Title

Rewiring of human lung cell lineage and mitotic networks in lung adenocarcinomas.

Authors

Kim, Il-Jin
To, Minh
Pham, Patrick
et al.

Publication Date

2013

DOI

10.1038/ncomms2660

Peer reviewed

HHS Public Access

Author manuscript

Nat Commun. Author manuscript; available in PMC 2015 May 31.

Published in final edited form as:

Nat Commun. 2013 ; 4: 1701. doi:10.1038/ncomms2660.

Rewiring of human lung cell lineage and mitotic networks in lung adenocarcinomas

Il-Jin Kim[1,2], **David Quigley**[2], **Minh D. To**[1,2], **Patrick Pham**[1], **Kevin Lin**[2], **Brian Jo**[2], **Kuang-Yu Jen**[2], **Dan Raz**[1], **Jae Kim**[1], **Jian-Hua Mao**[4], **David Jablons**[1,2], and **Allan Balmain**[2,3,*]

[1]Thoracic Oncology Laboratory, Department of Surgery, University of California, San Francisco, 2340 Sutter street, San Francisco, CA, 94115, USA

[2]Helen Diller Family Cancer Center, University of California, San Francisco, 1450 3rd street, San Francisco, CA, 94143, USA

[3]Department of Biochemistry and Biophysics, University of California, San Francisco, 600 16th street, San Francisco, CA, 94158, USA

[4]Life Sciences Division, Lawrence Berkeley National Laboratory (LBNL), 1 Cyclotron road, Berkeley, CA, 94720, USA

Abstract

Analysis of gene expression patterns in normal tissues and their perturbations in tumors can help to identify the functional roles of oncogenes or tumor suppressors and identify potential new therapeutic targets. Here, gene expression correlation networks were derived from 92 normal human lung samples and patient-matched adenocarcinomas. The networks from normal lung show that NKX2-1 is linked to the alveolar type 2 lineage, and identify PEBP4 as a novel marker expressed in alveolar type 2 cells. Differential correlation analysis shows that the NKX2-1 network in tumors includes pathways associated with glutamate metabolism, and identifies Vaccinia-related kinase (VRK1) as a potential drug target in a tumor-specific mitotic network. We show that VRK1 inhibition cooperates with inhibition of PARP signaling to inhibit growth of lung tumor cells. Targeting of genes that are recruited into tumor mitotic networks may provide a wider therapeutic window than that seen by inhibition of known mitotic genes.

*Corresponding author. Allan Balmain, Address: 1450 3rd Street, San Francisco, CA 94158, Tel: 415-502-4192, Fax: 415-502-6779, abalmain@cc.ucsf.edu.

Author contributions

I.J.K. designed the study, performed the experiments, analyzed the data, and wrote the manuscript. D.Q. analyzed the data and wrote the manuscript. M.D.T., K.L., J.H.M. contributed to the data analysis and a discussion. P.P., B.J., K.Y.J., D.R., J.K. performed the experiments and contributed to the data analysis and a discussion. D.J. provided human samples and contributed to the clinical data analysis. A.B. designed and supervised the study, performed data analysis, and wrote the manuscript.

Competing Financial Interests: The authors declare no competing financial interests.

Accession codes: Microarray data for normal lung and lung tumors has been deposited in the the Gene Expression Omnibus database under series accession code GSE32665

INTRODUCTION

Lung cancer remains the leading cause of cancer death in women and men in the U.S., taking more lives each year than breast, prostate, and colorectal cancers combined[1]. Despite major advances in our knowledge of lung cancer genetics, the 5 year survival rate for non-small cell lung cancer (NSCLC) patients is around 16%[2]. Lung cancer treatment is therefore moving rapidly towards an era of personalized medicine, where the molecular characteristics of a patient's tumor will dictate the optimal treatment modalities. NSCLC patients with *EGFR* mutations show significantly improved responses to treatment with Tyrosine kinase inhibitors, e.g., gefitinib or erlotinib that target this receptor kinase[3]. However, almost all of these patients eventually relapse due to rewiring of the genetic architecture of the tumors, either through secondary *EGFR* mutations, or other unknown genetic alterations[4]. This paradigm of tumor rewiring in the face of targeted treatment is evident from many clinical trials that have been reported or are in progress, suggesting that novel approaches to the identification of drug targets and/or methods for analysis of the rewiring process to find complementary targets, will be necessary to improve survival and the probability of long term cures.

We have previously shown that genetic analysis of gene expression can provide a network view of the molecular architecture of normal mouse skin[5], and of the changes in this architecture associated with transformation to carcinoma[6]. This approach allowed us to identify networks that control normal skin cellular components such as the hair follicles, muscle, and inflammatory cells, as well as the changes in these networks that accompany development of carcinomas. We have applied the same strategy to analysis of normal human lung, and matched lung adenocarcinomas from the same patients. The networks generated from a panel of 92 matched normal-tumor lung samples contain motifs representing major cell structural and functional components of the lung, as well as changes in mitotic and inflammation-associated networks in carcinomas. The availability of matched adenocarcinoma tissue from each patient enabled us to generate a coordinated network of genes that are up- or down-regulated in tumors compared to the normal tissue from the same individuals. This "progression network" revealed important contributions from stromal and infiltrating inflammatory cells due to influx of host cells into the tumor microenvironment, but also revealed tumor-specific changes in the cell cycle and mitotic networks that were exploited to identify *VRK1* (Vaccinia related kinase 1) as a mitotic target for the tumor cell cycle.

RESULTS

Gene expression networks in normal lung

Sets of genes that exhibit correlated expression patterns in normal tissue from a genetically heterogeneous population often share a common function or are part of the same physical structure[5,7–8]. We first carried out gene expression network analysis of 92 samples of normal unaffected human lung samples from lung cancer patients. Although previous studies of normal mouse skin clearly identified both structural and functional components of this tissue[5], application of the same approach to normal human lung identified a large

number of significant correlations that were not readily resolved into similar motifs (Supplementary Fig. S1).

Adenocarcinomas share many characteristics with cells of the normal Alveolar type 2 lineage, for which the best characterized markers are the genes encoding the surfactant proteins such as *SFTPB*. We therefore constructed a targeted gene expression network for *SFTA2* and *SFTPB* and the transcription factor *NKX2-1* from the normal lung samples, and noted that several surfactant proteins (*SFTPA1B, SFTPA2B, SFTA3*) and other alveolar type 2 markers (e.g. *LAMP3*) are represented in this correlation network (Fig. 1a).

NKX2-1 has previously been reported to be a lineage-specific oncogene that is amplified in a subset of NSCLC, in particular in adenocarcinomas[9–10]. Others have demonstrated by immunohistochemistry that the NKX2-1 protein is also mainly expressed in Alveolar type 2 cells[11–12]. A functional association of *NKX2-1* with surfactant protein genes is supported by the observation that mutation of *NKX2-1* results in interstitial lung disease in Brain-Lung-Thyroid Syndrome by surfactant protein promoter dysregulation[13]. Abnormal function of *NKX2-1* also affects respiratory distress syndrome of the newborn[14]. Taken together, these data support a role for *NKX2-1* in type 2 pneumocyte lineage determination both in normal lung and in lung adenocarcinomas[9–10].

While many of the genes in these networks are known markers of their respective lineages, this approach has the potential to identify completely novel markers for specific cell types within the lung and other tissues. We validated the network approach for identification of markers for Alveolar type 2 cells using antibodies to PEBP4, a RAF kinase inhibitory protein (RKIP) family member and a scaffold protein controlling myoblast and muscle cell differentiation through MEK and ERK signaling[15]. This protein has not previously been associated with Alveolar type 2 cells, but was significantly coexpressed with surfactant protein genes in normal human lungs. Immunohistochemical staining was carried out on normal human lung sections using antibodies specific for PEBP4 and two other type 2 pneumocyte markers, SFTPB and SFTPD. The results shown in Fig. 1b demonstrate that PEBP4 protein is localized in the same cell population as the known type 2 markers, thus providing support for the predictions made by the gene expression networks.

Rewiring of gene expression networks in lung tumors

Correlation analysis of lung tumors from the same patients revealed several major motifs representing cell growth and mitosis, inflammatory responses and stromal cells (Fig. 2a). In contrast to the normal lung network, the tumor gene expression correlations identified clear modules representing structural (Clara cells) and functional (inflammation, mitosis, translation) components in lung adenocarcinomas. We then produced a molecular view of the perturbations in gene expression between these normal lung tissues and matched adenocarcinomas by combining differential expression and correlation to identify functionally related sets of genes that are coordinately modified during tumorigenesis[6]. This "progression network" (Fig. 2b) was produced by subtracting gene expression results in patient's normal samples from the corresponding adenocarcinoma samples to produce a fold-change value for each gene (see **Methods**). We then identified gene pairs which were significantly correlated in their fold-change and significantly up- or down-regulated.

The progression network identified numerous structural and functional features of the normal lung that were altered either positively or negatively in matched lung tumors from the same patients. Motifs significantly enriched for genes with roles in mitosis, translation, tumor-stromal interactions and immune responses were significantly up-regulated in matched tumors compared to normal tissue. The collagen and stromal gene motif contained genes such as Fibroblast activation protein (*FAP*) and fibronectin type III domain containing 1 (*FNDC1*), which are associated with epithelial tumor growth[16–17]. Motifs associated with negative regulation of cell proliferation, cell adhesion, and focal adhesion assembly were coordinately down-regulated. Consistent with morphological changes that occur during tumorigenesis, the networks corresponding to the Alveolar Type 2 cells (surfactant protein genes) and Clara cells (secretoglobins *CC-10* or *SCG1A1*) formed distinct cliques that were significantly down-regulated in cancers.

Type 2 pneumocyte network rewiring in lung adenocarcinomas

NKX2-1 is significantly correlated with surfactant protein gene expression in tumors (Fig. 3), supporting a model whereby *NKX2-1* maintains a role in determination of the Alveolar type 2 cell lineage. However, the *NKX2-1* tumor correlation network provides evidence for significant rewiring of the *NKX2-1* gene network associated with transformation. Although several components of the normal *NKX2-1* network including surfactant protein genes were also significantly correlated with *NKX2-1* in tumors, the tumor network included many genes significantly correlated in tumors that showed minimal association with *NKX2-1* expression levels in normal lungs (see **Methods** and Table 1). Five of the 23 genes in this set are involved in the glutamine biosynthetic process (Glutamine synthetase 2: *GLS2*; Gammaglutamyl carboxylase: *GGCX*; gamma-glutamyltransferase 1 precursor: *GGT1*; gamma-glutamyltransferase 2 precursor: *GGT2*; Gamma-glutamyltransferase light chain 1: *GGTLC1*). This gene set is significantly enriched for genes with roles in glutathione metabolism and glutamate secretion (corrected P value $<2.4 \times 10^{-3}$, hypergeometric test). Glutamine metabolism is increased in cancer cells and has been linked to the Warburg effect associated with *Myc* deregulation[18]. These data indicate that the *NKX2-1* transcription factor pathway is substantially rewired in lung adenocarcinoma cells, and is now associated with pathways that are required for rapid cell growth and the metabolic requirements of cancer cells. The availability of small molecule inhibitors of *GLS2*[19] may provide a possible therapeutic approach to exploitation of this rewired network for treatment of adenocarcinomas that over-express *NKX2-1*.

Several additional genes not associated with normal type 2 pneumocytes were also linked to *NKX2-1* in this network (Fig. 3). One of these genes, *CLPTM1L* (*CLPTM1-like*) has been identified from Genome-Wide Association Studies (GWAS) as a candidate susceptibility gene for lung cancer[20–22]. *CLPTM1L* is physically located on chromosome 5p15 in close proximity to *TERT*, which is another strong candidate susceptibility gene in the region. Although the known functions of TERT in control of telomere length, as well as its other functions in growth control[23–24], make it a prime candidate, the strong coregulation of *CLPTM1L* with a gene known to play an important role in adenocarcinoma lineage commitment suggests that it may also be involved in lung cancer susceptibility.

VRK1 as a drug target in adenocarcinoma mitotic networks

The most strongly up-regulated motif in the adenocarcinoma progression network was significantly enriched for genes active in the M-phase of the mitotic cycle (corrected $P = 6.4 \times 10^{-15}$) (Fig. 2b). This network comprises many genes that are attractive drug targets for cancer therapy, including the kinases *AURKA, AURKB, CCNB2, TTK, MELK*, and *PLK4*[25–30]. Inhibitors of several of these kinases are presently being tested in clinical trials, however their essential roles in normal mitosis may preclude development as cancer therapeutics because of the toxicity that would be expected due to inhibition of dividing cells in the bone marrow, colon and other proliferative tissues. A major goal of this study was to use network analysis to identify "druggable" target genes that may function selectively in tumor as compared to normal lung tissue. As the cell cycle has been a rich source of cancer drug targets, we carried out network analysis of *AURKA* and *CCNB2*, both well-known cell cycle genes, in both normal lung and tumor samples. While a number of kinases exhibited significant correlations (r>0.6) in gene expression with *AURKA* and *CCNB2* in both normal and tumor tissue, *CDC2* and *VRK1* showed a clear cell cycle network that was specific for lung cancers. (Supplementary Fig. S2).

We decided to focus our functional studies on one of these targets, *VRK1*, which is a serine/threonine protein kinase originally characterized through homology to a vaccinia-related kinase[31]. Subsequent work determined that VRK1 phosphorylates the MDM2 binding site of TP53[32] and is in turn down-regulated by wild-type TP53[33] but not mutant TP53[34]. Ablation of *VRK1* in vitro blocks cell cycle progression at G1[35]. In normal lung tissue, *VRK1* expression is significantly correlated with a wide spectrum of different genes, possibly reflecting the diverse roles of this kinase in normal tissues, but this list is not significantly enriched in genes that control the cell cycle. In contrast, in the lung adenocarcinoma network shown in Fig. 2a, *VRK1* is significantly correlated with the mitotic cell cycle gene network (corrected *P* value for enrichment: 5×10^{-15}) (Fig. 4a). In contrast, no association with mitosis was found for VRK1 in normal lung gene expression networks (Supplementary Fig. S1). Although *VRK1* gene expression is frequently up-regulated in tumors compared to matched normal lung samples, the mean change in gene expression elevation at the RNA level was relatively modest (less than 2-fold) and substantially less than that seen for other known mitotic genes such as Aurora kinases and Cyclin B2 (Fig. 4b). For this reason, VRK1 did not appear in the progression network shown in Fig. 2B, which only shows genes that show the strongest up-regulation in tumors compared to matched normal tissue.

One possible explanation for the lack of association to mitosis in normal lung is that relatively few cells in normal lung are actually cycling. We therefore investigated the correlations for *Vrk1* in an independent data set derived from mouse skin, which is a proliferative self-renewing tissue[5]. No significant enrichment for *Vrk1* correlations with mitosis genes was found in these normal tissues by BiNGO analysis (Supplementary Fig. S3a), although previous analysis has clearly shown the presence of a strong normal mitotic network[5]. In contrast, significant numbers of mitosis-related genes correlating with *VRK1* were seen in tumor gene expression datasets from mouse carcinomas (Supplementary Fig.S3b), and in an independent set of human lung tumors[36] and two human breast tumor datasets[37–38] (data not shown).

Since *VRK1* has previously been implicated in *p53* signaling and DNA damage responses[33–34] we calculated correlation between *VRK1* and all other genes separately in adenocarcinomas with a wild-type *TP53* gene (N=54) and those with mutant *TP53* (N=23). We identified genes significantly correlated with *VRK1* in *TP53* wild type (WT) but not *TP53* mutant tumors, requiring a change in *r* statistic of at least 0.5. Of the nine genes matching these criteria, several are involved in DNA damage responses, including X-ray repair cross-complementing protein 3 (*XRCC3*), a *RAD51* family member involved in double-strand break repair, *FKBP3*, an *MDM2*-interacting protein that contributes to *TP53* regulation[39], and *CDC25A*, a key cell cycle checkpoint gene inhibited when *TP53* is activated during the DNA damage response[40]. This provides supporting evidence that *VRK1* influences the *TP53* DNA damage response pathway, but this influence is ablated in tumors where *TP53* has been mutated. We also performed *EGFR* and K-*ras* mutation screening and divided samples into mutant and wild-type for gene expression network analysis. There was no significantly different network based on either *EGFR* or K-*ras* mutation status, due to the small numbers of samples in each subset and the relative heterogeneity of tumors within a subset.

VRK1 co-localizes with cell cycle markers

Three lung cancer cell lines (A549, H1299, and H460) were transfected with vectors encoding *VRK1* shRNA, and immunofluorescence (IF) staining was used to verify correlation of *VRK1* expression with cell cycle markers such as *Ki-67*, and *CDC2*. *Ki-67* is a well known proliferation marker and CDC2 (CDK1) is a critical kinase involved in the G2/M phase of cell cycle progression. VRK1 and Ki-67 showed mainly nuclear staining in two lung cancer cell lines tested (A549 and H460) (Fig. 5a–b). Downregulation of VRK1 using shRNA in both cell lines led to depletion of nuclear staining for both Ki-67 and VRK1, with residual staining mainly in the cytoplasm or nuclear membrane. CDC2, which was correlated in gene expression with VRK1 in the tumor networks, showed both cytoplasmic and nuclear staining in A549 and H1299 and mainly nuclear staining in H460 (Fig. 5c–e). CDC2 expression was co-localized with VRK1 and was reduced in *VRK1* shRNA transfected cells (Fig. 5c–e). These data show that *VRK1* correlates with *CDC2* at both mRNA and protein levels.

We also checked the protein expression levels of VRK1 and CDC2 in normal lung and matched tumor tissue slides by IHC (Supplementary Fig. S4). VRK1 showed weak staining in mainly bronchioles and was very weak in other regions of normal lung, while it exhibited very strong staining in certain regions of tumors (Supplementary Fig. S4). Although *VRK1* mRNA expression increased by only around 1.5–2 fold in lung tumors compared to normal lung (Fig. 4), VRK1 protein levels were strongly increased in lung tumors (Supplementary Fig. S4). Staining patterns observed for CDC2 were similar to patterns for VRK1 in both normal and matched tumor slides (Supplementary Fig. S4). While bronchiole regions showed only weak but recognizable staining in normal lung slides, most of the tumors were stained by antibodies against CDC2.

Combined VRK1 and PARP inhibition alters lung cancer growth

We designed shRNA probes to down-regulate *VRK1* expression and transfected them in to lung cancer cell lines. After verifying that the shRNAs induced a significant VRK1 expression reduction (Fig. 6a), Flow Cytometry (FCM)/PI staining was done in vector and *VRK1* shRNA transfected lung cancer cells. G1 arrest was observed in *VRK1* shRNA transfected H460 lung cancer cell line while normal cell cycle progression was observed in vector transfected H460 (Student's t-test, $p < 0.0001$, Fig. 6b). Similar results were obtained with another lung cancer cell line H1299 (Fig. 6b). *VRK1* has previously been shown to have a *TP53*-mediated role in DNA damage repair, and *VRK1* gene expression is positively correlated with *BRCA1* in lung tumors (correlation $r = 0.59$). PARP (Poly (ADP-ribose) polymerase) inhibitors are novel anti-cancer drugs that are presently in clinical trials for treatment of breast cancers carrying *BRCA1* mutations[41]. PARP inhibitors have been used for ovarian, breast and prostate cancer but not in lung cancers. Thus, we checked whether *VRK1* shRNA transfected cells may be sensitized to treatment with PARP inhibitors (ABT-888). Significantly increased sensitivity to PARP inhibitor was observed only in *VRK1* shRNA cells compared to vector transfected cells (Fig. 6c). Around a 40% decrease of cell viability was observed after ABT-888 treatment of control cells compared to DMSO treatment, but a 70% decrease in cell viability was shown after ABT-888 treatment of the *VRK1* shRNA cell line (Student's t-test, $p= 0.001$, Fig. 6c). This suggests that *VRK1* could be a viable therapeutic target in lung cancer patients, especially in combination with PARP inhibitors. A colony formation assay was done in H1299 lung cancer cell line with *VRK1*shRNA and a control vector transfection. A significant reduction of colony number was identified in *VRK1*shRNA-transfected cells after PARP inhibitor (ABT-888) treatment (Fig. 6d). ABT-888 treatment results in two additional lung cancer cell lines (H2170 and A549) with *VRK1* shRNA are shown in Supplementary Fig. S5.

In order to check the effect of *VRK1* in normal lung cell lines, we transfected *VRK1*shRNA and a control vector into MRC-5 (normal human fetal lung fibroblasts) and BEAS-2B (normal human lung epithelial cells immortalized with an adenovirus 12-SV40 virus hybrid (Ad12SV40)). There was no effect of *VRK1* downregulation on growth of the MRC-5 cells (Student's t-test, $p=0.64$), although treatment with ABT-888 (PARP inhibitor, 20 μM) in both *VRK1*shRNA and control vector-transfected cells reduced cell growth. There was no synergistic effect of VRK1 downregulation and PARP inhibition (ABT-888 20 μM) in both MRC-5 (Student's t-test, $p=0.45$) and BEAS-2B (Student's t-test, $p=0.142$) cells, which supports our tumor-specific *VRK1* mitotic network data. We selected two genes, *Gli1* (control) and *AURKA* (a known mitotic cycle gene), to test the possibility that downregulation of any cell cycle gene may act together with PARP inhibition to reduce cell viability. SiRNAs for *Gli1*, *AURKA*, *VRK1*, or a scrambled control siRNA (Qiagen) were transfected into H2170 lung cancer cell line. ABT-888 (20 μM) was added to the cell line 24 hours after siRNA transfection and measured for cell viability using CellTiter-Glo® Luminescent Cell Viability Assay kit (Promega). No significant additive or synergistic effect of PARP inhibition and siRNA-mediated downregulation of *Gli1* or *AURKA* was observed while *VRK1* siRNA treatment showed a significant cell viability decrease with PARP inhibition (**Supplementary Fig. S6**).

DISCUSSION

Lung cancer is projected to increase substantially in the future due to the continued expansion of smoking habits particularly in Asia[42]. Substantial advances have been made in development of targeted treatments for lung cancers, in particular through the availability of inhibitors of *EGFR* (Cetuximab, Tarceva). These drugs have been used successfully for treatment of lung cancers carrying activating mutations in the *EGFR* kinase[3], but a consistent effect on long term survival has remained elusive, and most patients relapse due to development of resistance through various mechanisms. This paradigm, by which targeted therapies initially show success in subsets of patients followed by development of resistance, has been seen also for many other novel cancer drugs including Imatinib (Gleevec), and a series of *BRAF* kinase inhibitors[43–48]. There is therefore a major requirement for approaches that allow us to visualize the signaling networks that control both normal and tumor cell behavior, in order to identify signaling hubs that may allow development of synergistic combinations of inhibitors and chemotherapeutic drugs.

Network analysis of genetic control of gene expression has been previously carried out in lower organisms such as yeast[49] and has been applied to tissues from heterogeneous populations of mice[5, 50]. We have applied this network analysis approach to matched normal and tumor tissue from patients with lung adenocarcinomas. Analysis of the coordinated changes in gene expression between normal lungs and matched tumors resulted in a "progression network" that displays the main structural and functional components of the lung that show global changes in gene expression during adenocarcinoma development. Lung tumors contain invading immune inflammatory cells, tumor-associated stroma, and vascular cell populations. While we believe that motifs representing these cell types are relevant to the overall biology of lung adenocarcinomas, the proportion of epithelial to non-epithelial tissue in each sample is variable, and this cellular heterogeneity will have an impact on the normal and tumor networks. This approach allowed us to identify motifs associated with the major type 2 pneumocyte cell lineage, and the changes in these motifs that accompany transformation to adenocarcinomas. The network identified structural components of the cells in the type 2 pneumocyte lineage (e.g. the genes encoding the surfactant proteins) and the transcription factors such as *NKX2-1* that are linked to cell type-specific morphogenesis and lineage determination. The *NKX2-1* transcription factor is a classical marker for adenocarcinomas and although the surfactant protein gene expression motif is down-regulated in the tumors, *NKX2-1* remains linked to this network. However both the progression network, and differential correlation analysis that measures changes in the edges linking genes independently of alteration in actual levels, demonstrated rewiring of *NKX2-1* networks in tumors. Novel genes associated with *NKX2-1* only in tumors included several genes linked to glutamine and glutathione homeostasis that are critical for growth and survival of tumor cells. Several of these genes are enzymes that may be suitable for drug development to find combinations of small molecule inhibitors that coordinately affect this tumor-specific network. Elevated expression of *NKX2-1* may for example be a biomarker for tumors suitable for treatment with inhibitors of glutamine synthetase 2 (*GLS2*) that have already been identified.

NKX2-1 was originally implicated as an oncogene, showing amplification and over-expression in a subset of lung cancers[9–10], but more recent data from mouse models suggests that Nkx2-1 may act as a tumor suppressor[51]. High expression of NKX2-1 is linked to tumor metabolism and growth, but the tumors in this category also express high levels of differentiation markers linked to good prognosis[52–55]. NKX2-1 may therefore be neither an oncogene nor a tumor suppressor in the classical sense, but act as a mediator of alveolar type 2 lineage determination, loss of which allows development of less well differentiated tumors expressing new markers of progression. These data may help to rationalize recent controversy regarding the proposed disparate roles for NKX2-1 in lung cancer.[9,51] The testing of all possible drug targets or their combinations that can be identified by this network approach is not possible within the scope of the present study. We therefore sought to validate the network predictions, both in normal lungs and in lung tumors, by carrying out immunohistochemical studies on normal lung, primary tumors and cell lines to detect co-expression of network components at the protein level. For normal lung, this analysis identified *PEBP4* as a novel marker within the type 2 pneumocyte lineage in the lung. Although this gene is quite widely expressed and is therefore not uniquely expressed in lung-derived cells, it is nevertheless specifically expressed in type 2 pneumocytes both in normal lung and adenocarcinomas, and may therefore be a useful marker for tumors of this lineage. These data therefore demonstrate the value of whole tissue gene expression correlation analysis for identification of markers expressed only or preferentially in specific cell types within the tissue.

In order to identify potentially novel drug targets in lung tumors, we focused initially on the network motif associated with mitosis, which historically has proven to be a rich source of "druggable" genes for cancer therapy. Inhibitors of several mitotic kinases including *AURKA* and *AURKB*, *PLK1-4*, and *MELK* have been developed and several are presently in clinical trials[25–30]. Initial results using some of these inhibitors have shown that their utility may be limited by toxicity due to inhibition of the cell cycle in completely normal dividing cells. We therefore searched for kinases that are integral components of the cell cycle network in tumors, but are not linked to the same markers in normal tissue such as human lung or mouse skin, which is an example of a proliferative normal tissue. *VRK1* was a suitable candidate, as it was not part of the cell cycle network in normal tissue, but was strongly associated with cell cycle and DNA damage response markers in a range of cancer types including lung (Fig. 4 and Supplementary Fig. S2). *BRCA1* mutations are known to induce sensitivity to the effects of PARP inhibition.[56] We checked the '*BRCA1*' network in normal lung and matched tumor samples and found that *VRK1* and *BRCA1* showed a significant correlation only in tumors, and not in normal samples. Moreover, the *BRCA1* network showed significant enrichment for genes involved in the cell cycle, DNA repair, and DNA damage (cell cycle: $p= 1.1\times10^{-56}$, DNA repair: $p= 3.3\times10^{-19}$, and a response to DNA damage: $p= 4.8\times10^{-21}$)(p values from BiNGO) which is consistent with known functions of *BRCA1* and also confirms the robustness of our network analysis. The strong similarity between the *BRCA1* and *VRK1* networks suggested that down-regulation of VRK1 may be a good strategy for sensitization to PARP inhibition. Down-regulation of *VRK1* using shRNA resulted in significant growth arrest in lung cancer cells, particularly in combination with a small molecule inhibitor of poly-ADP ribose polymerase (PARP). While

additional studies will be required to assess *VRK1* as a drug target in a range of cancer types, we believe that this study validates the network approach both for identification of coordinated changes in gene expression motifs that are linked to cancer development, as well as for identification of genes within these motifs that may be suitable for drug development. The application of these bioinformatics tools to matched normal and cancer samples from human patients has the potential to elucidate mechanisms of network rewiring in cancer cells that are not dependent on changes in absolute levels of gene expression in tumors – traditionally a prerequisite for identification of a candidate drug target.

METHODS

Lung cancer patients

Paired normal/lung tumor (adenocarcinoma) RNA and DNA samples were obtained from patients undergoing surgery at University of California, San Francisco. All tumor samples in this study were specifically selected to allow a focused analysis of adenocarcinomas, and they contain at least 50% of tumor cells. All samples were obtained from patients with informed consent under the IRB approval of University of California, San Francisco.

Microarray analysis

RNA was extracted with Trizol followed by DNAse I treatment and purification (RNeasy RNA purification kit, Qiagen). The purified RNA was quantified with Nanodrop and checked by Bioanalyzer for the quality. Gene expression was measured with the Illumina Human Whole Genome 6 2.0 array (Illumina Inc., San Diego, CA, USA). Sample preparations, hybridization, and scanning were performed according to manufacturer's instructions. Briefly, 200ng of total RNA was prepared to make cRNA by using Illumina TotalPrep RNA Amplification Kit (Ambion). First strand cDNA was generated with T7 oligo (dT) primer and ArrayScript and then second cDNA was also made with DNA polymerase.

Biotin-NTP mix with T7 enzyme mix was used to make biotinylated cRNAs. The labeled cRNAs were purified, quantified, and checked again with Bioanalyzer. The labeled cRNA target was used for hybridization and scanning according to Illumina Human Whole Genome 6 2.0 array protocol. Probes with present/absent call ≥ 0.05 assessed by Illumina Beadstudio software were marked absent. Probes were re-annotated using data from[57].

Statistical analysis

Raw microarray data were quantile normalized and log2-transformed. Correlation analysis was performed using probes present in ≥ 90% of samples using Spearman rank correlation. Correlation was defined as significant at the 5% alpha level using the experiment-wise Genome Wide Error Rate permutation method as described in[58]. To calculate tumor progression networks, lung and adenocarcinoma microarrays were normalized together. Change in expression levels between normal and tumor tissue was assessed using a paired t test in SAM[59], retaining genes with FDR ≤ 5% and greater than two-fold change. The progression network was plotted using all differentially expressed gene pairs where correlation in change from normal to tumor was ≥ 0.7 and where each gene was a member

of a clique with size ≥ 3. The decision to use networks of degree 3 was made to enrich for sets of gene which were correlated with each other without eliminating all but the largest, densest sub-networks. We are using clique membership as a way to increase the likelihood of finding a biologically relevant signal. Gene expression networks were plotted in Cytoscape (www.cytoscape.org). Gene Ontology enrichment was calculated using BiNGO[60]. Meaningful decreases in correlation was defined as correlation at the GWER alpha-level for the target condition (e.g. tumor tissue) and a decrease in correlation magnitude of ≥ 0.5 in the other condition (e.g. normal tissue). In order to facilitate complete replication of the statistical findings presented in this manuscript from the raw data (**Supplementary Data 1–5**), a software script is available as **Supplementary Software**.

Mutation screening

PCR reactions were generally carried out in a volume of 25 μl containing 100 ng genomic DNA, 10 pmol of each primer, 250 μM each dNTP, 0.5 units of Taq polymerase and the reaction buffer provided by the supplier (QIAGEN, Hilden, Germany). Samples were denatured for 5 min at 94°C in a PCR machine (MJ Research, Inc. MA, USA), and then amplified by 35 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 1 min, with a final elongation of 10 min at 72°C. Hot spot areas of K-ras (codons 12 and 13) and TP53 (exons 4–8) were screened by bi-directionally sequencing using the Taq dideoxy terminator cycle sequencing kit and an ABI 3730 DNA sequencer (Applied Biosystems, Foster City, CA). *EGFR* mutation was screened by Sequenom OncoCarta panel (Sequenom).

VRK1 shRNA constructs

VRK1 shRNA sequences were designed by BLOCK-iT™ RNAi Designer (Invitrogen). A shRNA was cloned into pSuper retrovirus cassettes (OligoEngine). The shRNA sequence for *VRK1* is GGAGATATCAAGGCCTCAAAT (*VRK1*_sh526). The lung cancer cell lines (A549, NCI-H460 and NCI-H1299) were infected with retroviral stocks produced by the transfection of 293T Phoenix cells using lipofectamin 2000 (Invitrogen). After infection with supernatant medium of *VRK1* shRNA transfected cell, cells were selected with 1–5μg /ml of puromycin. A pSuper vector plasmid without shRNA sequences was used as a control. Cell culture was done with RPMI medium 1640 (Gibco) containing 10% FBS (Gibco), 100 U/ml of penicillin, 100 mg/ml of streptomycin, and 2 mM L-glutamine.

FACS analysis

Cells were trypsinized in a 5 mL petri-plate with 1 mL of trypsin. After three times PBS washing, 3ml of 70% cold EtOH were added to cells and mild vortexing was done. These cells were stored at least overnight at a dark cold room (4°C) with foil. After incubation at 4°C, three times PBS washings were done. Cells were then resuspended in 50 μL PBS with 50 μL RNAse A (Sigma, 5 mg/mL) and incubated at 37°C for 30 minutes. 500 μL of propidium iodide (PI, Calbiochem, 50 ug/mL) was added to the cells and incubated at 4°C for 30 minutes in the dark. The prepared cells were analyzed with BD FACSCalibur™ Flow Cytometer and BD CellQuest Pro software. The data was analyzed with FLOWJO software (Tree Star, Inc, OR).

Immunofluorescence staining

Cells were seeded and incubated for 24 hour in a chamber slides (4-Well Lab-Tek Chamber Slide, BioExpress) and then washed with PBS (Phosphate buffered saline) after removing a medium. A solution with acetone/methanol (1:1) was used for fixing cells for 20 minutes in −20°C followed by PBS washing three times. Slide blocking process was done with solution (10% normal goat serum plus 03% Triton ×100) for 30 minutes at room temperature. Primary antibodies in a blocking solution were incubated with slides for 2 hours at room temperature (dilution 1:500). VRK1 antibodies were used from either Sigma (Rabbit polyclonal, HPA000660) or 1F6 (mouse monoclonal, Cell signaling, #3307S). CDC2 antibody was used from Millipore (06-923). Slides were then washed three times with PBS and incubated with secondary antibody in a blocking solution (1:500 dilution). Alexa 555 goat anti-rabbit IgG (red, Invitrogen) and Alexa488 goat anti-mouse IgG (green, Invitrogen) with DAPI were used for secondary incubation. Slides were washed three times with PBS and checked in a fluorescence microscope.

Immunohistochemistry staining

The same antibodies which were used for IF staining (VRK1 and CDC2) were used for IHC staining. Antibodies for SFTPB (sc-133143, SantaCruz), SFTPD (sc-25324, SantaCruz), and PEBP4 (HPA025064, Sigma) were used in IHC staining. Sections (5μm in thickness) were hydrated in xylene and varying concentrations of ethanol and then steamed in citrate from Biogenex for 20 minutes. The slides were then incubated at room temperature with a blocking solution (10% goat serum in TBS and 1% BSA) for one hour. A blocking solution was removed and the primary antibody was applied. The slides were incubated at 4°C overnight. Following the overnight incubation, the slides were rinsed in 1× TBS containing 0.025% Triton, 1× TBS containing 0.3% H_2O_2, and 1× TBS. The slides were incubated using reagents from the Invitrogen Histostain Plus Broad Spectrum Kit (85-9643) according to the protocol provided by Invitrogen. The slides were rinsed in running water for 5 minutes and then dipped in Gills #2 hematoxylin stain for 1 minute. The slides were rinsed in running water for 1 minute and dehydrated in xylene and varying concentrations of ethanol. The slides were then mounted using organic mounting media and stored at 4°C.

PARP inhibitor treatment

ABT-888 (Chemie Tek) 20 μM was treated to cells in a serum-free condition. After 48 hours of ABT-888 treatment, cell viability was measured using Promega CellTiter-Glo® Luminescent Cell Viability Assay kit and Synergy HT Multi-Mode Microplate Reader (Biotek).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by grants from The Bonnie J. Addario Lung Cancer Foundation (to A.B. and D.J.), the National Cancer Institute U01 CA84244, CA176287-01 and RO1 CA111834-01 (to A.B.), and the UALC (Uniting Against Lung Cancer) foundation (to I.J.K.).

REFERENCES

1. Edwards BK, et al. Annual Report to the Nation on the Status of Cancer, 1975–2006, Featuring Colorectal Cancer Trends and Impact of Interventions (Risk Factors, Screening, and Treatment) to Reduce Future Rates. Cancer. 2010; 116:544–573. [PubMed: 19998273]
2. Delbaldo C, et al. Second or third additional chemotherapy drug for non-small cell lung cancer in patients with advanced disease. Cochrane Database Syst. Rev. 2007; 17 CD004569.
3. Lynch TJ, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N. Engl. J. Med. 2004; 350:2129–2139. [PubMed: 15118073]
4. Pao W, et al. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. PLoS Med. 2005; 2:e73. [PubMed: 15737014]
5. Quigley DA, et al. Genetic architecture of mouse skin inflammation and tumour susceptibility. Nature. 2009; 458:505–508. [PubMed: 19136944]
6. Quigley DA, et al. Network analysis of skin tumor progression identifies a rewired genetic architecture affecting inflammation and tumor susceptibility. Genome Biol. 2011; 12:R5. [PubMed: 21244661]
7. Lum PY, et al. Elucidating the murine brain transcriptional network in a segregating mouse population to identify core functional modules for obesity and diabetes. J Neurochem. 2006; 97:50–62. [PubMed: 16635250]
8. Mehrabian M, et al. Integrating genotypic and expression data in a segregating mouse population to identify 5-lipoxygenase as a susceptibility gene for obesity and bone traits. Nat. Genet. 2005; 37:1224–1233. [PubMed: 16200066]
9. Weir BA, et al. Characterizing the cancer genome in lung adenocarcinoma. Nature. 2007; 450:893–898. [PubMed: 17982442]
10. Kwei KA, et al. Genomic profiling identifies TITF1 as a lineage-specific oncogene amplified in lung cancer. Oncogene. 2008; 27:3635–3640. [PubMed: 18212743]
11. Kolla V, et al. Thyroid transcription factor in differentiating type II cells: regulation, isoforms, and target genes. Am. J. Respir. Cell Mol. Biol. 2007; 36:213–225. [PubMed: 16960125]
12. Khoor A, Whitsett JA, Stahlman MT, Olson SJ, Cagle PT. Utility of surfactant protein B precursor and thyroid transcription factor 1 in differentiating adenocarcinoma of the lung from malignant mesothelioma. Hum. Pathol. 1999; 30:695–700. [PubMed: 10374779]
13. Guillot L, et al. NKX2-1 mutations leading to surfactant protein promoter dysregulation cause interstitial lung disease in "Brain-Lung-Thyroid Syndrome". Hum. Mutat. 2010; 31:E1146–E1162. [PubMed: 20020530]
14. Doyle DA, Gonzalez I, Thomas B, Scavina M. Autosomal dominant transmission of congenital hypothyroidism, neonatal respiratory distress, and ataxia caused by a mutation of NKX2-1. J. Pediatr. 2004; 145:190–193. [PubMed: 15289765]
15. Garcia R, Grindlay J, Rath O, Fee F, Kolch W. Regulation of human myoblast differentiation by PEBP4. EMBO Rep. 2009; 10:278–284. [PubMed: 19197339]
16. Scanlan MJ, et al. Molecular cloning of fibroblast activation protein alpha, a member of the serine protease family selectively expressed in stromal fibroblasts of epithelial cancers. Proc. Natl. Acad. Sci. U. S. A. 1994; 91:5657–5661. [PubMed: 7911242]
17. Anderegg U, et al. MEL4B3, a novel mRNA is induced in skin tumors and regulated by TGF-beta and pro-inflammatory cytokines. Exp. Dermatol. 2005; 14:709–718. [PubMed: 16098131]
18. Gao P, et al. c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism. Nature. 2009; 458:762–765. [PubMed: 19219026]
19. Wang JB, et al. Targeting mitochondrial glutaminase activity inhibits oncogenic transformation. Cancer Cell. 2010; 18:207–219. [PubMed: 20832749]
20. McKay JD, et al. Lung cancer susceptibility locus at 5p15.33. Nat. Genet. 2008; 40:1404–1406. [PubMed: 18978790]
21. Rafnar T, et al. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. Nat. Genet. 2009; 41:221–227. [PubMed: 19151717]

22. Hsiung CA, et al. The 5p15.33 locus is associated with risk of lung adenocarcinoma in never-smoking females in Asia. PLoS Genet. 2010; 6:e1001051. [PubMed: 20700438]
23. Possemato R, et al. Suppression of hPOT1 in diploid human cells results in an hTERT-dependent alteration of telomere length dynamics. Mol. Cancer Res. 2008; 6:1582–1593. [PubMed: 18922974]
24. Lee J, et al. TERT promotes cellular and organismal survival independently of telomerase activity. Oncogene. 2008; 27:3754–3760. [PubMed: 18223679]
25. Girdler F, et al. Validating Aurora B as an anti-cancer drug target. J. Cell Sci. 2006; 119:3664–3675. [PubMed: 16912073]
26. Fu S, Hu W, Kavanagh JJ, Bast RC Jr. Targeting Aurora kinases in ovarian cancer. Expert Opin. Ther. Targets. 2006; 10:77–85. [PubMed: 16441230]
27. Warner SL, Bearss DJ, Han H, Von Hoff DD. Targeting Aurora-2 kinase in cancer. Mol. Cancer Ther. 2003; 2:589–595. [PubMed: 12813139]
28. Ko MA, et al. Plk4 haploinsufficiency causes mitotic infidelity and carcinogenesis. Nat. Genet. 2005; 37:883–888. [PubMed: 16025114]
29. Nakano I, et al. Maternal embryonic leucine zipper kinase is a key regulator of the proliferation of malignant brain tumors, including brain tumor stem cells. J. Neurosci. Res. 2008; 86:48–60. [PubMed: 17722061]
30. Kwiatkowski N, et al. Small-molecule kinase inhibitors provide insight into Mps1 cell cycle function. Nat. Chem. Biol. 2010; 6:359–368. [PubMed: 20383151]
31. Nezu J, Oku A, Jones MH, Shimane M. Identification of two novel human putative serine/threonine kinases, VRK1 and VRK2, with structural similarity to vaccinia virus B1R kinase. Genomics. 1997; 45:327–331. [PubMed: 9344656]
32. Lopez-Borges S, Lazo PA. The human vaccinia-related kinase 1 (VRK1) phosphorylates threonine-18 within the mdm-2 binding site of the p53 tumour suppressor protein. Oncogene. 2000; 19:3656–3664. [PubMed: 10951572]
33. Valbuena A, Vega FM, Blanco S, Lazo PA. p53 downregulates its activating vaccinia-related kinase 1, forming a new autoregulatory loop. Mol. Cell Biol. 2006; 26:4782–4793. [PubMed: 16782868]
34. Valbuena A, et al. Alteration of the VRK1-p53 autoregulatory loop in human lung carcinomas. Lung Cancer. 2007; 58:303–309. [PubMed: 17689819]
35. Valbuena A, Lopez-Sanchez I, Lazo PA. Human VRK1 is an early response gene and its loss causes a block in cell cycle progression. PLoS One. 2008; 3:e1642. [PubMed: 18286197]
36. Ding L, et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature. 2008; 455:1069–1075. [PubMed: 18948947]
37. Chin K, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell. 2006; 10:529–541. [PubMed: 17157792]
38. Fan C, et al. Concordance among gene-expression-based predictors for breast cancer. N. Engl. J. Med. 2006; 355:560–569. [PubMed: 16899776]
39. Ochocka AM, et al. FKBP25, a novel regulator of the p53 pathway, induces the degradation of MDM2 and activation of p53. FEBS Lett. 2009; 583:621–626. [PubMed: 19166840]
40. Demidova AR, Aau MY, Zhuang L, Yu Q. Dual regulation of Cdc25A by Chk1 and p53-ATF3 in DNA replication checkpoint control. J. Biol. Chem. 2009; 284:4132–4139. [PubMed: 19060337]
41. Turner NC, et al. A synthetic lethal siRNA screen identifying genes mediating sensitivity to a PARP inhibitor. EMBO J. 2008; 27:1368–1377. [PubMed: 18388863]
42. Chen ZM, Xu Z, Collins R, Li WX, Peto R. Early health effects of the emerging tobacco epidemic in China. A 16-year prospective study. JAMA. 1997; 278:1500–1504. [PubMed: 9363969]
43. Gorre ME, et al. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. Science. 2001; 293:876–880. [PubMed: 11423618]
44. Emery CM, et al. MEK1 mutations confer resistance to MEK and B-RAF inhibition. Proc. Natl. Acad. Sci. U. S. A. 2009; 106:20411–20416. [PubMed: 19915144]
45. Montagut C, et al. Elevated CRAF as a potential mechanism of acquired resistance to BRAF inhibition in melanoma. Cancer Res. 2008; 68:4853–4861. [PubMed: 18559533]

46. Comin-Anduix B, et al. The oncogenic BRAF kinase inhibitor PLX4032/RG7204 does not affect the viability or function of human lymphocytes across a wide range of concentrations. Clin. Cancer Res. 2010; 16:6040–6048. [PubMed: 21169256]
47. Tsai J, et al. Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. Proc. Natl. Acad. Sci. U. S. A. 2008; 105:3041–3046. [PubMed: 18287029]
48. Hoeflich KP, et al. Antitumor efficacy of the novel RAF inhibitor GDC-0879 is predicted by BRAFV600E mutational status and sustained extracellular signal-regulated kinase/mitogen-activated protein kinase pathway suppression. Cancer Res. 2009; 69:3042–3051. [PubMed: 19276360]
49. Zhu J, et al. Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks. Nat. Genet. 2008; 40:854–861. [PubMed: 18552845]
50. Wang K, Narayanan M, Zhong H, Tompa M, Schadt EE. Meta-analysis of inter-species liver co-expression networks elucidates traits associated with common human diseases. PLoS Comput. Biol. 2009; 5:e1000616. [PubMed: 20019805]
51. Winslow MM, et al. Suppression of lung adenocarcinoma progression by Nkx2-1. Nature. 2011; 473:101–104. [PubMed: 21471965]
52. Barlési F, et al. Positive thyroid transcription factor 1 staining strongly correlates with survival of patients with adenocarcinoma of the lung. Br. J. Cancer. 2005; 93:450–452. [PubMed: 16052216]
53. Martins SJ, Takagaki TY, Silva AG, Gallo CP, Silva FB, Capelozzi VL. Prognostic relevance of TTF-1 and MMP-9 expression in advanced lung adenocarcinoma. Lung Cancer. 2009; 64:105–109. [PubMed: 18801593]
54. Perner S, et al. TTF1 expression in non-small cell lung carcinoma: association with TTF1 gene amplification and improved survival. J. Pathol. 2009; 217:65–72. [PubMed: 18932182]
55. Anagnostou VK, Syrigos KN, Bepler G, Homer RJ, Rimm DL. Thyroid transcription factor 1 is an independent prognostic factor for patients with stage I lung adenocarcinoma. J. Clin. Oncol. 2009; 27:271–278. [PubMed: 19064983]
56. Farmer H, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature. 2005; 434:917–921. [PubMed: 15829967]
57. Barbosa-Morais NL, et al. A re-annotation pipeline for Illumina BeadArrays: improving the interpretation of gene expression data. Nucleic Acids Res. 2010; 38:e17. [PubMed: 19923232]
58. Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. Genetics. 1994; 138:963–971. [PubMed: 7851788]
59. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A. 2001; 98:5116–5121. [PubMed: 11309499]
60. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics. 2005; 21:3448–3449. [PubMed: 15972284]

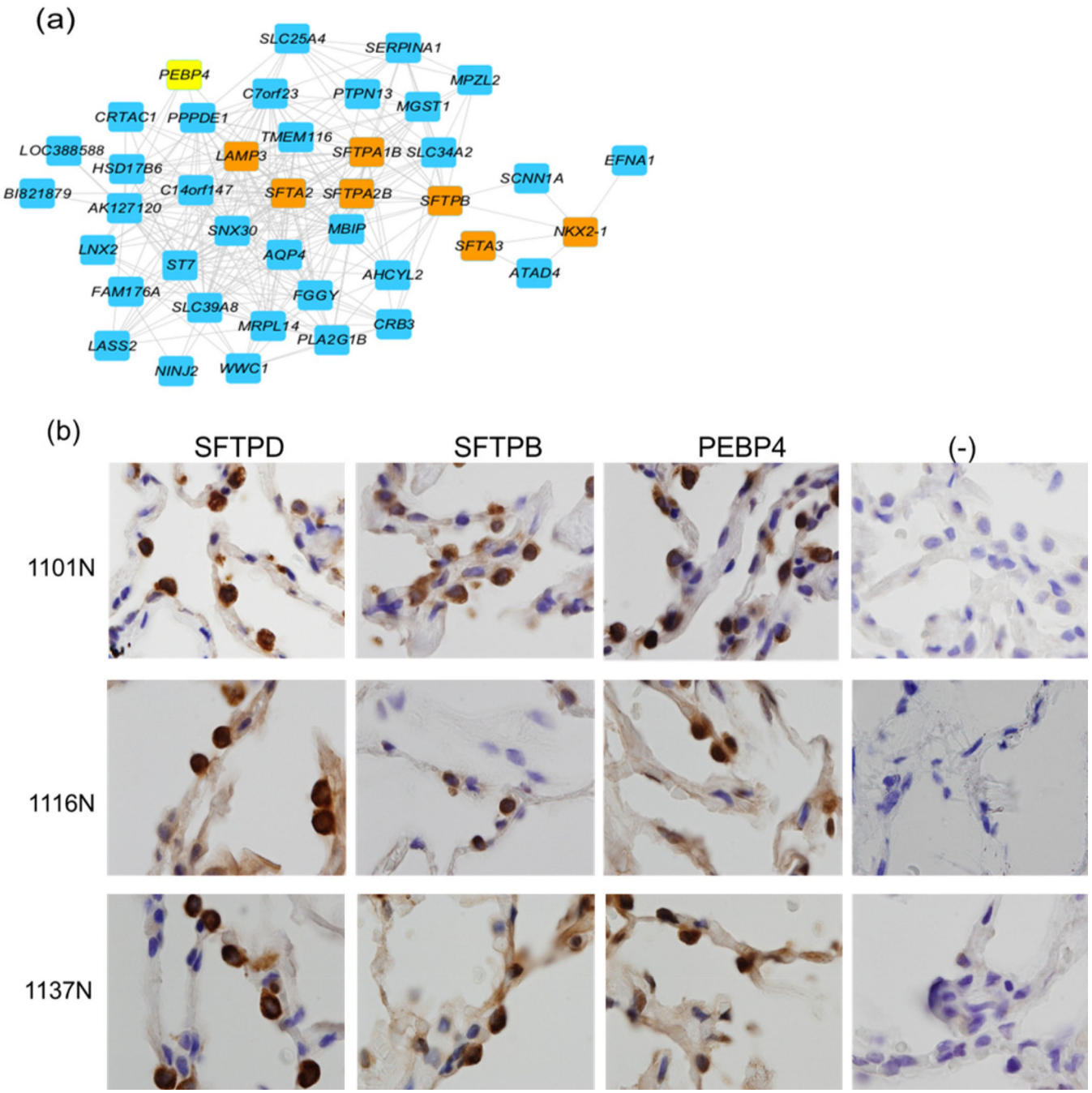

Figure 1. Gene expression network for alveolar type II markers in normal lung

(a) Gene co-expression networks associated with surfactant protein markers and NKX2-1 in normal human lung tissues. Gene pairs which are significantly correlated are drawn as blue boxes connected by a line; PEBP4 is shown in yellow. Other type 2 pneumocyte markers are shown in orange. (b) Adjacent sections of three normal slides from different lung cancer patients (1101N, 1116N, and 1137N) were stained for expression of SFTPD, SFTPB, PEBP4 protein, and a negative control (−) (microscope magnification:100×). Type 2 pneumocytes were stained for known markers SFTPD and SFTPB as well as PEBP4, which

was identified as a candidate marker for type 2 pneumocytes in gene expression networks. These slides indicate that the same cells express both surfactant proteins and PEBP4 protein.

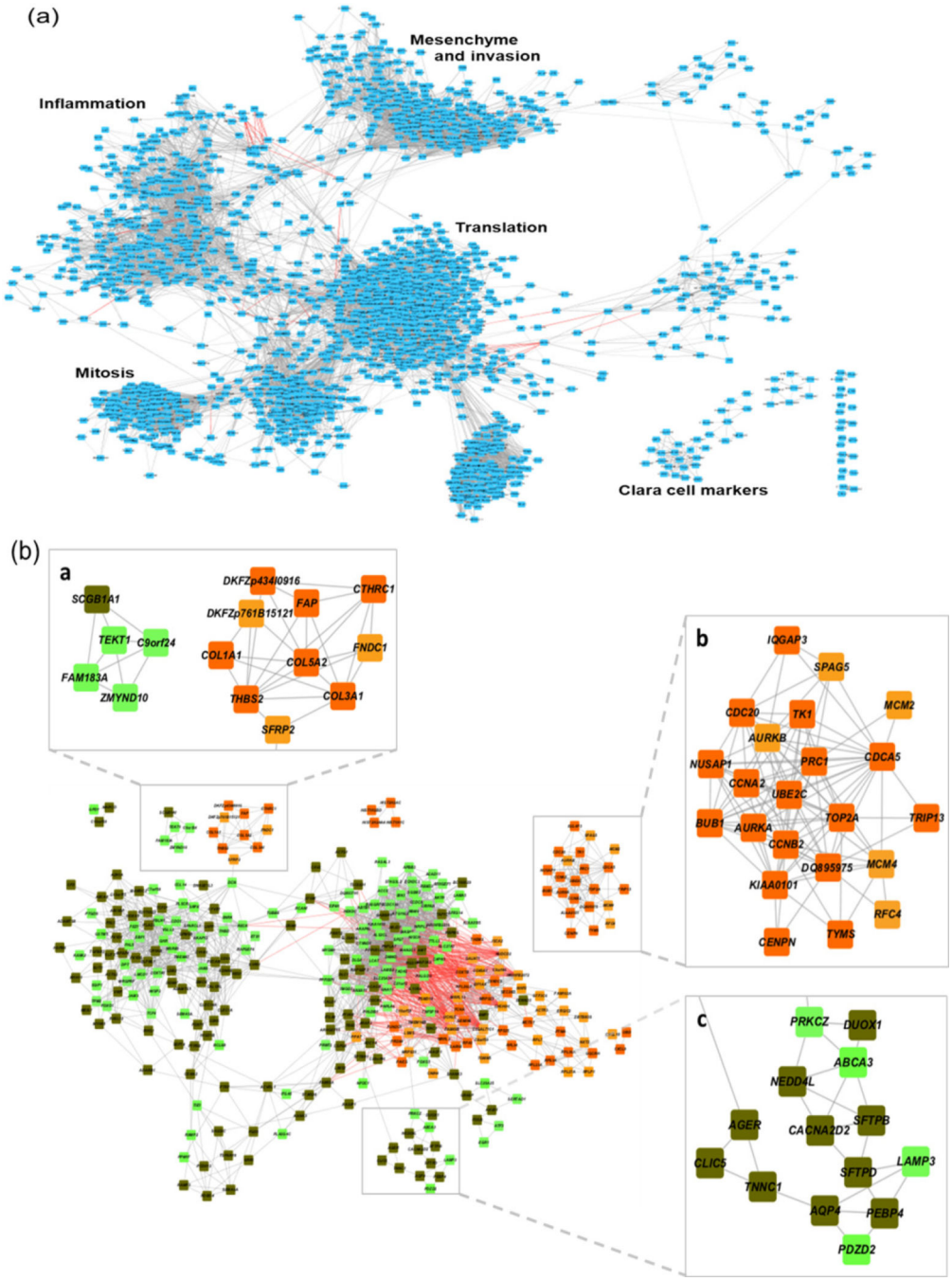

Figure 2. Human lung adenocarcinoma tissue gene expression network

(a) Significantly correlated gene pairs that were part of a three gene clique were plotted ($r \geq +/- 0.75$). This analysis identified network motifs associated with mitosis, protein translation, stromal cells/invasion, inflammation, and Clara cells.

(b) Progression network of lung adenocarcinoma. Gene pairs with significantly correlated change ($r \geq +/- 0.7$) and significant change in mRNA gene expression level between normal lung and adenocarcinomas are drawn as nodes (FDR < 5%, fold-change ≥ 2). Red nodes indicate increased expression and green nodes indicate decreased expression, with darker

color indicating more extreme change. Gray lines connect genes with significant direct correlated change and red lines indicate inverse correlation. a: clara cell and collagen motifs, b: cell cycle motif, c: type II pneumocyte motif

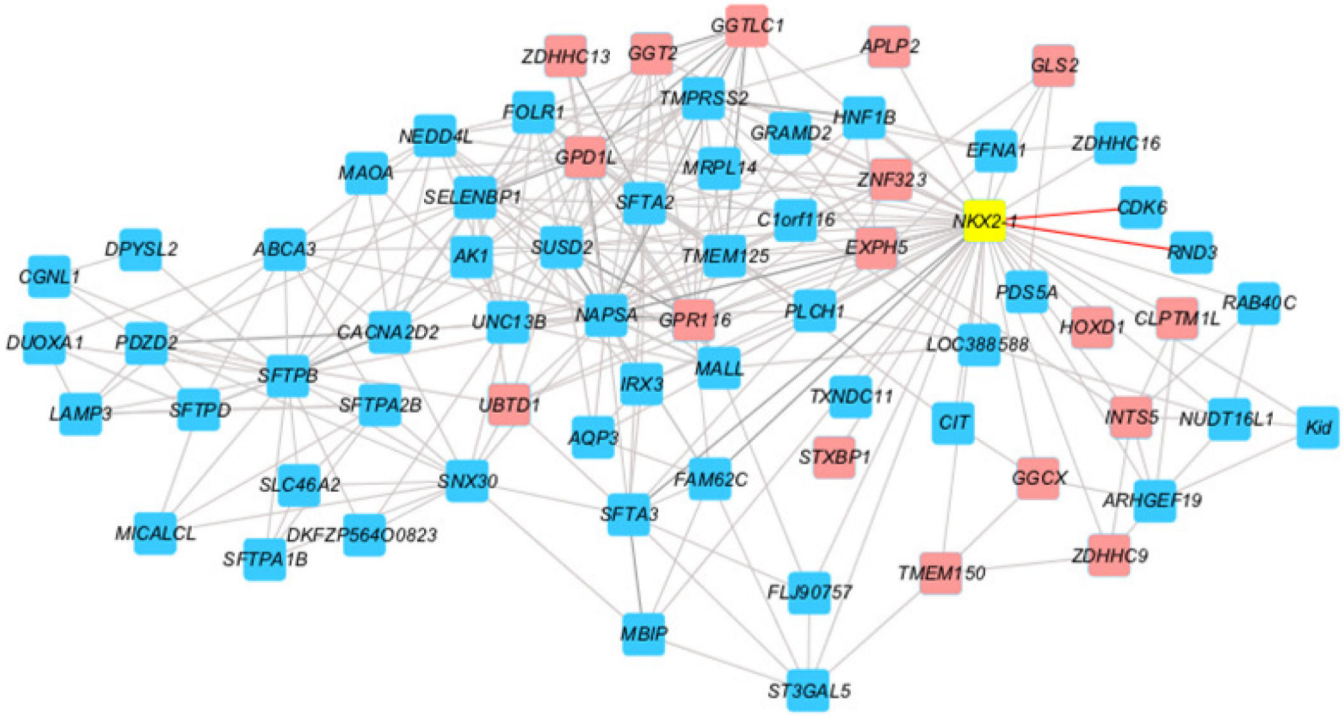

Figure 3. Type 2 pneumocyte network in lung adenocarcinomas

Gene co-expression networks associated with surfactant protein markers and *NKX2-1* in human lung adenocarcinomas. The *NKX2-1* network in tumors shows co-expression with genes relevant to tumor susceptibility (e.g. *CLPTM1L*) and tumor metabolism (e.g. *GLS2*, *GGT2*, *GGTLC1*, *GGCX*, *GPD1L*) that correlated at lower levels in the comparable normal lung network (colored in pink). Gray lines connect genes with significant direct correlation and red lines indicate inverse correlation.

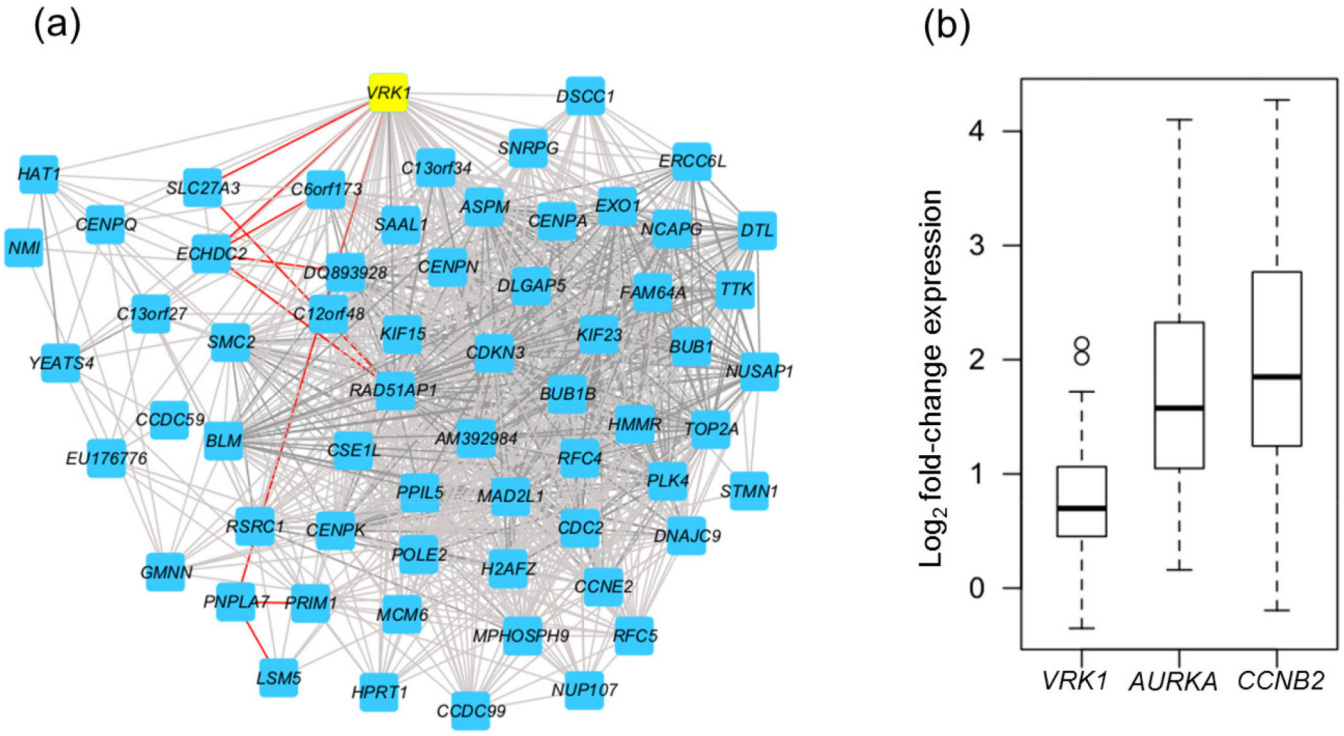

Figure 4. *VRK1* correlation network in adenocarcinomas

Gene pairs which are significantly correlated with *VRK1* (highlighted in yellow) are drawn as in figure 4. This network is significantly enriched for genes with which function in the M-phase of mitosis (a).

(b) mRNA gene expression of *VRK1*, *AURKA*, and *CCNB2* in normal lung and matched tumor tissues. Box plot displaying fold-change values for matched lung adenocarcinoma and normal lung tissue for three genes: *VRK1*, Aurora Kinase A (*AURKA*), and Cyclin B 2 (*CCNB2*). The mean increase in gene expression between normal and matched tumor samples is smaller in *VRK1* than in *AURKA* and *CCNB2*.

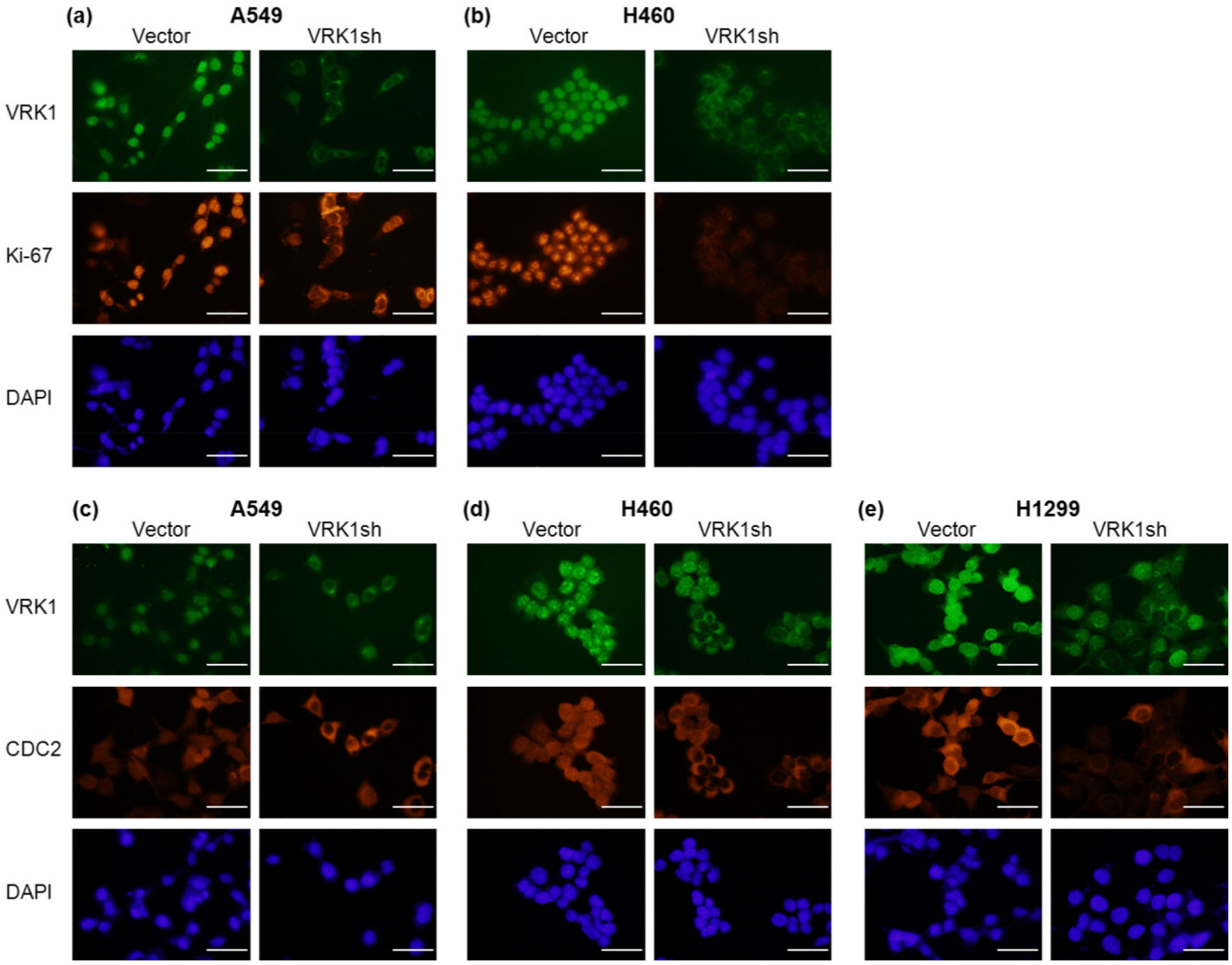

Figure 5. Immunofluorescence staining of VRK1 in lung cancer cell lines

VRK1 and K-67 was also co-stained with different fluorescent dyes (VRK1-green; Ki-67-red) in vector and VRK1 shRNA transfected cells (a–b). After shRNA knockdown of VRK1 in cell lines (A549 and H460), both VRK1 and KI-67 expression was significantly reduced (a–b).

VRK1 (green) and CDC2 (red) were co-stained in three lung cancer cell lines (A549, H460, and H1299). VRK1 and CDC2 were found to be mainly co-localized in VRK1 shRNA cell lines (c–e). Mouse monoclonal VRK1 antibody (1F6) was used for (a–e). Scale bar, 50 μm.

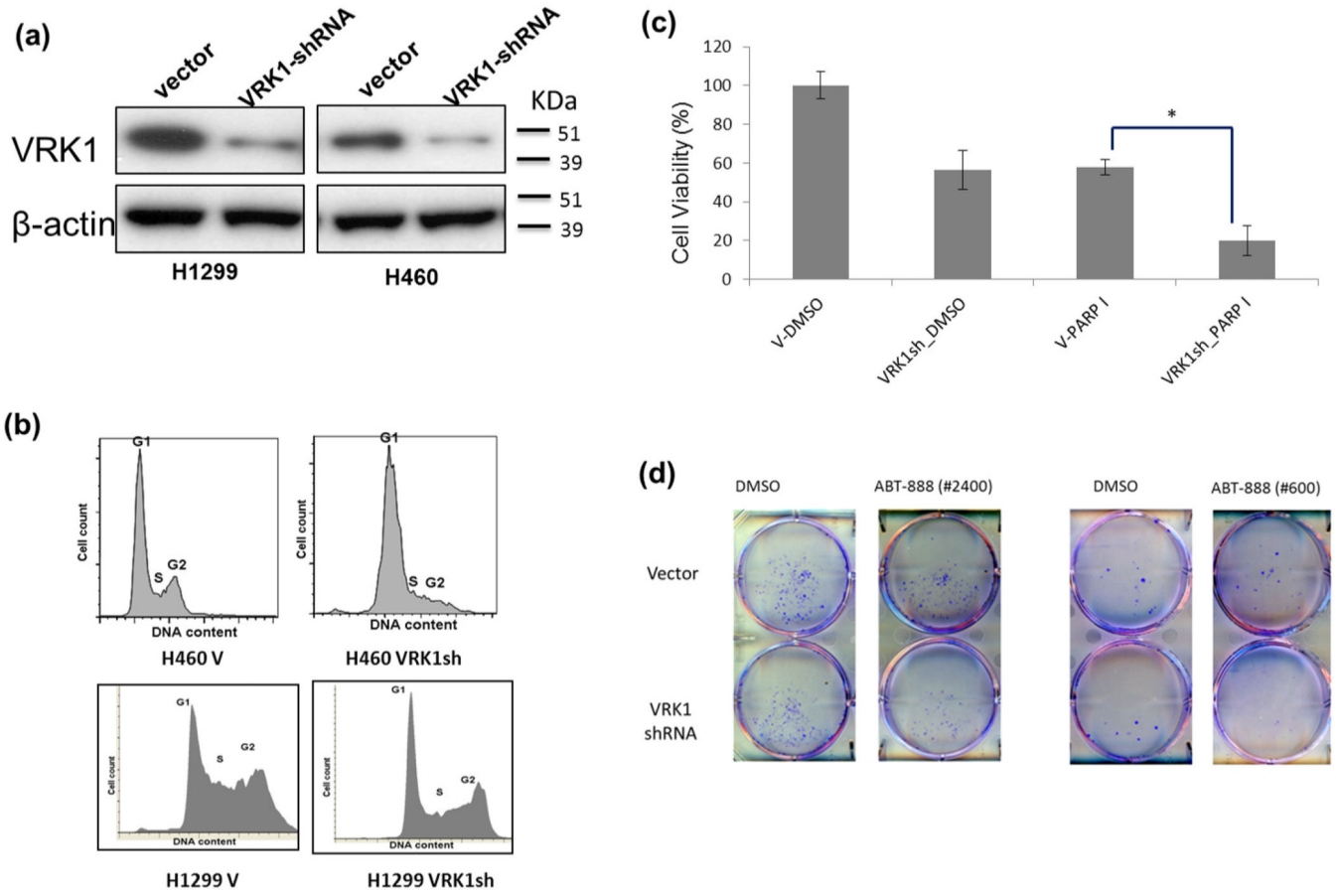

Figure 6. Down-regulation of *VRK1* causes G1 arrest and increases sensitivity to PARP inhibitor treatment

Lung cancer cell lines (H460 and H1299) were analyzed by Immunoblot assay to measure VRK1 protein levels after vector and VRK1 shRNA transfection. Reduced protein expression of VRK1 was observed in VRK1 shRNA cell lines. V: vector transfected cell lines, VRK1sh: VRK1 shRNA transfected cell line (a). The same cell lines were used for FACS analysis to check cell cycle status. While normal cell cycle progression was found in vector transfected cells, G1 arrest was observed in *VRK1*-shRNA transfected H460 cells ($p < 0.0001$) and H1299 cells ($p = 0.03$). Experiments were carried out three times independently (b). Vector and *VRK1* shRNA transfected H1299 cell lines were treated with PARP inhibitor (ABT-888, 20 μM) and DMSO to measure cell viability. A synergistic increase in sensitivity (decreased cell viability) was observed in PARP inhibitor treated *VRK1* shRNA cell lines compared to vector transfected cells ($p = 0.001$). Experiments were done three times independently. The data are represented as the mean ± s.d. (c). A colony formation assay was done in H1299 lung cancer cell line with *VRK1*shRNA and control vector transfection. H1299 cells (2400 cells per well in the left and 600 cells per well in the right panel) were treated with ABT-888 (PARP inhibitor, 20 μM) in serum-free medium for 48 hours. Regular medium was then given to cells after two times PBS washing. After a week with in regular medium, each plate was fixed with 100% methanol and stained with a 0.25% crystal violet staining solution. All experiments were done in triplicate. A synergistic reduction in colony number was identified in *VRK1*shRNA-transfected cells treated with the PARP inhibitor

(p=0.003 for 2400 cells and p= 0.003 for 600 cells/well) (d). The p-values were calculated using Student's t-test.

Table 1
Differential correlation analysis of NKX2-1

Listing of genes which fit the following criteria: 1) significantly correlated with *NKX2-1* in lung adenocarcinomas; 2) correlation of the same pair is ≥ 0.5 lower in normal lung tissue.

symbol	probe	rho_normal	rho_tumor	difference
BCAM	ILMN_1790455	−0.10	0.64	0.74
INTS5	ILMN_1796968	−0.02	0.72	0.74
ZNF323	ILMN_1655748	−0.07	0.64	0.71
GGCX	ILMN_1758232	0.06	0.72	0.66
GLS2	ILMN_1709771	0.08	0.73	0.65
ZNF323	ILMN_1679275	0.01	0.64	0.63
GGTLC1	ILMN_1692188	0.07	0.68	0.61
CDK6	ILMN_1802615	−0.11	−0.70	−0.59
GGT2	ILMN_1737387	0.09	0.68	0.59
GGT1	ILMN_1753340	0.05	0.64	0.59
CD163L1	ILMN_1661905	−0.05	−0.64	−0.59
STXBP1	ILMN_1728747	0.11	0.69	0.58
APLP2	ILMN_1710482	0.07	0.65	0.58
PRUNE	ILMN_1728914	0.08	0.64	0.56
CLPTM1L	ILMN_1752802	0.08	0.64	0.56
HOXD1	ILMN_1717381	0.10	0.66	0.56
GPD1L	ILMN_1694106	0.15	0.69	0.54
EXPH5	ILMN_1712066	0.18	0.72	0.54
GPR116	ILMN_1728785	0.14	0.68	0.54
ZDHHC9	ILMN_1734425	0.15	0.67	0.52
SALL2	ILMN_1740842	0.11	0.64	0.53
UBTD1	ILMN_1794914	0.15	0.67	0.52
TMEM150	ILMN_1765877	0.17	0.68	0.51

This list is enriched for genes with roles in glutamine biosynthetic process (highlighted in bold).