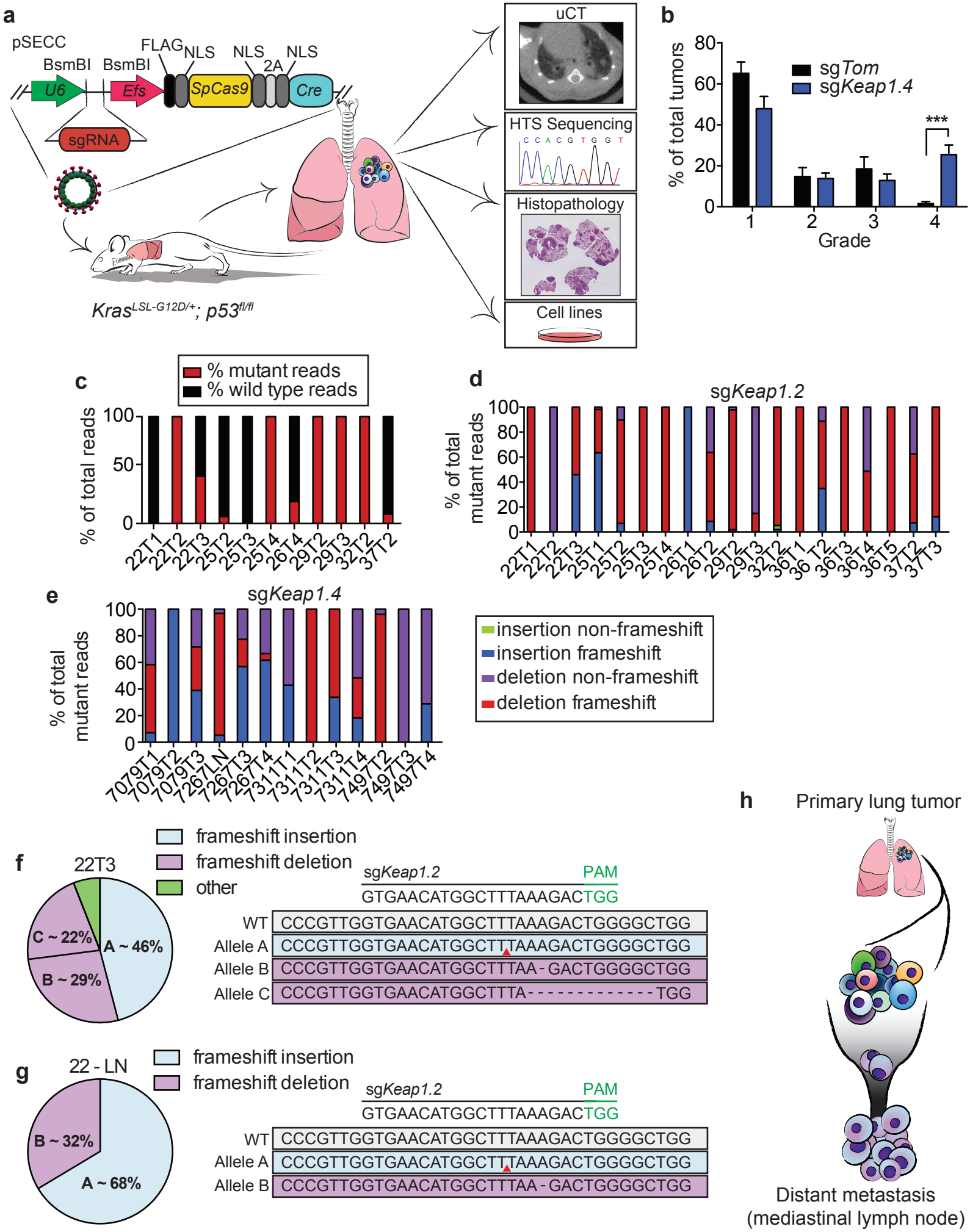
**Supplementary Table 4**: Human lung adenocarcinoma (TCGA) derived *KEAP1*-mutant gene expression signature (z-score table) and top enriched genes sets from MSigDB curated collections. Genes with increasingly positive scores are over-expressed in *KEAP1*-mutant samples whereas those with negative scores exhibit relatively lower expression in *KEAP1*-mutant samples (magnitude denotes strength of a gene's expression correlation with the signature).

**Supplementary Table 5**: Univariate and Multivariate Cox regression analyses for overall survival in the TCGA lung adenocarcinoma patient cohort.

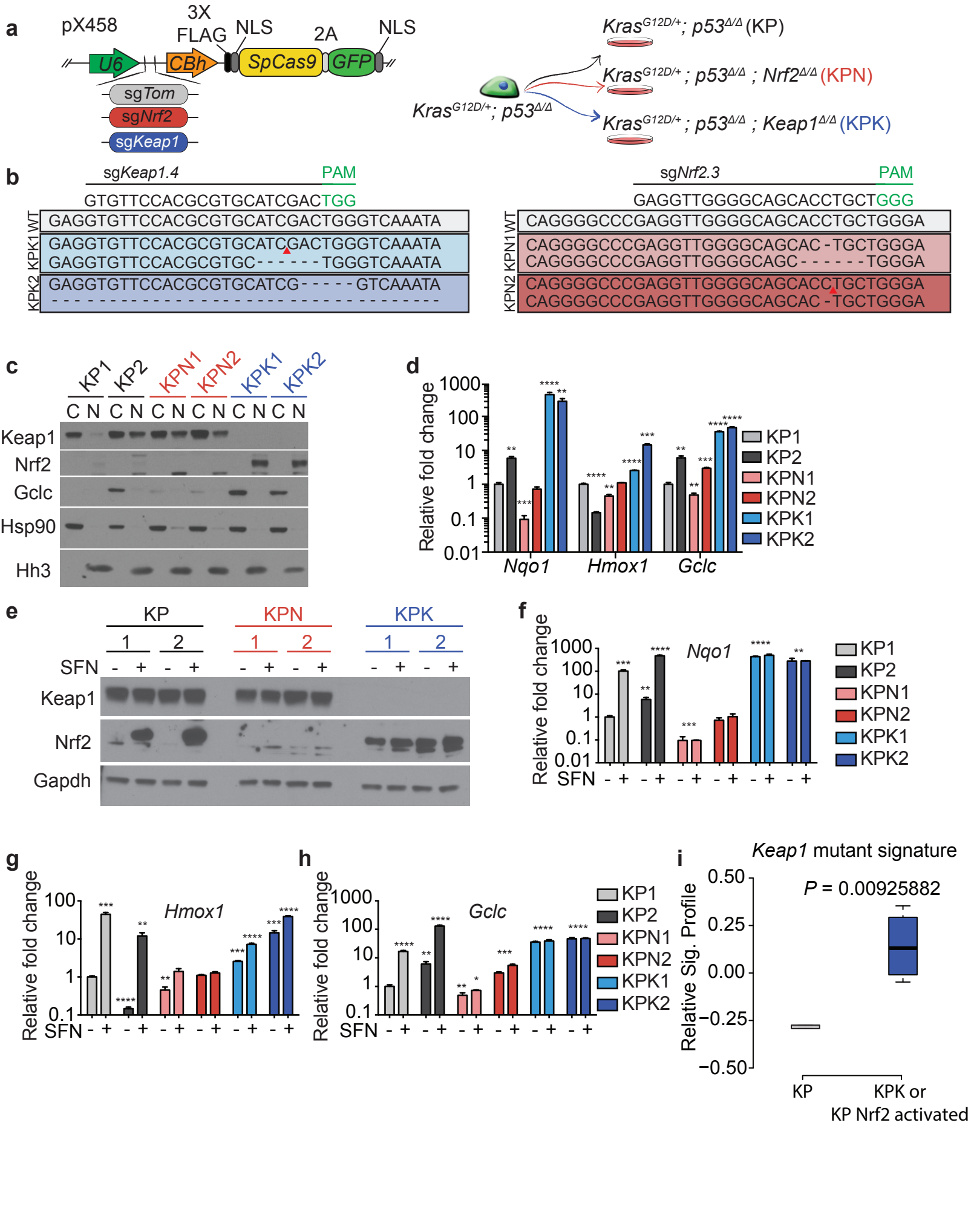
**Supplementary Table 6:** CRISPR/Cas9 Nrf2 transcriptional target screen containing sgRNA sequences, gene descriptions, and sgRNA scores.

**Supplementary Table 7:** Clinical and genetic features of PDX models

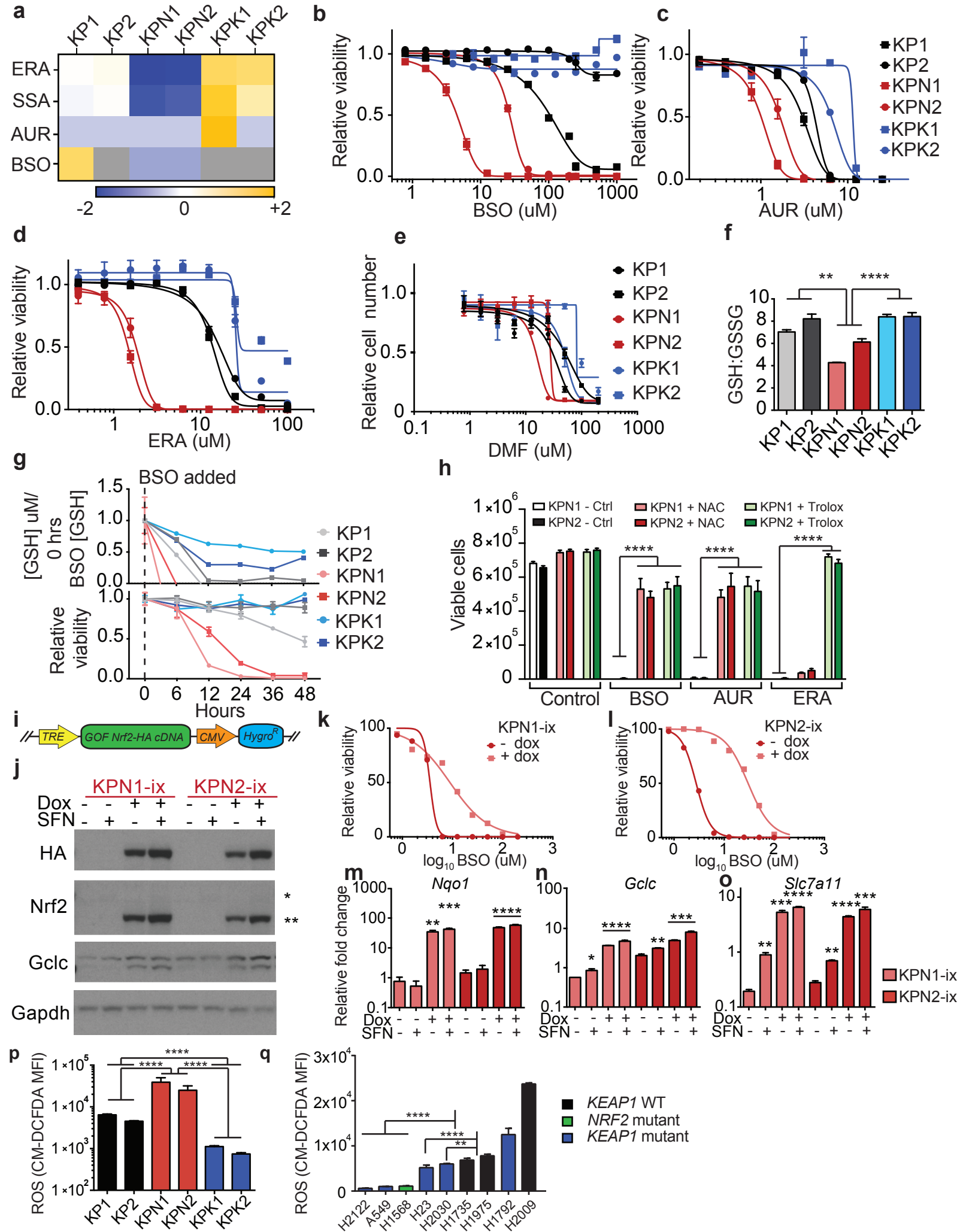
**Supplementary Figure 1**



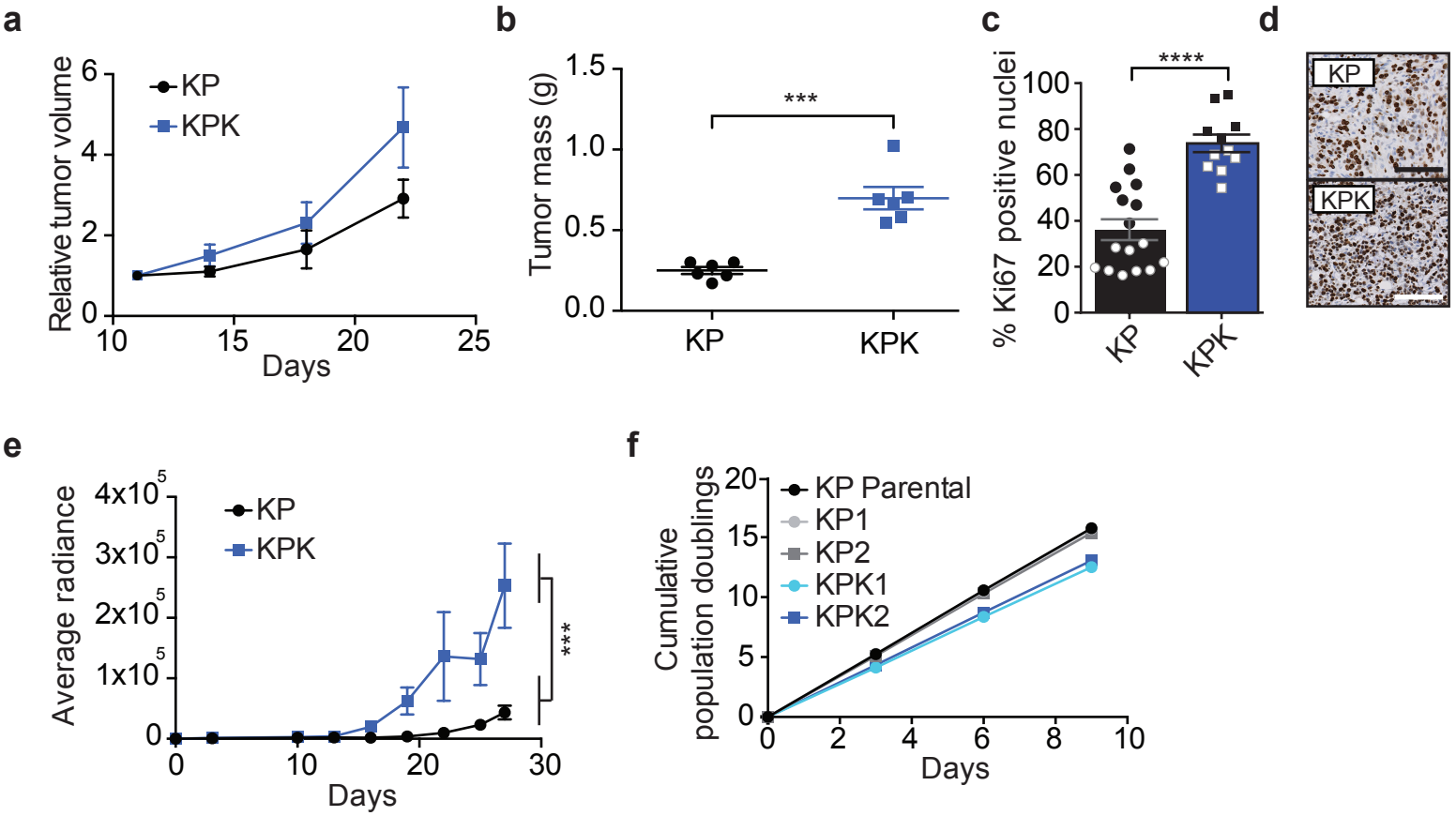
**Supplementary Figure 2**



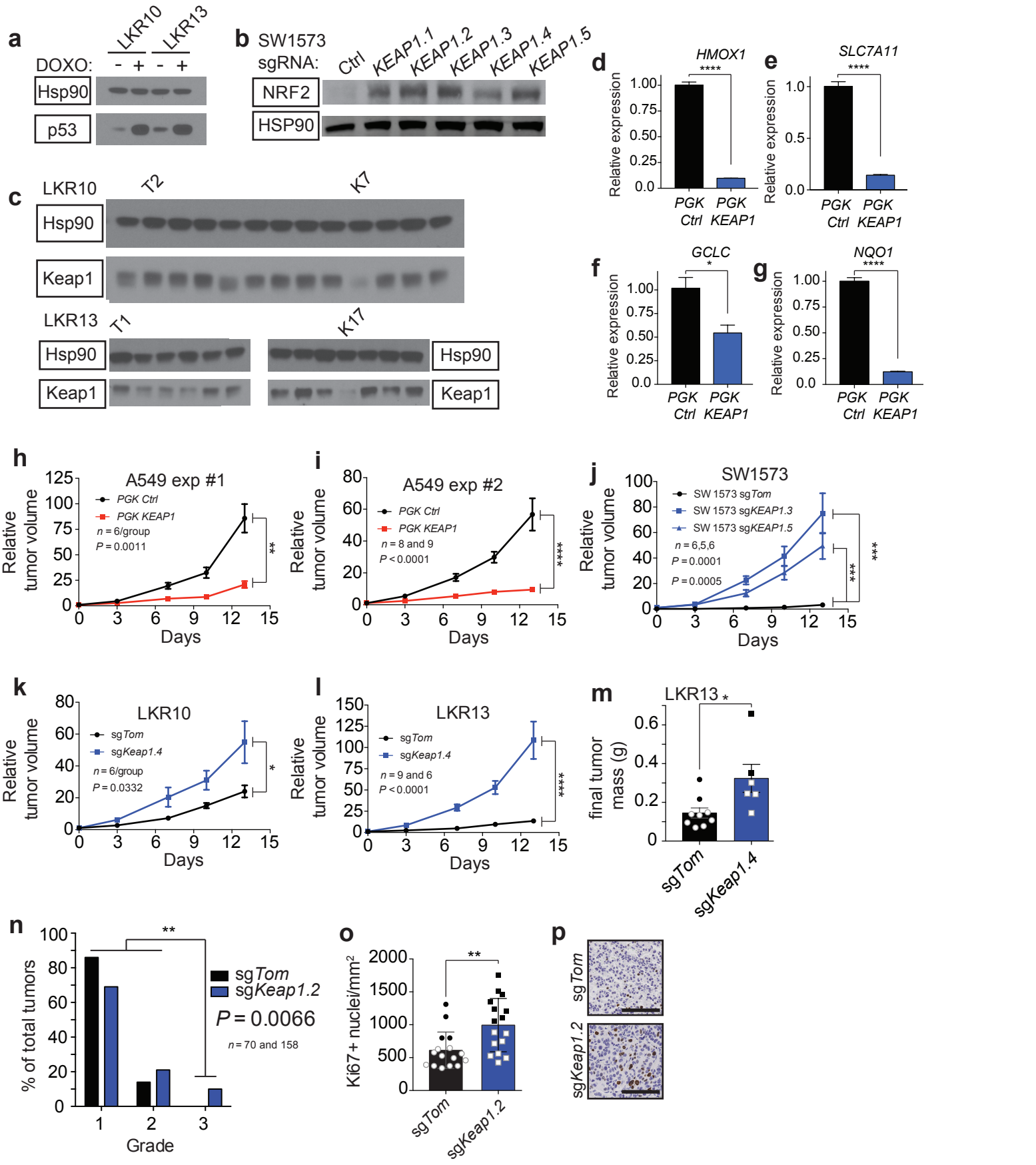
**Supplementary Figure 3**



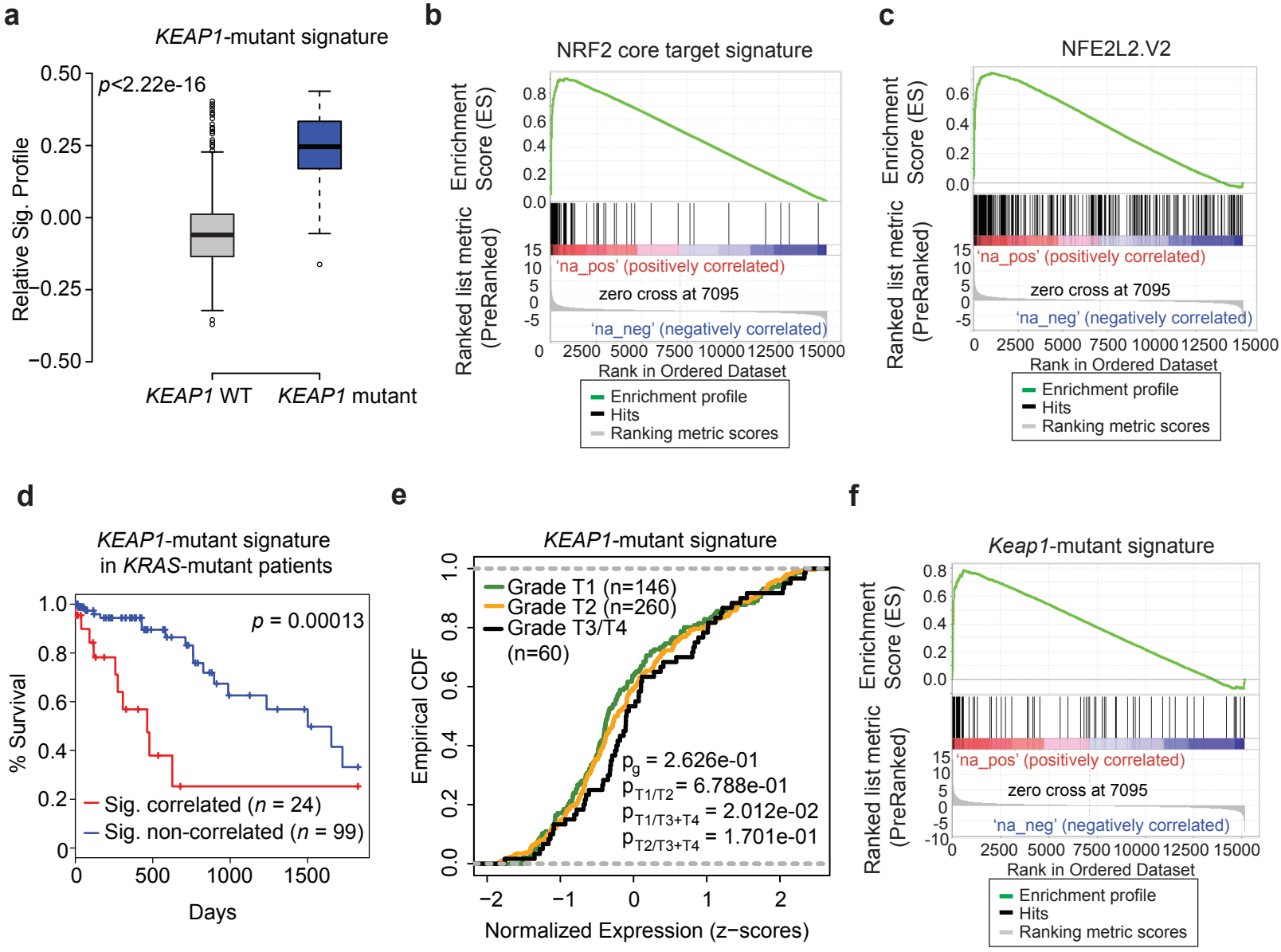
***Supplementary Figure 4***



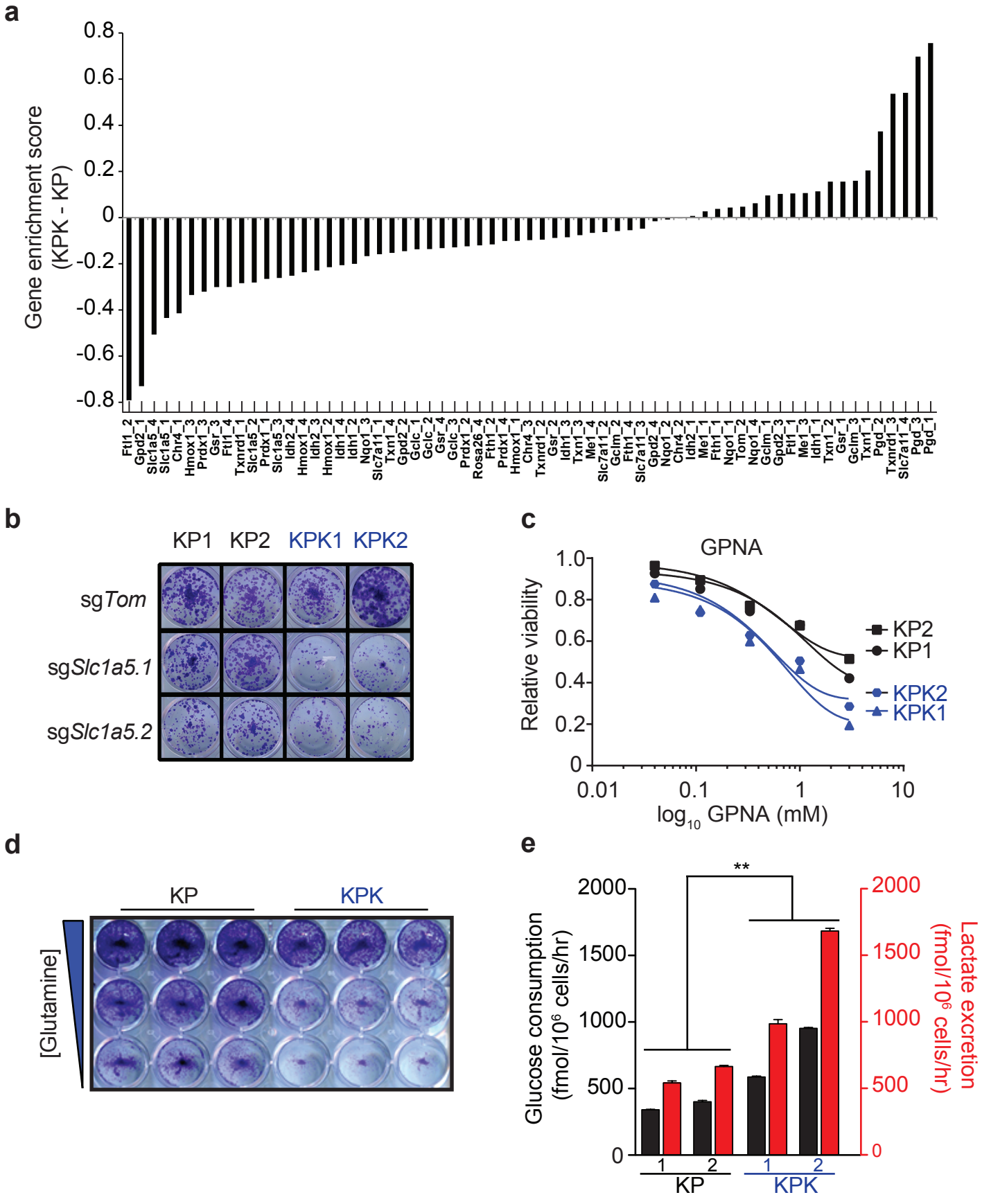
## *Supplementary Figure 5*



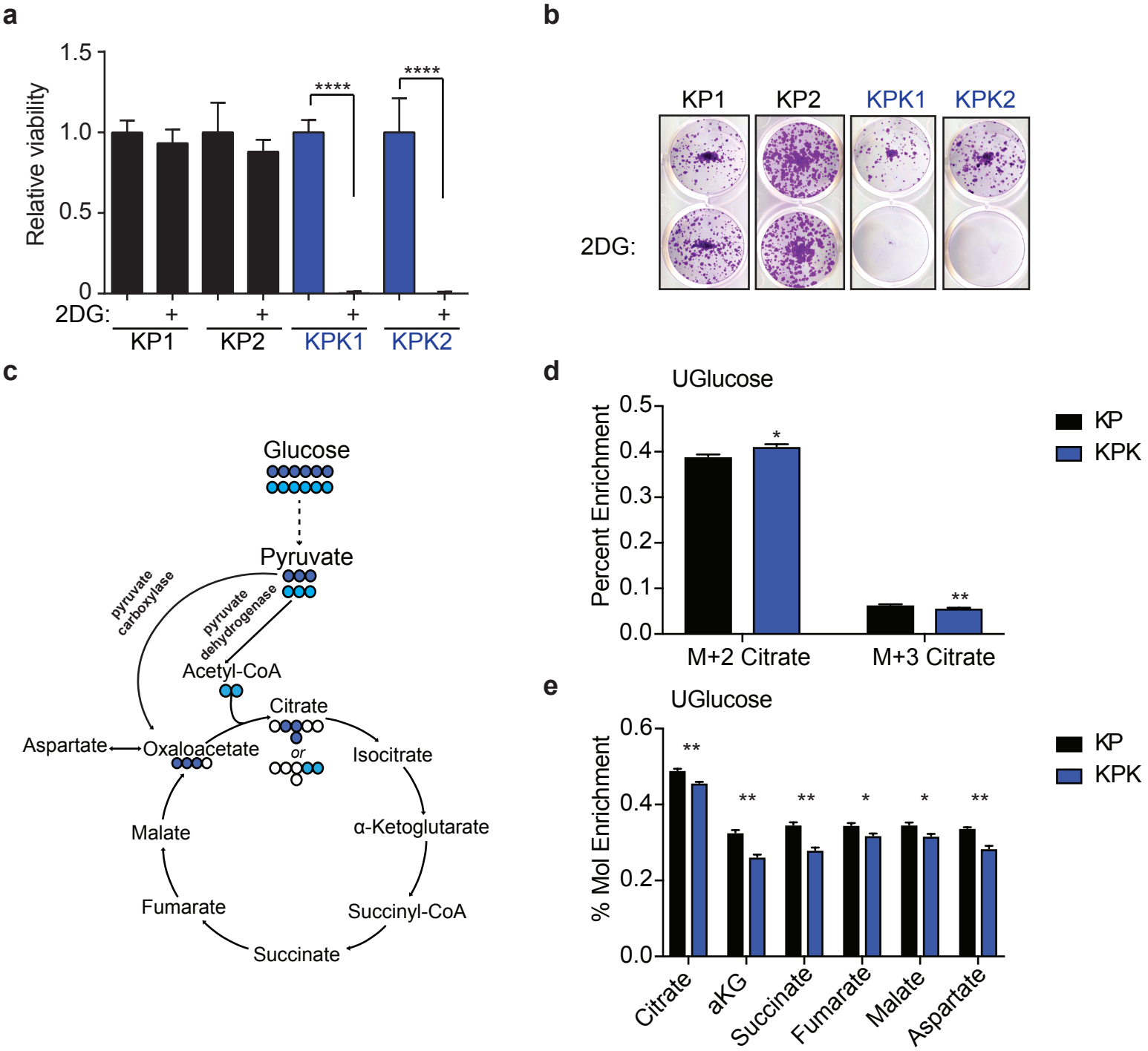
## *Supplementary Figure 6*



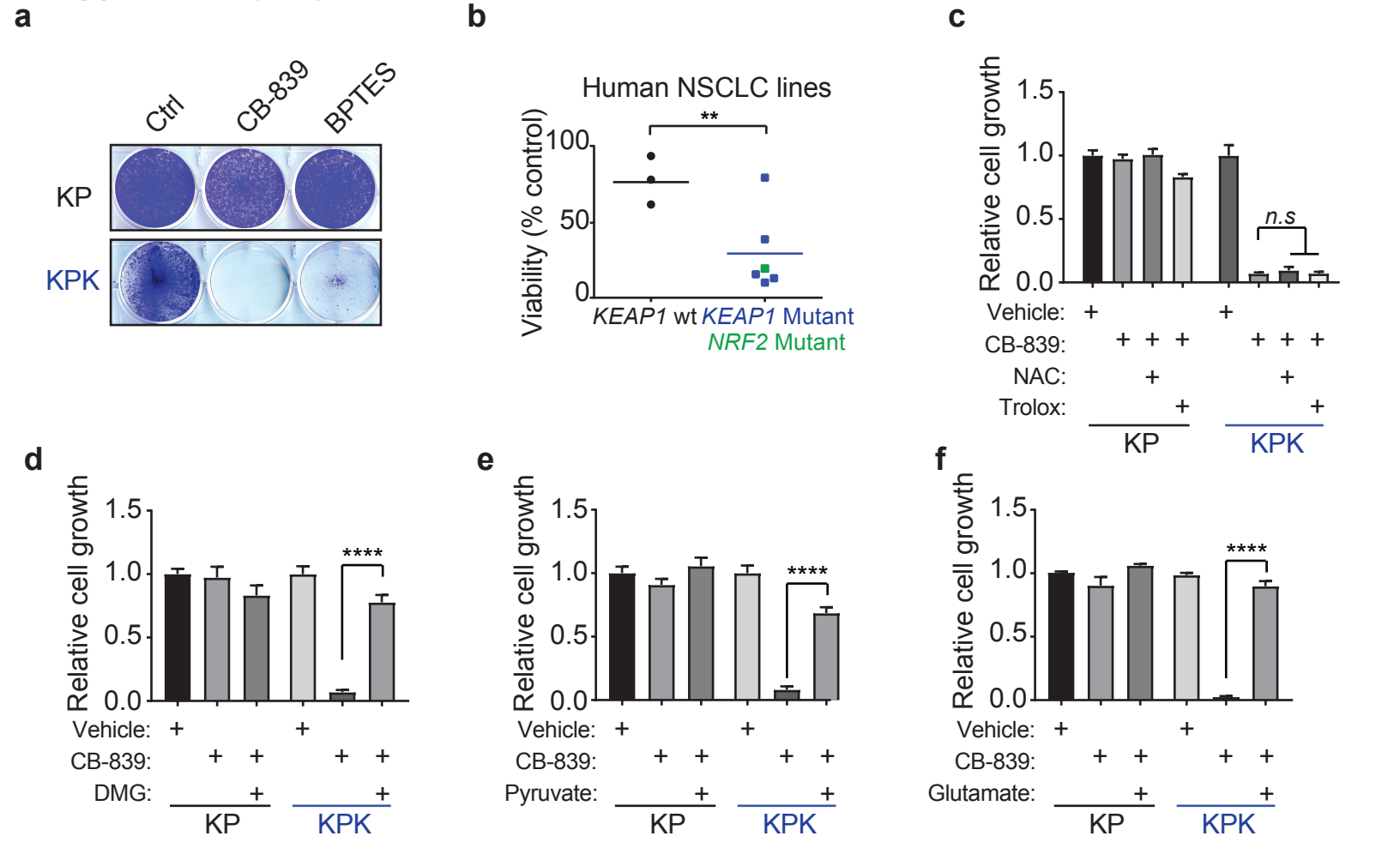
## *Supplementary Figure 7*



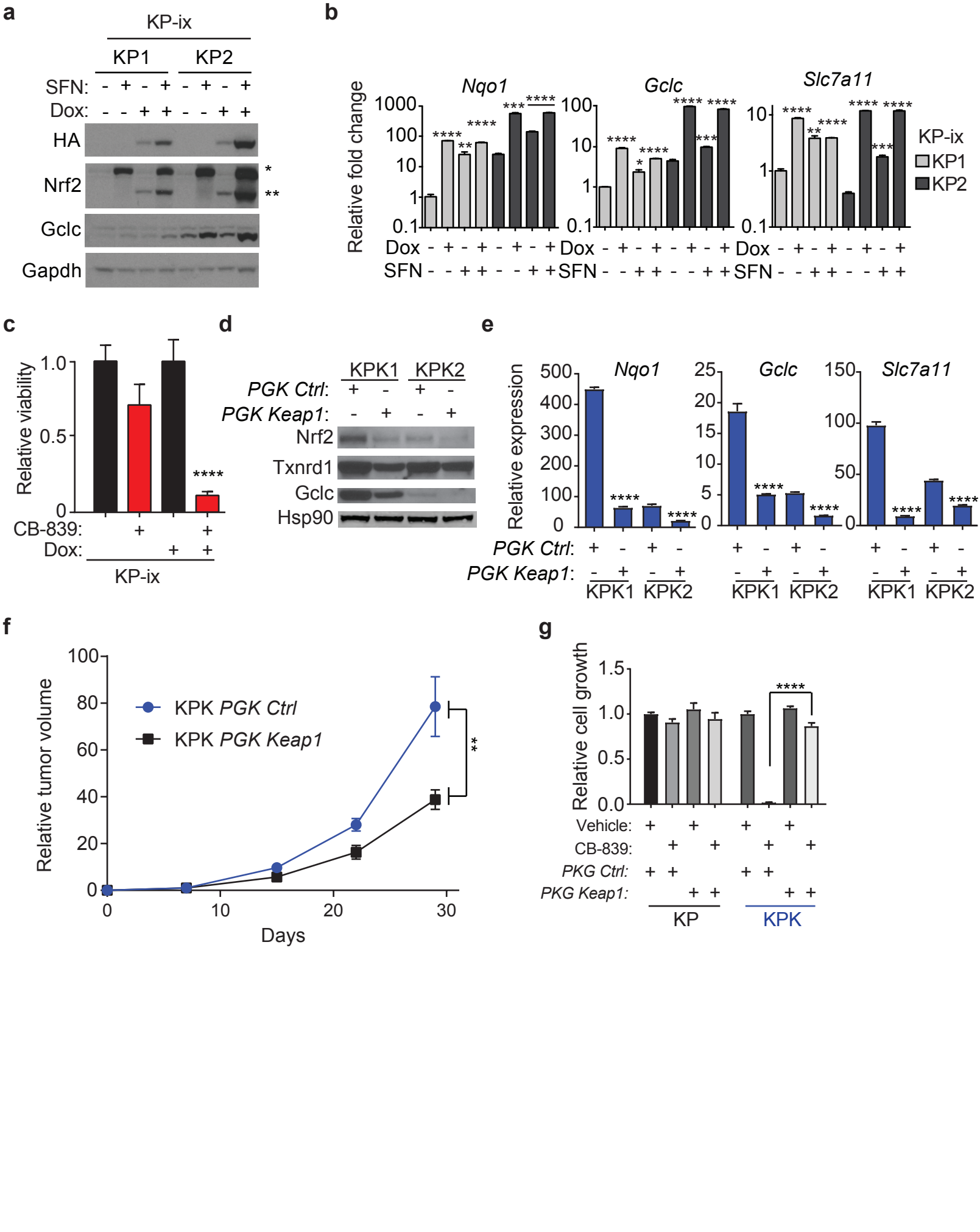
***Supplementary Figure 8***



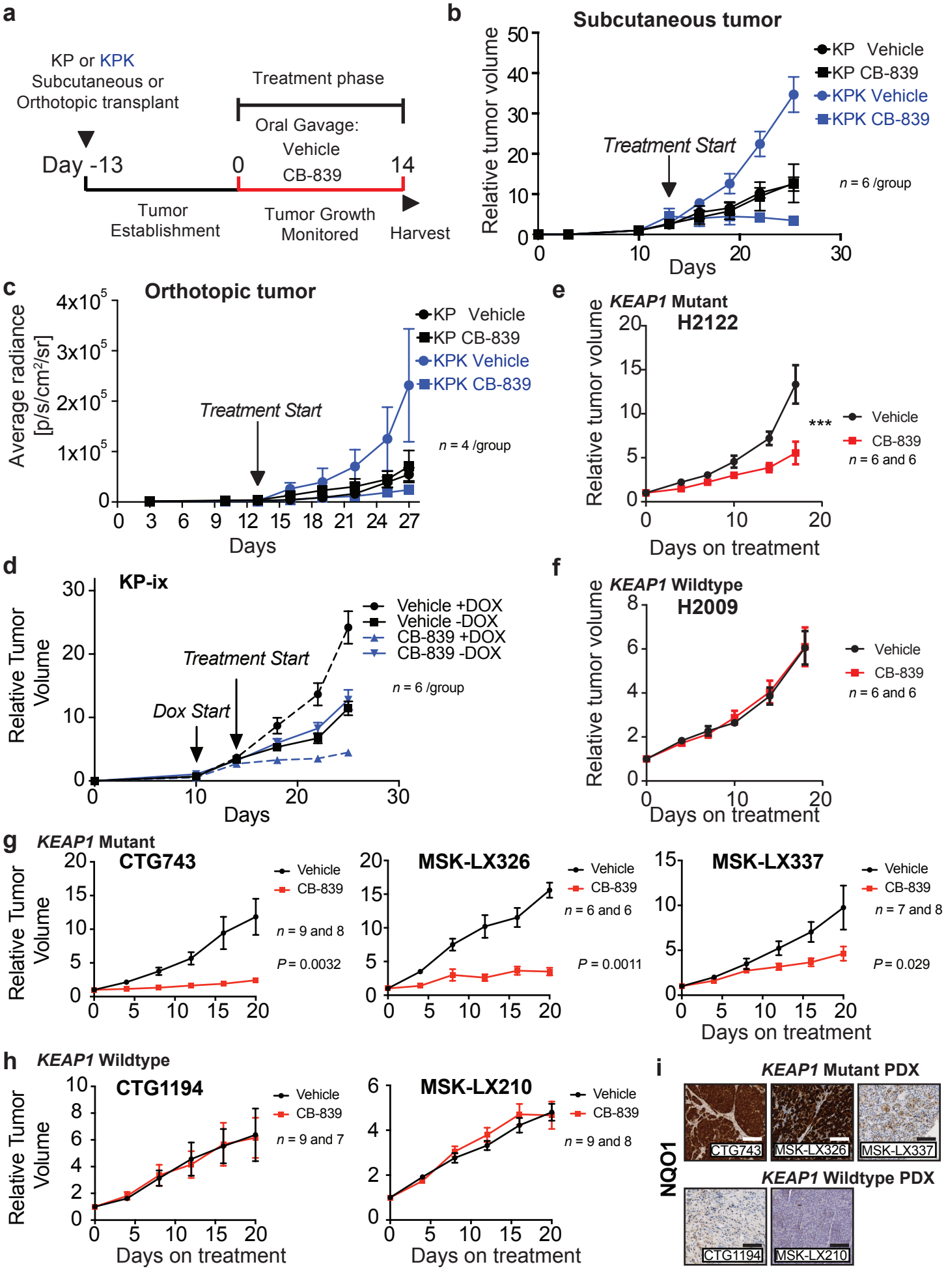
***Supplementary Figure 9***



***Supplementary Figure 10***

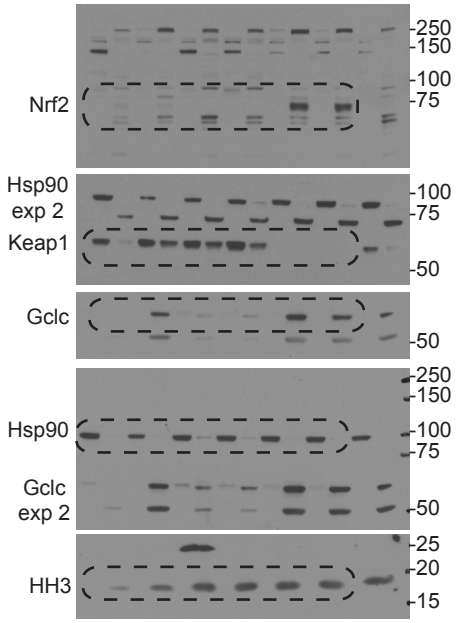


***Supplementary Figure 11***

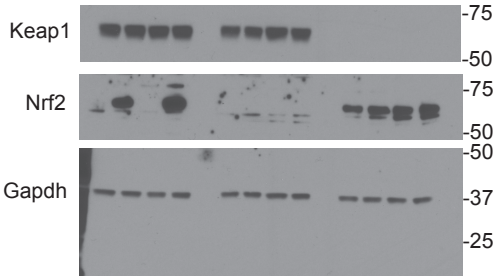


# Supplementary Figure 12

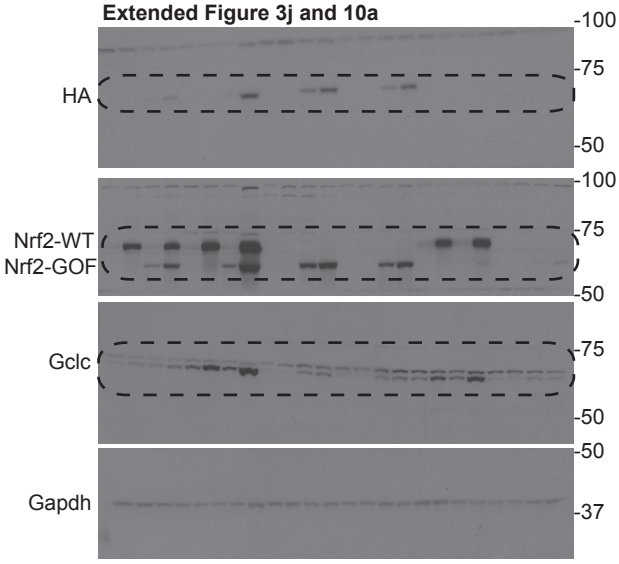
**Extended Figure 2c**



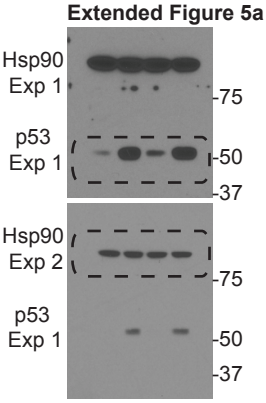

**Extended Figure 2e**



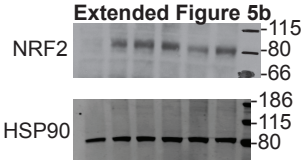

**Extended Figure 3j and 10a**



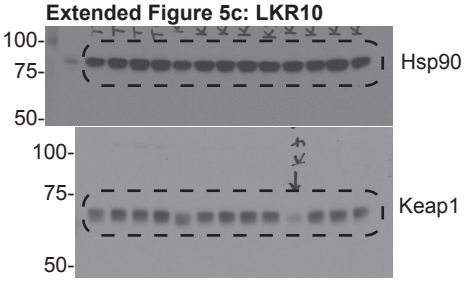

**Extended Figure 5a**



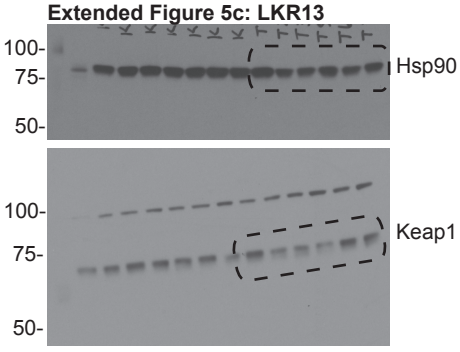

**Extended Figure 5b**



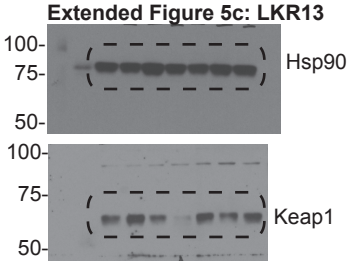

**Extended Figure 5c: LKR10**



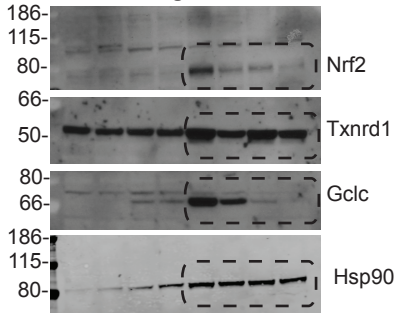

**Extended Figure 5c: LKR13**




**Extended Figure 5c: LKR13**




**Extended Figure 10d**



## Methods

### Mice

All animal studies described in this study were approved by the MIT Institutional Animal Care and Use Committee. *Kras*$^{LSL\text{-}G12D}$ and *Trp53*$^{flox}$ mice have already been described[41,42]. For all animal studies, >3 animals were used for each experimental cohort per specified genotype. All mice were maintained on a mixed C57BL/6:SV129 genetic background. Total burden, and grading analysis were conducted on >3 mice per genotype. No animals were excluded from analysis. Animals with the appropriate genotypes between the ages of 6-8 weeks were randomly selected to begin tumor initiation studies with pSECC-sgTom or pSECC-sgKeap1. Mice were infected intratracheally with lentiviruses as described[11]. Total lung area occupied by tumor was measured on hematoxylin and eosin (H&E) stained slides using NIS-elements software. All burden analysis and IHC was done in a blinded fashion, in which the researcher was unaware of which genotype the sample came from.

### Cell culture

Parental cell lines from KP[43] and LKR[18] mice were previously established and described. Human cell lines were acquired from ATCC. All lines were tested negative for mycoplasma. Cells were maintained in DMEM or RPMI supplemented with 10% Fetal Bovine Serum and gentamicin. Cell lines expressing rtTA were kept under Neomycin selection (400ug/mL). Cell lines expressing dox inducible Nrf2 constructs remained under Hygromycin selection (600ug/mL). Cells were treated with inhibitors D,L-Sulforaphane (SFN, EMD Millipore Calbiochem), Dimethyl fumarate (DMF, Sigma Aldrich), L-Buthionine-sulfoximine (BSO, Sigma Aldrich), Auranofin (AUR, TOCRIS bioscience), Erastin (ERA, Sigma Aldrich), L-Glutamic acid y-(p-nitroanilide)-hydrochloride (GPNA, Sigma Aldrich), 2-deoxy-D-glucose (2DG, Acros Organics), BPTES (Sigma Aldrich), CB-839 ( provided by Craig J. Thomas) and antioxidants or metabolites Trolox (Acros Organics), *N*-acetyl-L-cysteine (NAC, Sigma Aldrich), 6mM glutamate (Sigma Aldrich), 2mM pyruvate (Gibco), and 2mM dimethyl-2-oxoglutarate (DMG, Sigma Aldrich). Viability in the presence of all compounds was assessed

by cell titer glo (Promega #G7570) and trypan blue exclusion on a Countess II automated cell counter (Life Technologies). For clonogenic and low-density assays, cells were stained with Crystal Violet solution (25% Methanol 75% H20). For cell counts after DMF treatment cells were fixed with 4% paraformaldehyde for 15 minutes at 4°C. Cells were then washed in ice cold PBS and then stained with Hoechst DNA stain. Plates were quantified using a Tecan infinite M200 Pro plate reader or a SpectraMax M5 Microplate reader (Molecular Devices).

**Focused CRISPR/Cas9 genetic screen**

Oligonucleotides for sgRNAs were synthesized by Integrated DNA Technologies, annealed in vitro and inserted into lentiCRISPR-V2[44]. Cloned products were then transformed into E. coli 10G SUPREME electrocompetent cells (Lucigen). This plasmid pool was used to generate lentivirus-containing supernatants. The titer of lentiviral supernatants was determined by infecting targets cells at several amounts of virus in the presence of polybrene (8ug/mL; Millipore #TR-1003-G), counting the number of drug resistant infected cells after 3 days of selection. KP and KPK cells were infected at an MOI of ~0.5 and selected with puromycin (3ug/ml) 72 hours after infection. An initial pool of cells was harvested for genomic DNA extraction. The remaining cells were cultured for 14 doublings, after which cells were harvested for genomic DNA extraction. sgRNA inserts were PCR amplified, purified and sequenced on a MiSeq (Illumina) according to prior studies[44]. Sequencing reads were mapped and the abundance of each sgRNA was tallied. Gene score is defined as the median $\log_2$ fold change in the abundance between the initial and final population of all sgRNAs targeting that gene. The differential gene score is the difference between KP and KPK cell gene scores.

**Immunobloting**

Cells were lysed in 250 µL ice-cold RIPA buffer (Pierce, #89900) supplemented with 1× Complete Mini inhibitor mixture (Roche, #11 836 153 001) and mixed on a rotator at 4°C for 30 minutes. The protein concentration of the cell lysates was quantified using the Bio-Rad

DC Protein Assay (Catalog #500-0114). 50–80 μg of total protein was separated on 4–12% Bis-Tris gradient gels (Bio-Rad) by SDS-PAGE and then transferred to nitrocellulose membranes. The following antibodies were used for immunoblotting: anti-FLAG (Sigma, F1804, 1:1000), anti-GAPDH (Santa Cruz, sc-25778, 1:500), anti-Hsp90 (BD, #610418, 1:10,000), anti-Nrf2 (Santa-Cruz, sc-722, 1:200 and custom antibody provided by Edward Schmidt at 1:200), anti-Keap1 (CST, #8047, 1:1000), anti-Gclc (Santa Cruz, sc-22755, 1:200), anti-Slc1a5 (Santa Cruz, ASCT2 (M-63) sc-99003, 1:100), anti-Txnrd1 (Abcam, ab124954, 1:1000), and anti-p53 (CST, 2524S lot 12, 1:1000).

**Immunohistochemistry and immunofluorescence**

Mice were euthanized by carbon dioxide asphyxiation. Lungs were perfused through the trachea with 4% paraformaldehyde (PFA), fixed overnight, transferred to 70% ethanol and subsequently embedded in paraffin. Sections were cut at a thickness of four micrometers and stained with H&E for pathological examination. Chromogenic immunohistochemistry (IHC) was performed on a Ventana Medical Systems Discovery XT instrument with online deparaffinization using Ventana's reagents and detection kits and antigen retrieved in Ventana Cell Conditioner 1 or 2. The following antibodies were used for IHC: anti-phospho-Histone H3 (pHH3) (Ser10; Cell Signaling, 9701, 1:200), anti-Ki67 (Spring Bioscience, Cat# M3062, 1:400), anti-Nqo1 (Sigma Aldrich, HPA007308, 1:100), anti-Nrf2 (Provided by Edward E. Schmidt lab, 1:100), and anti-8-Hydroxydeoxyguanosine(8-oxo-dg) (Abcam, ab48508, N45.1, 1:200). Horseradish peroxidase (HRP) detection was used for NQO1, NRF2, pHH3 and Ki67. Alkaline phosphatase (AP) detection was used for 8-OXO and was visualized with Fast Red chromogen. NQO1 and NRF2 was antigen retrieved in Ventana Cell Conditioner 1 (Tris-Borate-EDTA). Antigen retrieval was performed with Ventana Cell Conditioner 2 (Citrate) for 8-OXO, Ki67 and pHH3. Pictures were obtained using a Nikon 80i microscope with a DS-U3 camera and NIS-elements software and with a digital whole slide scanner Leica SCN400F and Slidepath software version 4.0.8.

**Genomic DNA isolation**

Genomic DNA from entire snap-frozen left lung lobes or microdissected tumors was isolated using the High Pure PCR Template Preparation Kit (Roche #11796828001) following manufacturer guidelines. PCR products for MiSeq (Supplementary Sequences) were amplified using Herculase II Fusion DNA polymerase (Agilent #600679) (see Supplementary Sequences for primers used for genomic DNA isolation).

**Lentiviral production**

Lentiviruses were produced by co-transfection of 293 cells with lentiviral backbone constructs and packaging vectors (delta8.2 and VSV-G) using TransIT-LT1 (Mirus Bio #MR 2306). Supernatant was collected 48 and 72 hours post-transfection, concentrated by ultracentrifugation at 25,000 RPM for 90 minutes and resuspended in an appropriate volume of OptiMEM (Gibco #31985-062).

**Lentiviral vectors and sgRNA cloning**

pSECC lentiviral vector and cloning strategy was previously described[11].
For CRISPR experiments the lentiCRISPR-V2 lentiviral vector was used[45]. For sgRNA cloning, the lentiCRISPR-V2 vector was digested with BsmBI and ligated with BsmBI-compatible annealed oligos for sgRNAs (Supplementary Sequences).

**Tumor purity correction**

Lung lobe and microdissected tumor genomic DNA was used to perform real-time PCR based analysis to detect the relative levels of the un-recombined *lox-stop-lox Kras$^{G12D}$* allele using forward primer 5'-ctcttgcctacgccaccagctc-3' and reverse primer 5'-agctagcca ccatggcttgagtaagtctgc a-3'. To correct for DNA loading of each sample, we amplified chr5:10054507-10054621 using forward primer 5'-gaagaaattagagggcatgcttc-3' and reverse primer 5'-cttctcccagtgaccttatgta-3'. Real-time PCR reaction was performed using KAPA Fast SYBR master mix in a Roche LightCycler Real-Time PCR instrument. To calculate percent

purity we performed the following calculations for each sample: $\Delta Cp^{tumorX} = Cp^{Chr5} - Cp^{LSL\text{-}KrasG12D}$ to normalize for sample loading and then calculated $1/\Delta\Delta Cp = (\Delta Cp^{X} - \Delta Cp^{LungControl})$ for each sample.

**Transciptome analysis**

RNA was collected from cells as before[46] with RNeasy plus mini kit (Qiagen). For Real Time qPCR analysis, complementary DNA (cDNA) was synthesized from RNA with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems #4368814). **Genes** *Slc7a11*, *Gclc*, *Hmox1*, and *Nqo1* were analyzed by quantitative reverse transcription polymerase chain reaction on LightCycler 480 II (Roche). RT qPCR primers (Supplementary Sequences).

**Glutaminase Inhibitor**

Animals were treated as before with 200 mg/kg CB-839 or vehicle twice a day after tumor establishment phase. The vehicle contained 25% (w/v) hydroxypropyl-β-cyclodextrin in 10 mmol/L citrate (pH 2.0), and CB-839 was formulated at 20 mg/mL for a final dosing volume of 10 mL/kg.

**Extracellular Flux Measurements**

Extracellular flux measurements were calculated by extracting fresh and spent medium supernatant from tracing experiments after 24 hours of growth. Cells were assumed to grow exponentially over the culture period. Glucose, lactate and glutamine were measured using YSI biochemistry analyzer (Yellow Springs Instruments, Yellow Springs, OH).

**ROS and glutathione**

ROS in cultured cells were measured by incubating $1x10^6$ cells with 5 μM CM-H2DCFDA (C6827, Life Technologies) for 30 min at 37°C. DCF fluorescence was acquired on the Attune NxT (ThermoFisher) flow cytometer and analyzed using FlowJo software (Tree Star).

Reduced (GSH) and oxidized glutathione (GSSG) was measured with a GSH/GSSG-Glo Assay kit (V6611, Promega) for the indicated amount of time.

**Statistics**

For statistical analyses, we used GraphPad Prism software v.6.03, variance was similar between the groups that were compared: P-values were determined by Student's t-test for all measurements of tumor burden and IHC quantifications except for contingency tables, in which Fisher's exact test or Chi-square test were used. One-way analysis of variance (ANOVA) with Tukey's post hoc test were used for comparisons between multiple groups; for analysis between groups over multiple time measurements (growth curves) two-way ANOVA was used. Figure legends denominate statistical analysis used. Standardized IC50 values for the heatmap displayed in Extended data 3a were calculated as follows: $z=(X-\mu)/\sigma$, where z is the z-score. Patient co-occurrence was assessed using Fisher's exact test and enrichment was assessed using the Hypergeometric test. All error bars denote s.e.m. Detailed information on experimental design and reagents can be found online in the Life Sciences Reporting Summary accompanying this manuscript.

**Bioinformatic analysis of CRISPR-targeted loci**

For libraries prepared with the Nextera DNA library prep kit, Illumina MiSeq reads were trimmed to 120bp after reviewing base quality profiles, in order to drop lower quality 3′ends. Traces of Nextera adapters were clipped using the FAS TX toolkit (Hannon Lab, CSHL) and pairs with each read greater than 15bp in length were retained. Additionally, read pairs where either read had 50% or more bases below a base quality threshold of Q30 (Sanger) were dropped from subsequent analyses. For PCR amplicons (sequenced at the MGH sequencing facility), 142bp paired-end reads were used in downstream analyses. The reference sequence of the target locus was supplemented with 10bp genomic flanks and was indexed using an enhanced suffix array[47]. Read ends were anchored in the reference

sequence using 10bp terminal segments for a suffix array index lookup to search for exact matches. A sliding window of unit step size and a maximal soft-clip limit of 10bp was used to search for possible anchors at either end of each read. For each read, optimal Smith-Waterman dynamic programming alignment[48] was performed between the reduced state space of the read sequence and the corresponding reference sequence spanning the maximally distanced anchor locations. Scoring parameters were selected to allow for sensitive detection of short and long insertions and deletions while allowing for up to four mismatches and the highest scoring alignment was selected. Read pairs with both reads aligned in the proper orientation were processed to summarize the number of wild-type reads and the location and size of each insertion and deletion event. Overlapping reads within pairs were both required to support the event if they overlapped across the event location. Additionally, mutation events and wild-type reads were summarized within the extents of the sgRNA sequence and PAM site by considering read alignments that had a minimum of 20bp overlap with this region. Mutation calls were translated to genomic coordinates and subsequently annotated using Annovar[49]The alignment and post-processing code was implemented in C++ along with library functions from SeqAn[50] and SSW and utility functions in Perl and R (www.R-project.org). Mutation calls were subjected to manual review using the Integrated Genomics Viewer (IGV)[51].

**NYU Targeted Exome Capture Sequencing**

All protein-coding exons for the genes of interest were sequenced using the following methodology. 500ng of DNA from each sample were sheared to an average of 150 bp in a Covaris instrument for 360 seconds (Duty cycle - 10%; intensity - 5; cycles/Burst - 200). Barcoded libraries were prepared using the Kapa Low-Throughput Library Preparation Kit Standard (Kapa Biosystems). Libraries will be amplified using the KAPA HiFi Library Amplification kit (Kapa Biosystems) (8 cycles) and quantified using Qubit Fluorimetric Quantitation (Invitrogen) and Agilent Bioanalyzer. An equimolar pool of 24 barcoded libraries were used as input for hybridization-based capture using one reaction the XGen

Lockdown predesigned probes (Integrated DNA Technologies, IDT) target the coding exons of the genes of interest. Capture by hybridization was performed according to the IDT protocol. The final pooled capture libraries were quantified by Qubit (Invitrogen) and Bioanalyzer (Agilent) and sequenced in on an Illumina HiSeq 4000 as paired end 150 nucleotides.

**Methods for Lung Cancer Collection**

Eighty-eight primary lung cancers along with matching blood mononuclear cells and remote normal lung which were snap frozen in liquid nitrogen at the time of resection. DNA extractions were performed (Qiagen, Dusseldorf, Germany) for targeted exome capture, and matching formalin fixed paraffin embedded tissue used for NQ01 staining. All use of human tissue and body fluids was approved under NYU IRB protocol 8896.

**Human exome data analysis**

Sequencing results were demultiplexed and converted to FASTQ format using Illumina bcl2fastq software. The reads were adapter and quality trimmed with Trimmomatic [52]and then aligned to the human genome (build hg19/GRCh37) using the Burrows-Wheeler Aligner with the BWA-MEM algorithm [53]. Duplicate reads were removed using Sambamba [54]. Further local indel realignment and base-quality score recalibration and are performed using the Genome Analysis Toolkit (GATK) [55]. Single-nucleotide and small indel somatic variants were called with MuTect2 [56]. ANNOVAR [49]was used to annotate variants with functional consequence on genes as well as identifying presence in dbSNP, ExAC, 1000 Genomes project, and COSMIC.

**Allograft, Xenograft and Patient Derived Xenograft**

For allograft experiments, mouse lung tumor derived cells were transplanted subcutaneously ($1x10^6$ cells) under the skin or orthotopically ($2.5x10^5$ cells) in the lung of Nude or NSG mice. For Xenograft experiments human lung cancer cell lines ($1x10^6$ cells) were transplanted subcutaneously in to NSG mice with a 1:1 ratio of matrigel. Patient Derived Xenografts

(PDX) were revived and passaged ones in NSG mice, a live 2x2mm tumor piece was surgically transplanted subcutaneously in to recipient NSG mice under anesthesia. After 6 weeks of recovery and engraftment, tumors with volumes 25-100$mm^2$ were randomized to receive vehicle or CB-839 treatment (See section **Glutaminase Inhibitor**). Details on generation of PDX have been previously described in detail[57], MSK-IMPACT[58] was performed on xenografts after mouse stroma depletion and mutations were called against matched normal blood as previously detailed[59]. Subcutaneous tumor volumes were calculated according to the following formula: ($mm^3$; ($a^2$*$b$)*(π/6) where *a* is the smaller dimension and b is the larger dimension.

**Human clinical data analyses**

Genomic data for lung adenocarcinoma patient samples (*n* = 548) were obtained from the Cancer Genome Atlas (TCGA LUAD; http://cancergenome.nih.gov/). This included RNA-seq gene expression profiles of primary tumor patient samples (*n* = 488), mutation calls, and associated clinical data (*n* = 458 patients with RNA-seq data for primary tumors had associated survival data). Individual sample expression profiles were scored with gene expression signatures using ssGSEA[60,61]. Patients were stratified according to their correlation score, into equal top and bottom percentile sets (or top-scoring n% versus rest of the cohort). Kaplan-Meier survival analysis was conducted between these sets of patients and the log-rank test was used to assess significance. The murine-derived *Keap1*-mutant signature was similarly used to perform Kaplan-Meier survival analyses after translation of mouse gene names to human nomenclature (www.genenames.org). Additionally, the Cox proportional hazards regression model was used to analyze the prognostic value of the human-derived *KEAP1*-mutant signature across all patients within the TCGA LUAD cohort, in the context of additional clinical covariates. All univariate and multivariable analyses were conducted within a 5-year survival timeframe. The following patient and tumor-stage clinical characteristics were used: Signature (*KEAP1*-mutant signature strong vs. weak correlation); Gender (male vs. female); Age (years, continuous);

Smoking History (reformed > 15yrs vs. non-smoker, reformed < 15yrs vs. non-smoker, current smoker vs. non-smoker); Union for International Cancer Control (UICC) TNM Stage specification (Stage III/IV vs. I/II); UICC T score specification (T2 vs. T1, T3/T4 vs. T1); UICC N score specification (N1/N2 vs. N0).

Hazard ratio proportionality assumptions for the Cox regression model were validated by testing for all interactions simultaneously ($p$ = 0.723). Interactions between the *KEAP1* mutant signature and TNM stage, T score, and N score (significant covariates in the model) were tested using a likelihood ratio test (LRT) to contrast a model consisting of both covariates with another model consisting of both covariates plus an interaction term. No statistically significant difference was found between the two models (TNM: $p$ = 0.445, T score: $p$ = 0.455, N score: $p$ = 0.494; likelihood ratio test). To test for statistically significant associations between the *KEAP1* mutant signature correlation scores and TCGA LUAD TNM stage (stageI-IV) and grade levels (T-scores), the Kurskal-Wallis test was used to assess overall significance and the Kolmogorov-Smirnov test was used to assess pairwise differences. Results were visualized using Empirical Cumulative Distribution Function (ECDF) plots. All statistical analyses were conducted in R (www.R-project.org) and all survival analyses and were conducted using the survival package in R[47].

### *Nrf2* core target signature

Three published datasets were utilized to derive a high-confidence 108-gene signature of Nrf2-induced targets. Normalized microarray expression data for GSE38332[62] was downloaded from the Gene Expression Omnibus (GEO). Differential expression analysis using R/limma[63] was performed to identify genes that are differentially regulated between control and Nrf2-siRNA treated samples. A list of Nrf2-induced genes ($n$ = 433) was identified using FDR < 0.05 and fold-change < 1.5 thresholds. Nrf2 targets ($n$ = 345) derived by [7] from microarray analysis of A549 cells treated with Nrf2 siRNA versus a control siRNA were used as the second dataset (supplementary table 1 from[7]). High-confidence Nrf2 targets ($n$ = 244) derived by[20] from integrated analyses of microarray gene expression and

ChIP-seq data was used as the third dataset (supplementary table 5 from[20]). All non-human gene names were mapped to human gene equivalents where needed. Genes that overlapped between two or more datasets were included in the high-confidence Nrf2 core target signature and used in downstream analyses (see supplementary tables). Differential gene expression analyses were conducted in R.

**Gene expression signature analyses**

Illumina HiSeq 2000 50-nt single-ended reads were mapped to the UCSC mm9 mouse genome build (http://genome.ucsc.edu/) using RSEM[64]. Raw estimated expression counts were upper-quartile normalized to a count of 1000[65]. *Keap1*-mutant (*n* = 2), WT (*n* = 2), and WT plus SFN treated (*n* = 2) samples were jointly analyzed to derive a murine signature of *Keap1*-mutant gene expression changes. Given the complexity of the database in terms of a mixture of genotypes and treatment, a high-resolution signature discovery approach (Independent Component Analysis) was employed to characterize global gene expression profiles, as described previously[43,46,66]. This unsupervised blind source separation technique was used on this discrete count-based expression dataset to elucidate statistically independent and biologically relevant signatures. ICA is a signal processing and multivariate data analysis technique in the category of unsupervised matrix factorization methods. Conceptually, ICA decomposes the overall expression dataset into independent signals (gene expression patterns) that represent distinct signatures. High-ranking positively and negatively correlated genes in each signature represent gene sets that drive the corresponding expression pattern (in either direction). Each signature is thus two-sided, allowing for identification of up-regulated and down-regulated genes for each signature within each sample. Formally, utilizing input data consisting of a genes-samples matrix, ICA uses higher order moments to characterize the dataset as a linear combination of statistically independent latent variables. These latent variables represent independent components based on maximizing non-gaussianity, and can be interpreted as independent source signals that have been mixed together to form the dataset under consideration. Each

component includes a weight assignment to each gene that quantifies its contribution to that component. Additionally, ICA derives a mixing matrix that describes the contribution of each sample towards the signal embodied in each component. This mixing matrix can be used to select signatures among components with distinct gene expression profiles across the set of samples. The R implementation of the core JADE algorithm (Joint Approximate Diagonalization of Eigenmatrices)[67] was used along with custom R utilities. Statistical significance of biologically relevant signatures was assessed using the Mann-Whitney-Wilcoxon test (alpha = 0.05). A murine *Keap1*-mutant signature was derived from this analysis, identifying genes with a differential expression pattern between wild-type and all other samples. Genes from the resulting signature with |z-score| > 2 were used in subsequent Kaplan-Meier and Cox regression survival analyses. Similarly, expression profiles from the TCGA human lung-adenocarcinoma cohort were analyzed to derive a *KEAP1*-mutant gene expression signature. Utilizing mutation calls from TCGA (MAF files), patient primary tumor samples with protein altering mutations in *KEAP1* (n = 79) and wild-type *KEAP1* (n = 380) were identified. A combined dataset of these samples was analyzed (utilizing ICA) to detect a statistically significant expression pattern (Mann-Whitney-Wilcoxon test) separating mutant from wild-type samples. Genes from the resulting signature with |z-score| > 2 were used in subsequent Kaplan-Meier and Cox regression survival analyses. All RNA-seq analyses were conducted in the R Statistical Programming language (http://www.r-project.org/). Gene set enrichment analysis (GSEA) was carried out using the pre-ranked mode with default settings[61].

**GC/MS analysis of glucose traced metabolites**

$2x10^5$ cells were seeded in 2mL of RPMI-1640 in 6 well plates. Media was then replaced with 2mL of fresh RPMI-1640 containing 11mM of [U-$^{13}$C]-D-glucose. Cells were cultured for 24 hours to reach steady state labeling of TCA cycle intermediates. Cells were washed 1X in ice cold saline and then collected by scraping in 600uL of 80% (v/v) of ice cold methanol containing 1.4ug/mL norvaline (Sigma Aldrich). Samples were vortexed for 10 minutes at

4°C and then centrifuged at max speed for 10 minutes. Supernatant was transferred to fresh tubes and then dried under nitrogen. Dried and frozen metabolite extracts were then derivatized with 16uL of MOX reagent (Thermofisher) for 60 minutes at 37°C and *N*-tert-butyldimethylchlorosilane (Sigma Aldrich) for 30 minutes at 60°C. After derivatization, samples were analyzed by GC-MS using a DB-35MS column (Agilent Technologies) in an Agilent 7890A gas chromatograph coupled to an Agilent 5997B mass spectrometer. Helium was used as the carrier gas at a flow rate of 1.2mL/minute. One microliter of sample was injected in split mode (split 1:1) at 270°C. After injection, the GC oven was held at 100°C for 1 minute and then increased to 300°C at 3.5°C/minute. The oven was then ramped to 320°C at 20°C/minute and held for 5 minutes at 320°C.

The MS system operated under electron impact ionization at 70eV and the MS source and quaddrupole were held at 230°C and 150°C respectively. The detector was used in scanning mode, and the scanned ion range was 10-650m/z. Mass isotopomer distributions were determined by integrating appropriate ion fragments for each metabolite [68] using in-house software [69] that corrects for natural abundance using previously described methods [70].

**References**


41 Jackson, E. L. *et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev* **15**, 3243-3248, doi:10.1101/gad.943001 (2001).

42 Jackson, E. L. *et al.* The differential effects of mutant p53 alleles on advanced murine lung cancer. *Cancer Res* **65**, 10280-10288, doi:10.1158/0008-5472.CAN-05-2193 (2005).

43 Dimitrova, N. *et al.* Stromal Expression of miR-143/145 Promotes Neoangiogenesis in Lung Cancer Development. *Cancer Discov* **6**, 188-201, doi:10.1158/2159-8290.CD-15-0854 (2016).

44 Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* **11**, 783-784, doi:10.1038/nmeth.3047 (2014).

45 Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84-87, doi:10.1126/science.1247005 (2014).

46 Papagiannakopoulos, T. *et al.* Circadian Rhythm Disruption Promotes Lung Tumorigenesis. *Cell Metab* **24**, 324-331, doi:10.1016/j.cmet.2016.07.001 (2016).

47 Abouelhoda, M. I., Kurtz, S. & Ohlebusch, E. Replacing suffix trees with enhanced suffix arrays. *Journal of Discrete Algorithms* **2**, 53-86, doi:http://dx.doi.org/10.1016/S1570-8667(03)00065-0 (2004).

48 Smith, T. F. & Waterman, M. S. Identification of common molecular subsequences. *J Mol Biol* **147**, 195-197 (1981).

49 Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164, doi:10.1093/nar/gkq603 (2010).

50 Doring, A., Weese, D., Rausch, T. & Reinert, K. SeqAn an efficient, generic C++ library for sequence analysis. *BMC Bioinformatics* **9**, 11, doi:10.1186/1471-2105-9-11 (2008).

51 Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* **14**, 178-192, doi:10.1093/bib/bbs017 (2013).

52 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).

53 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).

54 Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics (Oxford, England)* **31**, 2032-2034, doi:10.1093/bioinformatics/btv098 (2015).

55 DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491-498, doi:10.1038/ng.806 (2011).

56 Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature biotechnology* **31**, 213-219, doi:10.1038/nbt.2514 (2013).

57 Gardner, E. E. *et al.* Chemosensitive Relapse in Small Cell Lung Cancer Proceeds through an EZH2-SLFN11 Axis. *Cancer Cell* **31**, 286-299, doi:10.1016/j.ccell.2017.01.006 (2017).

58 Cheng, D. T. *et al.* Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology. *J Mol Diagn* **17**, 251-264, doi:10.1016/j.jmoldx.2014.12.006 (2015).

59 Schneeberger, V. E., Allaj, V., Gardner, E. E., Poirier, J. T. & Rudin, C. M. Quantitation of Murine Stroma and Selective Purification of the Human Tumor Component of Patient-Derived Xenografts for Genomic Analysis. *PLoS One* **11**, e0160587, doi:10.1371/journal.pone.0160587 (2016).

60 Mootha, V. K. *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34**, 267-273, doi:10.1038/ng1180 (2003).

61 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

62 Singh, A. *et al.* Transcription factor NRF2 regulates miR-1 and miR-206 to drive tumorigenesis. *J Clin Invest* **123**, 2921-2934, doi:10.1172/JCI66353 (2013).

63 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47, doi:10.1093/nar/gkv007 (2015).

64 Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323, doi:10.1186/1471-2105-12-323 (2011).

65 Bullard, J. H., Purdom, E., Hansen, K. D. & Dudoit, S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* **11**, 94, doi:10.1186/1471-2105-11-94 (2010).

66 Li, C. M. *et al.* Foxa2 and Cdx2 cooperate with Nkx2-1 to inhibit lung adenocarcinoma metastasis. *Genes Dev* **29**, 1850-1862, doi:10.1101/gad.267393.115 (2015).

67 Biton, A., Zinovyev, A., Barillot, E. & Radvanyi, F. MineICA: Independent component analysis of transcriptomic data. (2013).

68 Lewis, C. A. *et al.* Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells. *Molecular cell* **55**, 253-263, doi:10.1016/j.molcel.2014.05.008 (2014).

69 Young, J. D., Walther, J. L., Antoniewicz, M. R., Yoo, H. & Stephanopoulos, G. An elementary metabolite unit (EMU) based method of isotopically nonstationary flux analysis. *Biotechnol Bioeng* **99**, 686-699, doi:10.1002/bit.21632 (2008).

70 Fernandez, C. A., Des Rosiers, C., Previs, S. F., David, F. & Brunengraber, H. Correction of 13C mass isotopomer distributions for natural stable isotope abundance. *Journal of mass spectrometry : JMS* **31**, 255-262, doi:10.1002/(sici)1096-9888(199603)31:3<255::aid-jms290>3.0.co;2-3 (1996).