Published in final edited form as:

Nat Cancer. 2020 June ; 1(6): 589–602. doi:10.1038/s43018-020-0071-1.

Keap1 mutation renders lung adenocarcinomas dependent on Slc33a1

Rodrigo Romero#[1,2], **Francisco J. Sánchez-Rivera**#[1,2], **Peter M.K. Westcott**[1], **Kim L. Mercer**[1,3], **Arjun Bhutkar**[1], **Alexander Muir**[1], **Tania J. González-Robles**[1], **Swanny A. Lamboy-Rodríguez**[2], **Laura Z. Liao**[2], **Sheng Rong Ng**[1,2], **Leanne Li**[1], **Caterina I. Colón**[1], **Santiago Naranjo**[1,2], **Mary Clare Beytagh**[2], **Caroline A. Lewis**[5], **Peggy P. Hsu**[1,6], **Roderick T. Bronson**[4,5], **Matthew G. Vander Heiden**[1,2,7], **Tyler Jacks**#,[1,2,3]

[1]Koch Institute for Integrative Cancer Research, Cambridge, MA, 02139, USA

[2]Massachusetts Institute of Technology Department of Biology, Cambridge, MA, 02139, USA

[3]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA

[4]Tufts University, Boston, Massachusetts, USA

[5]Harvard Medical School, Boston, Massachusetts, USA

[6]Massachusetts General Hospital Cancer Center, Boston, MA 02114, USA

[7]Dana Farber Cancer Institute, Boston, MA 02115, USA

# These authors contributed equally to this work.

Abstract

Approximately 20-30% of human lung adenocarcinomas (LUAD) harbor loss-of-function (LOF) mutations in Kelch-like ECH Associated-Protein 1 (*KEAP1*), which lead to hyperactivation of the nuclear factor, erythroid 2-like 2 (NRF2) antioxidant pathway and correlate with poor prognosis[1-3]. We previously showed that *Keap1* mutation accelerates KRAS-driven LUAD and produces a marked dependency on glutaminolysis[4]. To extend the investigation of genetic dependencies in the context of *Keap1* mutation, we performed a druggable genome CRISPR-Cas9 screen in *Keap1*-mutant cells. This analysis uncovered a profound *Keap1* mutant-specific dependency on solute carrier family 33 member 1 (*Slc33a1*), an endomembrane-associated protein with roles in autophagy regulation[5], as well as a series of functionally-related genes

---

#Corresponding author.

Author Contributions:

R.R., F.J.S.R., P.M.K.W., and T.J. designed the study; R.R., F.J.S.R., P.M.K.W., K.L.M., T.J.G.R. S.A.L.R., L.Z.L., and M.C.B. performed experiments; A.B. and P.M.K.W. performed bioinformatic analyses; A.M., P.P.H., and M.V.H. provided feedback and interpretation of metabolism data; S.R.N., L.L., and C.I.C. curated and generated the DGL; P.M.K.W. and R.T.B. trained Aiforia deep neural networks for histological assessment of lung tumor burden and grade; S.N. developed double-sgRNA cloning strategy; R.R., F.J.S.R., P.M.K.W, and T.J., wrote the manuscript with comments from all authors.

Declaration of Interest:

T.J. is a member of the Board of Directors of Amgen and Thermo Fisher Scientific. He is also a co-Founder of Dragonfly Therapeutics and T2 Biosystems. T.J. serves on the Scientific Advisory Board of Dragonfly Therapeutics, SQZ Biotech, and Skyhawk Therapeutics. None of these affiliations represent a conflict of interest with respect to the design or execution of this study or interpretation of data presented in this manuscript. T.J. laboratory currently also receives funding from the Johnson & Johnson Lung Cancer Initiative, but this funding did not support the research described in this manuscript. This work was supported by the Howard Hughes Medical Institute and Calico Life Sciences LLC.

implicated in the unfolded protein response. Targeted genetic and biochemical experiments using mouse and human *Keap1*-mutant tumor lines, as well as preclinical genetically-engineered mouse models (GEMMs) of LUAD, validate *Slc33a1* as a robust *Keap1*-mutant-specific dependency. Furthermore, unbiased genome-wide CRISPR screening identified additional genes related to *Slc33a1* dependency. Overall, our study provides a strong rationale for stratification of patients harboring *KEAP1*-mutant or NRF2-hyperactivated tumors as likely responders to targeted SLC33A1 inhibition and underscores the value of integrating functional genetic approaches with GEMMs to identify and validate genotype-specific therapeutic targets.

Cancer genome sequencing studies have identified Kelch-like ECH-Associated Protein 1 (*KEAP1*), a negative regulator of the antioxidant transcription factor nuclear factor, erythroid 2-like 2 (*NFE2L2;* hereafter *NRF2*), to be mutated in approximately 20-30% of all lung cancers, with loss-of-function (LOF) mutations spanning the full length of the KEAP1 protein[3,6-8]. NRF2 regulates a network of genes that together coordinate the cellular response to mitigate oxidative damage via detoxification of reactive oxygen species (ROS) and xenobiotic insults, as well as the regulation of glutathione (GSH) synthesis and recycling[2,9,10]. The high frequency of mutations in the *KEAP1/NRF2* pathway suggests a critical role for the oxidative stress pathway in lung tumorigenesis, and presents an opportunity to identify novel therapeutic strategies for this genetic subtype of LUAD.

Genetic inactivation of *Nrf2* in *Kras*-mutant genetically engineered mouse models (GEMMs) of pancreatic cancer and LUAD leads to decreased tumor initiation and tumor proliferation, suggesting that a subset of RAS-driven tumors are dependent on this transcription factor[11-13]. We recently showed that *Keap1* inactivation cooperates with *Kras* mutation in driving lung adenocarcinoma (LUAD) initiation and progression, and that these *Keap1*-mutant tumors are sensitive to glutaminase inhibitors[4]. Tumors with *KEAP1/NRF2* pathway mutations have also been shown to depend on non-essential amino acids, as well as exhibiting altered metabolism resulting from increased flux of the pentose phosphate pathway, as well as usage of glutamine, serine, and cysteine derived metabolites[14-18]. Importantly, *KEAP1* mutations are strongly correlated with decreased lung cancer mortality and shown to promote lung cancer metastasis in GEMMs of *Keap1*-mutant LUAD[19,20].

Given the large proportion of cancers with *KEAP1* pathway mutations and the inherent difficulty in therapeutically targeting LOF mutations, we investigated whether *Keap1*-mutant tumors (or otherwise NRF2-hyperactivated tumors) harbored additional genetic dependencies that could be therapeutically targeted. Towards this goal, we performed a CRISPR-Cas9-based screen of all druggable genes in *Keap1*-mutant isogenic LUAD cell lines. Cross validation with publicly available CRISPR screens and our druggable genome screen uncovered several novel dependencies, including a marked reliance on the endoplasmic reticulum (ER) resident gene, *Slc33a1*. Using multiple validation approaches, including *in vitro* and *in vivo* approaches, as well as preclinical GEMMs of *Kras*-driven LUAD, we confirmed *Slc33a1* as a *Keap1*-mutant-specific dependency. Finally, we demonstrate the power of genome-wide modifier genetic screens by identifying suppressors of *Slc33a1* dependency and highlight an underappreciated relationship between ER redox biology and *KEAP1* mutation. Altogether, these data implicate a unique vulnerability that could be therapeutically exploited in *KEAP1/NRF2*-mutant cancers with limited clinically approved targeted therapies.

Results:

Druggable genome CRISPR screen identifies *Keap1*-mutant dependencies

To uncover novel dependencies in a high-throughput manner, we performed pooled CRISPR-Cas9 screens in *Keap1*-mutant cell lines using a custom druggable-genome library (DGL) targeting 4,915 genes with high-confidence known or predicted druggability (Fig. 1a,b, Supplementary Fig. 1a,b, and Supplementary Table 1; see Methods)[21]. This library was used to transduce two isogenic pairs of Cas9-expressing murine LUAD cell lines with $Kras^{G12D/+}$; $p53^{-/-}$ (KP) and $Kras^{G12D/+}$; $p53^{-/-}$; $Keap1^{-/-}$ (KPK) genotypes[4] (Fig. 1a,b). Library representation was determined by deep sequencing of sgRNAs amplified from genomic DNA harvested early (t_0) and 8 population doublings post transduction (t_8; Fig.1b). To quantify the impact of gene targeting, we calculated gene scores (GS) as the median $\log_2$ fold-change in sgRNA abundance between t_8 and t_0. As expected, we observed significant genotype-independent depletion of sgRNAs targeting known essential genes, and enrichment of those targeting tumor suppressor genes at t_8 (Supplementary Fig. 1c and Supplementary Table 2).

To identify *Keap1*-mutant-specific vulnerabilities, we calculated differential GS by subtracting KP GS from KPK GS. Using thresholds defined by the distribution of control sgRNA “pseudogene” scores, 53 genes reached significance (FDR = 0.05) for differential depletion in isogenic pair 1 (KP1 and KPK1), and 25 in pair 2 (KP2 and KPK2; see Methods; Supplementary Fig. 1d,e). Of these genes, *Slc33a1*, *Arf4*, *Spcs2*, *Abcb7*, *Aifm1*, and *Rad51* reached significance in both isogenic pairs (Fig. 1c, Supplementary Fig. 1d,e, and Supplementary Table 2). Intriguingly, SLC33A1, ARF4, and SPCS2 are implicated in protein transport and the endoplasmic reticulum (ER)[22,23], while ER-stress has been shown to modulate levels of RAD51 and sensitivity to DNA damage[24,25]. Reinforcing the importance of these pathways, gene set enrichment analysis (GSEA) of the ranked list of differential GS identified significant enrichment of the ER stress and unfolded protein response (UPR) gene set (FDR = 0.0073; Fig. 1d and Supplementary Table 3).

Reliance *on Slc33a1* in *Keap1*-mutant cells is NRF2-dependent

To validate the most differentially depleted genes, *Slc33a1* and *Arf4,* as *Keap1* mutant-specific genetic vulnerabilities, we examined the top two scoring sgRNAs targeting these genes in fluorescence competition and colony formation assays (Supplementary Fig. 2a,b). We observed clone-specific depletion of sg*Arf4*-targeted KPK1 but not sg*Arf4*-targeted KPK2 cells, and thus deprioritized *Arf4* for further evaluation (Supplementary Fig. 2c). On the other hand, sg*Slc33a1*-targeted KPK1 and KPK2 cells exhibited a robust decrease in growth; while modest growth defects were observed in sg*Slc33a1*-targeted KP cells, the magnitude and kinetics of the effect was significantly less pronounced than in KPK cells (Fig. 1e,f and Supplementary Fig. 2d). We observed similar decreases in colony formation using inducible shRNAs or constitutive CRISPR-interference (CRISPRi; Supplementary Fig. 2e-k). Transduction with an sgRNA-resistant cDNA of *Slc33a1* fully rescued gene expression as well as the growth defect in KPK cells targeted with *Slc33a1* sgRNAs, demonstrating on-target dependency (Fig. 1g and Supplementary Fig. 3a). Collectively, these data validate the CRISPR screening results and further support *Slc33a1* as a potent target for inhibiting the proliferation of *Keap1*-mutant cancer cells.

To examine whether the *Slc33a1* dependency in *Keap1*-mutant cells is explained by increased NRF2 activity versus other *Keap1* clients[8,26], we targeted *Nrf2* in KP and KPK cell lines and examined the effect of targeting *Slc33a1*. Inactivation of *Nrf2* substantially rescued the growth defect caused by *Slc33a1* mutation in KPK cells (Fig. 1h and Supplementary Fig. 3b-f). These data indicate that NRF2-hyperactive cells are dependent on SLC33A1, and that decreased fitness following *Slc33a1* deletion is a consequence of the cellular programs controlled by the NRF2 transcription factor.

KEAP1-mutant human cancer cells are dependent on the ER-resident genes *SLC33A1*, *SUCO*, and *TAPT1*

To assess the effects of *SLC33A1* inactivation in human cell line studies, we analyzed whole-genome CRISPR-Cas9 screening results from the publicly available DepMap dataset encompassing 563 human cancer cell lines spanning 29 cell lineages[27] (see Methods). We first confirmed that *SLC33A1*, on average, does not behave as a core-essential gene across the large majority of human cell lines in the DepMap (Supplementary Fig. 4a). However, average *SLC33A1* CERES normalized GS (hereafter CERES scores) were ~3.5-fold lower in *KEAP1/NRF2*-mutant cancer cell lines (n = 63) compared to wild-type cancer cell lines (n = 500; Supplementary Fig. 4b and Supplementary Table 4).

To extend the analysis beyond the genes targeted by the DGL used in our studies, we explored previously published Cancer Coessentiality Networks, a meta-analysis of depletion screens across different cell types[28]. We identified *NRF2*, *SLC33A1*, *SUCO*, *TAPT1*, *DNAJB11, and ADPGK* in a single co-essential gene cluster using the DepMap dataset as input (Supplementary Fig. 4c). Importantly, analogous to the top hits identified in the DGL, this co-essential cluster is composed entirely of known endomembrane-associated genes. Similarly, given that *KEAP1* mutation results in increased NRF2 pathway activity[2], as well as data showing *NRF2* dependency in *KEAP1*-mutant human cell lines[29,30], we calculated the Pearson correlation coefficient between *NRF2* CERES scores in the DepMap compared to each other gene’s CERES scores as an alternative approach to identify genes that correlate most with *NRF2* dependence on a relative basis. Using this approach, we identified *SLC33A1* scores as having the second highest correlation to *NRF2* scores (r=0.361), followed by *MAFK* (r = 0.287), a known NRF2 transcriptional co-activator[31]. Moreover, the ER-resident proteins *SUCO* and *TAPT1*, previously established as the mammalian orthologs of the SLP1:EMP65 yeast complex[32], were also closely correlated with *NRF2* dependency in this analysis (Fig. 1i, Supplementary Fig. 4a,b, and Supplementary Table 4; r = 0.283 and r = 0.258, respectively). Importantly, *SLC33A1* dependency weakly correlated with activating transcription factor 4 (*ATF4*) dependency (r = −0.001), a downstream target of the UPR response known to exhibit cross-talk with the NRF2 pathway[15] (Fig. 1i, Supplementary Fig. 4d,e, and Supplementary Table 4).

To further explore these associations, we first probed the effects of *SLC33A1* mutation in a panel of human *KEAP1*-mutant cells by fluorescence competition assays (Fig. 1j). These data confirmed that *SLC33A1* represents a *KEAP1*-mutant vulnerability across genetically defined human cancer cell lines. Next, to address whether other dependencies identified in the human *NRF2*-*SLC33A1* co-essentiality network also validated, we targeted *Suco* or *Tapt1* in the murine KP and KPK isogenic cell line pairs. Supporting the bioinformatic predictions, *Suco* and *Tapt1* inactivation led to decreased fitness only in KPK cells (Supplementary Fig. 4f,g). These data support an increased dependence on endomembrane-associated protein function in the context of NRF2-hyperactivated cancer cells, corroborating the significant depletion of UPR target genes in the DGL by GSEA in KPK cell lines compared to KP cell lines (Fig. 1c,d and Supplementary Fig. 4c).

High NRF2-transcriptional signatures correlate with increased *SLC33A1* dependency

Given that tumor cells may upregulate the NRF2 response in the absence of *KEAP1/NRF2* mutation or *NRF2* amplification, we questioned whether the degree of NRF2 transcriptional activity could predict *SLC33A1* dependency. To that end, we used a previously established NRF2 core target gene set (hereafter NRF2 core) or *KEAP1*-mutant transcriptional signature as indicators of NRF2 activation state[4]. These signatures were used to stratify cell lines screened in the DepMap into high-scoring or low-scoring cohorts (NRF2 core high-scoring $n = 42$; low-scoring $n = 32$; *KEAP1*-mutant signature high-scoring $n = 43$; low-scoring $n = 21$; Supplementary Table 5; see Methods). By comparing the CERES scores between these high- and low-scoring groups, we identified 123 (NRF2 core, $p<0.05$) or 34 (*KEAP1*-mutant signature, $p<0.05$) statistically significant depleted genes in the high-scoring group, of which *NRF2*, *SLC33A1*, *SUCO*, and *TAPT1*, fell within the top 15 genes regardless of the signature used (Fig. 1k, Supplementary Fig. 5a, and Supplementary Table 5). Interestingly, although we could not validate a genotype-specific dependency for *ARF4* in murine KPK cell lines, we observed significant depletion of *ARF4* in the high-scoring groups of either signature. Moreover, despite the enrichment of *NRF2*- or *KEAP1*-mutant cell lines in either the NRF2 core or *KEAP1*-mutant high-scoring groups ($n = 27$, $p = 2.47e\text{-}18$ or $n = 18$, $p = 5.95e\text{-}8$, respectively; hypergeometric test), a substantial fraction did not harbor either *NRF2* or *KEAP1* mutations ($n = 15$, 35% and $n = 25$, 58%, respectively), demonstrating the importance of functionally classifying tumors by NRF2-transcriptional activity rather than *KEAP1* or *NRF2* mutational status. We observed a step-wise decrease in GS for *SLC33A1*, *SUCO*, and *TAPT1*, but not *RPA3* or *OR12D2* when grouping cell lines by *KEAP1*/*NRF2*-wild-type status ($n = 424$), mutant status ($n = 58$), or NRF2 core high-ranking cell lines that are genetically *KEAP1/NRF2*-wild-type ($n = 31$). However, we did not observe significant differences in CERES scores between wild-type and NRF2 core low-ranking cell lines that are genetically *KEAP1/NRF2*-wild-type ($n = 45$), as expected. Importantly. *KEAP1/NRF2* mutation displayed the largest fold change compared to wild-type cells (fold change *SLC33A1* = 2.49, *SUCO* = 1.41, *TAPT1* = 1.59), followed by NRF2 core high-ranking cells (fold change *SLC33A1* = 2.41, *SUCO* = 1.5, *TAPT1* = 1.4), further demonstrating that dependencies towards *SLC33A1*, *SUCO*, and *TAPT1* can be stratified by the magnitude of NRF2 transcriptional output independently of *KEAP1/NRF2* mutation status (Supplementary Fig. 5b,c).

Loss of *Slc33a1* results in induction of an unfolded protein response

SLC33A1 is an ER-localized transmembrane protein with reported acetyl-Coenzyme A (CoA) transport activity[5] (Fig. 2a). We observed that targeting of *Slc33a1* in KPK cells led to a striking increase in intracellular blebbing, consistent with previous reports of *Slc33a1* loss[5,33] (Fig. 2b). Given the reported ER localization of SLC33A1 and its significant KPK-specific depletion in our screens, we performed whole-transcriptome analysis (RNA-seq) on isogenic pair #1 following sg*Ctrl* (non-essential gene, *Olfr102*) or sg*Slc33a1* targeting to define a transcriptional signature of *Slc33a1* deficiency. We found that differentially-expressed genes resulting from *Slc33a1* loss were enriched in the UPR pathway, further supporting a role for *Slc33a1* in maintaining ER-homeostasis (NES=2.48; FDR=0; Fig. 2c and Supplementary Table 6).

Using previously defined, conserved UPR transcriptional signatures[34], we observed significant positive enrichment of gene sets associated with ER-chaperones (NES=2.13; FDR=0), aminoacyl tRNA synthetase (NES=1.91; FDR=0), ER/Golgi trafficking (NES=1.68; FDR=0.0146), and endoplasmic reticulum-associated protein degradation (ERAD; NES=1.61; FDR=0.025) in sg*Slc33a1*-targeted cells (Supplementary Fig. 6a-c and Supplementary Table 7). *Slc33a1* targeting results in increased p62 and LC3-II expression, PERK mobility shift, and an increased ratio of phospho-to-total eIF2A upon *Slc33a1*-targeting in KPK cells, with no change in GCN2 levels, all consistent with induction of autophagy and a UPR, respectively (Fig. 2e,f). Moreover, increased expression of UPR target genes was evident in KP- and KPK-sg*Slc33a1* cells by both protein (BiP/HSPA5, ATF4; Fig. 2e,f) and targeted gene expression analysis (*BiP/Hspa5*, *Chop/Ddit3*, *Atf4*, and *Xbp1*s; Fig. 2f). Notably, the level of induction of *Atf4*, *BiP* and *Chop/Ddit3* was greater in KPK-sg*Slc33a1* cells compared to KP-sg*Slc33a1* cells, suggesting that *Keap1* loss results in enhanced induction of a UPR program in the face of this form of ER-stress (Fig. 2e-g).

Slc33a1 dependency is rescued by pharmacological inhibition of GSH synthesis

Our previous studies showed that KPK cells potently upregulate the NRF2 pathway, resulting in increased glutathione (GSH) synthesis[4]. This is likely due to the upregulation of NRF2-dependent enzymes directly involved in GSH synthesis (glutamate cysteine ligase catalytic subunit; *Gclc*)[35] or uptake of L-Cystine that is necessary for GSH synthesis (system xCT encoded by *Slc7a11* and *Slc3a2*; Fig. 3a). Previous studies in yeast have demonstrated that enhanced extracellular GSH uptake results in growth inhibition and cell death proportional to the amount of GSH added to the media[36]. We hypothesized that increased NRF2-dependent GSH synthesis in combination with genetic perturbation of endomembrane-associated genes might induce an UPR, which may explain the increased dependency of *Keap1*-mutant cells on *Slc33a1*, *Suco*, and *Tapt1*. To test this hypothesis, we probed the GSH synthesis and recycling pathway at multiple steps using both genetic and small molecule approaches. Reduction of GSH synthesis via inhibition of GCLC (L-buthionine sulfoximine; BSO)[37] reduced total GSH within 48 hours of treatment (Supplementary Fig. 7a). Remarkably, treatment with BSO completely rescued the fitness defect of sg*Slc33a1*-, sg*Tapt1*-, and sg*Suco*-targeted KPK cells (Fig 3b, Supplementary Fig. 7b). Similarly, BSO treatment rescued the effects of shRNA-mediated *Slc33a1* knockdown (Supplementary Fig. 7c) and blunted the enrichment of *sgGclc* cells over time, further supporting its on-target activity in rescuing the viability defects caused by *Slc33a1* loss (Fig. 3c). Importantly, BSO treatment also rescued the fitness defects observed in a panel of *SLC33A1*-deficient *KEAP1*-mutant human LUAD cell lines (Fig. 3d). To determine whether additional NRF2-dependent metabolic changes could also modulate *Slc33a1* dependence, we treated cells with the xCT inhibitor Erastin[38], N-acetyl cysteine (NAC), or dimethyl 2-oxoglutarate (DMG; cell permeable alpha-ketoglutarate). Similar to BSO, Erastin fully rescued the fitness defect in sg*Slc33a1*-targeted cells (Fig. 3e and Supplementary Fig. 7b) and, importantly, neither agent was able to rescue sg*Rpa3*-targeted cells (Supplementary Fig. 4d). NAC and DMG were unable to rescue the fitness of *Slc33a1*-deficient KPK cells (Fig. 3e). These results suggest that *Slc33a1* dependency is not due to increased ROS, decreased cysteine availability, or defects in central carbon metabolism[14], respectively. However, these results do not rule out the possibility that organelle specific changes occur post *Slc33a1* loss. Perturbation of the GSH:GSSG ratio to favor GSSG accumulation via menadione-induced NADPH oxidation[39] fails to rescue *Slc33a1*-deficient KPK cells (Supplementary Fig. 7e). Moreover, treatment with the thioredoxin reductase-1 (Txnrd1) inhibitor, Auranofin[40], failed to rescue the effects of *Slc33a1* loss, suggesting that total pools or *de novo* synthesis of GSH, but not thioredoxin, mediate *Slc33a1* dependence (Fig. 3e). Collectively, these data demonstrate that inhibition of GSH synthesis is sufficient for rescuing the effects of *Slc33a1* loss.

To test whether induction of the UPR is prevented by GSH depletion, we treated KP- or KPK-sg*Slc33a1* and -sg*Ctrl* cells with BSO and quantified relative expression of a panel of UPR target genes. Consistent with recent observations, short-term treatment with BSO had no impact on UPR gene induction in sg*Ctrl* cells[41] (Fig. 3f). However, we observed increases in NRF2 pathway targets (*Gclc*, *Nqo1*) upon BSO treatment in KP and KPK cells, consistent with decreases in total GSH pools, which in turn lead to an oxidized environment that activates NRF2 signaling (Fig. 3f). Notably, BSO treatment rescued UPR gene induction in sg*Slc33a1* cells, suggesting that high GSH synthesis is a critical mediator in UPR induction following *Slc33a1* loss (Fig. 3f). To determine whether *Keap1* loss renders cells more sensitive to ER stress and UPR induction, we treated KP and KPK cells with tunicamycin and thapsigargin, commonly used small-molecule inducers of the UPR. On average, KPK cells displayed slightly increased sensitivity to treatment (Supplementary Fig. 8a-c). KP1 cells displayed similar sensitivity to tunicamycin compared to their isogenic KPK1 counterpart. We speculate that this is due to the increase in overall NRF2 pathway activation, as we previously detected increased GCLC protein in this line compared to other KP cell lines[4]. Importantly, co-treatment with BSO led to significantly increased thapsigargin resistance as measured by the area under the curve (AUC), with a larger magnitude of rescue in KPK cells (Supplementary Fig. 8d-f). Together, these data suggest that increased GSH synthesis promoted by *Keap1* deficiency leads to heightened sensitivity to genetic perturbations of ER-localized proteins (e.g. loss of *Slc33a1, Suco,* or *Tapt1*) or chemical perturbations (e.g. thapsigargin treatment).

Loss of *Slc33a1* is associated with widespread metabolic changes

Slc33a1 has been reported to directly transport acetyl-CoA into the ER lumen[42]. As such, mutation of *Slc33a1* could disrupt cellular metabolic processes that depend on acetyl-CoA transport. To characterize differences in polar metabolite levels and usage following *Slc33a1* loss, we performed liquid chromatography-mass spectrometry (LC-MS) in KP- and KPK-sg*Ctrl* or -sg*Slc33a1* cells maintained in BSO for 48 hours or 48 hours post BSO withdrawal (a time point at which there is no significant cell death; Supplementary Table 8). The steady state levels of 45/167 detected metabolites were significantly different across all samples (Supplementary Table 8). Unsupervised hierarchical clustering revealed three major clusters, with all technical replicates of vehicle-treated KPK-sg*Slc33a1* cells clustering separately from the rest (Fig. 4a). Consistent with earlier reports, we detected reduced TCA cycle intermediates in KPK-sg*Ctrl* compared to KP-sg*Ctrl* cells, with increased intracellular GSH, as well as concomitant decreases in intracellular α-ketoglutarate[4,14] (Supplementary Fig. 9a,b). Notably, TCA-derived metabolites were significantly increased in KPK cells targeted with sg*Slc33a1*, consistent with *Slc33a1* loss either boosting TCA cycle utilization or causing increases in TCA cycle components via autophagy-mediated breakdown (Supplementary Fig. 9a). Interestingly, in addition to increased total levels of GSH (Fig. 4b), we observed increased oxidized GSH (GSSG) levels, resulting in decreased GSH:GSSG ratios in KPK-sg*Slc33a1* cells compared to KPK-sg*Ctrl* cells (Supplementary Fig. 9b).

To assess whether the observed changes in intracellular metabolites were a result of changes in influx or secretion, we used gas-chromatography-mass spectrometry (GC-MS) to assess changes in metabolite consumption or secretion in tissue culture media. GC-MS-based detection of metabolites confirmed an increase in glutamine and cystine consumption and an increase in glutamate secretion in KPK-sg*Ctrl* versus KP-sg*Ctrl* cells, likely associated with increased demand for GSH synthesis resulting from *Keap1* deficiency (Supplementary Fig. 9c and Supplementary Table 9). Of note, the increase in glutamine and cystine consumption, as well as glutamate secretion, was further increased upon *Slc33a1* loss (Supplementary Fig. 9c).

Given that *Slc33a1* is reported to be involved in cytoplasmic-to-ER-acetyl CoA transport[5] and that steady state acetyl-CoA and CoA pools were lower in KPK-sg*Ctrl* cells relative to KP-sg*Ctrl* cells (Fig. 4c and Supplementary Fig. 9a), it is possible that KPK cells were more sensitive to *Slc33a1* loss due to limitations in the transport of acetyl-CoA (and other undefined metabolites) into the ER. Moreover, we detected further decreases of intracellular CoA and transsulfuration intermediates in KPK-sg*Slc33a1* cells relative to KPK-sg*Ctrl* cells (Supplementary Table 8). A potential explanation for this observation is that GSH or GSH synthesis may directly impact the availability of intracellular acetyl-CoA. Supporting this, BSO-treated KPK-sg*Slc33a1* cells clustered most closely with KP-sg*Ctrl* and KPK-sg*Ctrl* cells, and displayed an increased relative abundance of acetyl-CoA and CoA (Fig. 4c and Supplementary Fig. 9d,e). Altogether, these data demonstrate that *Slc33a1* loss results in dramatic changes in the metabolic profiles of murine *Kras*-mutant LUAD cell lines, particularly in combination with *Keap1* loss. However, whether acetyl-CoA is the relevant metabolite that explains *Slc33a1* dependency remains unknown.

Transplant models and GEMMs of *Keap1*-mutant LUAD are dependent on *Slc33a1*

To address the impact of *Slc33a1* loss *in vivo*, we first transplanted KP1 or KPK1 cell lines expressing sg*Ctrl* or sg*Slc33a1* into the flanks of immunocompromised mice and measured longitudinal growth, final subcutaneous tumor mass, and *in vivo* competitive fitness and confirmed a dependency for *Slc33a1* in transplanted KPK tumors (Fig. 5a-c). Moreover, inducible shRNA-mediated knockdown (Supplementary Fig. 2e-h) of *Slc33a1* post tumor initiation led to a significant decrease in tumor volume in KPK but not KP transplants compared to sh*Renilla* controls ($p = 0.0006$; Fig. 5d,e). Additionally, sg*SLC33A1*-targeted A549 xenografts (*KEAP1*-mutant) grew more slowly (A549 $p<0.001$) and resulted in smaller tumors (A549: fold change = 2.711, $p = 0.019$) compared to sg*CTRL* targeted tumors (Fig. 5f,g). Importantly, we do not observe decreases in tumor size in *KEAP1*-wild-type sg*SLC33A1*-targeted H2009 or H441 xenografts (Fig. 5f,g). These data confirmed a targetable dependency of *Keap1*-mutant cells on *Slc33a1* and prompted us to perform further validation in a genetically engineered mouse model (GEMM) of LUAD[43] that better recapitulates the differential nutrient availability between tumors and surrounding normal tissues, as well as the native lung microenvironment[44,45]. We utilized the $Kras^{LSL\text{-}G12D/+}$ (*LSL=loxP-STOP-loxP*); $p53^{fl/fl}$ GEMM of LUAD (hereafter KP)[46,47], in which tumors are induced upon intratracheal instillation of lentivirus expressing Cre recombinase[43,48]. KP mice were crossed to germline conditional alleles of *Keap1* (ref. 38) and *Cas9* ($R26^{LSL\text{-}Cas9}$; ref. 39) for NRF2 pathway activation and somatic editing of genes within initiated lung tumors, respectively, generating $Kras^{LSL\text{-}G12D/+}$; $p53^{fl/fl}$; $Keap1^{fl/fl}$; $R26^{LSL\text{-}Cas9/+}$ mice (hereafter KPC or KPKC). To assess the effects of *Slc33a1* targeting in *Keap1*-wild-type and *Keap1*-null tumors, we induced tumors in KPC or KPKC mice with pUSEC lentivirus[49] expressing Cre recombinase and sgRNAs targeting *Slc33a1*. To boost the efficiency of bi-allelic loss of function events[50], we generated pUSEC lentiviruses expressing two potent sgRNAs against *Slc33a1* or *Olfr102*, hereafter referred to as sg*Slc33a1* or sg*Ctrl*, respectively (Fig. 6a and Supplementary Fig. 10a).

Histological analysis revealed an increase in high grade adenocarcinomas, tumor burden, tumor number, and proportion of proliferating cells in KPKC-sg*Ctrl* mice compared to KPC-sg*Ctrl* mice ($p<0.05$; Fig. 6b-e), consistent with previous studies[4]. We did not detect significant differences in grade, tumor burden, or tumor number in KPC-sg*Slc33a1* compared to KPC-sg*Ctrl* mice (Fig. 6b-d). However, there was a significant reduction in high-grade adenocarcinomas (G2 fold change = 6.6, $p = 8.18e\text{-}9$; G3 fold change = 4.7, $p = 2.81e\text{-}6$; G4 fold change = 23, $p = 0.025$, tumor burden (fold change = 4.36, $p = 5.95e\text{-}7$), and number in KPKC-*sgSlc33a1* versus -sg*Ctrl* mice (Fig. 5b-d). We were unable to detect significant differences in the mitotic index of KPKC-sg*Slc33a1* tumors compared to KPKC-sg*Ctrl* tumors, as assessed by phosphorylated histone H3 immunohistochemistry (IHC; Fig. 6e and Supplementary Fig. 10b). These results suggest that the reduced tumor burden of KPKC-sg*Slc33a1* mice is likely due to decreased tumor initiation and/or a heightened barrier for transition to higher grade LUAD (Fig. 6e). Importantly, early lesions may have been enriched for incomplete *Keap1* recombination (resulting in decreased NRF2 activation) or incomplete deletion of *Slc33a1*.

Supporting the latter possibility, deep sequencing of the *Slc33a1* locus in micro-dissected tumors from sg*Slc33a1*-targeted mice revealed a significantly greater ratio of wild-type sequence reads and non-frameshift mutations (which may preserve gene function) in KPKC tumors compared to KPC tumors, indicating strong selection against combined deletion of *Keap1* and *Slc33a1* (Fig. 6f,g and Supplementary Fig. 10c-e). Furthermore, while *Keap1*-recombined tumors displayed enhanced expression of the NRF2 target gene, NQO1 (refs. 4, 42; Fig. 6h), we observed a nonsignificant trend towards decreased NQO1 positive tumors in KPKC-sg*Slc33a1* animals (Supplementary Fig. 10f). These results suggest a potential mechanism of tumor escape via failure to completely recombine *Keap1,* or alternative mechanisms to dampen the NRF2 program.

Collectively, these data validate *Slc33a1* as a potential target in *Keap1*-mutant LUAD *in vivo*. Furthermore, *Keap1*-wild-type tumors that have upregulated the NRF2 pathway via other genetic or epigenetic mechanisms may also show sensitivity to *Slc33a1* loss or inhibition.

Whole genome CRISPR screen identifies suppressors of *Slc33a1* dependency

To gain further insights into the underlying biology of *Slc33a1*, *Suco*, and *Tapt1* loss in a *Keap1*-mutant context, as well as identify patient genotypes that may not respond to SLC33A1 inhibition, we performed an unbiased whole genome CRISPR screen using a previously established library[51] (hereafter Brie; Fig. 7a). To that end, we leveraged BSO-mediated rescue to isolate KPK-*Slc33a1*, KPK-*Suco*, and KPK-*Tapt1* knockout clones (Fig. 7a and Supplementary Fig. 11a,b). Four KPK-*Slc33a1* single cell clones were pooled together for subsequent screening. We transduced the Brie library into KPK-*Slc33a1* cells and separated into BSO (3 replicates) or vehicle (6 replicates) treated conditions. Library representation was determined by deep sequencing of sgRNAs amplified from BSO or vehicle treated cells at early (t_0) and 8 population doublings post transduction (t_8; Fig.7b, Supplementary Fig. 11c and Supplementary Table 10). We assessed median GS within the vehicle treated groups (t_8-t_0) and identified *Nrf2*, *Gclc*, and *Gclm* as the top enriched genes, thus validating our previous data, now in a fully unbiased manner (Fig. 7c). Given that *Slc33a1* loss results in a potent growth-disadvantage in KPK cells, we hypothesized that any gene at t_8 with a positive sgRNA count across all replicate infections, and that was represented by ≥ 2 individual sgRNAs, would represent true biological hits capable of rescuing KPK-*Slc33a1* deficiency. Using these thresholds (see methods), we identified a number of additional enriched genes, including ER-resident oxidoreductase 1A (*Ero1l*), protein disulfide isomerase 1a (*Pdia*/*P4hb*), and the NRF2 target genes *Osgin1*, *Amotl2*, and *Pgd*[52] (Fig. 7c). GSs of the top enriched genes were decreased in BSO-treated samples (Supplementary Fig. 12a,b).

To validate the top enriched candidate genes and to further investigate underlying biology that might link the observed dependency associated with targeting *Slc33a1*, *Suco*, and *Tapt1*, we infected KPK-*Slc33a1*, KPK-*Suco*, and KPK-*Tapt1* knockout cells with lentiviruses encoding two independent top scoring sgRNAs from this screen, and assessed growth and phenotypic consequences post-BSO withdrawal. Remarkably, loss of the majority of the candidate genes rescued the growth of *Slc33a1* cells, but not *Suco* or *Tapt1* cells, as measured by cell numbers and colony formation assay (Fig. 7e and Supplementary Fig. 12c). Importantly, although *Nf2*, *Cul3*, *Cul5*, *Rnf7*, *Amotl2*, and *Eif2ak3* (Perk) sgRNAs rescued the growth of KPK-*Slc33a1* cells, *Ero1l*, *P4hb*, *Osgin1*, *Serp1*, and *Yipf5* sgRNAs rescued both the growth and blebbing phenotype associated with *Slc33a1* loss (Supplementary Fig. 13a). Moreover, *Ero1l*, *P4hb*, *Serp1*, and *Yipf5* are all known ER-resident genes, supporting the original observation that KPK cells have a dysfunctional ER. Additionally, loss of *Osgin1* rescued the growth and phenotype of *Slc33a1*-deficient cells, suggesting that other genes downstream of NRF2 might contribute to *Slc33a1* dependency (Fig. 7e, Supplementary Fig. 12c and Supplementary Fig. 13a). Interestingly, *Suco* and *Tapt1* knockout cells were only rescued by sgRNAs targeting *Nf2*, *Cul5*, *Rnf7*, *Osgin1*, and *Pgd*, suggesting that *Suco* and *Tapt1* dependency is biologically distinct from *Slc33a1* dependency, but overlap in redox biology via rescue through GCLC inhibition (Fig. 3b, Fig. 7e and Supplementary Fig. 12c). In summary, these data suggest that *Slc33a1* loss results in a dysfunctional ER state that significantly affects the growth of *Keap1*-mutant cells. ERO1L is positively regulated by GSH to generate oxidation events used by P4HB for proper protein-folding[53,54]. Therefore, loss of either gene is predicted to result in decreased ER-oxidation or decreased GSSG production and raises the possibility that SLC33A1 may be involved in regulating ER redox potentials.

Discussion:

We have demonstrated that *Keap1* mutation results in an increased NRF2-dependent vulnerability to *Slc33a1*, *Suco*, and *Tapt1* loss. Although several studies have identified a number of putative *KEAP1/NRF2*-mutant-specific synthetic lethal interactions that could be exploited for therapeutic purposes[4,14,16,55,56], all studies so far have been limited in scale (e.g. focused on interrogating canonical NRF2 target genes as candidate synthetic lethal factors) and do not rule out the possibility that more comprehensive and systematic queries using increasingly complex functional genomic approaches will uncover novel factors that could be targeted for therapeutic benefit. Importantly, although we focused this study on *Slc33a1*, we hypothesize that *Keap1*-mutant cells are universally sensitive to dysfunction of an ER-stress pathway of which *Slc33a1* loss imposes the largest selective disadvantage. In line with this observation, we demonstrate increased dependency for the endomembrane associated genes *SUCO* and *TAPT1*[32] in *KEAP1*-mutant human cancer cell lines observed in the DepMap, which we validated in our murine *Keap1*-mutant cell lines. While *SLC33A1* mutations are associated with developmental defects and childhood mortality, suggestive of a critical role in development[42], pooled genetic screens in embryonic and cancer cell lines suggest that *SLC33A1* is not a pan-essential gene[57,58]. Consistent with a non-essential role, we demonstrate that the *Slc33a1* dependence of *Keap1*-mutant cancer cells can be rescued by loss of a number of genes with roles in the integrated stress response network. Specifically, loss of the ER-resident genes *Erol1*, *P4hb*, *Serp1*, and *Yipf5* or the NRF2-target genes *Gclc*, *Gclm*, and *Osgin1* rescue both the growth disadvantage and cellular morphology associated with *Slc33a1* loss in NRF2-hyperactivated cells. Although the direct link between GSH and *SLC33A1*, *SUCO*, and *TAPT1* remains unclear, we hypothesize that the ER of *Keap1*-deficient cells have been shifted into a dysfunctional state that likely increases proteotoxic stress due to challenges in protein folding that can be rescued by decreasing ER-specific oxidation or ER-specific GSSG production. Consistent with this idea, we observe increased intracellular GSSG upon *Slc33a1* loss in *Keap1*-mutant cells. However, ER-specific accumulation of GSSG upon SLC33A1 inhibition remains to be experimentally validated, and it is possible that other mechanisms underlie the dependence of *KEAP1*-deficient cells on SLC33A1 and other ER-resident proteins.

Our results are particularly interesting in light of a recent meta-analysis in human cancer cell lines identifying a presumed connection between *SLC33A1* dependency, the NRF2 pathway, and high GSH[30]. Here, we demonstrate that *Slc33a1* dependency is rescued via inhibition of GSH synthesis and also behaves as a potent *Keap1*-mutant specific genetic dependency in a GEMM of *Kras*-mutant LUAD, validating the cell-autonomous requirement for *Slc33a1* function in *Keap1*-mutant LUAD *in vivo*. Overall, our study provides a strong rationale for substratification of patients harboring *KEAP1*-mutant or NRF2-hyperactivated tumors as likely responders to targeted *SLC33A1* inhibition and underscore the value of integrating functional genetic approaches with GEMMs to identify and validate genotype-specific therapeutic targets.

Materials and 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 $Trp53^{fl/fl}$ mice have already been described[46,47]. $Keap1^{fl/fl}$ mice were generated and shared by S. Biswal[59]. *Rosa26-LSL-Cas9* mice were obtained from Jackson labs[60]. 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. Animals lacking detectable tumors by histopathology were excluded from the analysis to ensure all animals were properly infected with pUSEC lentiviruses. Animals with the appropriate genotypes between the ages of 6-10 weeks were randomly selected to begin tumor initiation studies with pUSEC-sg*Olfr102* (sg*Ctrl*) or pUSEC-sg*Slc33a1*. Mice were infected intratracheally with 20,000 TU's of lentiviruses as described[43]. Total lung area occupied by tumor was measured on hematoxylin and eosin (H&E) stained slides using NIS-elements software. Histological quantification of mouse lung tumor burden by grade was performed by an automated deep neural network (unpublished) developed by Aiforia Technologies in collaboration with the Jacks lab, and in consultation with veterinarian pathologist Dr. Roderick Bronson. The algorithm performed consistently and with high correlation with human graders across multiple validation datasets independent of the training dataset. For murine cell sgRNA transplant experiments, a total of 1E6 pUSEPR-infected and puromycin selected cells were injected into the flanks of immunocompromised mice in a total volume of 100 uL of 50:50 PBS:Matrigel mix. For murine cell doxycycline inducible shRNA transplant experiments, a total of 500E5 LVt-TSTOP infected and puromycin selected cells were injected into the flanks of immunocompromised mice in a total volume of 100 uL PBS. Mice were placed on doxycycline chow when tumors reached 10-25mm^3. For xenograft transplant experiments (A549, H2009, H441), a total of 1E6 pUSEPR-infected and puromycin selected cells were injected into the flanks of immunocompromised mice in a total volume of 100 uL of 50:50 PBS:Matrigel mix. Subcutaneous tumor volumes were calculated according to the following formula: $mm^3 = (a^2 * b) * (\pi/6)$, where a is the smaller dimension and b is the larger dimension. 8-week female J/nu, athymic nude mice (Jackson 007850) were used for all transplant experiments.

Cell culture

HEK293T cells were maintained in DMEM (Corning #10-013-CV) supplemented with 10% fetal bovine serum, 2 mM L-glutamine (Gibco #25030) and 50 ug/mL gentamycin (Gibco #15710). Parental cell lines from KP mice[61] were previously established and described. Human cell lines were acquired from ATCC. All lines tested negative for mycoplasma and routinely tested every 3-4 months. Murine KP derivative cells were maintained in DMEM or RPMI supplemented with 10% Fetal Bovine Serum and 50 ug/mL gentamicin. Murine cell lines transduced with $Puro^R$-containing constructs were kept under 6ug/mL Puromycin selection. Cell lines transduced with lenti-Cas9-Blast (Addgene #52962) remained under Blasticidin selection (10ug/mL). Cells were treated with inhibitors: L-Buthionine-sulfoximine (BSO, Sigma Aldrich), Auranofin (AUR, TOCRIS bioscience), Erastin (ERA, Sigma Aldrich), *N*-acetyl-L-cysteine (NAC, Sigma Aldrich), and 2mM dimethyl-2-oxoglutarate (DMG, Sigma Aldrich) at concentrations denoted in figure legends. Viability in the presence of all compounds was assessed by Cell Titer Glo (Promega #G7570). Plates were quantified using a Tecan infinite M200 Pro plate reader or a SpectraMax M5 Microplate reader (Molecular Devices). For clonogenic and low-density assays, cells were stained with Crystal Violet solution (25% Methanol 75% H_2O). Crystal violet stains were analyzed by ImageJ software.

Design and cloning of the druggable genome sgRNA library

As previously described[11], the 4,915 genes included in the sgRNA library were selected from the Drug-Gene Interaction Database[62, 63]. Human gene names were first updated using the HUGO Gene Nomenclature Committee (HGNC) multi-symbol checker tool, and then converted to the corresponding mouse orthologs using the HGNC Comparison of Orthology Predictions tool (accessed 7/21/2016)[64,65]. sgRNA sequences (four per gene, 19,660 total) were obtained from the Brie whole-genome sgRNA library designed by the Broad Institute sgRNA Designer tool[51]. 500 non-targeting control sgRNA sequences from the Asiago whole-genome sgRNA library[51] were included, bringing the total library size to 20,160 sgRNAs. The sgRNA library was cloned into lentiGuide-Puro (Addgene #52963; lentiGuide-Puro was a gift from Feng Zhang), using a previously established protocol[66]. In brief, 5’ and 3’ flanking adapter sequences, corresponding to the U6 promoter and tracrRNA sequences, respectively, were appended to the sgRNA sequences. The oligo library was synthesized by Twist Bioscience, PCR amplified with Oligo-Fwd and Oligo-knockout-Rev primers using PfuUltra II Fusion HS DNA Polymerase (Agilent Technologies #600670), purified using the QIAquick PCR Purification Kit (QIAGEN #28104), then assembled into the BsmBI-digested lentiGuide-Puro vector using the Gibson Assembly Master Mix (New England Biolabs #E2611S). A total of 800 ng of assembled plasmid DNA was electroporated into Endura electrocompetent cells (Lucigen #60242), the plated onto LB-ampicillin plates overnight at 37C. The total number of colonies was quantified to ensure a representation of >100X (>2 million colonies for a 20,000 sgRNA library). Plasmids were isolated and purified using the QIAGEN Plasmid Plus Maxi Kit (QIAGEN #12963). To determine sgRNA distribution in the cloned library, the sgRNA target region was amplified with NGS-Lib-Fwd (primers 1-10) and NGS Lib-KO-Rev (primer 1) primers, size selected in a 2% agarose gel, purified using the QIAquick Gel Extraction Kit (QIAGEN #28704), then submitted for sequencing on an Illumina HiSeq 2500 system (100-nt single end reads) at the Whitehead Institute Genome Technology Core.

Infection of cells with the druggable genome library

All transductions of KP and KPK lines were performed by spin-infection. Cells were plated at a concentration of 2 million cells/well in 12-well plates, together with the appropriate amount of viral supernatant in a total volume of 2 mL per well, then centrifuged at 1,500 RPM for 2 hours at room temperature. Cells were then incubated 37C overnight.

To ensure that most cells harbor single integration events, a multiplicity of infection (MOI) of 0.3 was used for transduction of cell lines with the sgRNA library[66]. To determine the viral titer for each cell line, all cells were transduced with decreasing volumes of viral supernatant, selected with 6 ug/mL puromycin (Gibco #A11138) for 24 hours, then quantified by live cell numbers and normalized to a control, unselected sample.

For the actual screen, a minimum of 20 million cells was maintained at each step to ensure a coverage of >1000 cells per sgRNA. For each cell line, 200 million cells were transduced with the sgRNA library in 12-well plates at an MOI of 0.3, then pooled into 15 cm dishes the next day. 24 hours post re-plating, 100 million cells were snap frozen in liquid nitrogen, then stored at −80C (for the preselection t_0 population), while 100 million cells were selected with 6 ug/mL puromycin for 24 hours. Subsequently, cells were passaged every two days, and at least 20 million cells were plated after each passage. After 8 population doublings (t_0), 20 million cells were snap-frozen in liquid nitrogen, then stored at −80C.

Infection of cells with the Brie library

Brie library was obtained from Addgene (cat. #73633). KPK-*Slc33a1* knockout clones were pooled (5 western validated clones) and plated at a concentration of 2 million cells/well in 12-well plates, together with the appropriate amount of viral supernatant in a total volume of 2 mL per well, then centrifuged at 1,500 RPM for 2 hours at room temperature. Cells were then incubated 37C overnight. To ensure that most cells harbor single integration events, a multiplicity of infection (MOI) of 0.3 was used for transduction of cell lines with the Brie library[66]. A minimum of 5 million cells was maintained at each step to ensure a coverage of >50 cells per sgRNA (78,637 sgRNAs x 50X coverage = ~3.9E6 cells). Cells were transduced in technical triplicates in 50 uM BSO with the Brie library in 12-well plates at an MOI of 0.3, then pooled into 15 cm dishes containing 50 uM BSO the next day. 48 hours post infection, cells were selected in 6 ug/mL Puromycin for 48 hours. Following selection, 5E6 cells per triplicate infection were harvested for t_0 sgRNA representation. 5E6 cells were then plated into vehicle (water; 6 replicates) or 50 uM BSO (3 replicates) for the duration of the experiment. Subsequently, cells were passaged every two days or when confluency was reached, and at least 5 million cells were plated after each passage to maintain 50X library coverage. After 8 population doublings (t_8), 10 million cells were snap-frozen in liquid nitrogen, then stored at −80C.

Genomic DNA isolation

Cell pellets were resuspended in 6 mL NK cell lysis buffer (50 mM Tris HCl, 50 mM EDTA, 1% SDS, pH 8) per 30-50 million cells. 30 uL of 20 mg/mL proteinase K (QIAGEN #19131) was added to the sample, which was then rotated overnight at 55C. The next day, 30 uL of 10 mg/mL RNase A (QIAGEN #19101) was added to the sample, which was then inverted 25 times and incubated for 30 minutes at 37C. Samples were cooled on ice, then 2 mL of pre-chilled 7.5 M ammonium acetate was added. Samples were vortexed at high speed for 20 seconds, then centrifuged at 4,000 x g for 10 minutes. The supernatant was decanted into a new tube, 6 mL of 100% isopropanol was added to the cleared supernatant. The sample was inverted 50 times and centrifuged at 4000 x g for 10 minutes. The supernatant was discarded, leaving genome DNA as a small white pellet. 6 mL of 70% ethanol was added to wash the pellet, then the tube was inverted 10 times and centrifuged at 4000 x g for 1 minute. The supernatant was discarded, then the DNA pellet was re-dissolved in 500 uL of TE Buffer (Sigma #T9285) and incubated overnight at 55C.

Screen deconvolution and analysis (druggable genome library)

Integrated sgRNAs were amplified by PCR to attach sequencing adapters and barcodes, as described[51]. In brief, a mix of 1 P5_XPR/LKO1 forward primers and 15 P7-index reverse primers (see supplemental sequences) was used to amplify sgRNAs from each genome DNA sample, using TaKaRa Ex Taq DNA polymerase (RR001C) with 25 cycles at 62C annealing temperatures. Assuming that each cell contains approximately 6.6 pg of genome DNA[67], a total of 160 ug of genome DNA (corresponding to >20 million cells) was used as a template for t_0 samples, while 720 ug of genomic DNA (>100 million cells) was used for t_0 samples. Amplified PCR products of around 350 bp in length were size-selected on a 2% agarose gel, then purified using the QIAquick Gel Extraction Kit (QIAGEN #28704). All samples were submitted for sequencing on an Illumina NextSeq system (75-nt single end reads) at the MIT BioMicro Center.

Sequencing results were analyzed using R (Version 3.1.1, www.r-project.org). Flanking adapter sequences were stripped to retrieve sgRNA target sequences from FASTQ file. Target sequences were collapsed by identity and quantified in terms of exact matches to library sequences. Sequencing read counts per sample were normalized to the total number of reads to yield counts per million sequences, followed by a $\log_2$ transformation[51].

The $\log_2$ fold change (L2FC) score for each sgRNA (sgRNA score) was calculated as the difference between t_8 and t_0 $\log_2$ transformed counts. Gene scores (GS) were then calculated as the median of the four sgRNA scores associated with each gene[68]. Differential GS for each isogenic pair were calculated by subtracting KP line GS from the KPK line GS. To determine thresholds for significant differentially depleted genes in the KPK versus KP cell line pairs, the 500 control sgRNAs were randomly binned into groups of 4, median L2FC of t_8 versus t_0 sgRNA read counts calculated, and differential GS calculated for these “pseudogenes”. This process was repeated 10,000 times to define a confident range of possible pseudogene differential GS, and thresholds set at the lower 5% of scores (false-discovery rate = 0.05. −0.64 in pair #1, and −0.63 in pair #2). To filter out genes that substantially deplete in KP lines, only genes with scores above the median GS of all essential genes (defined as the union of two previously described essential gene sets[27,69, 70], Supplemental Table 1) in each KP line (−1.77 in KP1, −2.25 in KP2) were included. To ensure that identified genes are significantly depleted in KPK lines (rather than enriched in KP lines but neutral in KPK lines), the population of pseudogene scores determined previously were used to set thresholds for significant depletion in the KPK lines (FDR = 0.05. −0.43 in KPK1, −0.34 in KPK2). Genes with scores above these thresholds in the KPK lines were excluded. Finally, we excluded essential genes from the analysis[27,69, 70].

Screen deconvolution and analysis (Brie library)

gDNA extraction and library PCR amplification was performed as above. A total of 1 ug of genome DNA was used as a template for t_0 and t_8 samples. All samples were submitted for sequencing on an Illumina NextSeq system (75-nt single end reads) with 50% Phi-X DNA spike in at the MIT BioMicro Center. Sequencing results were analyzed using R (Version 3.1.1, www.r-project.org). Flanking adapter sequences were stripped to retrieve sgRNA target sequences from FASTQ file. Target sequences were collapsed by identity and quantified in terms of exact matches to library sequences. Sequencing read counts per sample were normalized to the total number of reads to yield counts per million sequences, followed by a $\log_2$ transformation[51]. The $\log_2$ fold change (L2FC) for each sgRNA (sgRNA score) was calculated as the difference between t_8 and t_0 $\log_2$ transformed counts for either vehicle treated or BSO treated samples. Gene scores (GS) were then calculated as the median of the four sgRNA scores associated with each gene across all replicates[68]. Differential GS were calculated by subtracting median BSO treated GS from the median vehicle treated GS. Thresholds set for screen analysis: For t_8 analysis, all sgRNAs with a normalized sgRNA count of 0 at t_0 across 2 or more replicates were removed from the analysis. Genes for which >2 gRNAs met these criteria across all replicates were used for further analysis. For GS analysis: To determine thresholds for significantly enriched genes in the t_8 samples, we isolated all control sgRNAs and binned them into “pseudogenes” as described above. With these control pseudogene scores, we defined the upper and lower boundaries as the 25th-percentile (GS = −3.445) and 75th-percentile (GS = −4.311). Genes falling above the 25th-percentile threshold were deemed significantly enriched in the dataset compared to the spread of the pseudogene scores. Finally, we excluded essential genes from the analysis[27,69, 70]

In vitro competition assays

For *in vitro* competition assays, lentiviral vectors expressing a target sgRNA together with mNeonGreen (pUSPmNG) or tagRFP-P2A-Puro (pUSEPR) was generated by Gibson assembly as described previously[71]. The U6-filler sequence was amplified by PCR using lentiGuide-puro as a template, the PGK promoter sequence was obtained from the GMAP collection[71], and the mNeonGreen sequence was obtained as a gBlock gene fragment (Integrated DNA Technologies). sgRNAs targeting *Slc33a1*, *Tapt1, Suco, Rpa3,* and *Olfr102* were cloned into the BsmBI-digested pUSPmNG or pUSEPR vectors, using a previously described protocol[72]. Dual sgRNA sequences were ordered as a gBlock gene fragment

(Integrated DNA Technologies) contained an sgRNA (#1) H1 sgRNA (#2) sequence flanked by BsmBI-digest sites. Dual sgRNA sequences were BsmBI-digested and cloned into BsmBI-digested pUSEPR vectors. Cas9-expressing cells were infected with pUSPmNG or pUSEPR vectors expressing each sgRNA at an infection efficiency of ~50%. The percentage of mNeonGreen/tagRFP-positive cells was quantified over time using a Guava easyCyte HT flow cytometer (Millipore). For each sgRNA, the percentage of mNeonGreen/tagRFP-positive cells was normalized to respective values at 2 days post infection (defined as day 2).

Colony formation assays Brie library

Cells were infected at a high MOI with pUSEPR-sgRNA constructs by spin-infection in a 12 well plate as described above in the screen format. Cells were selected and maintained in BSO as described above in the Brie library infection. After infection, cells were withdrawn from BSO and plated in 6 well plates at a density 50E5 cells. At 7 days post seed, cells were passaged into new 6 well plates at a density of 50E5 cells. At 14 days, pictures were taken using a Nikon Eclipse Ti2 microscope at 20X magnification. At day 15, plates were processed for crystal violet stain.

RNA knockdown studies

Lentiviral vectors encoding TRE3GS-mScarlet-miR30a-hPGK-TetOne-P2A-Puro (LVt-TSTOP) were generated by Gibson assembly as shown previously[71]. Briefly, TRE3GS promoter and TetOne (rtTA) sequences were PCR amplified from Clontech plasmid pLVX-TetOne inducible expression systems (Cat # 631844). mScarlet sequences containing miR30a cloning site in the mScarlet 3’-UTR were obtained as a gBlock gene fragment (Integrated DNA Technologies). PGK promoter sequence was obtained from the GMAP collection[71]. P2A-Puro sequence was PCR amplified from lentiCRISPR v2 (Addgene #52961). 97-mer sequences targeting *Renilla* and *Slc33a1* were obtained from http://www.splashrna.mskcc.org and cloned as previously described[73] (see supplemental sequences). Cells were infected at an M.O.I. ~0.5 and subsequently selected in puromycin (6ug/mL) for 72 hours. Cells were induced with doxycycline (Sigma #D9891; 10 ng/mL) for 7 days and continuously passaged in the presence or absence of doxycycline or BSO (50 uM). CRISPR-interference (CRISPRi) constructs were obtained Addgene (pLV hU6-sgRNA hUBC-dCas9-KRAB-T2A-Puro; #71236). Promoter targeting sgRNA’s for *Slc33a1* or *Olfr102* were obtained from previously published whole genome CRISPRi datasets.[74] For CRISPRi experiments, a dual sgRNA setup was used to obtain maximal promoter silencing. sgRNA binding sites were selected to be >50 bp apart in genomic space.

Immunoblotting

Cells were lysed in 250 μL ice-cold RIPA buffer (Pierce, #89900) supplemented with 1× Complete Mini inhibitor mixture (Roche, #11836153001) 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). 30–50 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-GAPDH (Santa Cruz, sc-25778, 1:500), anti-HSP90 (BD, #610418, 1:10,000), anti-NRF2 (Santa-Cruz, sc-722, 1:200), anti-KEAP1 (CST, #8047, 1:1000), anti-GCLC (Santa

Cruz, sc-22755, 1:200). anti-p62 (CST #23214, 1:1000), anti-LC3 (CST #12741, 1:1000), anti-BECLIN (CST #3495, 1:1000), anti-pS6K T421/424 (CST #9204, 1:1000), anti-BiP (CST 3177, 1:1000), anti-FLAG (CST #14793, 1:1000), anti-HH3 (CST #4499, 1:1000), anti-GCN2 (CST #65981S, 1:1000), anti-PERK (CST #5683S, 1:1000), anti-ATF4 (CST #11815S, 1:1000), anti-eIF2A (CST #5324S, 1:2000), anti-phospho-eIF2A S51 (CST #3398S, 1:1000), anti-SLC33A1 (Sigma HPA042430, 1:1000).

Immunohistochemistry and immunofluorescence

Mice were euthanized by carbon dioxide asphyxiation. Lungs were perfused through the trachea with PBS and then 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-NQO1 (Sigma Aldrich, HPA007308, 1:200). Horseradish peroxidase (HRP) detection was used for NQO1, and pHH3. NQO1 and pHH3 was antigen retrieved in Ventana Cell Conditioner 1 (Tris-Borate-EDTA). Pictures were obtained using a digital whole slide scanner Leica SCN400F and Slidepath software version 4.0.8.

Lentiviral production

Lentiviruses were produced by co-transfection of 293T cells with lentiviral backbone constructs and packaging vectors (psPAX2 and pMD2.G; Addgene #12260 and #12259) 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

pUSEC lentiviral vector and cloning strategy was previously described[49]. For *in vitro* CRISPR experiments, pUSEPR or pUSPmNG lentiviral vectors were used. For sgRNA cloning, all vectors were digested with BsmBI and ligated with BsmBI-compatible annealed oligos for sgRNAs (Supplementary Sequences). For dual sgRNA vectors, sgRNA dual sequences containing the H1 promoter were ordered as a gBlock gene fragment (Integrated DNA Technologies), BsmBI digested, and cloned as previously described[75].

DepMap analysis

All data analyzed were accessed from the Broad Institute (www.depmap.org) whole genome CRISPR screens (DepMap 19Q2 Public). Gene co-essentiality datasets were obtained from ref. 27. Ranking DepMap cell-lines by their transcriptional correlation with the *NRF2* core target gene set (and the *KEAP1*-mutant signature) was performed using ssGSEA[76]. Separately, in order to search for targets that correlate well with NRF2 dependence, we calculated the Pearson Correlation coefficient between select genes and every other gene. The genetic background of DepMap cell lines was assessed from mutational data made available by the CCLE. Results of these analyses were used to compare CERES scores between *KEAP1/NRF2*-mutant lines versus wild-type lines, as well as across the ranked list of cell lines based on their transcriptional correlation with the NRF2 core gene set.

Transcriptome analysis

RNA was collected from cells as before[4] 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 *Gclc*, *Hmox1*, *Nqo1*, *Atf4*, *Hspa5*, *Xpb1s*, *Ddit3* were analyzed by quantitative reverse transcription polymerase chain reaction on LightCycler 480 II (Roche). RT qPCR primers (Supplementary Sequences).

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. Metabolites were measured using YSI biochemistry analyzer (Yellow Springs Instruments, Yellow Springs, OH).

Liquid chromatography-mass spectrometry (LC/MS) analysis

Cells were seeded in the morning in DMEM supplemented with 10% dialyzed FBS at 1 mL/well containing 50 uM BSO in 6-well plates (100,000 cells/well). The following day, 2 mL of fresh DMEM was added to each well with or without 50 uM BSO. 48 hours after media replenishment, cell numbers were counted for another set of BSO-treated and untreated samples (3 wells per condition, each well counted twice). Samples in each well were washed twice with 4 mL of ice-cold Blood Bank Saline (Azer Scientific, #ES1244G), then extracted on ice using ice-cold methanol:water (80:20) containing norvaline as an internal standard. Extracts were collected, transferred to 1.5 mL tubes, vortexed at maximum speed for 15 min at 4°C, then centrifuged at 13,000 RPM for 10 min at 4°C. The supernatant was collected, transferred to a new 1.5 mL tube, then evaporated to dryness under nitrogen. Samples were analyzed by LC/MS (Whitehead Metabolite Profiling Core Facility, Cambridge, MA) as described previously[77].

Glutathione measurements

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 two-tailed Student's t-test for all measurements comparing untreated to treated samples of single time points. Mann Whitney U-test with Holm's multiple comparisons testing was used for tumor burden and grade quantifications. One-way analysis of variance (ANOVA) with Dunnet'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 with the appropriate multiple comparisons test listed in the figure legends. Figure legends denominate statistical analysis used. All error bars denote standard deviation, unless otherwise noted in the figure legend. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$.

Bioinformatic analysis of CRISPR-targeted loci

For PCR amplicons (sequenced at the MGH sequencing facility), 150-300 bp 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[78]. 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[79] 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[80]. The alignment and post-processing code was implemented in C++ along with library functions from SeqAn[81] 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[82] (IGV).

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[83] Raw estimated expression counts were upper-quartile normalized to a count of 1000[84]. *Keap1*-mutant ($n = 2$), WT ($n = 2$), and WT plus sg*Slc33a1* transduced ($n = 2$) samples were jointly analyzed to derive a murine signature of *Slc33a1*-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[4,61, 85]. 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 upregulated 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[86] (Joint Approximate Diagonalization of Eigenmatrices) 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 *Slc33a1*-mutant signature was derived from this analysis, identifying genes with a differential expression pattern between wild-type and all other samples. 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[87,88]. This *Slc33a1* knockout signature was compared to publicly available datasets[34].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We thank J. Settleman, D. Stokoe, T. Papagiannakopoulos, I. Harris, L. Sullivan, M. Sullivan, Z. Li, and G. DeNicola, for scientific discussions and feedback. S. Levine for massively parallel sequencing expertise; M. Griffin, M. Jennings and G. Paradis for FACS support; K. Cormier and the Hope Babette Tang (1983) Histology Facility for histology support; C. Lewis, B. Chan and the Whitehead Metabolomics core for LC/MS and metabolite analysis; K. Yee, A. Deconinck, J. Teixeira for administrative support; and the Swanson Biotechnology Center for excellent core facilities. This work was supported by the Howard Hughes Medical Institute, Calico Life Sciences LLC, and the Koch Institute Support (core) Grant P30-CA14051 from the National Cancer Institute. R.R. was supported by the National Science Foundation Graduate Research Fellowship under award number 1122374 and the National Cancer Institute of the National Institutes of Health under Award Number F31CA224796. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. T.J. is a Howard Hughes Medical Institute Investigator, David H. Koch Professor of Biology, Daniel K. Ludwig Scholar.

References:

1. Shen Ret al.Harnessing Clinical Sequencing Data for Survival Stratification of Patients With Metastatic Lung Adenocarcinomas. JCO Precis. Oncol1–9 (2019).
2. Jaramillo MC & Zhang DD The emerging role of the Nrf2-Keap1 signaling pathway in cancer. Genes Dev. 27, 2179–2191 (2013). [PubMed: 24142871]
3. LUAD TCGAet al.Comprehensive molecular profiling of lung adenocarcinoma. Nature511, 543–550 (2014). [PubMed: 25079552]
4. Romero Ret al.Keap1 loss promotes Kras-driven lung cancer and results in dependence on glutaminolysis. Nat. Med. 23, 1362–1368 (2017). [PubMed: 28967920]
5. Jonas MC, Pehar M & Puglielli L AT-1 is the ER membrane acetyl-CoA transporter and is essential for cell viability. J. Cell Sci 123, 3378–3388 (2010). [PubMed: 20826464]
6. Berger AHet al.High-throughput Phenotyping of Lung Cancer Somatic Mutations. Cancer Cell30, 214–228 (2016). [PubMed: 27478040]

7. Hammerman PSet al.Comprehensive genomic characterization of squamous cell lung cancers. Nature489, 519–525 (2012). [PubMed: 22960745]
8. Hast BEet al.Cancer-derived mutations in KEAP1 impair NRF2 degradation but not ubiquitination. Cancer Res. 74, 808–817 (2014). [PubMed: 24322982]
9. Rojo de la Vega M, Chapman E & Zhang DD NRF2 and the Hallmarks of Cancer. Cancer Cell 34, 21–43 (2018). [PubMed: 29731393]
10. Itoh K, Mimura J & Yamamoto M Discovery of the Negative Regulator of Nrf2, Keap1: A Historical Overview. Antioxid. Redox Signal 13, 1665–1678 (2010). [PubMed: 20446768]
11. Todoric Jet al.Stress-Activated NRF2-MDM2 Cascade Controls Neoplastic Progression in Pancreas. Cancer Cell32, 824–839.e8 (2017). [PubMed: 29153842]
12. Chio IICet al.NRF2 Promotes Tumor Maintenance by Modulating mRNA Translation in Pancreatic Cancer. Cell166, 963–976 (2016). [PubMed: 27477511]
13. DeNicola GMet al.Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. Nature475, 106–110 (2011). [PubMed: 21734707]
14. Sayin VIet al.Activation of the NRF2 antioxidant program generates an imbalance in central carbon metabolism in cancer. Elife6, 1–23 (2017).
15. DeNicola GMet al.NRF2 regulates serine biosynthesis in non-small cell lung cancer. Nat. Genet47, 1475–1481 (2015). [PubMed: 26482881]
16. Kang YPet al.Cysteine dioxygenase 1 is a metabolic liability for non-small cell lung cancer. Elife8, (2019).
17. Mitsuishi Yet al.Nrf2 Redirects Glucose and Glutamine into Anabolic Pathways in Metabolic Reprogramming. Cancer Cell22, 66–79 (2012). [PubMed: 22789539]
18. LeBoeuf SEet al.Activation of Oxidative Stress Response in Cancer Generates a Druggable Dependency on Exogenous Non-essential Amino Acids. Cell Metab. (2019).
19. Lignitto Let al.Nrf2 Activation Promotes Lung Cancer Metastasis by Inhibiting the Degradation of Bach1. Cell178, 316–329.e18 (2019). [PubMed: 31257023]
20. Shen Ret al.Harnessing Clinical Sequencing Data for Survival Stratification of Patients With Metastatic Lung Adenocarcinomas. JCO Precis. Oncol1–9 (2019).
21. Li Let al.Identification of DHODH as a therapeutic target in small cell lung cancer. Sci. Transl. Med11, eaaw7852 (2019). [PubMed: 31694929]
22. Pasetto Met al.Whole-genome RNAi screen highlights components of the endoplasmic reticulum/ Golgi as a source of resistance to immunotoxin-mediated cytotoxicity. Proc. Natl. Acad. Sci112, E1135–E1142 (2015). [PubMed: 25713356]
23. Cui Jet al.Competitive Inhibition of the Endoplasmic Reticulum Signal Peptidase by Non-cleavable Mutant Preprotein Cargos. J. Biol. Chem290, 28131–28140 (2015). [PubMed: 26446786]
24. Yamamori T, Meike S, Nagane M, Yasui H & Inanami O ER stress suppresses DNA double-strand break repair and sensitizes tumor cells to ionizing radiation by stimulating proteasomal degradation of Rad51. FEBS Lett. 587, 3348–3353 (2013). [PubMed: 24021650]
25. Liu Yet al.Activation of the Unfolded Protein Response via Inhibition of Protein Disulfide Isomerase Decreases the Capacity for DNA Repair to Sensitize Glioblastoma to Radiotherapy. Cancer Res. 79, 2923–2932 (2019). [PubMed: 30996048]
26. Cloer EWet al.p62-Dependent Phase Separation of Patient-Derived KEAP1 Mutations and NRF2. Mol. Cell. Biol38, (2018).
27. Meyers RMet al.Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. Nat. Genet49, 1779–1784 (2017). [PubMed: 29083409]
28. Kim Eet al.A network of human functional gene interactions from knockout fitness screens in cancer cells. Life Sci. Alliance2, e201800278 (2019). [PubMed: 30979825]
29. Singh Aet al.RNAi-Mediated Silencing of Nuclear Factor Erythroid-2-Related Factor 2 Gene Expression in Non-Small Cell Lung Cancer Inhibits Tumor Growth and Increases Efficacy of Chemotherapy. Cancer Res. 68, 7975–7984 (2008). [PubMed: 18829555]
30. Li Het al.The landscape of cancer cell line metabolism. Nat. Med25, 1–11 (2019). [PubMed: 30617338]

31. Igarashi Ket al.Regulation of transcription by dimerization of erythroid factor NF-E2 p45 with small Maf proteins. Nature367, 568–572 (1994). [PubMed: 8107826]

32. Hein MYet al.A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances. Cell163, 712–723 (2015). [PubMed: 26496610]

33. Pehar M, Jonas MC, Hare TM & Puglielli L SLC33A1/AT-1 Protein Regulates the Induction of Autophagy Downstream of IRE1/XBP1 Pathway. J. Biol. Chem 287, 29921–29930 (2012). [PubMed: 22787145]

34. Adamson Bet al.A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. Cell167, 1867–1882.e21 (2016). [PubMed: 27984733]

35. Meister A & Anderson ME Glutathione. Annu. Rev. Biochem 52, 711–760 (1983). [PubMed: 6137189]

36. Ponsero AJet al.Endoplasmic Reticulum Transport of Glutathione by Sec61 Is Regulated by Ero1 and Bip. Mol. Cell67, 962–973.e5 (2017). [PubMed: 28918898]

37. Griffith OW & Meister A Potent and specific inhibition of glutathione synthesis by buthionine sulfoximine (S-n-butyl homocysteine sulfoximine). 7558–7560 (1979).

38. Dolma S, Lessnick SL, Hahn WC & Stockwell BR Identification of genotype-selective antitumor agents using synthetic lethal chemical screening in engineered human tumor cells. Cancer Cell 3, 285–96 (2003). [PubMed: 12676586]

39. Criddle DNet al.Menadione-induced Reactive Oxygen Species Generation via Redox Cycling Promotes Apoptosis of Murine Pancreatic Acinar Cells. J. Biol. Chem281, 40485–40492 (2006). [PubMed: 17088248]

40. Marzano Cet al.Inhibition of thioredoxin reductase by auranofin induces apoptosis in cisplatin-resistant human ovarian cancer cells. Free Radic. Biol. Med42, 872–881 (2006). [PubMed: 17320769]

41. Harris ISet al.Deubiquitinases Maintain Protein Homeostasis and Survival of Cancer Cells upon Glutathione Depletion. Cell Metab. 29, 1166–1181 (2019). [PubMed: 30799286]

42. Peng Yet al.Deficient Import of Acetyl-CoA into the ER Lumen Causes Neurodegeneration and Propensity to Infections, Inflammation, and Cancer. J. Neurosci34, 6772–6789 (2014). [PubMed: 24828632]

43. DuPage M, Dooley AL & Jacks T Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. Nat. Protoc 4, 1064–1072 (2009). [PubMed: 19561589]

44. Davidson SMet al.Environment Impacts the Metabolic Dependencies of Ras-Driven Non-Small Cell Lung Cancer. Cell Metab. 23, 517–528 (2016). [PubMed: 26853747]

45. Muir A & Vander Heiden MG The nutrient environment affects therapy. Science. 360, 962–963 (2018). [PubMed: 29853672]

46. Jackson ELet al.Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. 3243–3248 (2001).

47. Jackson ELet al.The differential effects of mutant p53 alleles on advanced murine lung cancer. Cancer Res. 65, 10280–10288 (2005). [PubMed: 16288016]

48. Winslow MMet al.Suppression of lung adenocarcinoma progression by Nkx2–1. Nature473, 101–104 (2011). [PubMed: 21471965]

49. Tammela Tet al.A Wnt-producing niche drives proliferative potential and progression in lung adenocarcinoma. Nature545, 355–359 (2017). [PubMed: 28489818]

50. Erard N, Knott SRV & Hannon GJ A CRISPR Resource for Individual, Combinatorial, or Multiplexed Gene Knockout. Mol. Cell 67, 1080 (2017). [PubMed: 28938093]

51. Doench JGet al.Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol34, 184–191 (2016). [PubMed: 26780180]

52. Brennan MS, Matos MF, Richter KE, Li B & Scannevin RH The NRF2 transcriptional target, OSGIN1, contributes to monomethyl fumarate-mediated cytoprotection in human astrocytes. Sci. Rep 7, 42054 (2017). [PubMed: 28181536]

53. Tu BP & Weissman JS Oxidative protein folding in eukaryotes: mechanisms and. J. Cell Biol 164, 341–346 (2004). [PubMed: 14757749]

54. Sevier CS & Kaiser CA Ero1 and redox homeostasis in the endoplasmic reticulum. Biochim. Biophys. Acta - Mol. Cell Res 1783, 549–556 (2008).

55. McMillan EAet al.Chemistry-First Approach for Nomination of Personalized Treatment in Lung Cancer. Cell173, 864–878.e29 (2018). [PubMed: 29681454]

56. Bar-Peled Let al.Chemical Proteomics Identifies Druggable Vulnerabilities in a Genetically Defined Cancer. Cell171, 696–709.e23 (2017). [PubMed: 28965760]

57. Li Met al.Genome-wide CRISPR-KO Screen Uncovers mTORC1-Mediated Gsk3 Regulation in Naive Pluripotency Maintenance and Dissolution. Cell Rep. 24, 489–502 (2018). [PubMed: 29996108]

58. Hart Tet al.High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell163, 1515–1526 (2015). [PubMed: 26627737]

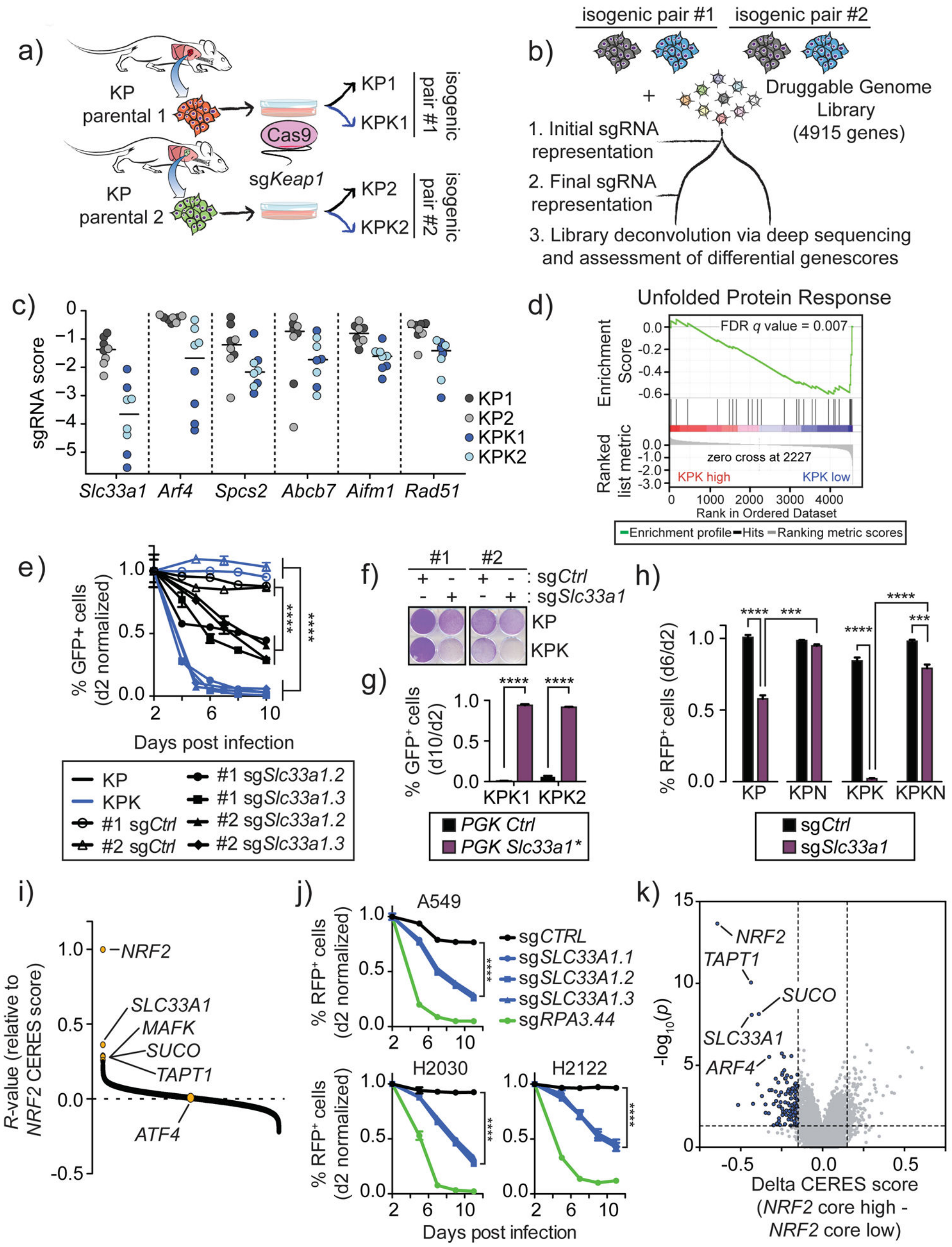

Figure 1: A genetic screen for *Keap1*-deficient genetic vulnerabilities yields *Slc33a1* as a potent dependency

(a) A schematic for the generation of two isogenic cell line pairs. Plucked tumors from two individual KP mice were used to generate parental KP tumor-derived cell lines. Each parental line was transfected *in vitro* with plasmids containing sg*Keap1*, Cas9 and GFP. Transfected lines were single cell sorted and validated for deficiency (KPK) or maintenance of two wild-type *Keap1* alleles (KP).

(b) CRISPR-Cas9 based screening strategy designed to identify *Keap1*-deficient genetic vulnerabilities. Two isogenic KP and KPK cell line pairs were infected with a Druggable Genome Library (DGL). Genomic DNA was collected at initial time points (t_0) and after 8 population doublings (t_8).

(c) Top common differentially depleted genes between each isogenic pair. Each dot represents the $\log_2$ fold change of each sgRNA (sgRNA score) targeting the indicating gene (4 sgRNAs per gene). Bar represents median gene score across all sgRNAs targeting the indicated gene per genotype.

(d) Gene set enrichment analysis of the rank ordered differential gene scores across the entire DGL of isogenic pair 1.

(e) Fluorescence competition assays of wild-type versus sg*Ctrl* or sg*Slc33a1* cells in KP and KPK isogenic cell lines. Plot displays day 2 normalized %GFP$^+$ (pUSPmNG) cells marking sg*Ctrl* or sg*Slc33a1* edited cells (mean ± SD; *n* = 3; *n* denotes technical replicate infections). Statistics derived from the comparison of the average KP- or KPK-sg*Ctrl* values versus the average KP- or KPK-sg*Slc33a1* values across all *Slc33a1* targeting sgRNAs.

(f) Colony formation of KP and KPK isogenic cells targeted with sg*Ctrl* or sg*Slc33a1* (pUSEPR).

(g) Overexpression of an sgRNA mutant-*Slc33a1* cDNA (*) or control cDNA (*mScarlet*) in both KPK cell lines. Plot displays %GFP$^+$ (pUSPmNG) cells 10 days post infection relative to day 2 (mean ± SD; *n* = 3; *n* denotes technical replicate infections).

(h) Fluorescence competition assays of KP, KP-*Nrf2* KO (KPN), KPK, KPK-*Nrf2* KO (KPKN) isogenic pairs. Plot displays %RFP$^+$ (pUSEPR) cells marking *sgCtrl* or *sgSlc33a1* infected cells 6 days post infection relative to day 2 (mean ± SD; *n* = 3; *n* denotes technical replicate infections).

(i) Waterfall plot of the rank ordered Pearson correlation coefficient between *NRF2* CERES scores versus all genes screened from the DepMap database version 19Q2.

(j) Fluorescence competition assay on a panel of *KEAP1*-mutant human LUAD cell lines. Plot displays day 2 normalized %RFP$^+$ (pUSEPR) cells marking sg*CTRL*, sg*SLC33A1*, or sg*RPA3* infected cells (mean ± SD; *n* = 3; *n* denotes technical replicate infections). Statistics derived from the comparison of the sg*CTRL* values versus the average sg*SLC33A1* values across all *SLC33A1* targeting sgRNAs.

(k) Differential CERES scores between core NRF2 core gene set high-ranking cell lines (*n* = 42) versus low-ranking cell lines (*n* = 32) from the DepMap database 19Q2. Horizontal dotted line represents *p*-value significance cut-off of $p<0.05$. Each dot represents the differential CERES score per gene. Blue dots represent genes that pass all set threshold values. Statistical analyses were performed using two-way ANOVA in (e) and (j) with Tukey's *post-hoc* multiple comparisons test. Student's two-tailed t-test was performed for (g), (h) and (k).

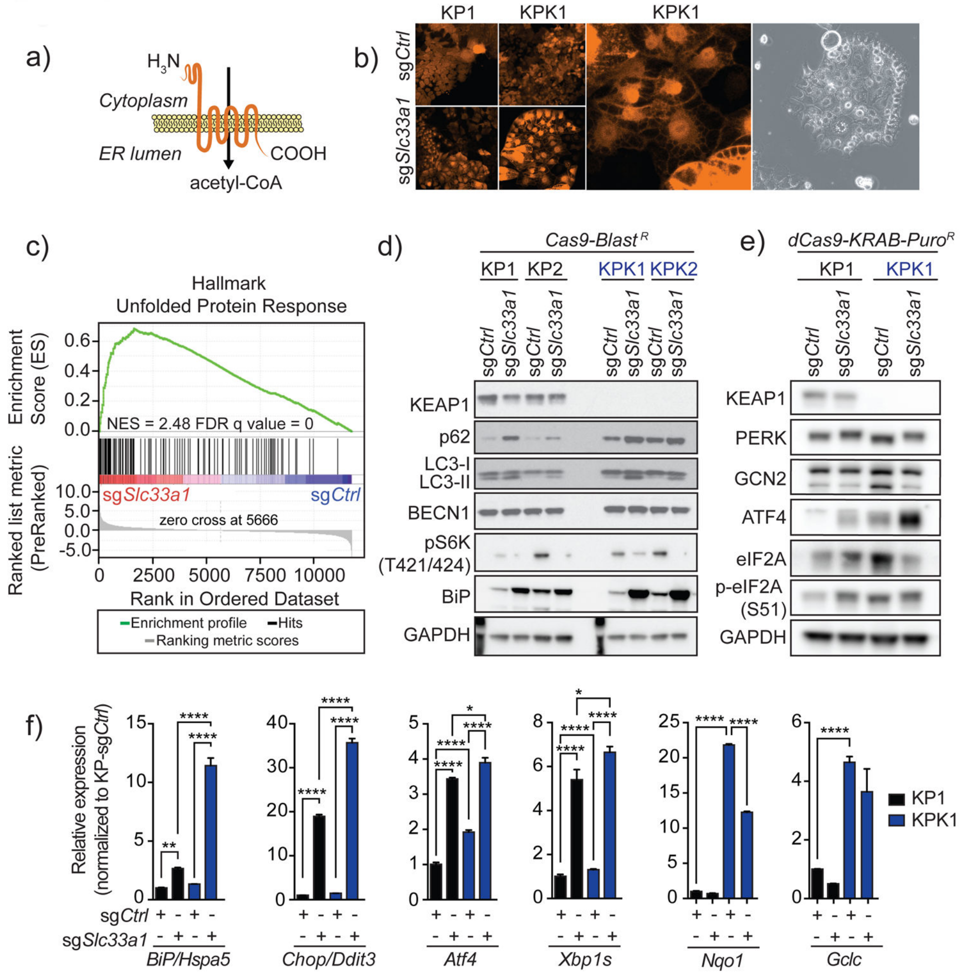

Figure 2: Loss of *Slc33a1* promotes the activation of an unfolded protein response

(a) Schematic of the reported function of SLC33A1.

(b) Confocal images of cytoplasmic-tagRFP (pUSEPR; see methods) labeled isogenic pair #1 infected with sg*Ctrl* or sg*Slc33a1*. Larger picture denotes increase magnification of KPK1-sg*Slc33a1* cells. Far right: Phase contrast microscopy of KPK-*Slc33a1* cells.

(c) GSEA enrichment plot of the Hallmark UPR signature enriched in sg*Slc33a1* targeted cells compared to sg*Ctrl* targeted cells.

(d) Western blot analysis of Cas9 expressing KP and KPK isogenic pairs targeted with *sgCtrl* or *s*g*Slc33a1* for autophagy and UPR protein markers.

(e) Western blot analysis of dCas9-KRAB expressing KP and KPK isogenic pairs targeted with *sgCtrl* or *s*g*Slc33a1* for autophagy and UPR protein markers.

(f) qPCR validation of UPR and NRF2 target genes in Cas9 expressing isogenic pair #1 targeted with sg*Ctrl* or sg*Slc33a1*. (mean ± SD; *n* = 4; *n* denotes technical replicates). Statistical analyses were performed using two-way ANOVA in (f) with Tukey's *post-hoc* test.

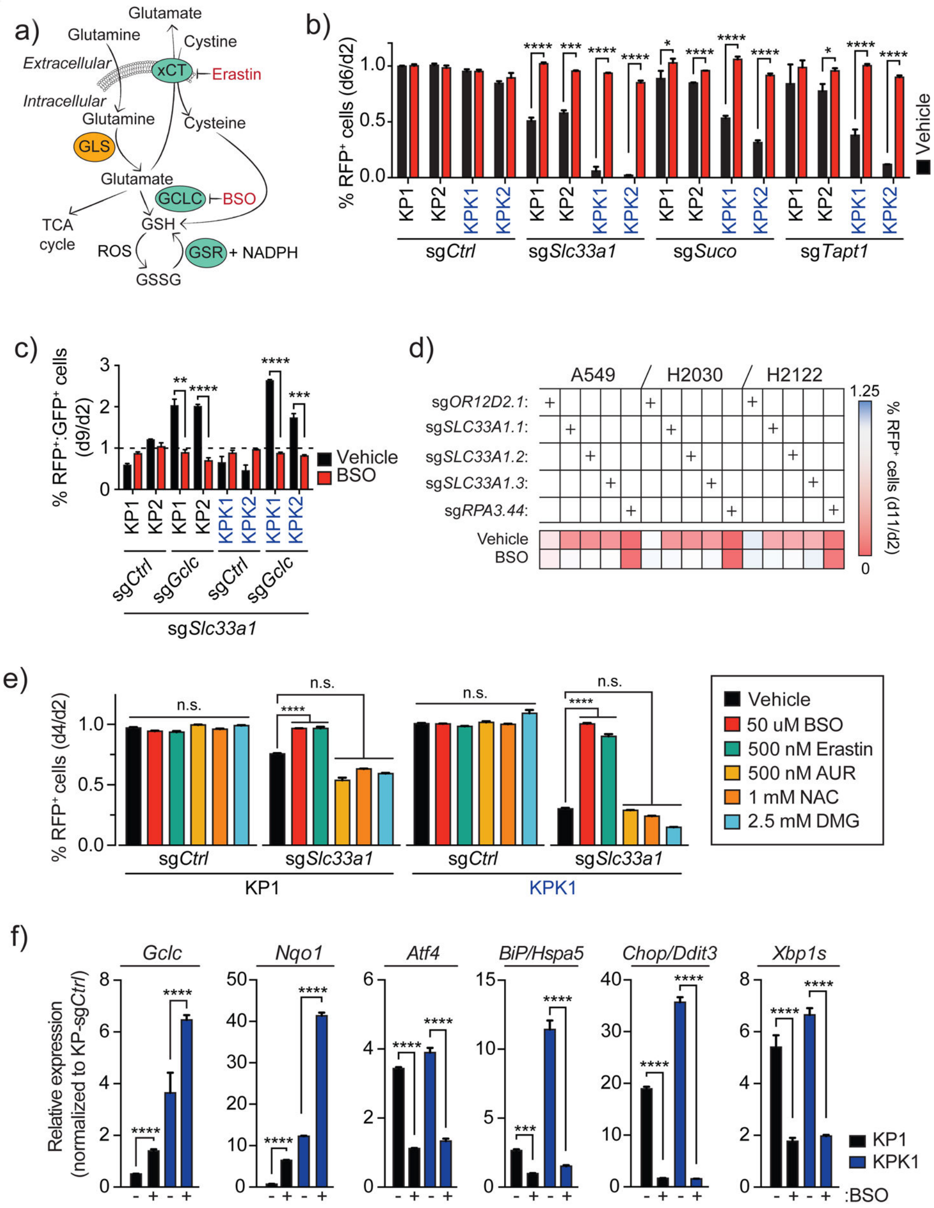

Figure 3: Loss of *Slc33a1* is rescued by inhibiting glutathione synthesis

(a) Schematic of the biosynthetic pathways contributing to GSH synthesis. Enzymes marked in blue circles are direct NRF2 target genes. Small molecule inhibitors are denoted by red font.

(b) Fluorescence competition assay in KP and KPK isogenic cell lines infected with pUSEPR lentiviruses expressing sg*Ctrl*, sg*Slc33a1*, sg*Suco*, or sg*Tapt1*. All cell and sgRNA combinations are treated with vehicle or 50 uM BSO. Relative %RFP is normalized to %RFP at day 2 post infection. Statistics are derived as the mean ± SD; *n* = 3; *n* denotes technical replicate infections.

(c) Fluorescence competition assay of RFP$^+$-*Slc33a1*-deficient cells (pUSEPR) targeted with sg*Ctrl* or sg*Gclc* containing pUSPmNG lentiviral vectors. Double positive %RFP:GFP values normalized to day 2 %RFP:GFP values. Cells were treated with vehicle or 50 uM BSO (mean ± SD; *n* = 3; *n* denotes technical replicate infections).

(d) Heat map of fluorescence competition assays in *KEAP1*-mutant human LUAD cell lines (pUSEPR) treated with vehicle or 50 uM BSO. Data normalization as in (b). Each block is derived as the mean ± SD; *n* = 3; *n* denotes technical replicate infections.

(e) Day 2 normalized %RFP (pUSEPR) marked cells in a fluorescence competition assay targeted with sg*Ctrl* or sg*Slc33a1*. Cells were treated with vehicle, 50 uM BSO, 500 nM Erastin, 500 nM Auronafin, 1 mM NAC, or 2.5 mM DMG (mean ± SD; *n* = 3; *n* denotes technical replicate infections and treatment).

(f) qPCR validation of UPR and NRF2 target genes in isogenic pair #1 targeted with sg*Ctrl* or sg*Slc33a1*. (mean ± SD; *n* = 4; *n* denotes technical replicates). Cells were treated with vehicle or 50 uM BSO for 48 hours post infection. Statistical analyses were performed using student's two-tailed t-test in (b) and (c) and two-way ANOVA in (e) and (f) with Tukey's multiple comparisons test. All error bars denote standard deviation.

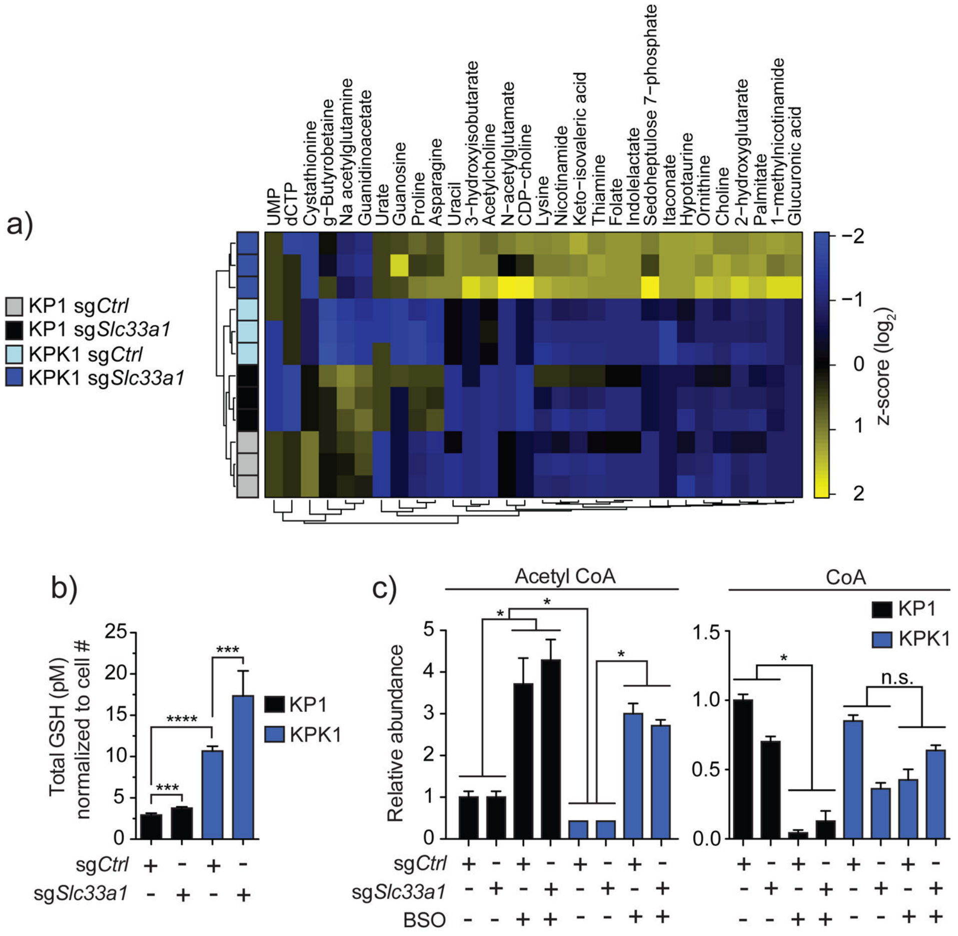

Figure 4: *Slc33a1* loss results in metabolic rewiring

(a) Unsupervised clustering of significantly differentially detectable polar metabolites. Each row represents a different metabolite. Each column represents a different sample. Each cell line condition was completed in triplicate. Data is normalized by cell counts for each cell line.

(b) GSH concentrations of KP and KPK cells 72 hours post transduction with the above labeled sgRNAs.

(c) Acetyl-CoA and CoA relative abundances from data represented in (a). Cells were treated with or without 50 uM BSO for 48 hours. Each cell line condition was completed in triplicate. Data is normalized by cell counts for each cell line. Statistical analyses were performed using Mann-Whitney U-test in (b) and (c).

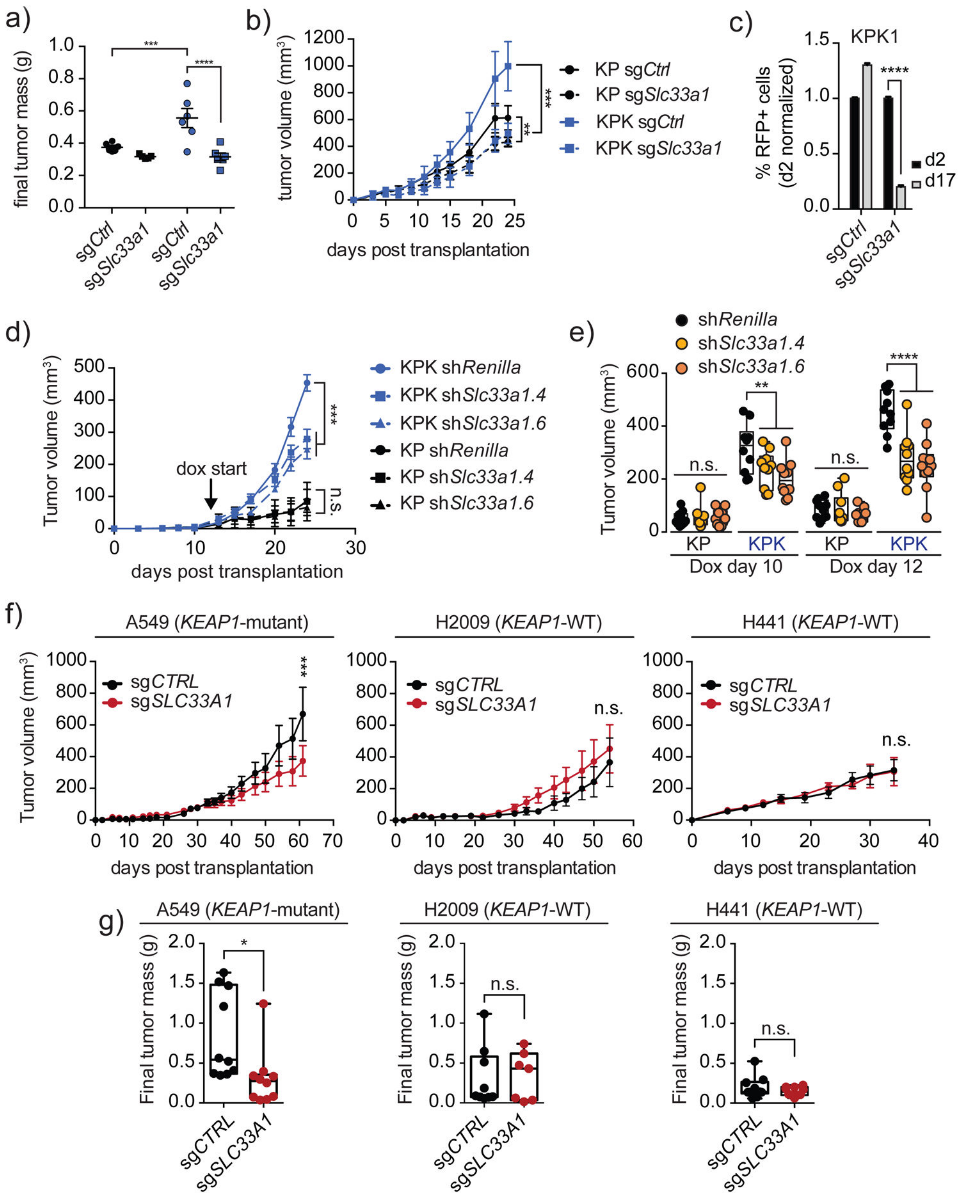

Figure 5: *Slc33a1* loss validates as a *Keap1*-mutant-specific vulnerability in transplant models

a) Final subcutaneous tumor masses obtained from the transplantation of indicated cell lines into immunocompromised mice.

b) Subcutaneous tumor volumes of KP and KPK cells transduced with the indicated sgRNAs injected into immunocompromised mice (n = 6 per group). KP sg*Ctrl* vs. KP sg*Slc33a1*, p = 0.0074. KPK sg*Ctrl* vs. KPK sg*Slc33a1*, p = 0.0003.

c) *in vivo* fluorescence based competition assay of GFP-labeled KPK cells injected into immunocompromised mice. KPK cells were transduced with pUSEPR constructs containing the indicated sgRNAs. Transduced cells representation was normalized to RFP$^+$ cells at d2 (pre-implantation).

(d) Subcutaneous tumor volumes of KP or KPK cells transduced with doxycycline-inducible shRNAs targeting *Renilla* or *Slc33a1* (LVt-TSTOP; see methods). Arrow indicates timepoint at which mice began doxycycline treatment. Error bars denote S.E.M. KPK-sh*Renilla* versus KPK-sh*Slc33a1* groups $p = 0.0006$. KP-sh*Renilla* versus KP-sh*Slc33a1* groups $p = 0.8467$).

(e) Tumor volumes related to (d) at the indicated time points. Day 10: KP, $p = 0.9387$, KPK, $p = 0.0088$. Day 12: KP, $p = 0.5126$, KPK, $p < 0.0001$.

(f) Xenografts of the *KEAP1*-mutant A549 or *KEAP1*-wild-type H2009 and H441 human lung cell lines infected with either sg*CTRL* or sg*SLC33A1* lentiviral vectors (pUSEPR). Error bars denote S.E.M. A549 ($n = 10$ per group): $p<0.001$. H2009 ($n = 9$ sg*CTRL*; $n = 10$ sg*SLC33A1*): $p = 0.292$, H441 ($n = 10$ sg*CTRL*; $n = 10$ sg*SLC33A1*): $p = 0.842$.

(g) Final subcutaneous tumor masses obtained from xenografts of indicated cell lines into immunocompromised mice. A549 ($n = 10$ sg*CTRL*; $n = 10$ sg*SLC33A1*): $p<0.0190$. H2009 ($n = 9$ sg*CTRL*; $n = 7$ sg*SLC33A1*): $p<0.8805$. H441 ($n = 10$ sg*CTRL*; $n = 10$ sg*SLC33A1*): $p<0.3061$. Statistics obtained by: Student's two-tailed t-test in (a), (c), and (g). Two-way ANOVA with Sidak's multiple comparisons testing in (b), and (f). Two-way ANOVA with Dunnet's multiple comparisons testing in (d). One-way ANOVA with Dunnet's multiple comparisons testing in (e).

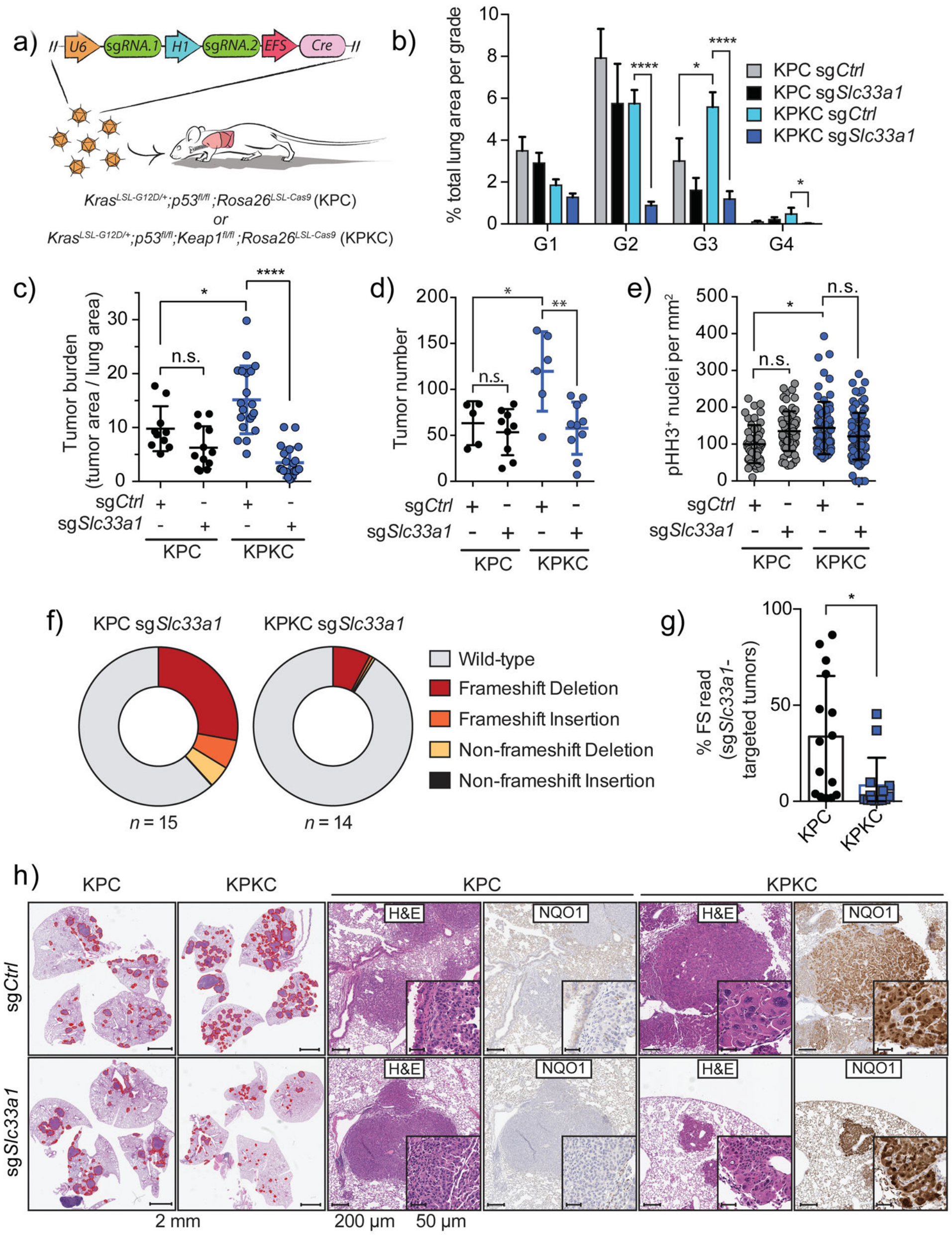

Figure 6: *Keap1*-deficient tumors harbor an increased dependency for *Slc33a1* in an autochthonous model of murine lung adenocarcinoma

(a) Schematic representation of $Kras^{LSL-G12D}/+$; $p53^{fl/fl}$; $Rosa26^{LSL-Cas9}$ (KPC) or $Kras^{LSL-G12D}/+$; $p53^{fl/fl}$; $Keap1^{fl/fl}$; $Rosa26^{LSL-Cas9}$ (KPKC) mice intratracheally infected with pUSEC lentiviruses containing dual sgRNAs targeting *Slc33a1* or *Olfr102* (sg*Ctrl*).

(b) Distribution of histological tumor grades (G1-G4) in KPC or KPKC mice 20 weeks after infection with pUSEC lentiviruses expressing control (sg*Ctrl*, KPC, *n* = 10 mice; KPKC, *n* = 20 mice) or sg*Slc33a1* (KPC, *n* = 12 mice; KPKC, *n* = 22 mice). Error bars denote S.E.M. (KP-sg*Ctrl* versus KPK-sg*Ctrl* *p* = 0.049; KPK-sg*Ctrl* versus KPK-sg*Slc33a1* G2 *p* = 8.2e-9; KPK-sg*Ctrl* versus KPK-sg*Slc33a1* G3 *p* = 2.81e-6; KPK-sg*Ctrl* versus KPK-sg*Slc33a1* G4 *p* = 0.025).

(c) Combined quantification of tumor burden (total tumor area/total lung area) in KPC or KPKC mice after infection with pUSEC lentiviruses. Mouse numbers equivalent to (b). (KP-sg*Ctrl* versus KP-sg*Ctrl* $p = 0.08$; KP-sg*Ctrl* versus KPK-sg*Ctrl* $p = 0.029$; KPK-sg*Ctrl* versus KPK-sg*Slc33a1* $p = 5.9\text{e-}7$).
(d) Combined quantification of average tumor number in KPC or KPKC mice after infection with pUSEC lentiviruses.
(e) Quantification of phosphorylated histone H3 (pHH3)-positive nuclei per squared millimeter of tumor for assessment of the mitotic index of tumor cells from lung tumors in KPC or KPKC mice at 20 weeks after infection with the indicated pUSEC lentivirus.
(f) Fraction of mutant and wild-type reads within individually plucked tumors from 20-week pUSEC-sg*Slc33a1* lentiviral infected KPC (mean of $n = 15$ tumors from 7 mice) and KPKC mice (mean of $n = 14$ tumors from 7 mice).
(g) Mean number of frameshift (FS) reads from USEC-sg*Slc33a1* lentiviral infected KPC (mean of $n = 15$ tumors from 7 mice) and KPKC mice (mean of $n = 14$ tumors from 7 mice) * $P = 0.0102$.
(h) Representative H&E and IHC staining of serial sections from lung tumors of KPC or KPKC mice 20 weeks after infection with pUSEC-sg*Ctrl* (top) or pUSEC-sg*Slc33a1* (bottom). First panels (KPC mice), overall lung tumor burden (scale bar = 2mm). Second panels, higher magnification H&E staining of representative tumors (scale bar = 100 um; insets depict higher-magnification images, scale bar = 50 um). Third panels, representative IHC staining for NRF2 target gene, NQO1. This is repeated for KPKC mice. Statistical analyses were performed using Mann-Whitney U-test with Holm’s multiple comparisons correction used in (b-e). Student’s two-tailed t-test was performed for (g).

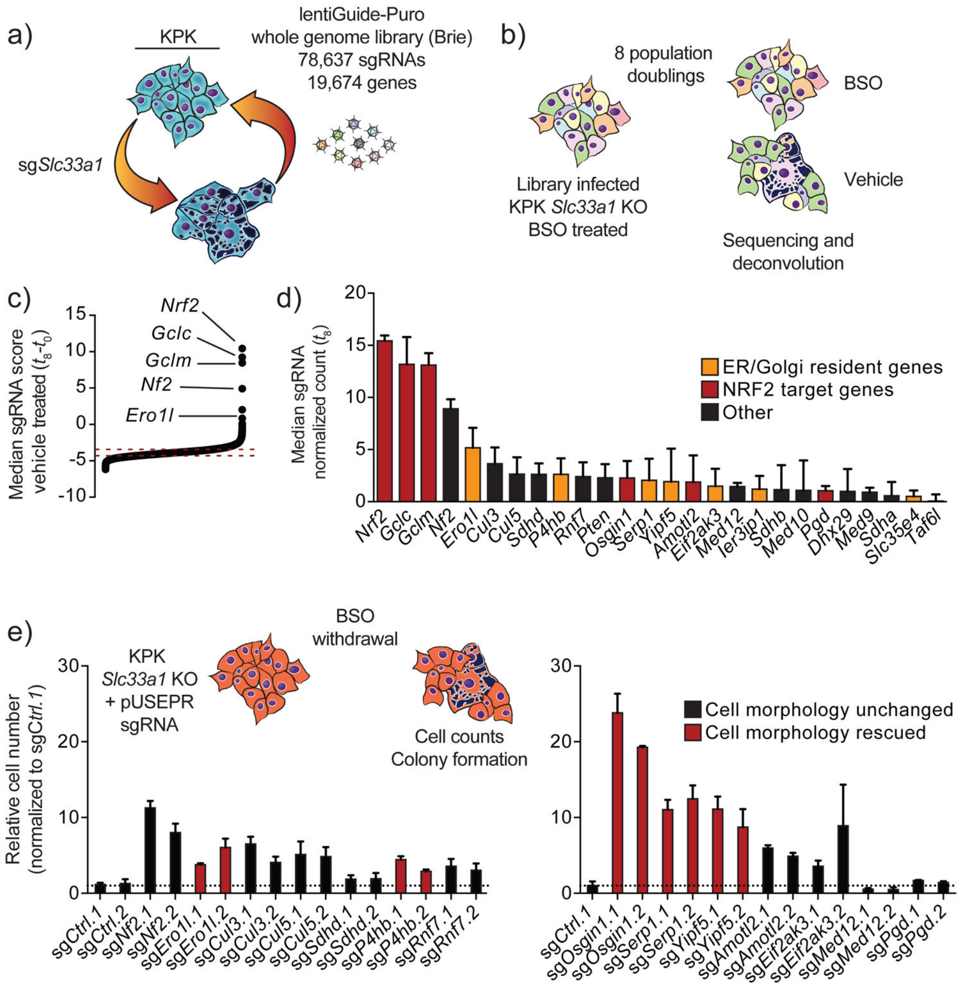

Figure 7: Genome-wide CRISPR screen identifies suppressors of *Slc33a1* deficiency

(a) Schematic of the screen conditions. KPK-*Slc33a1* knockout cells were generated and infected with the Brie library to identify hits that would rescue *Slc33a1* dependency.

(b) Schematic representing screen conditions. KPK-*Slc33a1* knockout cells were infected with Brie library in the presence of BSO and subsequently placed in BSO or vehicle treated conditions for 8 cumulative population doublings.

(c) Ranked median gene scores of genes in the Brie library that passed all thresholds (see methods). Dotted lines represent the spread of the control sgRNAs and mark the 25th and 75th gene score percentile boundaries.

(d) Waterfall plot of the normalized median sgRNA counts across all technical replicates at t_8. The median is plotted and the error bars denote the range of the data.

(e) Above: denotes the schematic for the validation pipeline. Below: Relative cell numbers post infection with pUSEPR vectors expressing the indicated sgRNA. Cell morphology related to Supplementary Fig. 13a.