nature genetics

NRF2 regulates serine biosynthesis in non–small cell lung cancer

Gina M DeNicola[1], Pei-Hsuan Chen[2,7], Edouard Mullarky[1,7], Jessica A Sudderth[2], Zeping Hu[2], David Wu[1], Hao Tang[3], Yang Xie[3], John M Asara[4], Kenneth E Huffman[5], Ignacio I Wistuba[6], John D Minna[5], Ralph J DeBerardinis[2] & Lewis C Cantley[1]

Tumors have high energetic and anabolic needs for rapid cell growth and proliferation[1], and the serine biosynthetic pathway was recently identified as an important source of metabolic intermediates for these processes[2,3]. We integrated metabolic tracing and transcriptional profiling of a large panel of non–small cell lung cancer (NSCLC) cell lines to characterize the activity and regulation of the serine/glycine biosynthetic pathway in NSCLC. Here we show that the activity of this pathway is highly heterogeneous and is regulated by NRF2, a transcription factor frequently deregulated in NSCLC. We found that NRF2 controls the expression of the key serine/glycine biosynthesis enzyme genes *PHGDH*, *PSAT1* and *SHMT2* via ATF4 to support glutathione and nucleotide production. Moreover, we show that expression of these genes confers poor prognosis in human NSCLC. Thus, a substantial fraction of human NSCLCs activates an NRF2-dependent transcriptional program that regulates serine and glycine metabolism and is linked to clinical aggressiveness.

Uniformly labeled [^{13}C]glucose ([U-^{13}C]glucose) is metabolized via the glycolytic intermediate 3-phosphoglycerate (3-PG) to serine M3 (with three ^{13}C-labeled atoms) and, subsequently, to glycine M2 (with two ^{13}C-labeled atoms) through the combined action of the serine biosynthetic pathway and serine hydroxymethyltransferase[4] (**Fig. 1a**). To profile the activity of the serine/glycine biosynthesis pathway in NSCLC, we labeled a panel of 79 human NSCLC cell lines with [^{13}C]glucose and quantified serine and glycine labeling via gas chromatography combined with mass spectrometry (GC/MS). We determined 6 and 24 h to be the optimal time points to detect labeling of serine and glycine (**Supplementary Fig. 1**). At 24 h, the fractional abundance of serine M3 and glycine M2 ranged from 0–40% (**Fig. 1b**,**c**), with significant correlation between the time points (**Supplementary Fig. 2**). Neither serine nor glycine labeling correlated with cellular doubling times (**Supplementary Fig. 3**). There was a significant correlation between serine M3 and glycine M2 labeling at both time points, indicating that the glycine produced from glucose was derived from serine (**Fig. 1d**,**e**). As has been reported previously[3], *de novo* serine synthesis conferred the ability to grow in the absence of extracellular serine (**Fig. 1f**,**g**). Thus, the serine biosynthesis pathway is not uniformly operant in NSCLC, and regulatory mechanisms exist for controlling the activation of this pathway in a subset of cell lines.

To identify a mechanism for increased serine biosynthesis in NSCLC, we correlated serine and glycine biosynthesis with gene expression using data from the Broad-Novartis Cancer Cell Line Encyclopedia (CCLE)[5]. We observed a significant correlation between serine and glycine labeling and the expression of the *PHGDH* gene encoding phosphoglycerate dehydrogenase (**Fig. 1h** and **Supplementary Table 1**), which catalyzes the first and rate-limiting step in serine biosynthesis[4]. 'Serine-high' cell lines (with a serine M3 *z* score >0.5 at 24 h) were sensitive to *PHGDH* silencing (**Fig. 1i**), and ectopic PHGDH rescued the proliferation of 'serine-low' cell lines (serine M3 *z* score <0.5 at 24 h) in serine-deficient medium (**Fig. 1j** and **Supplementary Fig. 4**). Previous studies have shown a role for *PHGDH* copy number gain in increasing serine biosynthetic activity[2,3]. However, we did not find evidence for this mechanism in NSCLC cell lines (**Supplementary Fig. 5** and **Supplementary Table 2**). To investigate alternative mechanisms of PHGDH regulation, we performed gene set enrichment analysis (GSEA)[6,7] on the genes whose expression positively correlated with serine and glycine biosynthesis. Interestingly, the gene set with targets of the transcription factor nuclear factor erythroid-2–related factor 2 (NRF2; encoded by *NFE2L2*) was the top hit (**Supplementary Fig. 6**), suggesting that NRF2 might be a regulator of PHGDH and the serine biosynthetic pathway.

Next, we examined NRF2 localization (**Supplementary Fig. 7**) and found a significant correlation of nuclear NRF2 with serine biosynthesis (**Fig. 2a**). Additionally, we ranked the cell lines according to expression of NRF2 target genes (**Supplementary Fig. 8** and

[1]Department of Medicine, Weill Cornell Medical College, New York, New York, USA. [2]Children's Medical Center Research Institute, University of Texas–Southwestern Medical Center, Dallas, Texas, USA. [3]Quantitative Biomedical Research Center, Department of Clinical Sciences, University of Texas–Southwestern Medical Center, Dallas, Texas, USA. [4]Division of Signal Transduction, Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts, USA. [5]Hamon Center for Therapeutic Oncology, University of Texas–Southwestern Medical Center, Dallas, Texas, USA. [6]Department of Translational Molecular Pathology, University of Texas MD Anderson Cancer Center, Houston, Texas, USA. [7]These authors contributed equally to this work. Correspondence should be addressed to L.C.C. (lcantley@med.cornell.edu).

Received 13 May; accepted 23 September; published online 19 October 2015; corrected after print 15 February 2016; doi:10.1038/ng.3421

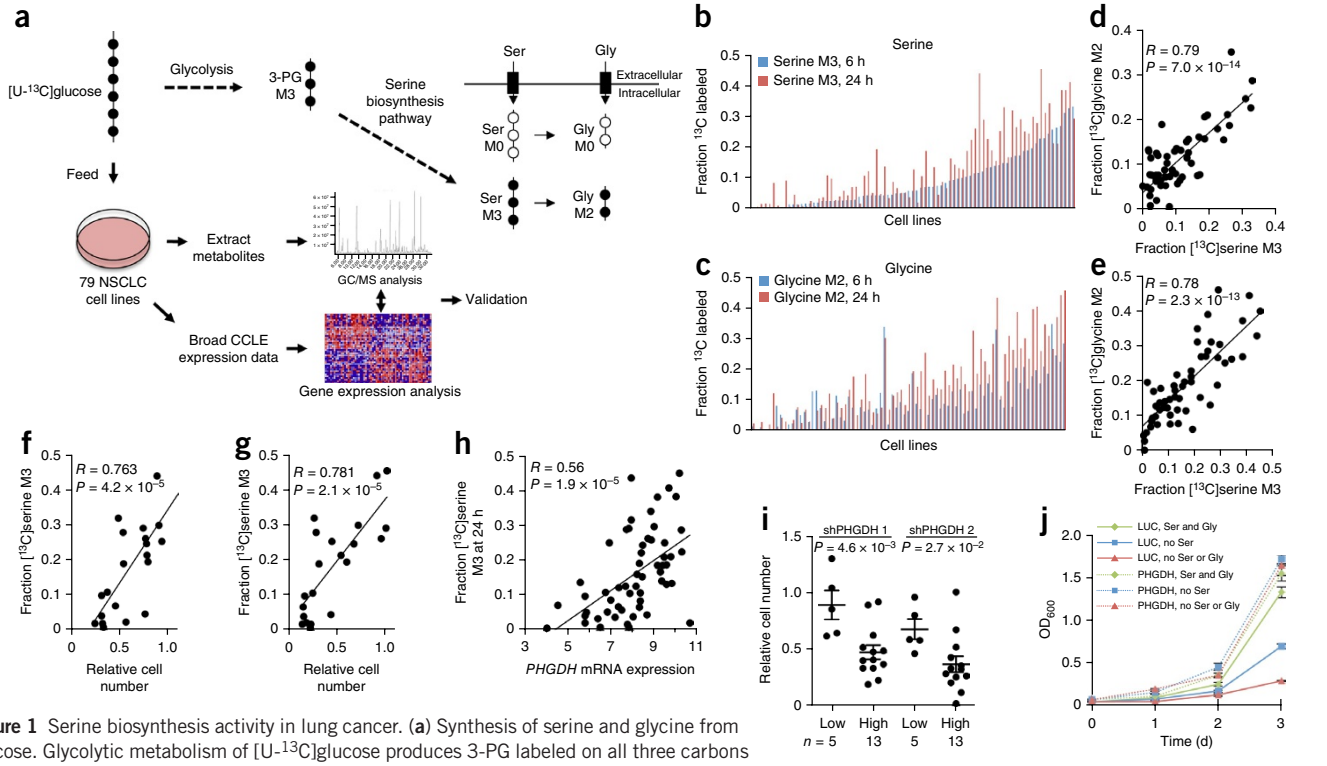

Figure 1 Serine biosynthesis activity in lung cancer. (**a**) Synthesis of serine and glycine from glucose. Glycolytic metabolism of [U-^{13}C]glucose produces 3-PG labeled on all three carbons (M3). The serine biosynthetic pathway produces serine (M3, three carbons labeled) and, subsequently, glycine (M2, two carbons labeled) from 3-PG. Unlabeled (M0) serine and glycine are contributed from the cellular medium. Filled circles, ^{13}C-labeled carbons; open circles, unlabeled carbons. (**b**,**c**) Fractions of labeled intracellular serine (**b**) and glycine (**c**) at 6 and 24 h in NSCLC cell lines. Cell lines are ordered according to serine M3 labeling at 6 h. (**d**,**e**) Correlation between serine and glycine labeling at 6 h (**d**) and 24 h (**e**). *R* is the Pearson correlation coefficient. *P* values were calculated by Student's *t* distribution with $n-2$ degrees of freedom. (**f**,**g**) Serine-high cell lines (with a serine M3 *z* score >0.5 at 24 h) are resistant to serine (**f**) or serine and glycine (**g**) starvation for 3 d. Each data point represents a cell line. (**h**) Serine labeling at 24 h correlates with *PHGDH* mRNA expression. mRNA expression data were obtained from CCLE. Each data point represents a cell line. (**i**) Serine-high cell lines demonstrate sensitivity to *PHGDH* knockdown that is not rescued by serine. Relative cell number was quantified 5 d after *PHGDH* knockdown. Each data point represents a cell line. shPHGDH, short hairpin RNA (shRNA) targeting *PHGDH*. (**j**) H1975 cells expressing luciferase (LUC) or PHGDH were assayed as in **f** and **g**. Results are the averages of three biological replicates. OD_{600}, optical density at 600 nm. All error bars, s.e.m.

Supplementary Table 3), which significantly correlated with NRF2 abundance (**Supplementary Fig. 7**). 'NRF2-high' cell lines had significantly higher serine M3 labeling at 6 and 24 h and glycine M2 labeling at 24 h (**Fig. 2b**,**c**). Next, we silenced NRF2 (*NFE2L2*) and found that it regulates the expression of serine/glycine biosynthetic pathway genes (*PHGDH*, *PSAT1*, *PSPH*, *SHMT1* and *SHMT2*) (**Fig. 2d**

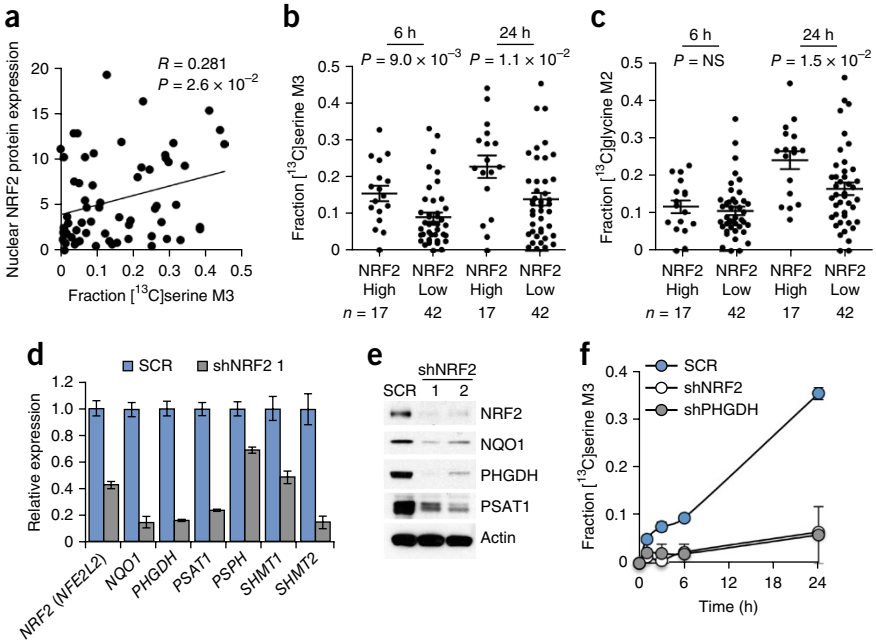

Figure 2 NRF2 regulates serine biosynthesis. (**a**) Nuclear NRF2 protein expression correlates with [^{13}C]serine M3 labeling at 24 h. Nuclear NRF2 protein expression data are presented in **Supplementary Figure 7**. (**b**) Cell lines with high NRF2 activity have significantly higher [^{13}C]serine M3 labeling at 6 and 24 h. Cell lines are grouped by NRF2 score into high-NRF2 (>1.4) and low-NRF2 (<1.4) lines. (**c**) NRF2-high cell lines have significantly higher [^{13}C]glycine M2 labeling at 24 h. *P* values in **b** and **c** were calculated using an unpaired, two-tailed Student's *t* test (NS, not significant). (**d**) mRNA expression in A549 cells expressing scrambled shRNA (SCR) or *NRF2* (*NFE2L2*) shRNA 1 (shNRF2). Decreased *NQO1* expression confirms that NRF2 activity was reduced upon knockdown. Results are the averages of three technical replicates. (**e**) Immunoblot analysis of serine pathway enzyme expression in lysates from A549 cells expressing scrambled shRNA or shNRF2 1 or 2. (**f**) Cell lines were grown in the presence of [U-^{13}C]glucose for the indicated times, metabolites were extracted and fractional ^{13}C labeling of serine was analyzed by liquid chromatography (LC)/MS. Results are the averages of three biological replicates. All error bars, s.e.m.

and **Supplementary Fig. 9**). Decreased *PHGDH* and *PSAT1* mRNA levels were accompanied by lower levels of the corresponding proteins (**Fig. 2e**). Furthermore, *NRF2* or *PHGDH* knockdown decreased the production of [^{13}C]serine from [^{13}C]glucose (**Fig. 2f** and **Supplementary Fig. 10**). These results demonstrate that the transcriptional regulation of serine biosynthetic enzymes by NRF2 controls the production of serine from glucose.

ATF4 transcriptionally activates serine biosynthetic genes in response to serine starvation in NSCLC cells[8]. Interestingly, nuclear ATF4 expression correlated with both serine labeling and NRF2 protein expression (**Supplementary Fig. 11**). ATF4 has been reported as both a direct transcriptional target[9,10] and heterodimerization partner[11,12] of NRF2. In agreement with transcriptional regulation of *ATF4* by NRF2, we observed marked reduction in *ATF4* mRNA expression, binding of RNA polymerase II to the *ATF4* promoter and the levels of newly synthesized *ATF4* mRNA upon *NRF2* silencing (**Fig. 3a** and **Supplementary Figs. 12a** and **13a**,**b**), with no effect on *ATF4* mRNA stability (**Supplementary Fig. 13c**,**d**). Furthermore, ATF4 protein levels were decreased upon *NRF2* knockdown (**Fig. 3b**), but NRF2 did not regulate *ATF4* translation[13] (**Supplementary Fig. 14**). Notably, *ATF4*

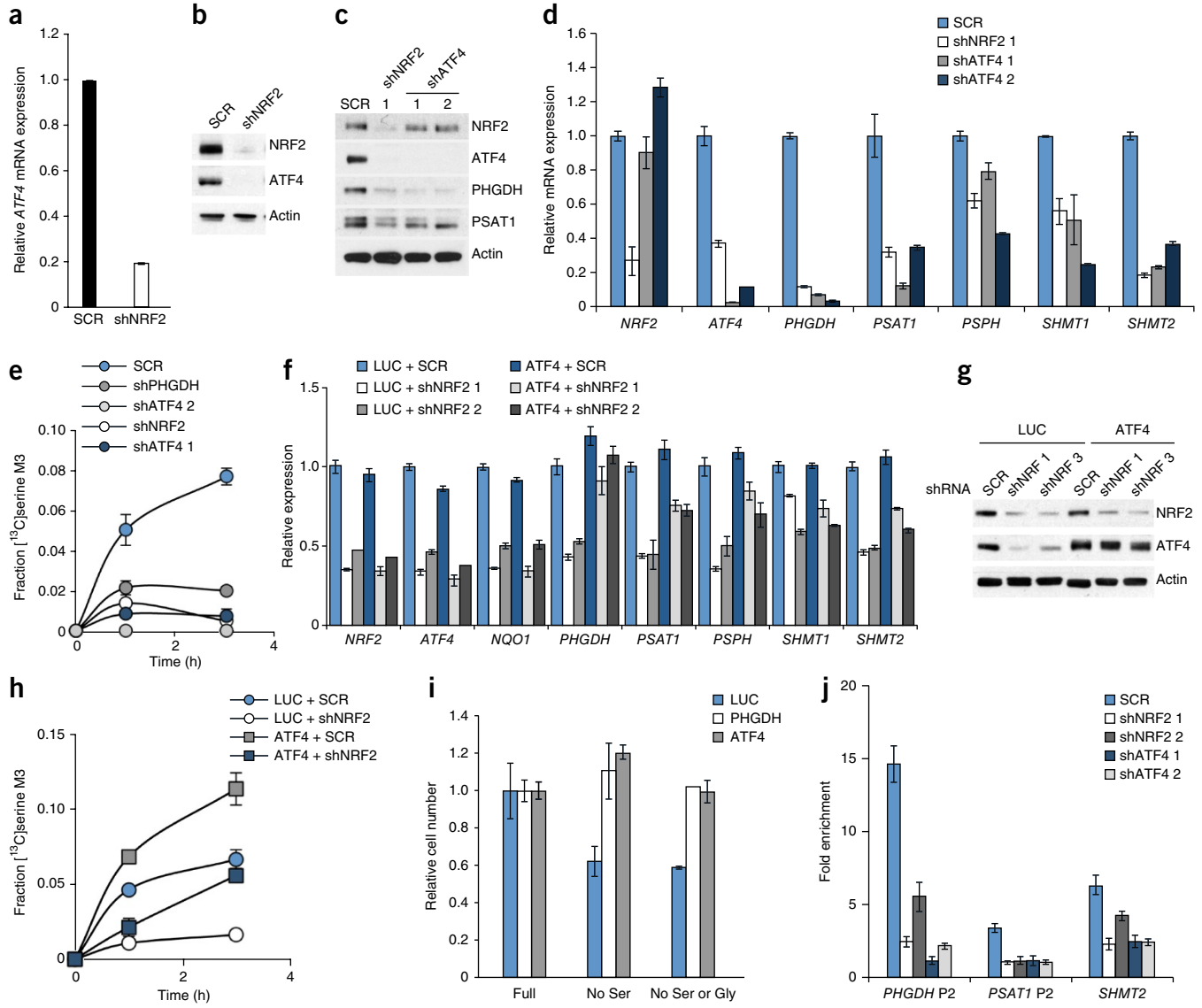

Figure 3 NRF2 regulates the expression of serine/glycine biosynthesis genes through ATF4. (**a**) *ATF4* mRNA expression in A549 cells expressing scrambled shRNA or shNRF2 1. (**b**) Immunoblot analysis of NRF2, ATF4 and actin expression in the cells from **a**. (**c**) Immunoblot analysis of NRF2, ATF4 and serine pathway enzyme expression in lysates from A549 cells expressing scrambled shRNA, shNRF2 1 or *ATF4* shRNA (shATF4) 1 or 2. (**d**) mRNA expression in the cells from **c**. (**e**) *ATF4* knockdown impairs serine biosynthesis. Cell lines were grown in the presence of [U-^{13}C]glucose for the indicated times, metabolites were extracted and fractional ^{13}C labeling of serine was analyzed by LC/MS. (**f**) ATF4 rescues serine biosynthesis enzyme expression following *NRF2* knockdown. A549 cells were infected with lentivirus encoding mouse ATF4 (mATF4) before infection with lentivirus encoding scrambled shRNA or shRNA targeting *NRF2*. (**g**) Immunoblot analysis of NRF2, ATF4 and actin expression in the cells from **f**. (**h**) ATF4 rescues the serine biosynthesis defect in A549 cells with shRNA-mediated knockdown of *NRF2*. Cells were assayed as in **e**. (**i**) ATF4 rescues the growth of H1975 cells in serine-deficient medium. Cells expressing luciferase, PHGDH or ATF4 were grown in the indicated media for 3 d, and cell number was normalized to that for cells grown in full medium. (**j**) Chromatin immunoprecipitation with antibody to ATF4 of the *PHGDH*, *PSAT1* and *SHMT2* promoters. Binding was normalized to that in immunoprecipitation with IgG control. P2 is a binding site (**Supplementary Fig. 16**). Results are the average of three technical (**a**,**d**,**f**,**j**) or biological (**e**,**h**,**i**) replicates. All error bars, s.e.m.

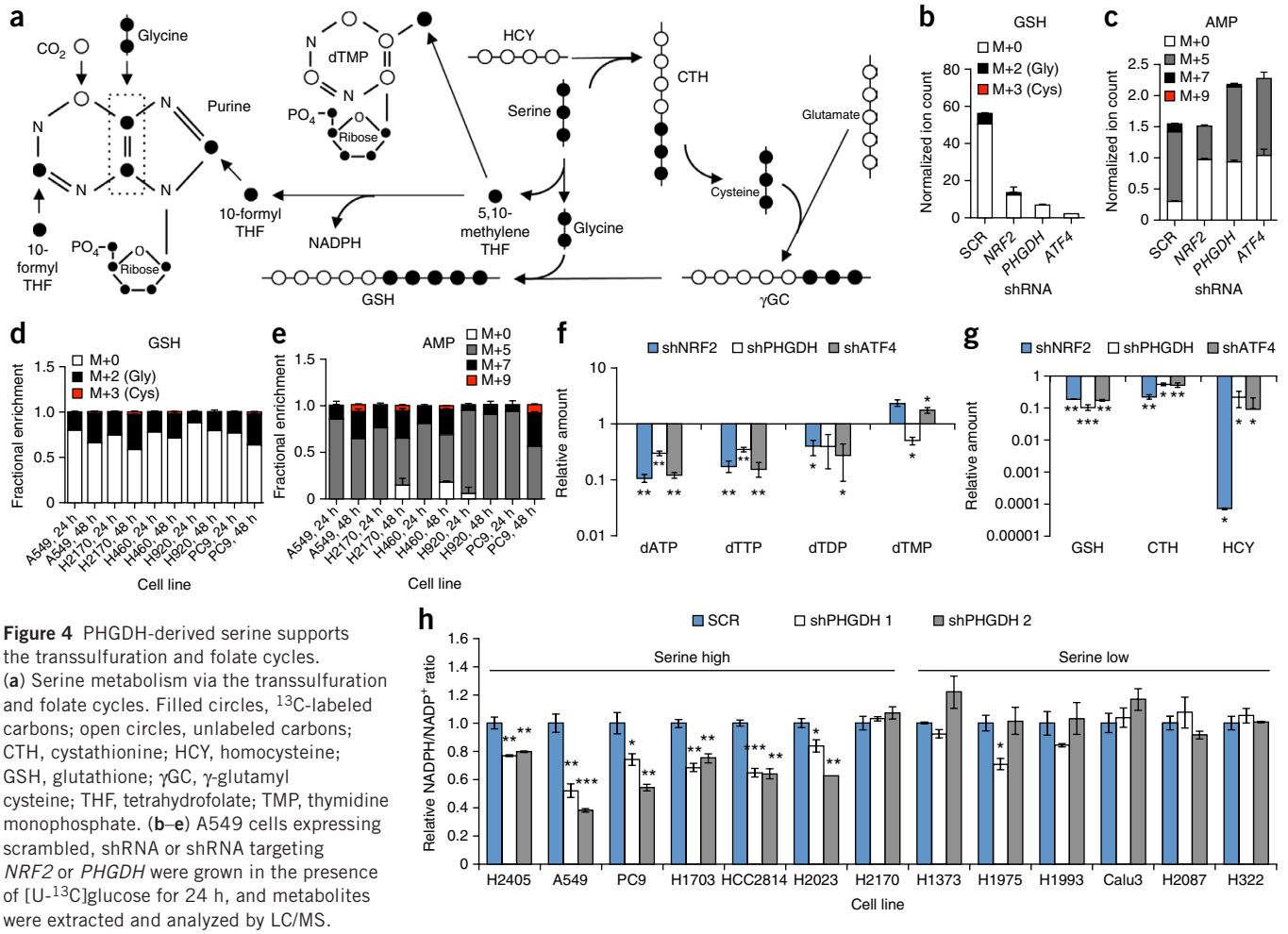

Figure 4 PHGDH-derived serine supports the transsulfuration and folate cycles. (**a**) Serine metabolism via the transsulfuration and folate cycles. Filled circles, ^{13}C-labeled carbons; open circles, unlabeled carbons; CTH, cystathionine; HCY, homocysteine; GSH, glutathione; γGC, γ-glutamyl cysteine; THF, tetrahydrofolate; TMP, thymidine monophosphate. (**b**–**e**) A549 cells expressing scrambled, shRNA or shRNA targeting *NRF2* or *PHGDH* were grown in the presence of [U-^{13}C]glucose for 24 h, and metabolites were extracted and analyzed by LC/MS. (**b**) Analysis of ^{13}C labeling of the glycine component of glutathione. (**c**) Analysis of ^{13}C labeling of the purine metabolite AMP. (**d**,**e**) Analysis of glutathione (**d**) and AMP (**e**) labeling in serine-high cell lines. Cells were labeled with [^{13}C]glucose for 24 or 48 h, as indicated. (**b**,**d**) PHGDH-derived serine is incorporated into glutathione through the generation of glycine and cysteine. M+0, no carbons labeled; M+2, labeling of the glycine moiety; M+3, labeling of the cysteine moiety; M+5, labeling of both glycine and cysteine. M+5 labeling was not observed. (**c**,**e**) M+5 labeling occurs following ribose-5-phosphate labeling via the pentose phosphate pathway. M+7 labeling is the result of ribose labeling plus either glycine or formyl-THF labeling on the purine ring. M+9 labeling is the result of ribose labeling plus glycine and formyl-THF labeling on the purine ring. M+0, no labeled carbons. (**f**,**g**) LC/MS analysis of total metabolite levels in the nucleotide (**f**) and transsulfuration (**g**) pathways. (**h**) NADPH/NADP$^+$ ratios 4 d after *PHGDH* knockdown. Results are the averages of three biological replicates. All error bars, s.e.m.; $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

knockdown reduced the expression and activity of the serine pathway components to an extent similar to that seen with *NRF2* knockdown, whereas it did not affect NRF2 expression (**Fig. 3c**–**e**). Ectopic expression of NRF2 in *NRF2*-depleted cells partially rescued *NQO1*, *ATF4* and serine biosynthesis enzyme expression (**Supplementary Fig. 12c**) and induced expression of these genes in the serine-low cell line H1975 (**Supplementary Fig. 12d**), confirming the regulation of these genes by NRF2. Similarly, ectopic ATF4 expression rescued the effects of *NRF2* silencing on serine biosynthesis enzyme expression (**Fig. 3f**,**g** and **Supplementary Fig. 12b**) and serine labeling from glucose at early time points (**Fig. 3h**), although defects in serine production were observed at later time points (**Supplementary Fig. 15**). Furthermore, ectopic ATF4 rescued the growth of H1975 cells in serine-deficient medium (**Fig. 3i**). We identified the ATF4 binding sites in the *PHGDH*, *PSAT1* and *SHMT2* promoters (**Supplementary Fig. 16**) and found that, although NRF2 itself did not bind to these sites (**Supplementary Fig. 16**), *NRF2* silencing significantly decreased binding of ATF4 (**Fig. 3j**). These results demonstrate that NRF2 regulates serine biosynthesis gene expression through ATF4.

We examined how PHGDH-derived serine contributes to downstream metabolism (**Fig. 4a**). *PHGDH*, *NRF2* or *ATF4* silencing decreased the incorporation of glucose-derived serine into cystathionine (**Supplementary Fig. 17a,b**) and the incorporation of glucose-derived glycine into glutathione (**Fig. 4b** and **Supplementary Fig. 17c**), without loss of cell viability (**Supplementary Fig. 18**). Metabolism of serine to glycine results in the production of one-carbon units via the folate cycle that are used for purine and thymidine synthesis (**Fig. 4a**). We observed a decrease in the PHGDH-derived labeling (M+7) of purines, including inosine monophosphate (IMP), adenosine monophosphate (AMP), adenosine diphosphate (ADP) and inosine, following *PHGDH* or *ATF4* silencing, whereas *NRF2* silencing decreased both the ribose (M+5) and PHGDH-derived (M+7) labeling of these purines (**Fig. 4c** and **Supplementary Fig. 17d**–**f**). Labeling in serine-high cell lines of glutathione and purines was significantly higher at 48 h than at 24 h (**Fig. 4d**,**e** and **Supplementary Fig. 19a**,**b**). Furthermore, the majority of the ^{13}C-labeled serine and glycine had escaped the cell by 24 h of labeling with [^{13}C]glucose owing to equilibration with unlabeled amino acids in the medium, suggesting

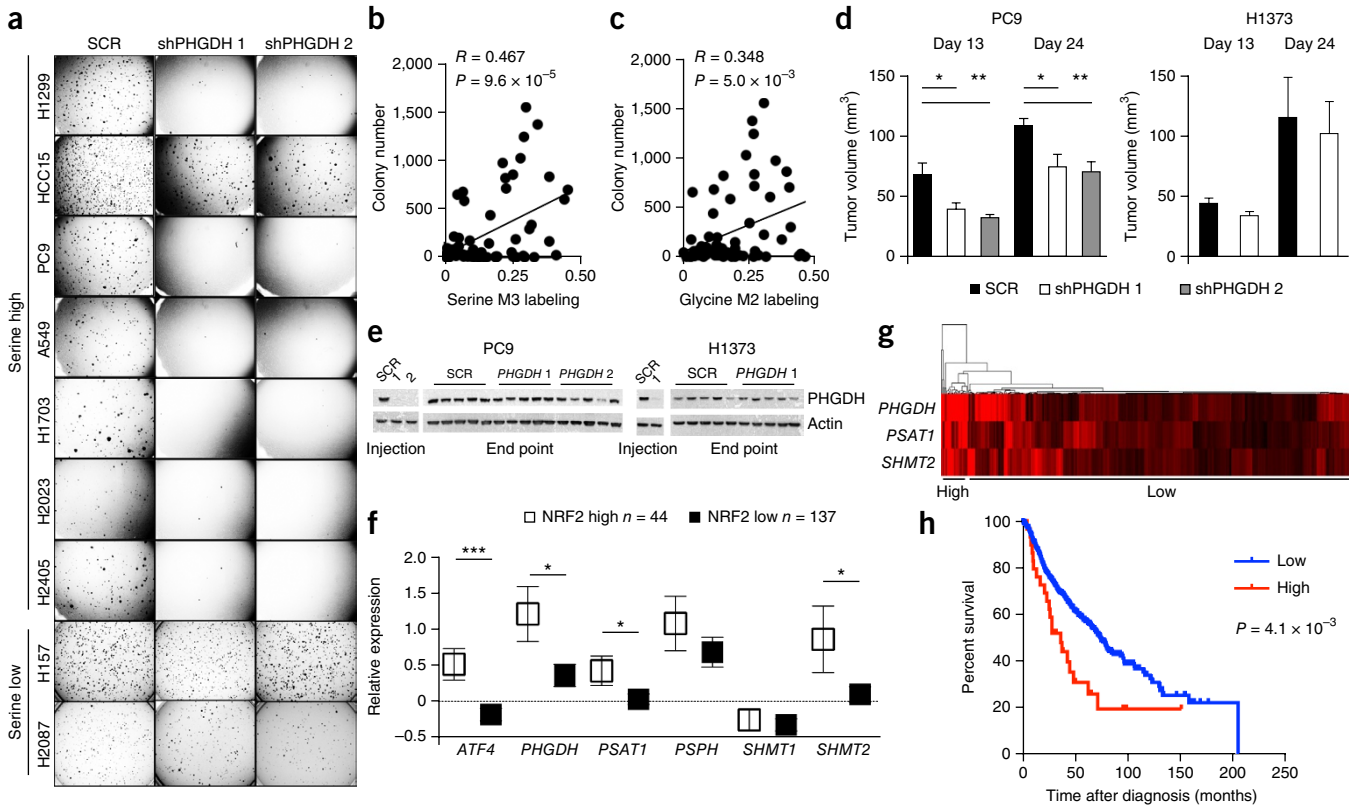

Figure 5 Activation of the serine biosynthesis pathway promotes tumorigenesis in NSCLC. (**a**) *PHGDH* knockdown impairs the soft agar growth of serine-high but not serine-low cell lines. (**b**,**c**) Soft agar growth correlates with serine (**b**) and glycine (**c**) labeling at 24 h. Each cell line was plated at 5,000 cells/well, and the number of colonies was determined after 14 d. *P* values were calculated by Student's *t* distribution with $n-2$ degrees of freedom. (**d**) *PHGDH* knockdown impairs the xenograft growth of a serine-high cell line (PC9; left) but not a serine-low cell line (H1373; right). Results are the averages for five tumors. (**e**) Immunoblot analysis of PHGDH expression of the cell lines and xenografts from **d** upon injection and at the end point. (**f**) Patients with high NRF2 protein expression (*z* score >0.5) demonstrate elevated serine pathway gene expression in samples from The Cancer Genome Atlas (TCGA) lung adenocarcinoma cohort. Boxes represent mean values, and error bars represent s.e.m. (**g**) Gene expression of *PHGDH*, *PSAT1* and *SHMT2* in the Director's Consortium lung adenocarcinoma data set clusters patients into cohorts with high and low expression. (**h**) Kaplan-Meier survival analysis of patients with high ($n = 29$; red) or low ($n = 414$; blue) expression of *PHGDH*, *PSAT1* and *SHMT2* based on the patient clustering from **g**. Median survival was 36 months (high expression) versus 73.2 months (low expression). The *P* value was calculated using the Mantel-Cox test. All error bars, s.e.m.; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

that the fractional labeling at 24 h was underestimating the total contribution of PHGDH to these metabolite pools (**Supplementary Fig. 20** and **Supplementary Note**). In support of this notion, *PHGDH*, *NRF2* or *ATF4* silencing resulted in significant decreases in the total levels of purines and thymidine nucleotides (**Fig. 4f**) as well as glutathione, cystathionine and homocysteine (**Fig. 4g**). In contrast, we did not observe differences in the S-adenosyl methionine (SAM)/S-adenosyl homocysteine (SAH) ratio, which is also modulated by the folate cycle, or in histone or DNA methylation (**Supplementary Fig. 21**). Furthermore, we observed a significant decrease in the nicotinamide adenine dinucleotide phosphate reduced/oxidized ($NADPH/NADP^+$) ratio in serine-high cells following *PHGDH* silencing that was not observed in serine-low cells (**Fig. 4h**). These results demonstrate that the serine biosynthesis pathway supports glutathione and nucleotide production in NSCLC.

We next asked whether this pathway promotes tumorigenesis. *PHGDH* silencing significantly impaired the soft agar growth of serine-high but not serine-low cell lines (**Fig. 5a**). Interestingly,

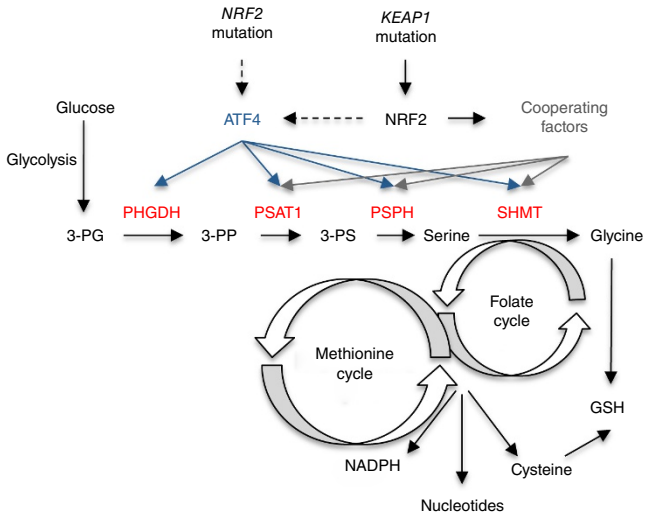

Figure 6 Model of the regulation of serine/glycine biosynthesis by NRF2. An ATF4 transcriptional program, indirectly activated by *NRF2* or *KEAP1* gene mutations, regulates the expression of serine/glycine biosynthesis enzymes. These enzymes produce serine and glycine from the glycolytic intermediate 3-PG and funnel the carbon into glutathione and nucleotides via the folate and transsulfuration cycles. NRF2- and ATF4-regulated enzymes are shown in red. 3-PG, 3-phosphoglycerate; 3-PP, 3-phosphohydroxy pyruvate; 3-PS, 3-phosphoserine; GSH, glutathione.

we observed a significant correlation between serine and glycine labeling and colony number in soft agar (**Fig. 5b**,**c**). Furthermore, *PHGDH* silencing impaired the xenograft growth of the serine-high cell line PC9 (**Fig. 5d**, left) but not the serine-low cell line H1373 (**Fig. 5d**, right). We observed that all tumors re-expressed PHGDH at the end point (**Fig. 5e**), suggesting that PHGDH was required for tumor formation. Next, we examined whether serine pathway gene expression correlated with the overall survival of human patients. Human tumors with high NRF2 protein expression displayed significantly higher expression of *ATF4*, *PHGDH*, *PSAT1* and *SHMT2* mRNAs (**Fig. 5f** and **Supplementary Fig. 22a**). Furthermore, we found that high expression of *PHGDH*, *PSAT1* and *SHMT2* conferred a significantly poorer prognosis (**Fig. 5g**,**h**) and was associated with higher tumor grade (**Supplementary Fig. 22b**–**d**). These results demonstrate that, in human NSCLC, NRF2 regulates the expression of serine biosynthetic enzymes, which correlates with poor prognosis.

We have demonstrated a striking heterogeneity in the activity of the serine biosynthetic pathway in NSCLC. Notably, intracellular amino acid labeling from [^{13}C]glucose is underestimated because of rapid exchange with ^{12}C-containing amino acids from the medium, which is likely mediated by amino acid antiporters (**Supplementary Note**). Caution should be used when interpreting data from ^{13}C-labeling experiments because of these exchange mechanisms.

Heterogeneity in the activity of metabolic or signaling pathways is a common phenomenon across tumors, cell lines and even between cells from the same tumor. Here, by systematically analyzing the serine/glycine biosynthesis pathway in a large, highly annotated panel of NSCLC cell lines, we identified NRF2 as the molecular driver of this pathway. NRF2 is frequently deregulated in NSCLC through somatic mutations that disrupt the NRF2-KEAP1 interaction to constitutively activate NRF2 (refs. 14–18). Cancers with high NRF2 levels are associated with poor prognosis[19,20], resistance to therapy and rapid proliferation[18,21]. Notably, NRF2 ablation in various tumor models results in elevated reactive oxygen species (ROS) and the suppression of tumor growth *in vivo*[21–23]. However, the precise mechanism by which NRF2 promotes tumorigenesis is unclear. Recent studies have shown that, in addition to genes that promote ROS detoxification, NRF2 regulates genes involved in anabolic metabolism[24,25]. Here we demonstrate that NRF2 regulates a serine biosynthesis metabolic program via ATF4 and PHGDH to supply the substrates for glutathione and nucleotide production (**Fig. 6**), with regulation of this program synergizing with regulation of the pentose phosphate pathway to supply ribose for nucleotides[24,25]. NRF2 indirectly regulates ATF4 transcriptional activity via unknown mechanisms. Additionally, although ATF4 completely rescues *PHGDH* expression following *NRF2* depletion, the partial rescue of other serine biosynthesis genes suggests that NRF2 regulates these genes combinatorially through ATF4 and additional factors. Furthermore, these findings suggest that multiple NRF2-regulated pathways coordinately contribute to tumorigenesis. Our work encourages the integration of metabolite tracing on large panels of cancer cell lines with gene expression analysis. This approach is a powerful tool for determining the mechanisms responsible for the differential regulation of metabolic pathways and may identify additional links between the activity of metabolic pathways and genetic alterations in cancers.

URLs. The Cancer Genome Atlas Research Network, http://cancergenome.nih.gov/.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank G. Poulogiannis for bioinformatics advice and H. Abbasi, C. Klimko and M. Yuan for technical support with mass spectrometry experiments. This work was supported by US National Institutes of Health grants P01 CA117969 and R01 GM041890 (L.C.C.), R01 CA157996-01 (R.J.D.), 5R01 CA152301 (Y.X.) and P50 CA70907 (J.D.M., I.I.W., Y.X. and K.E.H.) and by Cancer Prevention Research Institute of Texas (CPRIT) funding to J.D.M., Y.X., I.I.W. and K.E.H. (RP110708 and RP120732) and R.J.D. (RP130272). P.-H.C. was supported by a grant from the Welch Foundation to R.J.D. (I-1733). The mass spectrometry work was partially supported by US National Institutes of Health grants 5P30 CA006516 and 5 P01 CA120964 (J.M.A.). G.M.D. was the Malcolm A.S. Moore Hope Funds for Cancer Research Fellow and is supported by the PanCAN/AACR Pathway to Leadership grant.

AUTHOR CONTRIBUTIONS

G.M.D., R.J.D. and L.C.C. designed the study. G.M.D. and E.M. performed molecular biology experiments. G.M.D., P.-H.C., E.M., J.A.S., Z.H. and J.M.A. performed metabolomics and isotope labeling and analyzed the data. D.W. performed xenograft experiments. H.T. and Y.X. performed bioinformatics analysis. K.E.H., I.I.W. and J.D.M. contributed highly annotated lung cancer cell lines. G.M.D., E.M. and L.C.C. wrote the manuscript. All authors commented on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Vander Heiden, M.G., Cantley, L.C. & Thompson, C.B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–1033 (2009).
2. Locasale, J.W. *et al.* Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. *Nat. Genet.* **43**, 869–874 (2011).
3. Possemato, R. *et al.* Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476**, 346–350 (2011).
4. Mullarky, E., Mattaini, K.R., Vander Heiden, M.G., Cantley, L.C. & Locasale, J.W. *PHGDH* amplification and altered glucose metabolism in human melanoma. *Pigment Cell Melanoma Res.* **24**, 1112–1115 (2011).
5. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
6. Mootha, V.K. *et al.* PGC-1α–responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
7. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
8. Ye, J. *et al.* Pyruvate kinase M2 promotes *de novo* serine synthesis to sustain mTORC1 activity and cell proliferation. *Proc. Natl. Acad. Sci. USA* **109**, 6904–6909 (2012).
9. Miyamoto, N. *et al.* Transcriptional regulation of activating transcription factor 4 under oxidative stress in retinal pigment epithelial ARPE-19/HPV-16 cells. *Invest. Ophthalmol. Vis. Sci.* **52**, 1226–1234 (2011).
10. Afonyushkin, T. *et al.* Oxidized phospholipids regulate expression of ATF4 and VEGF in endothelial cells via NRF2-dependent mechanism: novel point of convergence between electrophilic and unfolded protein stress pathways. *Arterioscler. Thromb. Vasc. Biol.* **30**, 1007–1013 (2010).
11. Ye, P. *et al.* Nrf2- and ATF4-dependent upregulation of xCT modulates the sensitivity of T24 bladder carcinoma cells to proteasome inhibition. *Mol. Cell. Biol.* **34**, 3421–3434 (2014).
12. He, C.H. *et al.* Identification of activating transcription factor 4 (ATF4) as an Nrf2-interacting protein. Implication for heme oxygenase-1 gene regulation. *J. Biol. Chem.* **276**, 20858–20865 (2001).
13. Harding, H.P. *et al.* Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol. Cell* **6**, 1099–1108 (2000).
14. Hayes, J.D. & McMahon, M. *NRF2* and *KEAP1* mutations: permanent activation of an adaptive response in cancer. *Trends Biochem. Sci.* **34**, 176–188 (2009).
15. Kim, Y.R. *et al.* Oncogenic *NRF2* mutations in squamous cell carcinomas of oesophagus and skin. *J. Pathol.* **220**, 446–451 (2010).
16. Konstantinopoulos, P.A. *et al. Keap1* mutations and Nrf2 pathway activation in epithelial ovarian cancer. *Cancer Res.* **71**, 5081–5089 (2011).
17. Seng, S. *et al.* NRP/B mutations impair Nrf2-dependent NQO1 induction in human primary brain tumors. *Oncogene* **28**, 378–389 (2009).
18. Zhang, P. *et al.* Loss of Kelch-like ECH-associated protein 1 function in prostate cancer cells causes chemoresistance and radioresistance and promotes tumor growth. *Mol. Cancer Ther.* **9**, 336–346 (2010).

19. Shibata, T. *et al.* Cancer related mutations in *NRF2* impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy. *Proc. Natl. Acad. Sci. USA* **105**, 13568–13573 (2008).
20. Solis, L.M. *et al.* Nrf2 and Keap1 abnormalities in non–small cell lung carcinoma and association with clinicopathologic features. *Clin. Cancer Res.* **16**, 3743–3753 (2010).
21. Singh, A. *et 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).
22. DeNicola, G.M. *et al.* Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. *Nature* **475**, 106–109 (2011).
23. Ohta, T. *et al.* Loss of Keap1 function activates Nrf2 and provides advantages for lung cancer cell growth. *Cancer Res.* **68**, 1303–1309 (2008).
24. Mitsuishi, Y. *et al.* Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming. *Cancer Cell* **22**, 66–79 (2012).
25. Singh, A. *et al.* Transcription factor NRF2 regulates miR-1 and miR-206 to drive tumorigenesis. *J. Clin. Invest.* **123**, 2921–2934 (2013).

ONLINE METHODS

Animals. Male nude mice (CrTac:NCr-Foxn1nu) were obtained from Taconic and maintained under pathogen-free conditions. Experiments were performed according to Institutional Animal Care and Use Committee (IACUC) guidelines. Mice were injected at 8 weeks of age with 2×10^6 NSCLC cells on each flank. shRNAs were randomized so they were evenly distributed across mice. As the animal study was exploratory, no statistical test was used to determine adequate sample size. No mice were excluded from the analysis. The study did not use blinding.

Cell culture. All NSCLC cell lines used in this study were obtained from the Hamon Cancer Center Collection (University of Texas–Southwestern Medical Center). Cells were maintained in RPMI-1640 (Life Technologies) supplemented with 5% or 10% FCS without antibiotics at 37 °C in a humidified atmosphere containing 5% CO_2 and 95% air. All experiments were performed in media containing serine and glycine except where otherwise noted. All cell lines have been DNA fingerprinted using the PowerPlex 1.2 kit (Promega) and tested for mycoplasma by the e-Myco (Boca Scientific) or MycoAlert (Lonza) kit. Although NCI-H157 is listed in the database of commonly misidentified cell lines, it was originally derived by J.D.M. and fingerprinted before use.

CellTox Green staining. Cells were incubated in RPMI supplemented with 10% FCS containing 1× CellTox Green (Promega) and 5 μM Syto 17 (Life Technologies) for 20 min, washed in PBS and imaged on an EVOS FL cell imaging system (Life Technologies).

Chromatin immunoprecipitation. Cells (5×10^6) were fixed at 37 °C in RPMI with 1% formaldehyde for 10 min, lysed in 1% SDS, 10 mM EDTA and 50 mM Tris-HCl, pH 8.1, plus protease inhibitors and sonicated with a probe tip until DNA was an average of 1 kb in size. Input was saved, and lysate was diluted in immunoprecipitation buffer (1% Triton, 2 mM EDTA, 150 mM NaCl and 20 mM Tris-HCl, pH 8.1) and mixed with beads (Dynal Protein A, Invitrogen) that were prebound overnight with antibodies to Nrf2 (H-300, Santa Cruz Biotechnology), ATF4 (11815, Cell Signaling Technology) and RNA polymerase II (pSer5, ab5131, Abcam) or with rabbit IgG (sc-2027, Santa Cruz Biotechnology). Chromatin was immunoprecipitated overnight, and beads were washed six times with RIPA buffer (50 mM HEPES, pH 7.6, 1 mM EDTA, 0.7% sodium deoxycholate, 1% NP-40 and 0.5 M lithium chloride) and twice with TE buffer. Beads were incubated with solution containing 1% SDS and 0.1 M $NaHCO_3$ for 30 min at room temperature, and cross-links were then reversed for both the input and the immunoprecipitate by heating overnight in a 65 °C water bath. DNA was purified with a QIAquick spin kit (Qiagen), and quantitative PCR was performed in triplicate with Fast SYBR Green Master Mix on a Step One Real-Time PCR system (all from Life Technologies). Primer sequences are listed in **Supplementary Table 4**.

DNA methylation. DNA was extracted from cells in lysis buffer (10 mM Tris, 100 mM NaCl, 10 mM EDTA, 0.5% SDS and 0.4 μg/ml proteinase K, pH 8.0) by heating overnight at 65 °C. DNA was purified by precipitating protein with sodium chloride and precipitated with isopropanol. DNA (2 μg) was denatured in solution with 0.4 M NaOH and 10 mM EDTA at 95 °C for 10 min, and samples were neutralized by the addition of an equal volume of 2 M ammonium acetate (pH 7.0). DNA (200 ng) was spotted on a nylon membrane (GE Healthcare), cross-linked twice on Optimal mode in a Stratalinker oven and blocked in 5% milk for 1 h. The membrane was incubated overnight with antibody to 5-methylcytosine (5meC; 33D3, Epigentek; 1:1,000 dilution), followed by incubation with a horseradish peroxidase (HRP)-conjugated secondary antibody and chemiluminescence visualization. To ensure equal loading, the membrane was stained with 0.02% methylene blue in 0.3 M sodium acetate (pH 5.2). The intensity for 5-meC was normalized to methylene blue staining intensity.

Gas chromatography and mass spectrometry metabolite tracing. All NSCLC cell lines were cultured under identical conditions to identify cell-autonomous differences in glucose metabolism. Cells were seeded into 60-mm culture dishes and grown until 70–80% confluent. Each dish was then rinsed in warm PBS and overlaid with 4 ml of RPMI-1640 containing 5% dialyzed FCS, 4 mM unlabeled glutamine and 10 mM [U-^{13}C]glucose (Cambridge Isotope Laboratories). (Note that RPMI-1640 contains unlabeled serine and glycine.) After 6 or 24 h, the medium was aspirated and cells were rapidly rinsed in cold normal saline solution. Cells were lysed in 0.5 ml of cold 50% methanol, with three rapid cycles of freeze-thawing between liquid nitrogen and a 37 °C water bath. The lysates were cleared of cellular debris by centrifugation, and metabolites in the supernatant were evaporated under blown air or by centrifugation under vacuum. Derivatization, mass spectrometry and mathematical correction for natural-abundance isotopes were performed according to published methods[26,27]. The following fragments were monitored, which represent derivatized species: serine: mass-to-charge ratio (m/z) 306 (M+0) and 309 (M+3); glycine: m/z 276 (M+0) and 278 (M+2). Every cell line was analyzed in biological replicates where $n \geq 3$, and no two replicates were performed on the same day.

Lentivirus production and infection. Lenti-X 293T cells (Clontech) were transfected at 90% confluence with Lipofectamine 2000 (Invitrogen). The packaging plasmids pCMV-dR8.2 dvpr (Addgene, 8455) and pCMV-VSV-G (Addgene, 8454) were used with the vectors listed in **Supplementary Table 5**. Viral supernatants were collected at 48 and 72 h and added to target cells with 8 μg/ml polybrene for 3 h. Cells were selected in 1 μg/ml puromycin where appropriate.

Liquid chromatography and mass spectrometry measurement of total and ^{13}C-labeled metabolites in A549 cells. Cells were plated the day before labeling at 2×10^6 cells/10-cm dish, and the medium was changed to glucose-free RPMI containing 10% dialyzed serum and 10 mM [U-^{13}C]glucose for the indicated time points. Metabolites were extracted in ice-cold 80% methanol and analyzed by targeted LC-MS/MS via selected reaction monitoring (SRM), as described[28]. The following precursor ions were monitored: serine: m/z 106 (M+0) and 109 (M+3); cystathionine: m/z 223 (M+0) and 226 (M+3); glutathione: m/z 308 (M+0); glutathione (glycine): m/z 310 (M+2); homocysteine: m/z 136; IMP: m/z 349 (M+0), 354 (M+5) and 356 (M+7); inosine: m/z 267 (M+0), 272 (M+5) and 274 (M+7); AMP: m/z 348 (M+0), 353 (M+5) and 355 (M+7); dATP: m/z 490; dTTP: m/z 481; dTDP: m/z 401; dTMP: m/z 323. The indicated amino acid for glutathione indicates which constituent amino acid was ^{13}C labeled. Samples were analyzed in triplicate. Data represent median-normalized values.

Luciferase assays. Luciferase assays were performed with the Dual-Glo Luciferase Assay System (Promega) according to the manufacturer's instructions.

NADPH/NADP$^+$ ratios. NADP$^+$ and NADPH concentrations were determined with the NADP/NADPH-Glo Assay kit (Promega) according to the manufacturer's protocol.

NRF2 score calculation. Cell lines were grouped into NRF2-high and NRF2-low groups on the basis of the expression of five classic NRF2 target genes: *NQO1*, *GCLC*, *GLCM*, *SLC7A11* and *AKR1C1*. The expression of each gene was normalized to the median across the cell lines, and the results for these five genes were then averaged together. NRF2 high versus NRF2 low was defined as the top quartile versus the bottom three quartiles. The top 20 overexpressed genes in the NRF2-high cell lines in comparison to the NRF2-low cell lines were used to calculate the NRF2 score: *AKR1C1*, *AKR1C2*, *SPP1*, *ALDH3A1*, *LOC344887*, *AKR1C3*, *OSGIN1*, *PGD*, *CYP4F11*, *AKR1B10*, *KIAA0319*, *SRXN1*, *NR0B1*, *SLC7A11*, *LOC100292680*, *ABCC2*, *CABYR*, *JAKMIP3*, *KYNU* and *PTGR1*. Gene expression values were normalized to the median, and the results for these 20 genes were then averaged together to obtain the NRF2 score. Individual gene expression values and NRF2 scores are listed in **Supplementary Table 3**, and NRF2-high versus NRF2-low clustering is shown in **Supplementary Figure 8**.

Patient samples and survival analysis. The National Cancer Institute Director's Challenge Consortium study (Director's Consortium)[29] and TCGA lung adenocarcinoma data were used in this study to evaluate the prognostic performance of gene signatures. The Director's Consortium data set collected 442 resected lung adenocarcinomas at four US institutions[29], and the TCGA

doi:10.1038/ng.3421

Research Network data include 203 patient samples for which gene expression and survival data are available. Unsupervised cluster analysis was used to group patients on the basis of the expression of *PHGDH*, *PSAT1* and *SHMT2*, using average linkage clustering with the Spearman's rank correlation distance metric. Clustering was performed with Cluster 3.0. Heat-map visualization was performed with JavaTreeView. Kaplan-Meier survival curves were used to determine the survival rate as a function of time, and survival differences were analyzed by a log-rank Mantel-Cox test using GraphPad Prism.

Proliferation assays. Cells were seeded at 500–10,000 cells/well in 96-well plates on day −1 and infected on day 0 with lentivirus. Alternatively, cells were switched into RPMI supplemented with 10% dialyzed FBS containing full amino acids, lacking serine, or lacking serine and glycine. Plates were fixed on the indicated days with 4% paraformaldehyde, stained with crystal violet, washed and dried. Crystal violet was solubilized in 10% acetic acid, and OD_{600} was measured.

Pulse labeling with 4-thiouridine. Cells were labeled in 10-cm dishes at 70% confluence with 200 μM 4-thiouridine (4sU, Sigma-Aldrich) for 60 min as described[30]. RNA was extracted with TRIzol. 4-thiouridine–containing mRNA molecules were biotinylated with biotin-HPDP (EZ-Link Biotin-HPDP, Pierce, 21341) and purified with the μMACS streptavidin kit (Miltenyi Biotec). RNA was eluted in 100 mM DTT, cDNA was synthesized and *ATF4* levels were quantified and normalized to *ACTB* levels.

Reagents. Serine and glycine were purchased from Sigma-Aldrich. RPMI media lacking serine and/or glycine were custom prepared by Life Technologies.

RT-PCR. RNA was isolated using an RNeasy kit (Qiagen). cDNA was synthesized using Superscript VILO Master Mix and analyzed by quantitative PCR using Fast SYBR Green Master Mix on a Step One Real-Time PCR system. Target gene expression was normalized to *ACTB* expression and is shown relative to expression in control samples. Primer sequences are listed in **Supplementary Table 4**.

siRNA transfection. 100,000 cells/well were reverse transfected in 800 μl of growth medium in 12-well dishes. DharmaFECT Duo (2 μl) was combined with 100 pmol of siRNA in a final volume of 200 μl, according to the manufacturer's instructions, which was then added to the cells. Cells were analyzed after 2 d. Dharmacon ON-TARGETplus non-targeting siRNA (D-001810-10) and NFE2L2 (L-003755-00-0005) pools were used.

Soft agar assays. Soft agar assays were performed in triplicate in six-well dishes. A 1-ml base layer of 0.8% agar in RPMI was plated and allowed to solidify, and 5,000 cells/well were then plated in 0.4% agar on top. RPMI (1 ml) was added the following day to each well, and the medium was changed as needed. (Note that RPMI contains serine and glycine.) Soft agar assays were stained with 0.01% crystal violet in 4% paraformaldehyde in PBS and imaged in a ChemiDoc system (Bio-Rad). Colonies were quantified with ImageJ software.

Statistical analysis. Data were analyzed using a two-sided unpaired Student's *t* test, and the Mantel-Cox test was used for survival analyses. For all statistical analyses, GraphPad Prism 6 software was used, and values of $P < 0.05$ were considered statistically significant ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$). The mean ± s.e.m. of at least three independent experiments performed in triplicate is reported. For all experiments, similar variances between groups were observed. Normal distribution of samples was not determined.

Immunoblotting. Protein lysates were prepared using RIPA lysis buffer and separated on 4–12% NuPAGE gels (Invitrogen), transferred onto a nitrocellulose membrane (Millipore) and incubated with the following antibodies: monoclonal antibodies to actin (ab6276) and NRF2 (EP1808Y) (both Abcam) and antibodies to NQO1 (HPA007308) and PHGDH (HPA021241) (both Sigma), PSAT1 (PA5-22124, Pierce) and ATF4 (11815, Cell Signaling Technology). Alternatively, nuclear extracts were prepared as described[31]. Histone extracts were prepared with the Histone Extraction kit according to the manufacturer's instructions (Abcam, ab113476). Histone extracts were probed with antibodies to the following histones: H3K4me3 (9727, Cell Signaling Technology), H3K27me3 (07-449, Millipore) and total histone H3 (4499, Cell Signaling Technology).

26. Cheng, T. *et al.* Pyruvate carboxylase is required for glutamine-independent growth of tumor cells. *Proc. Natl. Acad. Sci. USA* **108**, 8674–8679 (2011).
27. Mullen, A.R. *et al.* Reductive carboxylation supports growth in tumour cells with defective mitochondria. *Nature* **481**, 385–388 (2012).
28. Yuan, M., Breitkopf, S.B., Yang, X. & Asara, J.M. A positive/negative ion-switching, targeted mass spectrometry–based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nat. Protoc.* **7**, 872–881 (2012).
29. Shedden, K. *et al.* Gene expression–based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat. Med.* **14**, 822–827 (2008).
30. Rädle, B. *et al.* Metabolic labeling of newly transcribed RNA for high resolution gene expression profiling of RNA synthesis, processing and decay in cell culture. *J. Vis. Exp.* doi:10.3791/50195 (8 August 2013).
31. Meylan, E. *et al.* Requirement for NF-κB signalling in a mouse model of lung adenocarcinoma. *Nature* **462**, 104–107 (2009).

Erratum: NRF2 regulates serine biosynthesis in non–small cell lung cancer

Gina M DeNicola, Pei-Hsuan Chen, Edouard Mullarky, Jessica A Sudderth, Zeping Hu, David Wu, Hao Tang, Yang Xie, John M Asara, Kenneth E Huffman, Ignacio I Wistuba, John D Minna, Ralph J DeBerardinis & Lewis C Cantley
***Nat. Genet.* 47, 1475–1481 (2015); published online 19 October 2015; corrected after print 15 February 2016**

In the version of this article initially published, the colors of the lines in the key in the top right corner of Figure 5h were incorrect. The line labeled "High" should be red and the line labeled "Low" should be blue. The error has been corrected in the HTML and PDF versions of the article.