Compound haploinsufficiency of *Dok2* and *Dusp4* promotes lung tumorigenesis

Ming Chen, … , Robin Plevin, Pier Paolo Pandolfi

*J Clin Invest.* 2018. [https://doi.org/10.1172/JCI99699](https://doi.org/10.1172/JCI99699).

Recurrent broad-scale heterozygous deletions are frequently observed in human cancer. Here we tested the hypothesis that compound haploinsufficiency of neighboring genes at chromosome 8p promotes tumorigenesis. By targeting the mouse orthologs of human *DOK2* and *DUSP4* genes, which were co-deleted in approximately half of human lung adenocarcinomas, we found that compound-heterozygous deletion of *Dok2* and *Dusp4* in mice resulted in lung tumorigenesis with short latency and high incidence, and that their co-deletion synergistically activated MAPK signaling and promoted cell proliferation. Conversely, restoration of *DOK2* and *DUSP4* in lung cancer cells suppressed MAPK activation and cell proliferation. Importantly, in contrast to downregulation of *DOK2* or *DUSP4* alone, concomitant downregulation of *DOK2* and *DUSP4* was associated with poor survival in human lung adenocarcinoma. Therefore, our findings lend in vivo experimental support to the notion that compound haploinsufficiency, due to broad-scale chromosome deletions, constitutes a driving force in tumorigenesis.

Find the latest version:

[http://jci.me/99699/pdf](http://jci.me/99699/pdf)
Compound haploinsufficiency of Dok2 and Dusp4 promotes lung tumorigenesis

Ming Chen,1 Jiangwen Zhang,2 Alice H. Berger,1 Moussa S. Diolombi,1 Christopher Ng,1 Jacqueline Fung,1 Roderick T. Bronson,3 Mireia Castillo-Martín,4,5 Tin Htwe Thin,4 Carlos Cordon-Cardo,4 Robin Plevin,6 and Pier Paolo Pandolfi1

1Cancer Research Institute, Beth Israel Deaconess Cancer Center, Department of Medicine and Pathology, Beth Israel Deaconess Medical Center (BIDMC), Harvard Medical School, Boston, Massachusetts, USA. 2School of Biological Sciences, The University of Hong Kong, Hong Kong, China. 3Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA. 4Department of Pathology, Icahn School of Medicine at Mount Sinai, New York, New York, USA. 5Department of Pathology, Champalimaud Center for the Unknown, Lisbon, Portugal. 6Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, Glasgow, United Kingdom.

Recurrent broad-scale heterozygous deletions are frequently observed in human cancer. Here we tested the hypothesis that compound haploinsufficiency of neighboring genes at chromosome 8p promotes tumorigenesis. By targeting the mouse orthologs of human DOK2 and DUSP4 genes, which were co-deleted in approximately half of human lung adenocarcinomas, we found that compound-heterozygous deletion of Dok2 and Dusp4 in mice resulted in lung tumorigenesis with short latency and high incidence, and that their co-deletion synergistically activated MAPK signaling and promoted cell proliferation. Conversely, restoration of Dok2 and DUSP4 in lung cancer cells suppressed MAPK activation and cell proliferation. Importantly, in contrast to downregulation of DOK2 or DUSP4 alone, concomitant downregulation of DOK2 and DUSP4 was associated with poor survival in human lung adenocarcinoma. Therefore, our findings lend in vivo experimental support to the notion that compound haploinsufficiency, due to broad-scale chromosome deletions, constitutes a driving force in tumorigenesis.

Introduction

Human cancers are characterized by recurrent somatic copy number alterations (SCNAs), which include both focal and broad arm-level amplifications and deletions (1). How these alterations contribute to tumorigenesis remains a central question in cancer research. Specifically, large heterozygous deletions are thought to produce a “second hit” toward the inactivation of a tumor suppressor gene (TSG), hence fulfilling Knudson’s two-hit hypothesis (2). However, TSGs that fulfill this paradigm are often not identified in chromosomal loci displaying consistent loss of heterozygosity (LOH) (1, 3), calling into question whether bona fide TSGs actually reside in those regions. One possible explanation is that these deletions arise from genomic instability and are inconsequential to tumor formation, but another plausible explanation is that haploinsufficiency of 2 or more TSGs in the deleted regions cooperates to promote tumorigenesis (4–9).

Large deletions of chromosome 8p are common in a variety of tumor types, including breast, colon, liver, and lung cancer and leukemia (9–12). A region of 8p loss, spanning cytobands 8p12–8p23, is frequently observed in human lung adenocarcinoma (11, 13), and multiple putative TSGs have been proposed in this region (9, 13, 14). Using a knockout mouse model, we previously identified the 8p21 gene DOK2, encoding an adaptor protein that modulates the RAS/RTK signaling pathway, as a human lung tumor suppressor (14). Human genetics and in vitro functional data have also identified DUSP4, encoding a dual-specificity phosphatase that functions as part of a negative feedback loop to terminate the activity of MAPKs, as another candidate lung TSG on 8p (11). Notably, deletions of 8p in human lung adenocarcinoma frequently encompass both DOK2 and DUSP4, and are enriched in tumors with EGFR mutations (11, 15).

Given the genomic localization of DOK2 and DUSP4 and their convergent function in suppression of MAPK signaling, we hypothesized that compound haploinsufficiency of DOK2 and DUSP4 promotes lung tumorigenesis. Here we show that although heterozygous or complete loss of Dusp4 is not sufficient to initiate lung tumorigenesis in mice, heterozygous loss of Dusp4 markedly accelerates lung tumorigenesis in Dok2 heterozygous mice. In agreement with these findings, we find that concomitant downregulation of DOK2 and DUSP4 is associated with poor survival in human lung adenocarcinoma.

Results

Dusp4 and Dok2 heterozygosity cooperate to promote lung tumorigenesis in vivo. Human genetics and in vitro functional studies previously identified DUSP4 as a putative 8p12 TSG in human lung adenocarcinoma (11). To study the impact of Dusp4 loss on lung tumorigenesis in mice, we analyzed Dusp4-KO mice generated by a targeted deletion strategy (16). Unlike Dok2 mutant mice (14), Dusp4+− and Dusp4−− mice did not develop lung cancer within a 12- or 18-month follow-up (Figure 1, A and B, and Table 1), indicating that loss of Dusp4 alone in mice is not sufficient to initiate lung tumorigenesis in vivo.
Dok2+/− Dusp4+/− mice were adenocarcinomas with solid and papillary growth patterns (Supplemental Figure 1B), similar to lung tumors in Dok-mutant mice (14). However, they had a significantly elevated frequency of mitotic cells positive for Ki-67 staining when compared with Dok2+/− lung tumors (Figure 2B). Given the role of DOK2 and DUSP4 in the regulation of RTK signaling, we conducted IHC to determine the levels of phosphorylated MAPK/Erk in the murine tumors. Dok2+/− lung tumors showed moderate staining for phosphorylated Erk (Figure 2C), while Dok2+/− Dusp4+/− lung tumors showed stronger staining for phosphorylated Erk than Dok2+/− lung tumors (Figure 2C), which was confirmed by Western blot analysis (Figure 2D). In contrast, lung tissues from WT and Dusp4-mutant mice showed low or undetectable staining for Ki-67 and phosphorylated Erk (Supplemental Figure 1C). Thus, in contrast to heterozygous loss of Dok2, heterozygous loss of Dusp4 did not initiate lung tumorigenesis, but instead accelerated the progression of tumors initiated by Dok2 heterozygosity, possibly through augmentation of MAPK activation.

Importantly, IHC staining showed that protein expression of both Dok2 and Dusp4 was retained in the Dok2+/− Dusp4+/− lung tumors, indicating that complete genetic loss of either Dok2 or Dusp4 is not required for tumor development (Figure 2E).

We previously identified DOK2 as a candidate human lung TSG (14). 30% of Dok2 heterozygous (Dok2+/−) mice develop lung adenocarcinoma, albeit with a long latency of 15–19 months. Tumors from these mice do not lose the WT Dok2 allele, suggesting that Dok2 is a haploinsufficient TSG. Thus, we sought to determine whether haploinsufficiency of Dusp4 impacts lung tumorigenesis initiated by Dok2 haploinsufficiency in vivo. We crossed the Dusp4+/− mice with Dok2+/− mice to generate mice compound-heterozygous for Dok2 and Dusp4 (Dok2+/− Dusp4+/− mice) and compared the incidence of lung adenocarcinoma in single- and compound-mutant mice at 12 and 18 months of age.

At 12 months, Dok2+/− Dusp4+/− mice were the only genotype of mice that developed lung adenocarcinoma at moderate (21%) penetrance (Figure 1A and Table 1). Further examination of lungs of the younger Dok2+/− Dusp4+/− mice revealed that as early as 9 months of age, 20% of Dok2+/− Dusp4+/− mice developed lung tumors (Supplemental Figure 1A, n = 10). At 18 months, an analysis of 157 mice showed that Dok2+/− Dusp4+/− mice had a significantly higher incidence of lung adenocarcinoma than Dok2+/− mice (Figure 1B and Table 1).

Compared with Dok2+/− lungs, Dok2+/− Dusp4+/− lungs contained significantly larger tumor nodules (Figure 2A). Tumors in Dok2+/− Dusp4+/− mice were adenocarcinomas with solid and papillary growth patterns (Supplemental Figure 1B), similar to lung tumors in Dok-mutant mice (14). However, they had a significantly elevated frequency of mitotic cells positive for Ki-67 staining when compared with Dok2+/− lung tumors (Figure 2B). Given the role of DOK2 and DUSP4 in the regulation of RTK signaling, we conducted IHC to determine the levels of phosphorylated MAPK/Erk in the murine tumors. Dok2+/− lung tumors showed moderate staining for phosphorylated Erk (Figure 2C), while Dok2+/− Dusp4+/− lung tumors showed stronger staining for phosphorylated Erk than Dok2+/− lung tumors (Figure 2C), which was confirmed by Western blot analysis (Figure 2D). In contrast, lung tissues from WT and Dusp4-mutant mice showed low or undetectable staining for Ki-67 and phosphorylated Erk (Supplemental Figure 1C). Thus, in contrast to heterozygous loss of Dok2, heterozygous loss of Dusp4 did not initiate lung tumorigenesis, but instead accelerated the progression of tumors initiated by Dok2 heterozygosity, possibly through augmentation of MAPK activation.

Importantly, IHC staining showed that protein expression of both Dok2 and Dusp4 was retained in the Dok2+/− Dusp4+/− lung tumors, indicating that complete genetic loss of either Dok2 or Dusp4 is not required for tumor development (Figure 2E).

Table 1. Incidence of lung adenocarcinoma in cohorts of WT, Dok2-, Dusp4-, and compound-mutant mice

<table>
<thead>
<tr>
<th>Age of mice</th>
<th>Genotype</th>
<th>12 months</th>
<th></th>
<th></th>
<th>18 months</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n&lt;sub&gt;cancer&lt;/sub&gt;</td>
<td>n&lt;sub&gt;total&lt;/sub&gt;</td>
<td>% with cancer</td>
<td>P versus Dok2+/− Dusp4+/−</td>
<td></td>
<td>n&lt;sub&gt;cancer&lt;/sub&gt;</td>
<td>n&lt;sub&gt;total&lt;/sub&gt;</td>
</tr>
<tr>
<td>WT</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>0.031</td>
<td></td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>Dusp4+/−</td>
<td>0</td>
<td>22</td>
<td>0</td>
<td>0.028</td>
<td></td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Dok2+/−</td>
<td>0</td>
<td>29</td>
<td>0</td>
<td>0.010</td>
<td></td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Dok2+/− Dusp4+/−</td>
<td>6</td>
<td>28</td>
<td>21%</td>
<td>–</td>
<td></td>
<td>7</td>
<td>26</td>
</tr>
<tr>
<td>Dok2−/−</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>0.027</td>
<td></td>
<td>17</td>
<td>30</td>
</tr>
<tr>
<td>Dusp4−/−</td>
<td>0</td>
<td>28</td>
<td>0</td>
<td>0.025</td>
<td></td>
<td>14</td>
<td>28</td>
</tr>
</tbody>
</table>

The incidence of lung adenocarcinoma in Dok2 and Dusp4 mutant mice at 12 and 18 months of age, respectively, was compared with that of compound Dok2 and Dusp4 heterozygous mice using Fisher’s exact test; P < 0.05 was considered significant.
Therefore, Dok2 and Dusp4 are haploinsufficient in their tumor-suppressive function and their ability to suppress MAPK activation. DOK2 and DUSP4 cooperate to suppress MAPK activation and cell proliferation. To assess the tumor-suppressive role of DOK2 and DUSP4 in human lung cancer, we stably overexpressed DOK2 or DUSP4 in 3 lung cancer cell lines, H1650, H1975, and H3255, all characterized by EGFR mutation and loss of one copy of DOK2 and DUSP4 (Supplemental Table 1; data from Cancer Cell Line Encyclopedia hosted on cBioPortal) (17). Overexpression of either DOK2 or DUSP4 significantly suppressed MAPK activation,
and combined expression of DOK2 and DUSP4 suppressed MAPK to the greatest degree in H1650 and H3255 cell lines (Figure 3, A–C). Consistent with these findings, in a growth curve assay, cells expressing either DOK2 or DUSP4 displayed an impaired growth rate compared with cells transfected with empty vector (Figure 3, D–F). Combined expression of DOK2 and DUSP4 further impaired cell growth compared with either DOK2 or DUSP4 overexpression (Figure 3, D–F). These data further support a cooperative role of DOK2 and DUSP4 in lung tumor suppression.

Compound heterozygosity of DOK2 and DUSP4 confers enhanced sensitivity to MEK inhibition. The observation that DOK2 and DUSP4 converge on the suppression of MAPK signaling led us to investigate whether cells with one copy loss of DOK2 and DUSP4 might exhibit enhanced sensitivity to MEK inhibition compared with cells with intact copies of DOK2 and DUSP4. To avoid confounding effects of MAPK activity induced by mutant EGFR or KRAS, we examined the response to MEK inhibition in cells with WT EGFR and KRAS with either one copy loss of DOK2 and DUSP4 (ChagoK1 and H1299) or no loss of DOK2 and DUSP4 (H1648) (Supplemental Table 1). We found that both ChagoK1 and H1299 cells exhibited higher ERK activation and were more sensitive to MEK inhibition than H1648 cells (Figure 3G). These
in 49% of cases, and compound-heterozygous loss of DOK2 and DUSP4 accounted for 90% of these co-deletion cases (Figure 4B).

Copy number loss of DOK2 or DUSP4 was correlated with low expression of their respective mRNAs (Supplemental Figure 2B), suggesting that copy number alterations deregulate the level of DOK2 and DUSP4 in lung adenocarcinoma.

To evaluate DOK2 and DUSP4 protein expression in human lung tumors, we performed tissue microarray (TMA) analysis in 57 lung tumors (Supplemental Table 2). DOK2 expression was at a low or undetectable level in the majority of tumor samples (Supplementary Figure 1A). Representative images of IHC staining for DOK2 and DUSP4 protein from normal lung tissue or 2 pairs of tumor samples with moderate or low expression of DOK2 and DUSP4 protein. Arrows indicate positive staining of DOK2 or DUSP4 protein in lung epithelial cells. Scale bars: 50 μm. (E) Kaplan-Meier survival curve of lung adenocarcinoma patients stratified by DOK2 and DUSP4 expression. P values were determined by log-rank test. (F) Models of compound haploinsufficiency of DUSP4 and DOK2 promoting lung tumorigenesis. DUSP4-heterozygous cells: The activation of RTK signaling is tightly controlled by a negative feedback loop involving the recruitment of DOK2 and RASA1. The remaining WT DUSP4 allele is sufficient to suppress the low level of ERK activity transduced by RTK signaling. DUSP4 and DOK2 compound-heterozygous cells: Loss of DOK2 results in decreased recruitment of RASA1, allowing enhanced RTK and downstream ERK activation. The remaining WT DUSP4 allele is insufficient to suppress the high level of ERK activity and leads to increased lung tumorigenesis. White rectangles indicate genomic deletions; gray text indicates low expression of proteins of interest.

data suggest that compound haploinsufficiency might serve as a biomarker to predict responsiveness to anticancer therapies.

Compound loss of DOK2 and DUSP4 predicts poor survival in human lung adenocarcinoma. To assess the relevance of these findings to human lung adenocarcinoma, we analyzed copy number of DOK2 and DUSP4 in 513 human lung adenocarcinoma samples from The Cancer Genome Atlas (TCGA). Copy number loss of DOK2 and DUSP4 was observed in 52% and 50% of cases, respectively, the majority of which were single copy losses (Figure 4A and Supplemental Figure 2A). Co-loss of DOK2 and DUSP4 occurred in 49% of cases, and compound-heterozygous loss of DOK2 and DUSP4 accounted for 90% of these co-deletion cases (Figure 4B). Copy number loss of DOK2 or DUSP4 was correlated with low expression of their respective mRNAs (Supplemental Figure 2B), suggesting that copy number alterations deregulate the level of DOK2 and DUSP4 in lung adenocarcinoma.

To evaluate DOK2 and DUSP4 protein expression in human lung tumors, we performed tissue microarray (TMA) analysis in 57 lung tumors (Supplemental Table 2). DOK2 expression was at a low or undetectable level in the majority of tumor samples (Sup-
plemental Table 2). Likewise, DUSP4 protein expression was significantly lower in tumor tissue compared with the noninvolved normal lung (Figure 4C and Supplemental Table 2). Tumors with low expression of both DOK2 and DUSP4 protein accounted for 63% (34 of 54) of tumor samples examined (Supplemental Table 2 and Figure 4D).

To determine whether compound loss of DOK2 and DUSP4 impacts clinical outcome, we analyzed whether downregulation of DOK2 and DUSP4 is associated with survival outcome in human lung adenocarcinoma. We analyzed gene expression data from 764 lung cancer patients with available survival data (18-22). We found that although low expression of either DOK2 or DUSP4 alone had no association with survival outcome (Figure 4E), concomitant low expression of DOK2 and DUSP4 was significantly associated with poor survival (Figure 4E). Thus, compound loss of DOK2 and DUSP4 may be a predictor of aggressive disease in human lung cancer. Consistent with these findings, the association between copy number loss of DOK2 or DUSP4 alone and patient survival only approached significance (Supplemental Figure 2C). Taken together, these data indicate that concomitant downregulation of DOK2 and DUSP4 promotes lung tumorigenesis.

Discussion

Our study provides direct evidence for a model in which compound haploinsufficiency contributes to tumor progression as a result of heterozygous chromosomal deletions (Figure 4F). We found that in mice, loss of Dusp4 alone is insufficient to initiate cancer in the lung, which could be due to the redundancy in the Dusp protein family. However, heterozygous loss of Dusp4 did accelerate the onset and incidence of lung adenocarcinoma initiated by Dok2 heterozygosity. This work highlights the utility of mouse models and functional assays in disentangling the genetic effects of SCNAs. Future efforts should be therefore directed toward defining the combinatorial genetic interactions in recurrent cancer-associated somatic alterations, facilitated by TALEN or CRISPR/Cas9 technology that can now inactivate multiple genes simultaneously in heterozygosity (8, 23).

This study lends further support to the notion that recurrent broad-scale deletions in human cancer, including 3p, 5q, 8p, 9p, and 17p (1), play a critical role in tumorigenesis. Our work adds to a body of evidence supporting our hypothesis that loss of one allele, or reduced expression, of a critical TSG is functionally important either intrinsically or in combination with mutation of other genes (4, 5, 24).

Our results demonstrate that compound haploinsufficiency can lead to activation of the convergent pathway and confer enhanced sensitivity to drugs inhibiting the aberrantly activated pathway. While much work is needed to translate the therapeutic implications of compound haploinsufficiency to the clinic, the high frequency of these genomic events in human cancer warrants further investigation. Finally, as it is a recently identified lung TSG, the genetic interactions of DOK2 with its various effectors, including RASAI/RASGAP (Figure 4F), should be further explored. For instance, loss or mutation of RASAI is observed in 2% of lung adenocarcinoma, which would be predicted to disrupt DOK2’s ability to suppress MAPK signaling, hence defining a novel subtype of lung cancer that is highly sensitive to MEK inhibition (25). Therefore, understanding the relationship between 8p genes and their effectors might also help facilitate patient stratification and the development of more effective cancer treatments.

Methods

Mice and histopathological analyses. All mouse strains have been previously described (16, 26). Dusp4-KO mice originally generated in a mixed C57BL/6 and 129/Sv genetic background were backcrossed to a 129/Sv background by successive mating of Dusp4-heterozygous male mice to 129/Sv female mice and then maintained on a 129/Sv background as Dok2-KO mice. Total body necropsy and histopathological analyses were performed on cohorts of male and female mice at 9, 12, and 18 months of age. Mouse tissues were fixed in 4% paraformaldehyde, embedded in paraffin, and sectioned and stained with H&E for pathological analyses.

Cell lines, Western blotting, and immunohistochemistry. All cell lines were obtained from ATCC and checked for mycoplasma by use of a MycoAlert Mycoplasma Detection Kit (Lonza). Cells and tissues were lysed with RIPA buffer (50 mM Tris [pH8], 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% NP40, 1 mM EDTA, and protease and phosphatase inhibitor cocktail [Roche]). For Western blotting, the following antibodies were used. Anti-phospho-ERK (catalog 9101) and anti-ERK (catalog 9102) were from Cell Signaling Technology. Anti-DUSP4 (S-18) and anti-HSP90 (H-114) were from Santa Cruz Biotechnology. Anti-β-actin (A3853) was from Sigma-Aldrich. Densitometry quantification was performed with NIH ImageJ. For immunohistochemistry, anti-phospho-ERK (20G11) was from Cell Signaling Technology. Anti-Ki-67 (SP6) was from Thermo Fisher Scientific. Anti-DUSP4 (10739-1-AP) was from Proteintech. Anti-DOK2 for both applications was provided by R. Clarkson (Memorial Sloan Kettering Cancer Center, Molecular Pharmacology and Chemistry Program, New York, New York, USA). See complete unedited blots in the supplemental material.

Generation of lung cancer cell lines stably overexpressing DOK2 or DUSP4, and cell growth analysis. Retrovirus was generated by transfection of pCL-Ampho and pBabe-puro-DOK2 or pBabe-hygro-DUSP4 into 293T cells. Forty-eight and 72 hours after transfection, culture supernatants were filtered, supplemented with Polybrene, and added to lung cancer cell lines. Forty-eight hours after the last infection, cells were selected for 2 days in media containing 2 μg/ml puromycin or 100 μg/ml hygromycin before use in experiments. For growth curve analysis, 1.5 × 10^4 cells per well were plated into 12-well plates. At the indicated times, cells were washed with PBS, fixed in 10% formalin, and rinsed with distilled water. Cells were stained with 0.1% crystal violet, rinsed extensively, and dried. Cell-associated dye was extracted with 1.0 ml of 10% acetic acid and then diluted, and the optical density (OD) at 595 nm was determined. Within an experiment, each point was determined in triplicate. The growth curve experiment was repeated 3 times.

TMA analysis. Sections of formalin-fixed, paraffin-embedded lung surgical specimens from 57 patients with lung cancer were stained with H&E and reviewed to identify viable, morphologically representative areas of normal lung and tumor areas from which needle core samples were taken. From each specimen, triplicate tissue cores with diameters of 1 mm were punched and arrayed onto a recipient paraffin block using a precision instrument (Beecher Instruments). Five-micrometer sections of these TMA blocks were stained with H&E for morphology assessment and further used for IHC analysis.
Copy number, gene expression, and survival analysis. For the copy number variation analysis, we downloaded data from the TCGA database. The cutoff thresholds we used were −0.3 for deletions and 0.6 for amplifications; those lower than −0.8 were considered as homozygous deletions and higher than 0.8 as 2-copy amplifications. GISTIC (Genomic Identification of Significant Targets in Cancer) analysis was conducted to identify the copy number alterations of the chromosome 8p genomic region in TCGA lung adenocarcinoma samples. Deletion of many regions of chromosome 8p was detected across TCGA lung adenocarcinoma samples. The estimated FDR (q values) for chromosome 8p deletion was drawn along the chromosome coordinates (27). A t test was used to compare mRNA expression of DOK2 or DUSP4 against its copy number change among the homozygous, heterozygous, and no-alteration groups. For the gene expression and survival analysis of human samples, we collected lung cancer microarray raw datasets from the NCBI GEO, including GSE31210, GSE37745, GSE19188, GSE29013, and GSE3141. All of these data were from the Affymetrix U133 plus 2.0 platform. We conducted gene expression data normalization together. Patient samples were ranked by the expression of the DOK2 gene from low to high. The first one-third of patients were considered as DOK2-low, the last one-third were considered as DOK2-high. A similar classification of patient samples was conducted for the DUSP4 gene. The “co-low” group (n = 96 samples) was derived from the overlap of DOK2-low (n = 164 samples) and DUSP4-low (n = 164); the “co-high” group (n = 65 samples) from overlap of DOK2-high and DUSP4-high. Survival analysis was conducted with the CRAN R packages “survminer” and “survival.” We performed pairwise comparisons in survival analysis among 4 groups, with corrections for multiple testing.

Statistics. For analysis of continuous data, datasets were compared using unpaired 2-tailed Student’s t tests. For analysis of categorical data (for example, cancer incidence), 2 × 2 contingency tables were constructed, and datasets were compared using Fisher’s exact test. For analysis of cell growth data, datasets were compared using 2-way ANOVA with Bonferroni’s post hoc tests. P values less than 0.05 were considered statistically significant. All statistical tests were executed using GraphPad Prism software or the statistical software R (version 3.1.2).

Study approval. Mouse studies were approved by the BIDMC IACUC under protocol 082-2014. The TMA used in this study was constructed at the Icahn School of Medicine at Mount Sinai following an IRB-approved protocol. Written informed consent was obtained from patients for use of pathologic tissue.

Author contributions
MC, MSD, CN, JF, MCM, and THT performed the experiments. MC, AHB, and PPP conceived and designed the experiments. RP, CCC, and PPP supervised the study. JZ performed all bioinformatic analyses. RTB conducted pathology analyses of mouse tissues. MC, JZ, MCM, and PPP analyzed the data. MC, JZ, AHB, and PPP wrote the manuscript. All authors critically discussed the results and the manuscript.

Acknowledgments
We would like to thank all the members of the Pandolfi laboratory for critical comments, and Lauren Southwood and Elizabeth Stack for editing the manuscript. We thank the BIDMC Histology Core and the Biorepository and Pathology Core of the Icahn School of Medicine at Mount Sinai for histological and immunohistochemical analyses. This work was supported by NIH grant R01CA142787 to PPP. MC was supported in part by the Lung Cancer Research Foundation. AHB was supported by a postdoctoral fellowship from the American Cancer Society and an NIH/NCI Pathway to Independence award.

Address correspondence to: Pier Paolo Pandolfi, Beth Israel Deaconess Medical Center, CLS Building, Room 401, 330 Brookline Avenue, Boston, Massachusetts 02215, USA. Phone: 617.735.2121; Email: ppandolf@bidmc.harvard.edu.

MC’s present address is: Department of Pathology, Duke University School of Medicine, and Duke Cancer Institute, Duke University, Durham, North Carolina, USA.

AHB’s present address is: Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.

20. Xie Y, et al. Robust gene expression signature


