

# Pioglitazone-mediated reversal of elevated glucose metabolism in the airway epithelium of mouse lung adenocarcinomas

Donghai Xiong,<sup>1,2</sup> Jing Pan,<sup>1,2</sup> Qi Zhang,<sup>1,2</sup> Eva Szabo,<sup>3</sup> Mark Steven Miller,<sup>4</sup> Ronald A. Lubet,<sup>4</sup> Yian Wang,<sup>1,2</sup> and Ming You<sup>1,2</sup>

<sup>1</sup>Cancer Center and <sup>2</sup>Department of Pharmacology & Toxicology, Medical College of Wisconsin, Milwaukee, Wisconsin, USA. <sup>3</sup>Lung and Upper Aerodigestive Cancer Research Group and <sup>4</sup>Chemopreventive Agent Development Research Group, Division of Cancer Prevention, National Cancer Institute, Rockville, Maryland, USA.

Airway epithelial cells are prone to the damage caused by lung cancer risk factors, such as cigarette smoking. Little is known about surrogate biomarkers in the bronchial airway epithelium that can be used to assess the effect of potential chemoprevention drugs on lung adenocarcinoma formation/progression. Pioglitazone has been suggested as a chemoprevention drug for lung cancer. To study the mechanisms underlying the role of pioglitazone in lung cancer prevention, we performed transcriptome sequencing (RNA-Seq) and found that *Kras* signaling was repressed by pioglitazone treatment in the airway epithelial cells of mice with lung adenocarcinoma (FDR  $q = 9.8E-04$ ). It was also found that glucose metabolic pathways were elevated in the airway epithelium of mice with lung adenocarcinomas and inhibited by pioglitazone treatment (FDR  $q = 0.01$ ). Downregulation of glucose metabolism genes was also observed in lung tumors of mice treated with pioglitazone. The high-risk expression signature of elevated glucose metabolism was associated with poor survival outcome in multiple lung adenocarcinoma patient populations ( $P$  values ranging from  $1.0E-9$  to  $5.5E-5$ ). Our results suggest that the role of pioglitazone in preventing lung adenocarcinoma may depend on inhibiting *Kras* signaling and glucose metabolism, which may serve as biomarkers of agent action in the airway epithelium.

## Introduction

Adenocarcinomas (ADCs) account for about 40% of lung cancer, which is the leading cause of cancer-related deaths around the world (1). Although 85%–90% of lung cancer cases can be attributed to exposure to cigarette smoke, the genetic alterations that lead to lung cancer have still not been identified in a significant portion of lung cancer cases (2–4). Correspondingly, the diagnosis and cure rate remains poor for lung ADCs, although limited progress has been made in areas such as imaging and surgery. Because of the “field-of-cancerization” effect, as a result of exposure to environmental toxicants, the gene expression patterns of airway epithelial cells in lung cancer patients or mouse models have been investigated as a surrogate and a more readily accessible tissue that can be used for monitoring lung cancer risk or therapeutic effect (5–7). One advantage of studying airway epithelial cells rather than tumors per se is that the cell samples from the extended injured area of the airway can be collected in a less invasive manner by bronchial brushing. We previously identified airway gene expression signatures that could be used as biomarkers of lung squamous cell carcinoma (SCC) (7). The activity of key cancer-related pathways was found to be modulated by chemoprevention agents in the airways of mice with lung SCC, which mimicked the corresponding gene expression changes in lung SCC tumors (7). These results suggested that there is positive correlation between airway epithelial cells and lung tumor cells in terms of the oncogenic gene expression patterns, which may facilitate the diagnosis and treatment of lung cancer.

Pioglitazone is a PPAR $\gamma$  agonist used for the clinical treatment of diabetes. Previous work from our (8, 9) and other laboratories (10) has shown that pioglitazone treatment significantly prevented the formation of or inhibited the progression of lung ADCs. However, the underlying mechanisms remained unclear. We have developed a carcinogen-induced lung ADC mouse model that allows for the collection of airway epithelial

**Authorship note:** D. Xiong, J. Pan, and Q. Zhang contributed equally to this work.

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

**Submitted:** March 24, 2017

**Accepted:** May 30, 2017

**Published:** July 6, 2017

**Reference information:**

*JCI Insight.* 2017;2(13):e94220.

<https://doi.org/10.1172/jci.insight.94220>.

insight.94220.

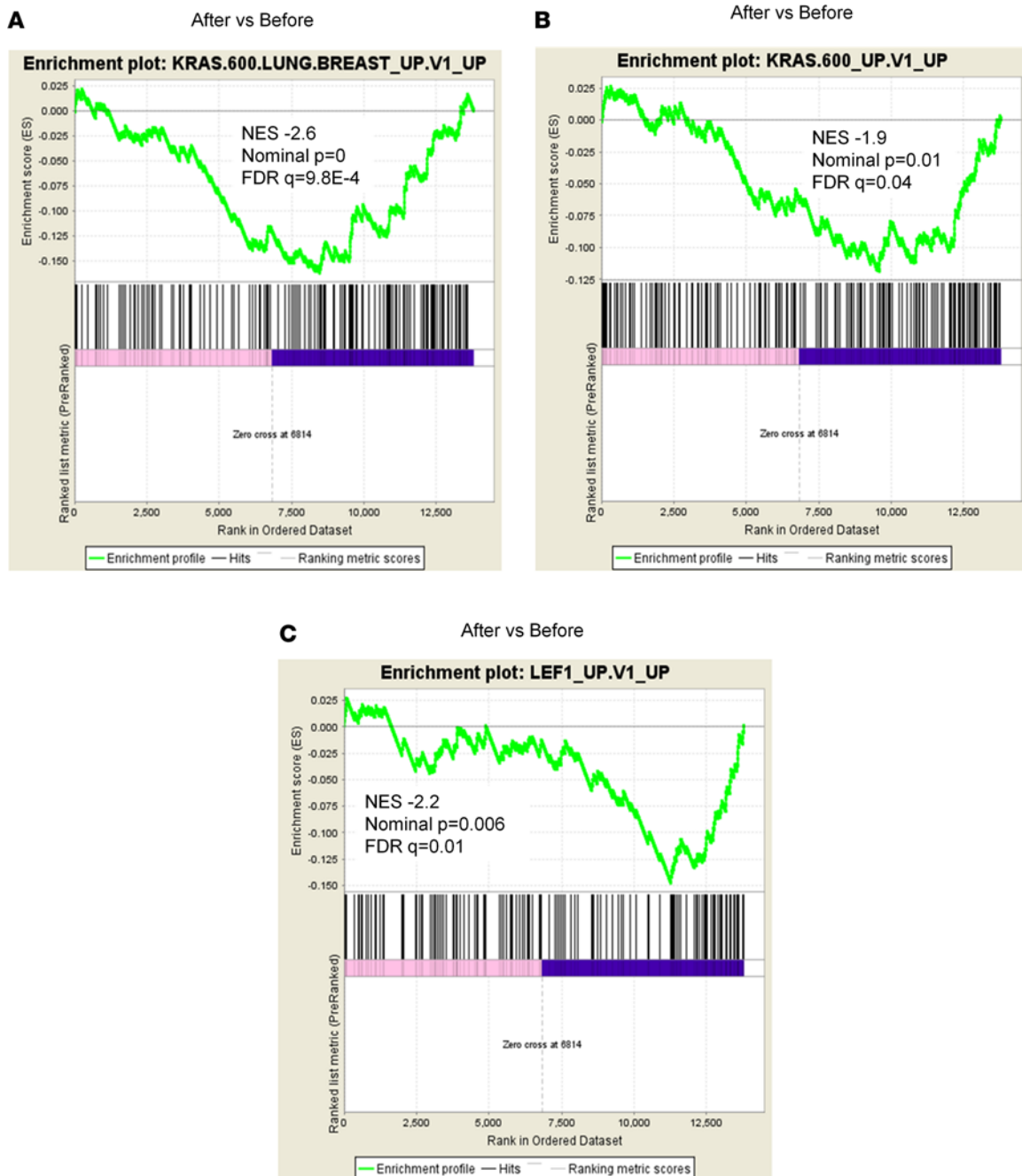
cells before and after pioglitazone treatment. In this study, we performed transcriptome sequencing of airway epithelial cells sampled from mice with lung ADCs before and after pioglitazone treatment, with the aim of identifying deregulated pathways that could be hallmarks of lung ADC tumor. Further, we also determined whether treatment with pioglitazone altered the expression of genes in these pathways in the airway epithelial cells to establish the potential for airway epithelial cells to serve as biomarkers of lung tumor formation and the chemoprevention/therapeutic activity of pioglitazone effects.

## Results

*Kras oncogenic gene sets were repressed by pioglitazone treatment in cytologically normal bronchial airway epithelial cells.* RNA-sequencing (RNA-seq) analysis of carcinogen-induced mouse lung adenoma and ADC tumors revealed significantly elevated oncogenic Kras signaling activity compared with the normal mouse lung tissues (Supplemental Figure 1A; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.94220DS1>) due to the oncogenic Kras mutation Q61R (11). Comparing the airway gene expression patterns between benzo[a]pyrene-treated [B(a)P-treated] mice and normal mice also showed upregulated oncogenic Kras pathway activity (Supplemental Figure 1B), indicating a field cancerization effect in the lung. To find the airway biomarkers that may be associated with pioglitazone's chemoprevention effect, the oncogenic pathway activity changes in the cytologically normal bronchial airway epithelial cells extracted from B(a)P-treated mice were analyzed. The bronchial brushing samples were obtained from 8 mice subjected to 20 weeks of B(a)P exposure followed by 4 weeks of pioglitazone treatment (15 mg/kg by gavage). Total RNAs were extracted from the brush samples of each of the 8 mice before and after pioglitazone treatment. The pairwise comparison of transcriptome profiling was conducted to identify the oncogenic pathways that could be inhibited by pioglitazone. After the preprocessing of raw mRNA-seq data, we further conducted the following analyses. First, we filtered out genes with low expression (median expression values <4 counts per million mapped reads). Then, the statistical R package edgeR (12, 13) was used to detect the differentially expressed genes in the paired bronchial airway samples from the mice before and after pioglitazone treatment. A total of 4,013 genes (FDR < 0.05) was detected as differentially expressed, with 2,020 genes upregulated and 1,993 genes downregulated in the bronchial airway samples of the mice after pioglitazone treatment relative to the same mice before the treatment (Supplemental Tables 1 and 2).

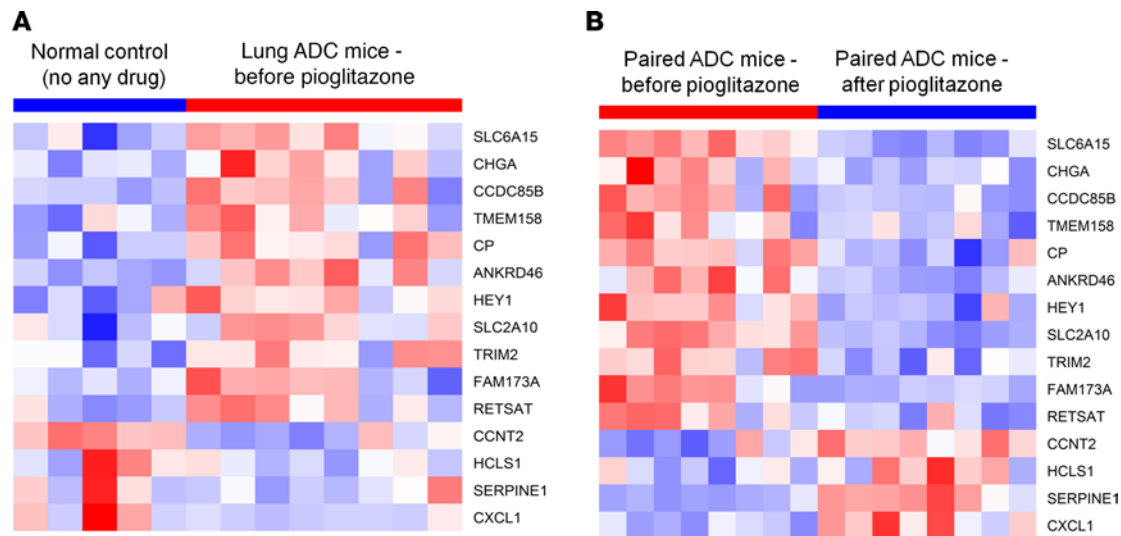
Based on the results from initial differential expression analysis, we next sought to explore the oncogenic pathways that were reversed by pioglitazone treatment in the upper airway cells of mice harboring ADCs. We conducted gene set enrichment analysis (GSEA) and filtered the multiple oncogenic pathways archived in collection C6: oncogenic signatures of the Molecular Signatures Database (MSigDB) (14) according to the criteria of (a) upregulation in the cancer cells that overexpressed the corresponding oncogenes, as described by the associated studies deposited in MSigDB, and (b) downregulation in the airway epithelium after pioglitazone treatment with a normalized enrichment score <-2.0 and FDR  $q < 0.05$  (for multiple testing correction) when testing the signatures with our data. Only KRAS and LEF1 oncogenic signatures met these criteria, suggesting that pioglitazone treatment had significant inhibitory effects on the expression of the oncogenes involved in these signaling pathways (Figure 1). According to the description in MSigDB, the tested Kras gene set includes the signature genes upregulated in lung epithelial cancer cell lines overexpressing the oncogenic form of KRAS. The LEF1 gene set includes genes upregulated in DLD1 cells (colon carcinoma) overexpressing LEF1. However, only the Kras signature gene sets included the signatures related to lung epithelial cells, so they were more relevant to the bronchial airway epithelial cells we studied here and we, thus, solely focused the analysis of Kras signatures. The follow-up heatmap analyses further revealed the KRAS signature genes whose expression was significantly changed in the airway epithelium of lung ADC mice compared with the normal mice (Figure 2A) and such changes were reversed by pioglitazone treatment in the brush airway samples from mice harboring ADC lesions (Figure 2B).

*Effect of Kras signaling downregulation by pioglitazone on glucose metabolism.* GSEA comparing the transcriptome profiling of brush airway samples from mice before pioglitazone treatment [only exposed to B(a)P for 20 weeks] to the samples from the normal control group (no carcinogen or drug treatment) was conducted on the C2: curated gene sets of MSigDB (14). The C2 collection includes the gene sets collected from various sources, such as online pathway databases, publications in PubMed, and knowledge of domain experts. The analysis revealed that the glucose metabolism pathways were activated in the bronchial airway samples of B(a)P-treated mice containing ADCs (Figure 3, carcinogen-treated vs. untreated mice). To figure out the cellular metabolic pathways inhibited by pioglitazone treatment,



**Figure 1.** GSEA identified the significant changes of the activity of oncogenic pathways in the bronchial airways of B(a)P-treated mice after pioglitazone treatment. The two Kras oncogenic signature gene sets (A and B) and one LEF1 signature gene set (C) were significantly downregulated by pioglitazone treatment in mice ( $n = 8$  per group).

we performed a GSEA comparison of RNA profiling between bronchial airway cells with low Kras signaling activity following pioglitazone treatment and untreated samples with high Kras signaling activity (Figure 1, A and B). These studies showed that the activated glucose metabolism pathways can be inhibited as a result of Kras signaling downregulation by pioglitazone treatment (Figure 3, after vs. before pioglitazone treatment). Heatmap analysis showed that expression of the majority of the glucose metabolic genes that were significantly changed was upregulated in the untreated bronchial airway epithelial cells with high Kras signaling activity, whereas treatment with pioglitazone downregulated glucose metabolic genes in conjunction with reduced Kras signaling (Figure 4). These results suggest that downregulation of Kras signaling by pioglitazone treatment may cause the inhibition

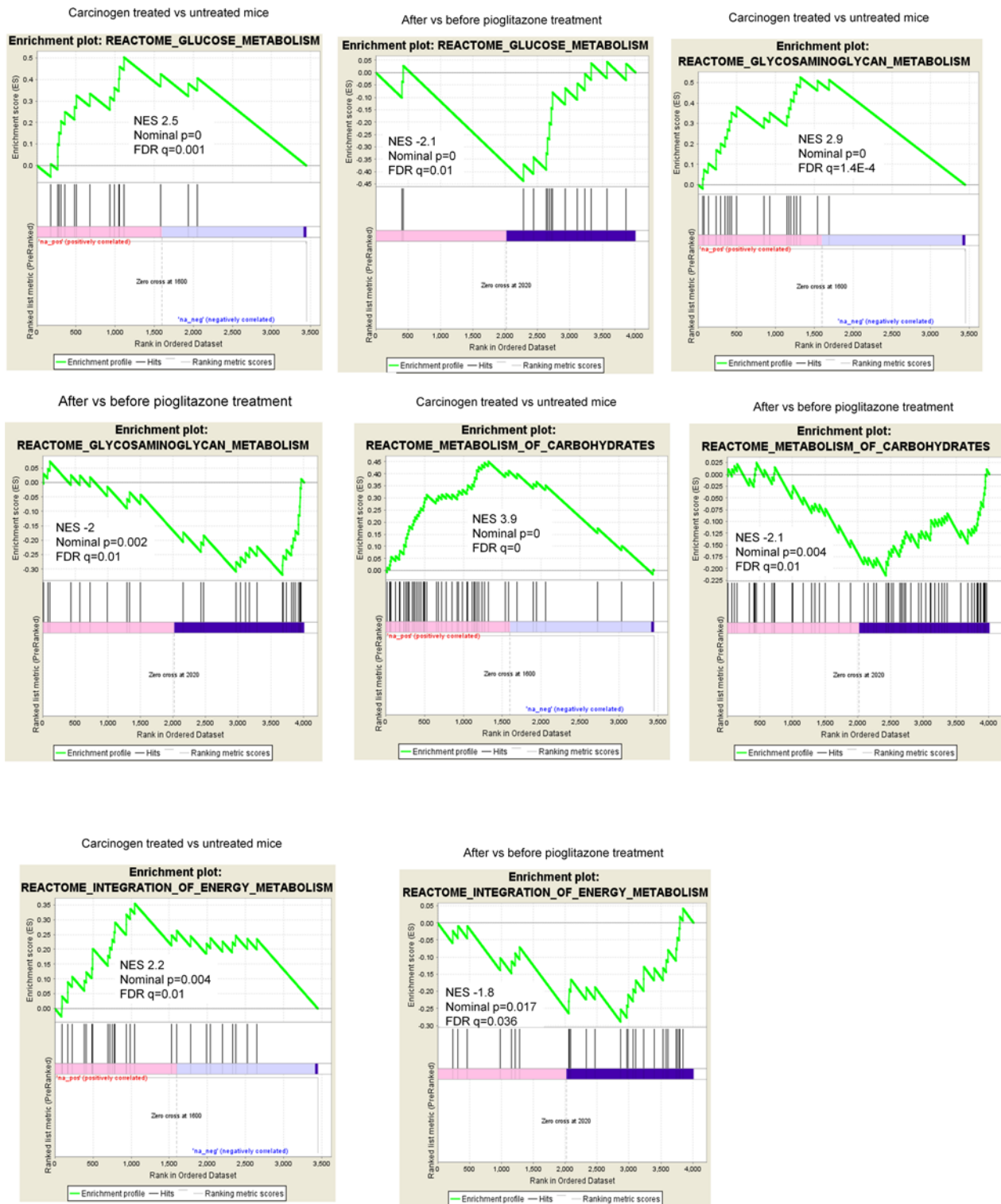


**Figure 2. Heatmap of the Kras oncogenic gene signatures.** Heatmap of the Kras oncogenic gene signatures whose gene expression was (A) significantly altered (mostly upregulated) in the airway epithelium samples of lung ADC mice [B(a)P induced] before pioglitazone treatment in the pioglitazone group ( $n = 8$ ) relative to the normal control group ( $n = 5$ ) and (B) reversed to the normal levels in after pioglitazone treatment airway epithelium samples compared with the before treatment samples obtained from the same mice in the pioglitazone group ( $n = 8$  per group). Mouse IDs are 1 to 8 from left to right for both before and after treatment airway samples.

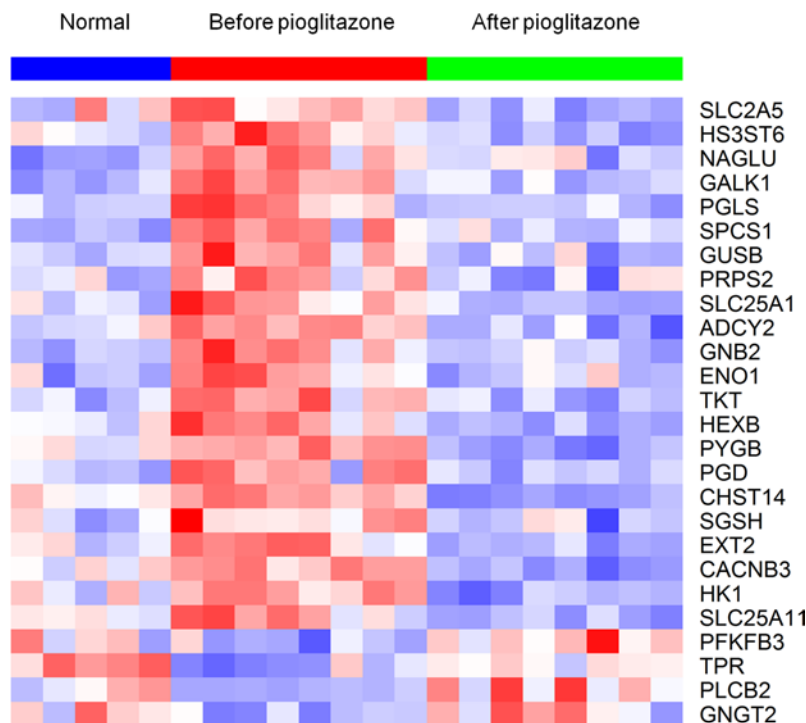
of glucose metabolism, which can be detected in the bronchial airways of mice bearing lung ADC lesions. This finding is consistent with the recent studies showing that the oncogenic Kras tumor cells exhibited increased glucose metabolism in lung and pancreatic cancers (15, 16).

*Expression of key glycolytic enzymes was downregulated by pioglitazone treatment in bronchial airway cells.* One of the most important hallmarks of cancer development is the use of the glycolytic pathway to generate ATP production. The set of differentially expressed glucose metabolism genes was further analyzed using the R package, Pathview (17). It was found that expression of a number of key glycolytic enzymes was significantly downregulated by pioglitazone treatment in the bronchial airways of mice harboring ADC lesions (Figure 5). These dysregulated enzymes include phosphoglucomutase 1 (*PGM1*), fructose biphosphatase 1 (*FBP1*), phosphofructokinase (*PFKM*), aldolase, fructose-bisphosphate C (*ALDOC*), phosphoglycerate mutase 2 (*PGAM2*), enolase 1 (*ENO1*), enolase 3 (*ENO3*), and phosphoenolpyruvate carboxykinase 2 (*PCK2*). *PGM1* catalyzes the transfer of phosphate between the 1 and 6 positions of glucose. *FBP1* catalyzes the hydrolysis of fructose 1,6-bisphosphate to fructose 6-phosphate and inorganic phosphate. *PFKM* catalyzes the phosphorylation of fructose-6-phosphate to fructose-1,6-bisphosphate. *ALDOC* catalyzes the reversible aldol cleavage of fructose-1,6-bisphosphate and fructose 1-phosphate to dihydroxyacetone phosphate and either glyceraldehyde-3-phosphate or glyceraldehyde, respectively. *PGAM2* catalyzes the reversible reaction of 3-phosphoglycerate (3-PGA) to 2-phosphoglycerate (2-PGA) in the glycolytic pathway. *ENO1* and *ENO3* were isoforms of a glycolytic enzyme that catalyzes the reversible conversion of 2-PGA to phosphoenolpyruvate. *PCK2* is a mitochondrial enzyme that catalyzes the conversion of oxaloacetate to phosphoenolpyruvate in the presence of GTP. The schematic plot of their functions in the glycolysis process is shown in Supplemental Figure 2.

*Downregulation of glucose metabolism genes by pioglitazone in bronchial airways and lung ADCs.* To figure out whether the bronchial airway gene expression signatures can reflect the similar gene expression pattern changes in the primary lung ADC tumors, we also studied transcriptome profiling in mouse lung ADC tumors. The BD-Func program to detect oncogenic pathway activity changes (18) was applied to the RNA-seq data of the lung ADC tumor samples from mice with or without pioglitazone treatment. As can be seen in Supplemental Figure 3, the oncogenic KRAS signaling (gene set: KRAS.LUNG\_UP.V1 from MSigDB) was significantly upregulated in the lung ADCs, while the pioglitazone treatment reversed the activity of oncogenic KRAS signaling to a normal level, similar to the bronchial airway expression results (Figure 1). We further studied the regulation of glucose metabolism gene expression in the bronchial airway samples and the tumor samples from mice treated with pioglitazone. In addition to the discovery set of mouse bronchial airway samples, for



**Figure 3.** The GSEA plots of the activity changes of the glucose metabolism REACTOME pathways in the bronchial airways of lung ADC mice induced by B(a)P. Activity changes of the glucose metabolism REACTOME pathways in the bronchial airways of lung ADC mice induced by B(a)P were downregulated in the after pioglitazone treatment brush samples (low Kras signaling) compared with the before pioglitazone treatment brush samples (high Kras signaling) ( $n = 8$  per group).



**Figure 4. Heatmap showing that the expression of the majority of the significantly changed glucose metabolic genes was upregulated in the B(a)P-exposed bronchial airway epithelial cells before pioglitazone treatment and downregulated after pioglitazone treatment.** Normal, the control mice not exposed to any carcinogen or drug ( $n = 5$ ); Before pioglitazone, the B(a)P-exposed mice with high oncogenic Kras signaling activity in their airway samples before pioglitazone treatment ( $n = 8$ ); After pioglitazone, the B(a)P-exposed mice with low oncogenic Kras signaling activity in their airway samples after pioglitazone treatment ( $n = 8$ ).

validation purposes, we utilized an independent set of pioglitazone-treated mice with airway epithelium samples extracted before and after treatment. We also used the data from mouse lung ADC tumors extracted from mice with and without pioglitazone treatment. The data showed that of the 30 initially discovered differentially expressed genes involved in glucose metabolism (Supplemental Figure 4), 14 genes were validated to be significantly downregulated in multiple independent sample sets by pioglitazone, i.e., the discovery airway epithelium data set (Figure 6A), the validation airway epithelium data set (Figure 6B), and the lung ADC data set (Figure 6C). These data suggested that the anti-gly-

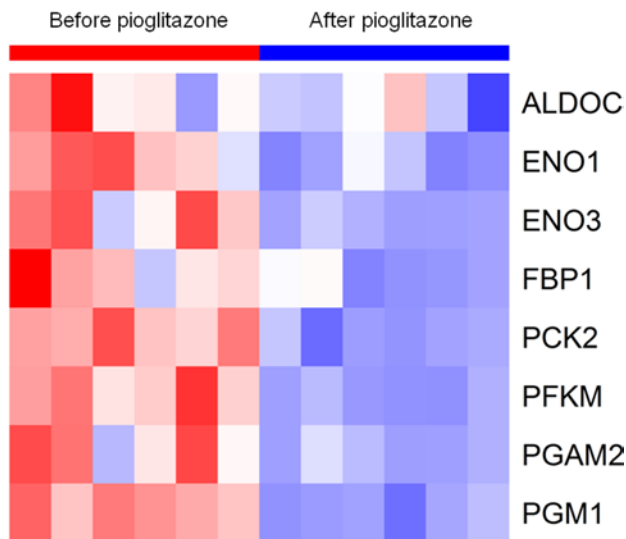
colysis effect of pioglitazone on lung ADC tumors could be recapitulated in the bronchial airway epithelium of lung ADC mice. We also checked public databases to see the effect of pioglitazone on the expression of these 14 glucose metabolism genes in other epithelial cell populations. There is only one publicly available gene expression data set (GSE68852) studying the effects of pioglitazone treatment on the transcriptomic profile of intestinal epithelial cells (Caco-2 cells) (19). We compared our results with those of that study and found that only 2 of the 14 validated lung brushing glucose metabolism genes were downregulated in the Caco-2 cells by pioglitazone treatment; these were NUP107 (fold change =  $-1.4$ ,  $P = 0.01$ ) and VCAN (fold change =  $-1.9$ ,  $P = 3.5 \times 10^{-5}$ ). These results suggested that most of the significant glucose metabolism gene expression changes we identified and validated were specific to the lung.

The expression signature of the pioglitazone-modulated glucose metabolism genes was associated with poor outcome for overall survival of lung ADC in patient populations.

Pioglitazone treatment significantly repressed lung ADC phenotypes, such as tumor size and multiplicity, in mice, as shown previously by our laboratory (8, 9) and others (10). However, the underlying mechanism remained obscure. We found that pioglitazone treatment decreased expression of glucose metabolism genes in the bronchial airway epithelium and tumors of lung ADC mice (Supplemental Figure 4). We sought to explore whether the expression signature of this gene set (called “Piog\_Glu\_LungADC”) was associated with lung ADC survival in the human populations. Four published lung ADC gene expression data sets with lung ADC survival data were analyzed using the program SurvExpress (20) to test whether the pioglitazone treatment-regulated glucose metabolism gene set (“Piog\_Glu\_LungADC”) was associated with overall survival of the patients (21–24). It can be seen from Supplemental Figure 5 that the high-risk gene expression signature of the “Piog\_Glu\_LungADC” gene set was significantly associated with poor overall survival outcome across all of the 4 major representative human lung ADC data sets available from the public data sets described in the Methods. This suggested that decreasing the overall expression of this set of glucose metabolism genes by chemopreventive agents, such as pioglitazone, may benefit the overall survival outcome of lung ADC patients.

## Discussion

Identifying the transcriptome changes in bronchial airway epithelial cells that are able to detect the risk for lung cancer and the therapeutic effects of chemoprevention agents remain appealing subjects of research. Airway gene expression studies in humans may serve as an alternative tissue biomarker for alterations that occur in the peripheral cells of the lung lower airways. For example, PI3K pathway



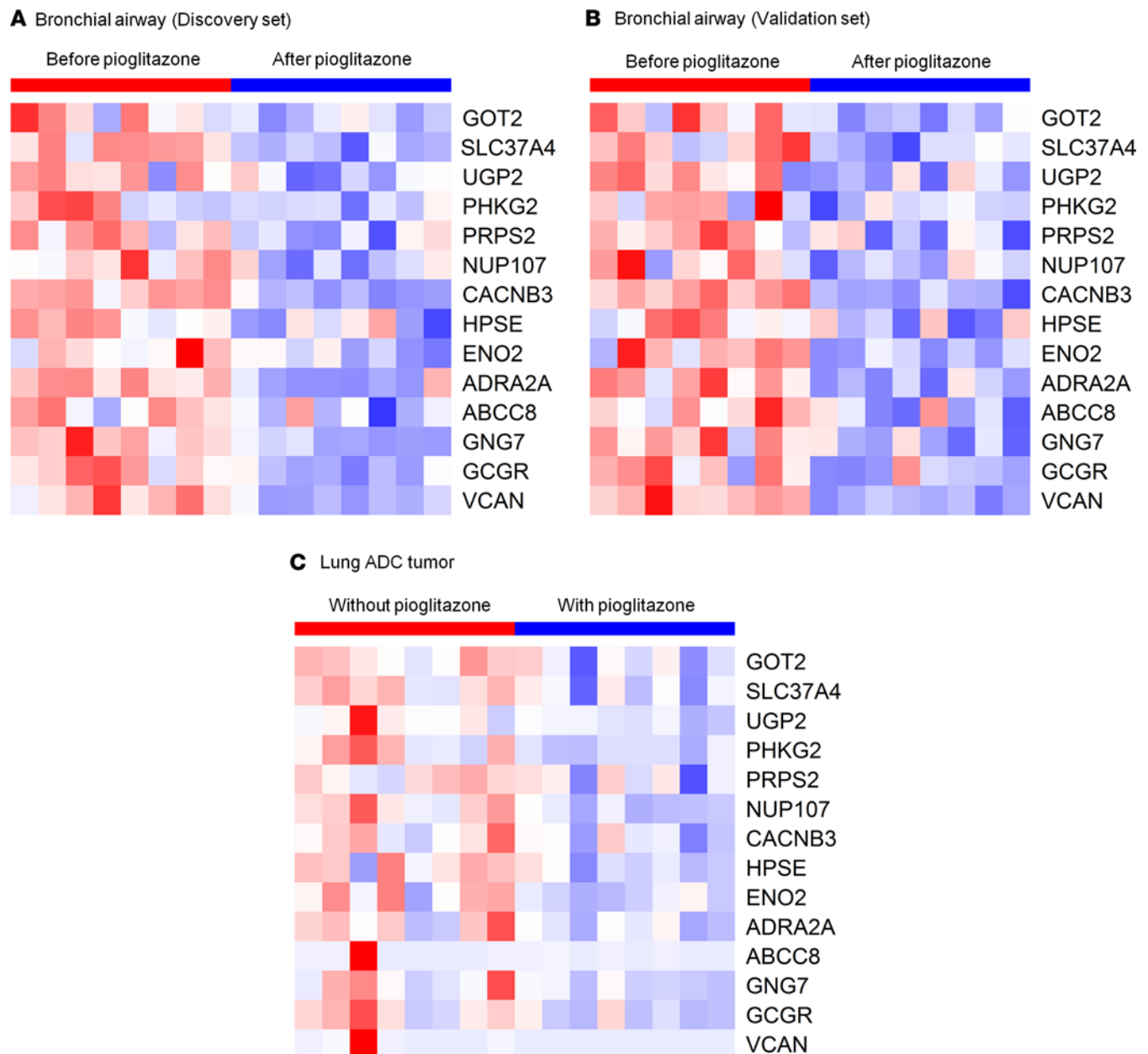
**Figure 5. The downregulation of key glycolytic enzymes by pioglitazone treatment in the bronchial airways of mice harboring ADC lesions.** Heatmap showed the downregulation of key glycolytic enzymes in the bronchial airway of pioglitazone-treated mice ( $n = 6$  per group).

activity has been found to be significantly increased in the bronchial airway of smokers with lung cancer but decreased in the airway of patients who had significant regression of dysplasia after treatment with myo-inositol (5). Using the lung SCC mouse model that was induced by the carcinogen *N*-nitroso-tris-chloroethylurea (NTCU), we demonstrated that the PI3K/AKT and Myc signaling pathways were activated in the normal appearing bronchial airway epithelial cells of NTCU mice and that this can be reversed by treatment with XL-147 (a PI3K inhibitor) and pioglitazone, respectively (7). However, similar studies focusing on the bronchial airway epithelial cells in the lung ADC mice is lacking.

The oncogenic *Kras* signaling pathway is frequently upregulated in patients with non-small-cell lung cancer (NSCLC), especially lung ADC patients (15, 25, 26). We observed similar changes in the tumor and airway brushing samples of mice harboring ADCs (Supplemental Figure 1). Due to the field-of-cancerization effect and the fact that the entire airway tissue, including lungs, was exposed to B(a)P, it was expected that a proportion of *Kras*-mutated cells was present in the airway epithelial cell populations of lung ADC mice. These cells containing *Kras* mutations may account for the elevated *Kras* signaling in the lung ADC mice model induced by B(a)P compared with the normal mice without carcinogen exposure. Interestingly, we found that pioglitazone treatment significantly reversed the oncogenic *Kras* signaling activity in the airway samples of these mice (Figures 1 and 2). Moreover, we found that the airway expressions of the glucose metabolism pathway genes were significantly upregulated in the lung ADC mice compared with the normal mice (Figure 3). Such glucose metabolism activation in the airway was repressed by pioglitazone treatment and such inhibition coupled with the suppression of *Kras* signaling (Figures 3 and 4). We also analyzed the glucose metabolism gene expression changes, comparing the lung tumors collected from the vinyl carbamate-induced (VC-induced) lung ADC mice model and the normal lung tissues collected from the control healthy mice. Among the 77 significantly changed glucose metabolism genes (FDR < 0.05), 70 genes were upregulated and 7 were downregulated (Supplemental Figure 6). The data indicated the overall elevation of glucose metabolism in the lung ADC tumors and were consistent with the field effect observed in the bronchial brushings.

In pancreatic and lung cancers, it has been shown that oncogenic *Kras* serves a vital role in controlling tumor metabolism through stimulation of glucose uptake and enhanced glucose metabolism (15, 16). However, the metabolic effect of oncogenic *Kras* signaling on other cell types and, more importantly, its interaction with chemoprevention drugs is unclear. Our study is the first one to our knowledge showing that, in the bronchial airway epithelial cells of mice with lung ADCs, suppression of oncogenic *Kras* signaling by treatment with pioglitazone appeared to result in inhibition of the expression of genes involved in glucose metabolism. These findings suggested that the oncogenic pathway-related metabolic status and its modulation by chemoprevention drugs in lung ADC tumor cells can be recapitulated by the bronchial airway epithelial cells. Thus, the airway cells obtained from bronchial brushing samples may prove useful for early screening for lung cancer in high-risk populations as well as for monitoring of the effects of chemoprevention drugs.

We identified the upregulation of a large number glycolytic enzymes and proteins involved in glucose metabolism and their repression by pioglitazone treatment in the airways of tumor-bearing mice (Figures 4–6). Our results support a recent proteomics study conducted in the airway epithelial cells of bronchial brushing specimens from lung cancer risk-stratified individuals, which revealed enrichment of carbohydrate metabolic enzymes in high-risk individuals (27). Therefore, data from both humans and mice provide evidence for a dramatic glucose metabolic activity upregulation in the bronchial airway epithelium. Increased glucose metabolism and glycolysis is a hallmark of human cancers and has been known to be a consequence of tumorigenic mutations and epigenetic alterations essential for malignant transformation (28). However, the timing of the glucose metabolism alterations during lung



**Figure 6. Heatmap showing the 14 glucose metabolism genes.** Heatmap showing the 14 glucose metabolism genes (A) downregulated in the discovery set of airways of tumor-bearing mice after pioglitazone treatment; (B) similarly downregulated in the validation set of airways of tumor-bearing mice after pioglitazone treatment; and (C) similarly downregulated in the lung ADCs of mice treated with pioglitazone compared with the tumors of nontreated mice ( $n = 8$  per group for each data set).

ADC development remained unknown. The data we obtained from the bronchial airway epithelial cells and reported herein suggested that the lung cancer–associated metabolic derangements may contribute significantly to the early molecular events involved in lung ADC tumorigenesis, although more and larger independent samples are needed to validate this hypothesis.

This study is timely in that it demonstrated that pioglitazone could be an effective chemoprevention agent for lung ADC by significantly reducing the expression of key enzymes and proteins important for overall glucose metabolism and glycolysis. More importantly, we discovered that there was significant overlap between the results obtained in airway epithelial cells and lung ADC tumors, in terms of the repression of glucose metabolic gene expression by pioglitazone treatment (Figure 6 and Supplemental Figure 4).



The overlapped gene set was named “Piog\_Glu\_LungADC,” and many of the glucose metabolism genes in this gene set have been implicated in cancers. For example, phosphoribosyl pyrophosphate synthetase 1 (PRPS1) was reported as a key enzyme involved in the pentose phosphate pathway, which is of particular importance for glucose metabolism in cancer cells (29). PRPS1 was upregulated significantly at the mRNA and protein levels in colorectal cancer (CRC) tissues, and the high expression of PRPS1 predicted a poor prognosis of CRC patients (29). These findings were similar to those we observed in this study of murine lung ADCs. Glutamate oxaloacetate transaminase 2 (GOT2) encodes a key enzyme involved in the glutamine-dependent pathway of cytosolic nicotinamide adenine dinucleotide phosphate (NADPH) production, which maintains redox homeostasis and supports cellular proliferation; this gene was highly upregulated in pancreatic ductal ADCs (30). GOT2 acetylation promotes pancreatic cell proliferation and tumor growth through stimulating NADPH production to suppress ROS and to protect cancer cells from oxidative damage (31). Solute carrier family 37 member 4 (SLC37A4) regulates glucose-6-phosphate (G6P) transport from the cytoplasm to the lumen of the endoplasmic reticulum in order to maintain glucose homeostasis and is a prognostic biomarker for ovarian cancer (32). UDP-glucose pyrophosphorylase 2 (UDP-UGP2) encodes the enzyme important for glucose synthesis, and its expression has been found to be associated with poor outcome of a number of cancers (33, 34). Dipeptidyl peptidase 4 (DPP4) functions as an ectopeptidase that can inactivate incretins, catalyze the cleavage of chemokines, promote cell migration, and activate lymphocytes; its expression is linked to the carcinogenesis of many malignant tumors (35). Calcium channel  $\beta$  3 (CACNB3) was one member gene of a 4-gene prognostic model exhibiting potential clinical utility for risk stratification of stage I to stage III NSCLC patients (36). The hyaluronan-mediated motility receptor (HMMR) encodes a cell surface oncogenic protein that is widely upregulated in human cancers and plays a role in promoting cell motility and invasion. HMMR was detectable in the majority of lung ADC samples and was significantly elevated in patient plasma compared with controls and thus was considered as a biomarker for circulating tumor cells (37). The glucagon receptor is important in controlling blood glucose levels and was reported as a prognostic biomarker for lung SCC (38). Glucose transporter 3 (GLUT3, also known as SLC2A3) is responsible for glucose transport and is one of the key players in the glycolysis and pentose phosphate pathways. Its overexpression was significantly correlated with lung ADCs (39) and is an indicator of poor prognosis in NSCLC (40).

We further tested for the association of the “Piog\_Glu\_LungADC” gene set with the survival phenotypes in lung cancer patient populations. Interestingly, although this gene expression signature was developed from mouse lung ADCs, it can be readily applied to human tumor samples. This gene set identified the high-risk group and was significantly associated with poor overall survival across all four large human lung ADC data sets tested in this study (Supplemental Figure 5). This leads us to hypothesize that pioglitazone treatment may be able to improve the overall survival in lung ADC patients through downregulation of this set of glucose metabolism genes, a hypothesis that warrants formal testing.

Taken together, this study supported the potential to detect early events associated with lung ADC tumorigenesis in airway epithelium and important molecular responses to interventions in lung cancer. It also suggested that the role of pioglitazone in preventing lung ADCs may depend on inhibiting Kras signaling and glucose metabolism that can be detected in the airway epithelial cells. Finally, a glucose metabolism-related airway gene expression signature was developed, which can be modulated by pioglitazone treatment and can serve as a prognostic biomarker for overall survival of lung ADC patients.

## Methods

*Bronchial brushing and lung ADC sample collection.* B(a)P (99% pure) was purchased from Sigma-Aldrich and was prepared in tricapylin immediately before use in animal bioassays. Female A/J mice at 6 weeks of age were obtained from Jackson Laboratories for the bronchial brushing experiment. Three groups of A/J mice were utilized to determine the effects of pioglitazone on airway epithelial cells. The negative control group consisted of A/J mice that were not subjected to any treatment (“normal” group consisted of 5 mice). In the two remaining groups, i.e., the B(a)P and pioglitazone treatment groups, mice were treated with a single i.p. dose of B(a)P of 100 mg/kg body weight (8 mice per group) at 7 weeks of age to induce lung ADCs. Twenty-two weeks after B(a)P injection, the B(a)P group was left untreated (only gavaged with the vehicle control [corn oil]) for an additional 4 weeks, while the treatment group was administered 15 mg/kg pioglitazone by oral gavage for 4 additional weeks. Pioglitazone was dissolved in corn oil before use. For sample collection, bronchial brushing was performed once for the normal group and twice for the treatment group. Specifically,

for the mice in the treatment group, the brushing before pioglitazone treatment was done at 20 weeks after the carcinogen-B(a)P induction. After 2 weeks of recovery time, the pioglitazone treatment was started at the time point of 22 weeks after B(a)P induction and then the brushing after pioglitazone treatment was done at 26 weeks after B(a)P induction. The mice were treated daily with pioglitazone for 5 days/week for a total of 4 weeks. The procedures for brushing techniques to obtain mouse airway epithelial cell samples are described in detail in Supplemental Figure 7. Particularly, based on the H&E staining, over 95% of brushed cells were confirmed to be bronchial epithelial cells. In addition, most of the harvested cells were shown to have intact epithelial cell membrane and morphology (Supplemental Figure 7B). Therefore, it is reasonable to assume that brushings of exfoliated cells can adequately recapitulate lining epithelial cells in terms of gene expression. For validation purposes, we utilized an independent set of pioglitazone-treated mice with airway epithelium samples extracted before and after treatment (8 mice with 16 airway epithelium samples obtained before and after pioglitazone treatment on the same mice). The treatment procedures were the same between the validation and discovery sample sets. A strength of this approach is that the murine airway samples were collected using the same techniques that can be used in clinical trials, making this a powerful technique for validating agents in animal models before moving to clinical trials. In addition, we set up a different cohort consisting of lung ADC mice and normal control mice and collected primary lung tumor samples following euthanasia from ADC-bearing mice, either untreated or treated with pioglitazone, as well as normal lung tissues from control mice (each group had 8 mice). In this cohort, the lung ADCs in A/J mice was induced by VC purchased from Toronto Research Chemicals Inc. Original A/J mice were obtained from the Jackson Laboratory. Except for the normal control group, the mice of the VC and pioglitazone treatment groups were administered two doses of VC at 7 weeks of age by i.p. injection once a week for 2 consecutive weeks (0.35 mg per injection in 0.2-cc sterile saline without adjustment of pH). The treatment group was administered 15 mg/kg pioglitazone by oral gavage for 4 weeks, starting at 22 weeks after VC induction. Timing of the experiments and sample collections in different groups of mice are shown in Supplemental Figure 8.

*RNA processing and RNA-seq experimentation.* Total RNA samples were extracted from the airway brush samples or lung tissue samples using a Qiagen RNeasy Mini Kit according to the manufacturer's instructions. The quality of the total RNA samples obtained was high, with RNA integrity number values in the range of 9–10. We used the TruSeq RNA Library Preparation Kit v2 to construct the RNA-seq libraries. The sequencing of these RNA-seq library samples was performed in the Medical College of Wisconsin Human and Molecular Genetics Center Sequencing Core using the HiSeq 2500 platforms (Illumina). The reads generated were single end and 50 nucleotides in length. The qualities of the RNA-seq and miRNA-seq reads were analyzed using the FastQC program (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). The coverages ranged from 15 million to 32 million reads per RNA-seq sample. The quality scores of >95.3% of all the bases of each sample are >30, averaging around 40, greatly exceeding the threshold of 20.

*RNA-seq read alignment and differential expression analysis.* The preprocessed sample RNA-seq reads were aligned to the mm9 mouse genome (UCSC version, July 2007) using Bowtie-TopHat (version 2.0.4, segment length 29-nt, 1 mismatch in segment permitted for maximum sensitivity, coverage search performed; refs. 41, 42). Read counts were obtained using HTseq (43). Batch effects were adjusted using the R package RUVSeq (44). Data normalization and differential expression analysis were performed using the statistical algorithms implemented in the statistical R package edgeR (12, 13). FDR-corrected *P* values of less than 0.05 were used as criteria for significantly regulated genes. We used FastQC (45) to check the quality of raw sequences after the miRNA-seq based on the HiSeq 2500 platform. Cutadapt (46) was used to remove adapter and unwanted sequences from raw data of the RNA-seq reads. Then, FastQC was used again to recheck the quality of the preprocessed reads.

*Oncogenic pathway analysis and activation analysis.* The list of differentially expressed genes in the bronchial airway samples was analyzed for pathway enrichment using GSEA (14). GSEA analysis was performed for the preranked differentially expressed genes using the GseaPreranked option. In GSEA, 1,000 permutations were used to calculate significance. A gene set was considered to be significantly enriched in one of the two groups when the FDR was lower than 0.05 for the corresponding gene set. The BD-Func program to detect oncogenic pathway activity changes (18) was applied to the RNA-seq data of the lung ADC tumor samples from mice treated or untreated by pioglitazone. The set of differentially expressed glucose metabolism genes was further analyzed and visualized by using the R package Pathview (17). The software SurvExpress (20) was used to test whether the pioglitazone-modulated gene set ("Piog\_Glu\_LungADC") was associated with overall survival of the human lung cancer

patients. These human patient data sets were obtained from the following sources: (a) The Cancer Genome Atlas LUAD data set, with 475 lung ADC samples from GDC Data Portal (<https://gdc-portal.nci.nih.gov/>); (b) Director's Challenge Consortium NCI Lung, with 443 lung ADC samples (21); (c) Okayama Kohno Lung GSE31210, with 226 lung ADC samples (22, 23); and (d) Chitale Lung data set, with 185 lung ADCs (24). Heatmaps were plotted by using the heatmap.3 function implemented in R 3.1.2 programming environment (<https://www.r-project.org/>).

**Data availability.** The RNA-seq data from bronchial airway samples from mice before and after pioglitazone treatment in both the discovery and validation cohorts have been deposited in the Sequence Read Archive under accession SRP106671.

**Statistics.** All data are expressed as mean  $\pm$  SEM. For RNA-seq expression analysis, batch effects were adjusted using the R package RUVSeq (44). Data normalization and differential expression analysis were performed using the statistical algorithms implemented in the statistical R package edgeR (12, 13). FDR-corrected *P* values of less than 0.05 were used as criteria for significantly regulated genes. For Supplemental Figure 5, Kaplan-Meier survival analyses were implemented to estimate the survival functions after the samples were classified into two risk groups according to their risk scores. Differences of the survival risk between the two risk groups were assessed using the Mantel-Haenszel log-rank test.

**Study approval.** The study protocols involving mice were approved by the Animal Care and Use Committees at the Medical College of Wisconsin and conducted in accordance with the NIH guidelines for care and use of animals.

## Author contributions

DX completed RNA-seq-related experiments, data analyses, interpretation of data, and manuscript writing. JP and QZ did the animal work and bronchial airway sample collection. ES, MSM, RAL, and YW contributed to the writing and editing of the manuscript. MY designed the study, supervised the project, reviewed the data, and edited the manuscript.

## Acknowledgments

We thank the outside reviewers for manuscript suggestions and revisions. This work was supported by NIH grants N01CN201200015 and R01CA134682.

Address correspondence to: Ming You or Yian Wang, National Cancer Institute, 9609 Medical Center Drive, Rockville, Maryland 20850, USA. Phone: 414.805.8228; Email: [myou@mcw.edu](mailto:myou@mcw.edu) (M. You). Phone: 414.955.8364; Email: [yiwang@mcw.edu](mailto:yiwang@mcw.edu) (Y. Wang).

- Ridge CA, McErlean AM, Ginsberg MS. Epidemiology of lung cancer. *Semin Intervent Radiol.* 2013;30(2):93–98.
- Sy SM, et al. Distinct patterns of genetic alterations in adenocarcinoma and squamous cell carcinoma of the lung. *Eur J Cancer.* 2004;40(7):1082–1094.
- Yakut T, et al. Assessment of molecular events in squamous and non-squamous cell lung carcinoma. *Lung Cancer.* 2006;54(3):293–301.
- Zöchbauer-Müller S, Minna JD. The biology of lung cancer including potential clinical applications. *Chest Surg Clin N Am.* 2000;10(4):691–708.
- Gustafson AM, et al. Airway PI3K pathway activation is an early and reversible event in lung cancer development. *Sci Transl Med.* 2010;2(26):26ra25.
- Spira A, et al. Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat Med.* 2007;13(3):361–366.
- Xiong D, et al. Bronchial airway gene expression signatures in mouse lung squamous cell carcinoma and their modulation by cancer chemopreventive agents. *Oncotarget.* 2017;8(12):18885–18900.
- Wang Y, et al. Chemopreventive effects of pioglitazone on chemically induced lung carcinogenesis in mice. *Mol Cancer Ther.* 2010;9(11):3074–3082.
- Fu H, et al. Chemoprevention of lung carcinogenesis by the combination of aerosolized budesonide and oral pioglitazone in A/J mice. *Mol Carcinog.* 2011;50(12):913–921.
- Seabloom DE, et al. Safety and preclinical efficacy of aerosol pioglitazone on lung adenoma prevention in A/J mice. *Cancer Prev Res (Phila).* 2017;10(2):124–132.
- Westcott PM, et al. The mutational landscapes of genetic and chemical models of Kras-driven lung cancer. *Nature.* 2015;517(7535):489–492.
- Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.
- Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene

- expression data. *Bioinformatics*. 2010;26(1):139–140.
14. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–15550.
15. Kerr EM, Gaude E, Turrell FK, Frezza C, Martins CP. Mutant Kras copy number defines metabolic reprogramming and therapeutic susceptibilities. *Nature*. 2016;531(7592):110–113.
16. Ying H, et al. Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell*. 2012;149(3):656–670.
17. Luo W, Brouwer C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics*. 2013;29(14):1830–1831.
18. Warden CD, Kanaya N, Chen S, Yuan YC. BD-Func: a streamlined algorithm for predicting activation and inhibition of pathways. *PeerJ*. 2013;1:e159.
19. Choteau L, et al. Role of mannose-binding lectin in intestinal homeostasis and fungal elimination. *Mucosal Immunol*. 2016;9(3):767–776.
20. Aguirre-Gamboa R, et al. SurvExpress: an online biomarker validation tool and database for cancer gene expression data using survival analysis. *PLoS One*. 2013;8(9):e74250.
21. Director's Challenge Consortium for the Molecular Classification of Lung Adenocarcinoma, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med*. 2008;14(8):822–827.
22. Okayama H, et al. Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. *Cancer Res*. 2012;72(1):100–111.
23. Yamauchi M, et al. Epidermal growth factor receptor tyrosine kinase defines critical prognostic genes of stage I lung adenocarcinoma. *PLoS One*. 2012;7(9):e43923.
24. Chitale D, et al. An integrated genomic analysis of lung cancer reveals loss of DUSP4 in EGFR-mutant tumors. *Oncogene*. 2009;28(31):2773–2783.
25. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511(7511):543–550.
26. Ding L, et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature*. 2008;455(7216):1069–1075.
27. Rahman SM, et al. The airway epithelium undergoes metabolic reprogramming in individuals at high risk for lung cancer. *JCI Insight*. 2016;1(19):e88814.
28. DeBerardinis RJ, Thompson CB. Cellular metabolism and disease: what do metabolic outliers teach us? *Cell*. 2012;148(6):1132–1144.
29. Qiu Z, et al. MicroRNA-124 reduces the pentose phosphate pathway and proliferation by targeting PRPS1 and RPIA mRNAs in human colorectal cancer cells. *Gastroenterology*. 2015;149(6):1587–1598.e11.
30. Chakrabarti G, et al. Targeting glutamine metabolism sensitizes pancreatic cancer to PARP-driven metabolic catastrophe induced by  $\beta$ -lapachone. *Cancer Metab*. 2015;3:12.
31. Yang H, et al. SIRT3-dependent GOT2 acetylation status affects the malate-aspartate NADH shuttle activity and pancreatic tumor growth. *EMBO J*. 2015;34(8):1110–1125.
32. Willis S, et al. Single gene prognostic biomarkers in ovarian cancer: a meta-analysis. *PLoS One*. 2016;11(2):e0149183.
33. Wang Q, et al. SHP2 and UGP2 are biomarkers for progression and poor prognosis of gallbladder cancer. *Cancer Invest*. 2016;34(6):255–264.
34. Tan GS, et al. Novel proteomic biomarker panel for prediction of aggressive metastatic hepatocellular carcinoma relapse in surgically resectable patients. *J Proteome Res*. 2014;13(11):4833–4846.
35. Liang PI, et al. DPP4/CD26 overexpression in urothelial carcinoma confers an independent prognostic impact and correlates with intrinsic biological aggressiveness. *Oncotarget*. 2017;8(2):2995–3008.
36. Mitra R, et al. Prediction of postoperative recurrence-free survival in non-small cell lung cancer by using an internationally validated gene expression model. *Clin Cancer Res*. 2011;17(9):2934–2946.
37. Man Y, et al. Newly identified biomarkers for detecting circulating tumor cells in lung adenocarcinoma. *Tohoku J Exp Med*. 2014;234(1):29–40.
38. Li J, Wang J, Chen Y, Yang L, Chen S. A prognostic 4-gene expression signature for squamous cell lung carcinoma. *J Cell Physiol*. 2017; 9999:1–12.
39. Choi WH, Yoo JeR, O JH, Kim TJ, Lee KY, Kim YK. Is the Glut expression related to FDG uptake in PET/CT of non-small cell lung cancer patients? *Technol Health Care*. 2015;23 Suppl 2:S311–S318.
40. Younes M, Brown RW, Stephenson M, Gondo M, Cagle PT. Overexpression of Glut1 and Glut3 in stage I nonsmall cell lung carcinoma is associated with poor survival. *Cancer*. 1997;80(6):1046–1051.
41. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25.
42. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009;25(9):1105–1111.
43. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166–169.
44. Risso D, Ngai J, Speed TP, Dudoit S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat Biotechnol*. 2014;32(9):896–902.
45. Krull KW, et al. Quality Control for RNA-Seq (QuaCRS): an integrated quality control pipeline. *Cancer Inform*. 2014;13(Suppl 3):7–14.
46. Chen C, Khaleel SS, Huang H, Wu CH. Software for pre-processing Illumina next-generation sequencing short read sequences. *Source Code Biol Med*. 2014;9:8.