LKB1 LOSS INDUCES CHARACTERISTIC PATHWAY ACTIVATION IN HUMAN TUMORS AND CONFERS SENSITIVITY TO MEK INHIBITION ASSOCIATED WITH ATTENUATED PI3K-AKT-FOXO3 SIGNALING

By

JACOB KAUFMAN

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Cancer Biology

December, 2013

Nashville, Tennessee

Approved:

Professor William Pao

Professor Alissa M. Weaver

Professor Ethan Lee

Professor Pierre P. Massion

Professor David P. Carbone

This work is dedicated, with love, to my amazing wife, Elizabeth

and

To our three wonderful daughters, Claire, Caroline, and Mary

ACKNOWLEDGMENTS

This work would not have been possible without the financial support from Vanderbilt's Medical Scientist Training Program and the Strategic Partnering to Evaluate Cancer Signatures grant NCI U01CA114771. It is a privilege to have had such resources available to fund my training and education.

I am deeply grateful for the incredible support of my mentor, David Carbone. Dr. Carbone has shared not only his inexhaustible knowledge of biology and cancer, but also his dedication to improving the lives of others, and his sincere compassion both for patients and colleagues. Working with him has given me a model of what being a physician scientist should be, and I hope to emulate his character and compassion throughout my career. This work could not have been completed without the help of my co-authors, especially Joe Amann and Paul Park, as well as Horton Li, Rajesh Arasada and Yu Shyr. I have also benefited immeasurably from many faculty members with whom I've had the honor of working and interacting, especially Yu Shyr, Pierre Massion and David Johnson, and my committee members William Pao, Alissa Weaver, and Ethan Lee. Thank you all for everything you have done for me.

Most of all, I am forever indebted to my family for their constant love, support, and encouragement. My parents have supported me in more ways than I can count, and throughout my life have been models of goodness, compassion, love, generosity, and good doctoring. My three wonderful daughters bring a smile to my face every day. I am so blessed by each of them. And to my amazing wife, Elizabeth, you mean more to me than anything. Thank you, from the bottom of my heart, for all the help, love, support,

and friendship that you have sustained me with through our years, and for everything you are to me. I love you, E.

TABLE OF CONTENTS

Page

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF FIGURES viii

LIST OF TABLES ix

Chapter

I. INTRODUCTION 1

Discovery of the LKB1 Tumor Suppressor 1
Structure, function, and regulation of LKB1 2
LKB1 regulation of AMPK 7
LKB1 and the mTOR pathway 10
Regulation of other processes by LKB1 12
Somatic LKB1 loss in human tumors 14
Mouse models of LKB1 loss 17
Murine model of LKB1/KRAS mutant lung cancer 20
Gene expression analysis 23

II. GENE EXPRESSION SIGNATURE OF LKB1 LOSS 29

Introduction 29
Methods and materials 31
Results 36
LKB1 loss results in consistent gene expression changes in human tumors 36
A 16-Gene LKB1-loss classifier accurately predicts mutational and non-mutational loss of LKB1 in resected human lung Adenocarcinomas 41
The LKB1-loss classifier accurately predicts mutational and non-mutational loss of LKB1 in NSCLC cell lines 50
The LKB1-loss classifier is associated with LKB1 mutations in cell lines derived from non-lung primary cancers 51
Association of LKB1-loss signature with loss of LKB1 in resected, non-lung primary tumors 54
Development of an assay for LKB1-loss score suitable for analysis of clinical samples using the nanoString n-Counter assay 58
Discussion 61

III. INTERPRETING THE BIOLOGICAL SIGNIFICANCE OF THE LKB1 SIGNATURE 65

Introduction 65
Methods and materials 67
Results 73
Genes associated with LKB1 loss correspond to particular tumor phenotypes and transcription factors 73
Activation of the NRF2 pathway is common among LKB1-deficient tumors 74
NRF2 association with LKB1 is partially explained by deletion events affecting LKB1 and KEAP1 78
Gene set enrichment analysis of the mitochondria/mTOR cluster and the down-regulated cluster 80
Gene set enrichment analysis of the LKB1-loss cluster identifies association with FOXO3, FOXA2, and CREB Transcription Factors 83
Association between LKB1 loss and prevalence of other mutations 85
Clinical phenotypes of LKB1-deficient lung adenocarcinomas 88
Protein, microRNA, and copy number alterations associated with LKB1 status 90
Wild-type LKB1 decreases the expression of the LKB1-associated signature genes 93
Discussion 96

IV. LKB1 LOSS IS ASSOCIATED WITH SENSITIVITY TO MEK INHIBITION AND ALTERATIONS IN PI3K-AKT-FOXO3 SIGNALING 105

Introduction 105
Methods and materials 108
Results 110
LKB1-deficient cell lines show increased susceptibility to MEK inhibition 110
Restoring LKB1 expression in LKB1-mutant cell lines induces resistance to MEK inhibitors. 113
The expression of the 16-gene signature is independent of MEK/ERK signaling 114
Restoring LKB1 to cell lines in vitro induces phosphorylation of AKT and FOXO3 116
PI3K/AKT signaling is attenuated in resected human tumors with LKB1 loss 117
Discussion 120

V. CONCLUSION 125

Appendix

A. DATA SOURCES AND STATISTICAL COMPARISONS 135
B. INITIAL GENES IN FOUR TRANSCRIPTIONAL CLUSTERS 139
C. TOP 200 GENES ASSOCIATED WITH EACH CLUSTER 140

REFERENCES 143

LIST OF FIGURES

Figure Page

1.1. Structure of LKB1 and disrupting mutations observed in cancer 4
1.2. Schematic of LKB1 interactions with AMPK family members and their downstream effects 8
1.3. Schematic of LKB1 and AMPK interactions with mTOR pathway 10
2.1. LKB1 loss produces a characteristic pattern of gene expression 38
2.2. Gene Set Enrichment Analyses of LKB1 associated genes 40
2.3. LKB1 loss scores exhibit a bimodal distribution 43
2.4. LKB1-loss signature is predictive of mutations and non-mutational loss of LKB1 45
2.5. Receiver operating characteristics for LKB1-loss score in resected lung adenocarcinomas and NSCLC cell lines 46
2.6. Comparison of LKB1 loss scores derived from two different training sets 46
2.7. Comparison of protein and gene expression differences associated with known LKB1 mutations or associated with predicted LKB1 loss among LKB1 WT tumors 48
2.8. Decreased LKB1 mRNA is associated with LKB1-loss signature in resected breast cancer, lung squamous cell carcinoma, and cervical cancer 58
2.9. Performance of nanoString platform to assess LKB1 loss signature in cell lines and clinical samples 60
3.1. Association of NRF2 activation cluster with KEAP1 mutations and large deletions of chromosome 19 78
3.2. TGF-beta mRNA expression is decreased in tumors with LKB1 loss, and TGF-beta can cause down-regulation of the CREB/FOXO3 transcriptional cluster 82
3.3. Association between patient outcome and LKB1 loss 89
3.4. Restoring wild-type LKB1 in cell lines harboring mutations slows growth and attenuates the expression of the LKB1-deficient gene signature 94
3.5. Expression of wild-type LKB1 in A549, H2122, or HeLa cell lines decreases the expression of the genes in the CREB transcriptional node 95
4.1. LKB1 loss score is associated with sensitivity to MEK inhibitors 112
4.2. Restoring LKB1 confers resistance to MEK inhibition 114
4.3. Influence of MEK inhibition on gene expression for three previously published MEK signatures and the LKB1 loss signature 115
4.4. Restoring LKB1 alters AKT and FOXO3 phosphorylation 117
4.5. Signaling through MEK/ERK and PI3K/AKT pathways regulates mTOR activation 119
4.6. Signaling through MEK/ERK and PI3K/AKT pathways regulates apoptosis 120

LIST OF TABLES

Table Page

2.1. Association of LKB1 mutations with LKB1 loss score in cell lines excluding NSCLC 53
2.2. Association of LKB1 mutations with LKB1 loss score in TCGA cancer cohorts other than lung adenocarcinoma 55
2.3. Association of very low LKB1 expression with LKB1 loss score in TCGA cancer cohorts other than lung adenocarcinoma 57
3.1. Results from Gene Set Enrichment Analysis of NRF2-associated gene cluster 77
3.2. Results from Gene Set Enrichment Analysis of Mitochondria/mTOR-associated gene cluster 81
3.3. Results from Gene Set Enrichment Analysis of CREB/FOXO-associated gene cluster 84
3.4. Association of LKB1 loss with other somatic mutations in lung adenocarcinoma 87
3.5. Association of LKB1 loss with clinical variables in lung adenocarcinoma 90
3.6. Association of LKB1 loss with copy number alterations in lung adenocarcinoma 91
3.7. Association of LKB1 loss with differences in microRNA expression in lung adenocarcinoma 92
3.8. Association of LKB1 loss with differences in protein expression and phosphorylation in lung adenocarcinoma 93

CHAPTER I

INTRODUCTION

Discovery of the LKB1 Tumor Suppressor

LKB1 was first identified as a tumor suppressor in 1998 when it was found that mutations in LKB1 were associated with the rare inherited disease Peutz-Jeghers syndrome (PJS). This syndrome was first described by Johannes Peutz (Peutz, 1921) and again in 1949 by Harold Jeghers et al (Jeghers et al., 1949), who documented families in which several children were affected by widespread polyposis throughout the gastrointestinal tract, including large and small intestines and the nasopharynx; affected individuals also showed characteristic mucocutaneous pigmentations on the face, lips and mouth, as well as on the hands and in the rectal mucosa. The polyps are best characterized as benign hamartomous lesions, representing disorganized polyclonal proliferation of the multiple cell types normally present in the gastrointestinal mucosa. This syndrome exhibits an autosomal dominant pattern of inheritance and is associated with increased risk of malignant cancers within the GI tract, pancreas, breast, and lung (Hearle, 2006). Analysis of copy number changes within PJS hamartomas revealed frequent deletion of portions of chromosome 19, and these specifically affected the chromosome inherited from the parent unaffected by PJS (Hemminki et al., 1998). Further linkage analysis showed that the affected region of the chromosome – 19p13 – included the gene STK11/LKB1, which was affected by frame shift mutations, nonsense mutations, truncations, and point mutations across multiple PJS families (Hemminki et

al., 1998). This established LKB1 as a novel tumor suppressor, which was subsequently found to be inactivated by somatic mutations in approximately 30% of lung cancer (Sanchez-Cespedes et al., 2002) and approximately 20% of cervical cancer (Wingo et al., 2009), and in multiple other cancer types at lower prevalence. At the time of these original discoveries, LKB1 had been identified as a serine-threonine kinase, but little was known about its biological functions.

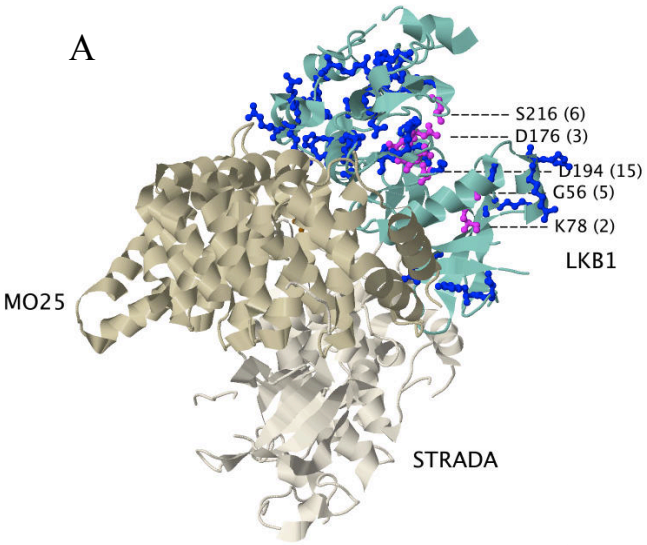
A breakthrough in the understanding of LKB1 function came in 2003 and 2004, when three independent groups found that LKB1, in a complex with the pseudo-kinase STRADA and the adaptor protein MO25, acts as an upstream regulator of the adenylate monophosphate-activated protein kinase (AMPK) (Hawley et al., 2003; Shaw et al., 2004a; 2004b; Woods et al., 2003). AMPK is a key regulator of cellular as well as organismal metabolism. Thus, the identification of LKB1 as the key kinase upstream of AMPK disclosed a novel link between metabolic dysregulation and cancer. A large body of literature over the past ten years focuses on the importance of LKB1 and AMPK in cancer, and much has been learned about the specific pathways and phenotypes regulated by these important genes.

Structure, Function, and Regulation of LKB1

LKB1 is a serine-threonine kinase that has been evolutionarily conserved throughout animal species, with significant homology to Xenopus embryologic gene XEEK1 and the polarity-regulating gene Par-4 in C. Elegans (Shackelford and Shaw, 2009). In contrast to many kinases, especially the tyrosine kinases, which require autophosphorylation to become active, LKB1 appears to require the formation of a

heterotrimeric complex with a scaffolding protein, MO25, and STRAD-alpha, a pseudo-kinase, in order to achieve catalytic kinase activity (Zeqiraj et al., 2009). The crystal structure of this complex was published in 2009 (Zeqiraj et al., 2009), revealing that the heterotrimer adopts a globular conformation with interactions between each of the three proteins (Fig. 1.1a). A catalytic cleft in LKB1 binds ATP and allows docking of substrates for phosphorylation.

In Peutz-Jeghers syndrome, hamartoma formation occurs with haplo-insufficiency of LKB1 and the remaining wild-type allele is infrequently disrupted. However, in human tumors somatic loss of both LKB1 alleles appears to be required, and western blot analysis of cell lines with LKB1 loss typically reveals either no protein expression or expression of mutated LKB1 protein exclusively. Disruption of LKB1 can take place by a variety of mechanisms. Loss of LKB1 expression can occur due to homozygous deletion (Gill et al., 2011; Matsumoto et al., 2007) or methylation (Esteller et al., 2000). Mutations affecting LKB1 can result in a loss of expression with splice site, nonsense, or frame shift mutations (Fig. 1.1b). Intragenic deletions affecting one or more exons have also been demonstrated (Matsumoto et al., 2007). Somatic missense mutations occur in lung cancer and are found throughout the entire protein. Because of the existence and prevalence of multiple mechanisms of tumor suppressor inactivation, I will refer to these collectively throughout this work as instances of 'LKB1 loss,' or 'LKB1-deficient tumors,' with the presumption – supported by our later findings – that these alterations yield similar tumor phenotypes.



B

Type of mutation	Affected tumors n (%)
Frameshift	84 (32)
Nonsense	61 (24)
Missense	72 (28)
Splice site	15 (6)
Other	27 (10)

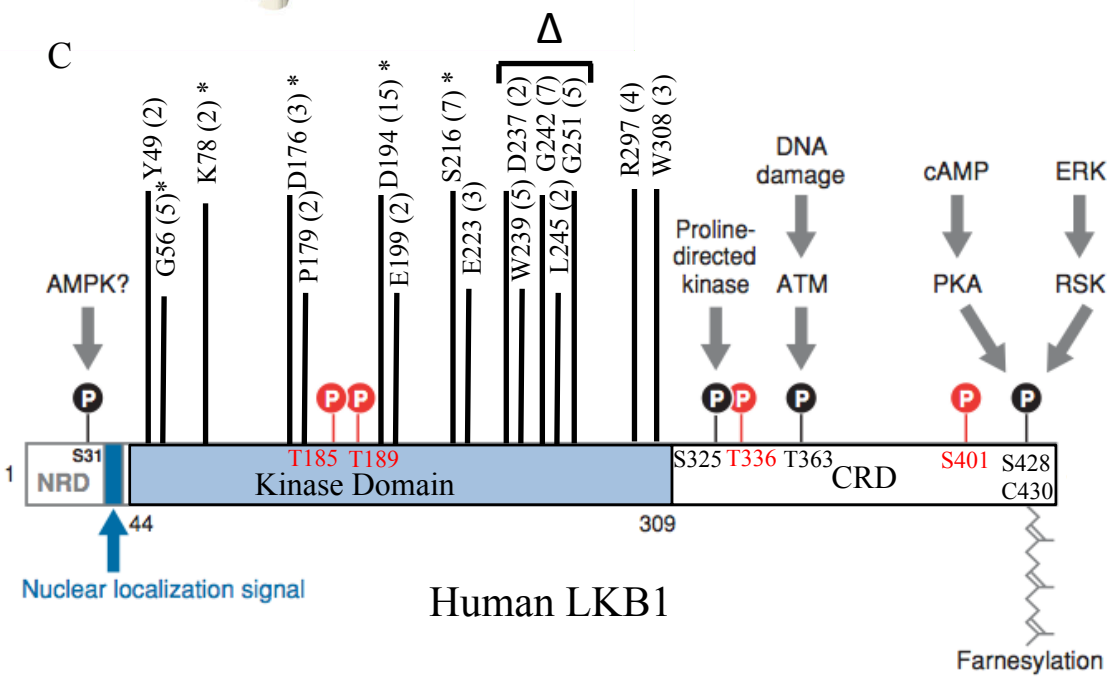


Figure 1.1 Structure of LKB1 and disrupting mutations observed in cancer.
A, Crystal structure of LKB1 in trimeric complex with MO25 and STRAD-alpha (Zeqiraj et al, 2009). Amino acids altered in cancer are marked, with blue coloration for sites altered in a single case, and magenta marking recurrently altered sites affecting the catalytic cleft. **B,** Types of mutations affecting LKB1 in NSCLC, by mutation class and prevalence. **C,** Schematic of LKB1 representing protein domains, phosphorylation sites, and locations of all sites mutated in more than one case of lung cancer or Peutz-Jeghers syndrome (modified from Alessi et al, 2006). Numbers in parentheses indicate number of mutations affecting site. CRD: C-terminal regulatory domain; NRD: N-terminal regulatory domain. Asterisk indicates known sites known to disrupt catalytic activity. Delta indicates a cluster of mutations in a region that disrupts protein stability and/or STRAD/MO25 complex formation (Zeqiraj et al, 2009).

In my analysis of published LKB1 sequencing studies of human tumors, roughly 30% of discovered mutations were frameshift, 30% nonsense, 30% missense, and 10% affected splice sites or were not fully described (Fig. 1.1b). However, intragenic deletions, and epigenetic silencing were not directly assessed in these studies. Chromosomal loss is also assessed by copy number profiling, but in contrast to discrete alterations observed in DNA sequence, this is a continuous numeric variable, and should not be used to conclusively determine homozygous deletion without additional evidence. Thus, I will refrain from giving an estimate of the frequency of deletions based on genomic data. Gill et al examined loss of the LKB1 locus using chromogenic in situ hybridization and concluded that single copy loss of LKB1 was present in more than half of lung adenocarcinomas, and homozygous deletion affected 28% of samples (Gill et al., 2011).

Missesnse mutations affecting the catalytic cleft are quite common, with aspartate 194, which binds a catalytic magnesium ion, being the single most common site of mutation, representing the site of 15% of LKB1 missense mutations (Zeqiraj et al., 2009) (Fig. 1.1a,c). Such catalytic mutations result in decreased kinase activity and attenuated phosphorylation of downstream targets such as AMPK, but with preservation of the heterotrimeric complex. Other recurrent sites of mutation affect the stability of the LKB1-STRADA-MO25 complex, as shown by the resultant loss of downstream AMPK activity and an inability of mutated LKB1 to immunoprecipitate STRAD-alpha and MO25 (Zeqiraj et al., 2009).

LKB1 is also regulated by phosphorylation at a number of amino acid residues, which is known to affect nuclear/cytoplasmic localization, but could also affect specificity of substrate binding, heterotrimer stability, or kinase activity (Fig. 1.1c). The tumor suppressive effects of LKB1 and its ability to activate AMPK appear to be dependent on cytoplasmic localization. Phosphorylation of serine 307 by protein kinase C (PKC) has been found to induce nuclear localization and function (Xie et al., 2009). Similarly threonine 428 in the carboxyl terminal tail of LKB1 is phosphorylated by multiple kinases, including ERK1/2, RSK, PKC, and PKA, representing a point at which a number of signaling pathways can impact the function of LKB1 (Alessi et al., 2006; Sapkota, 2001; Xie et al., 2008). Oncogenic BRAF has also been shown to induce phosphorylation of LKB1 at this site due to elevated ERK1/2 activity, and this results in attenuation of LKB1 activity and decrease in the activation of AMPK in response to energy stress (Bin Zheng et al., 2009; Esteve-Puig et al., 2009). Interestingly, the carboxyl terminus of LKB1 has a number of positively charged amino acids that are affected by somatic mutations, suggesting that these mutations could result in electrostatic effects similar to phosphorylation of threonine 428.

Other potential modes of regulation have also been demonstrated. Transcriptional regulation of the LKB1 gene itself affects its expression and could affect its level of functional activation; this has been shown to be dependent in part on the transcription factors SP1 and FOXO3 (Lützner et al., 2012a; 2012b). Activity of LKB1 can be regulated by altering its interactions with binding partners STRAD-alpha and MO25-apha. The expression MO25-alpha is regulated in part by mir-451, which has been shown to inhibit MO25-alpha expression and thus downregulate LKB1 activity in glioma to

allow adaptation to metabolic stress (Godlewski et al., 2010). Experimentally induced mutations in STRAD-alpha and MO25 can also result in loss of complex formation and LKB1 activity (Zeqiraj et al., 2009). Homozygous deletion of STRAD-alpha results in a rare genetic disorder – polyhydramnios, megalencephaly and symptomatic epilepsy – resulting from loss of LKB1 activity in neurons and the developing brain (Orlova et al., 2010; Puffenberger et al., 2007). Although this would appear to be a potential additional mechanism for loss of LKB1 activity, somatic loss of STRADA or MO25 has not been reported in cancer. Mutations in downstream AMPK family members can be observed in exome sequencing data from the Cancer Genome Atlas (TCGA), but there is no evidence that they are functionally significant and our analysis does not show them to induce a similar phenotype as LKB1 loss (data not shown). The multiple levels at which LKB1 can be regulated reflect the importance of this gene in a variety of pathways, but also show the complexity that can arise in understanding its effects. Many of these intricacies are still poorly understood.

LKB1 Regulation of AMPK

LKB1 exerts its actions within the cell by phosphorylating a family of downstream kinases in the AMPK family, resulting in activation (Fig. 1.2). One of the most studied targets of LKB1 is AMPK, which was known as an important regulator of cellular metabolism and glucose homeostasis prior to its connection with LKB1 (Hardie and Alessi, 2013). The finding that LKB1 was the main upstream activator of this protein established a new link between cancer and metabolism. AMPK is activated by increases in AMP and ADP concentrations within the cell in response to energy stress, such as

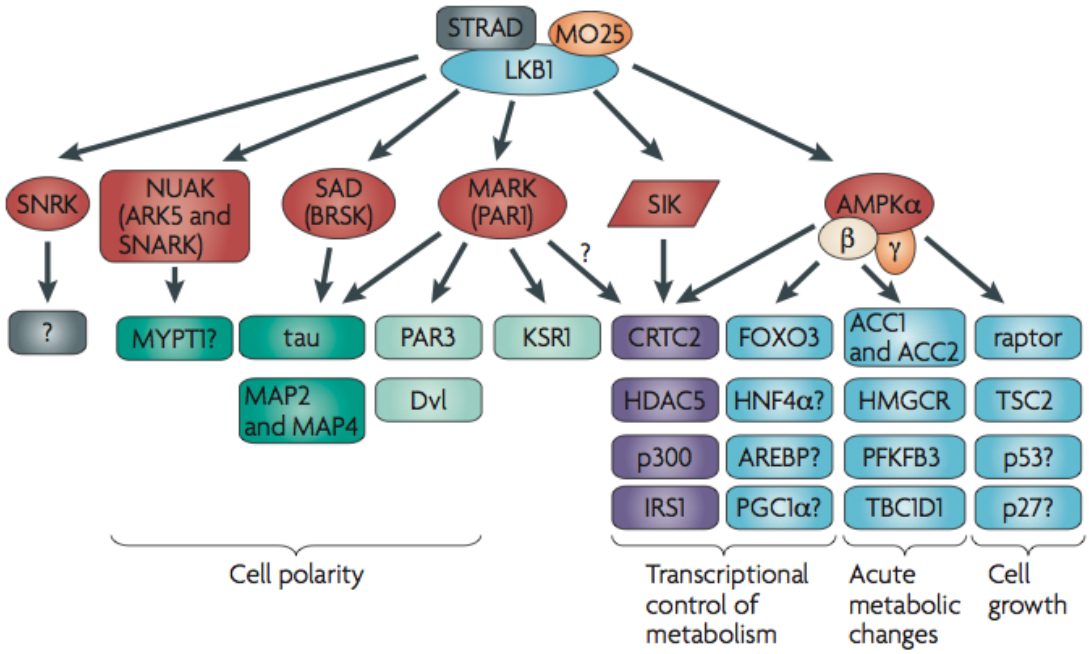


Figure 1.2 Schematic of LKB1 interactions with AMPK family members and their downstream effects.
LKB1 induces the activation of 14 kinases in the AMPK protein kinase family. Only the activation of AMPK itself is affected by nutrient deprivation. AMPK inhibits mTOR and also regulates metabolic effects both acutely and through longer term transcriptional mechanisms. LKB1's effects on organismal development, cell polarity and motility are thought to be regulated by downstream kinases other than AMPK. (modified from Shackelford et al, 2009)

starvation. The mechanism by which AMP levels regulate AMPK activation has been elucidated (Hardie and Alessi, 2013; Xiao et al., 2007; 2011). This involves four adenine nucleotide binding sites within the gamma subunit of the protein, which bind ATP, ADP, or AMP. Nucleotide occupancy for two of the sites varies according to changes in the intracellular concentrations of these species. Within a stressed cell, exchange of ATP for ADP or especially AMP causes conformational changes of the AMPK protein structure that both increase phosphorylation and decrease dephosphorylation of threonine 172 within the activation loop of the protein (Hardie and Alessi, 2013; Xiao et al., 2007; 2011). This activation induces downstream phosphorylation of the targets mentioned

above, leading to a concerted effort to restore the energy balance to homeostasis by attenuating energy-consuming anabolic pathways, increasing catabolism, and slowing cell growth.

AMPK phosphorylates a number of substrates, including metabolic enzymes, such as acetyl-CoA carboxylase (ACC) and 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR), as well as transcriptional regulators of metabolism such as peroxisome proliferator-activated receptor gamma, coactivator 1 alpha (PGC1A), carbohydrate-responsive element-binding protein (ChREBP), and sterol regulatory element-binding proteins (SREBP) (Alessi et al., 2006; Hardie and Alessi, 2013; Shackelford and Shaw, 2009). These alterations have been shown to affect energy balance in response to stress; lacking this adaptive response can leave cells susceptible to death in such conditions. The LKB1 mutant lung cancer cell line, A549, undergoes approximately 80% cell death after 24 hours of glucose deprivation, which is almost eliminated when LKB1 is restored. This effect was mediated in part by reactive oxygen species (ROS), and cell death could be decreased either by knocking down the AMPK targets ACC1 or ACC2, or by ROS scavengers. Inhibition of mTOR did not affect this phenotype, showing that the metabolic and growth signaling effects of LKB1 are distinct and have different cellular consequences. In cancer, loss of AMPK phosphorylation and activity is a defining characteristic of LKB1-deficient tumors, and this results in activation of the oncogenic mTOR pathway (Gwinn et al., 2008; Shaw and Cantley, 2006a), but also confers greater susceptibility to metabolic stress, for instance by the mitochondrial inhibitor phenformin (Jeon et al., 2012; Shackelford et al., 2013). Interactions between LKB1, AMPK, mTOR and other pathways are complex and are summarized briefly in Fig. 1.3.

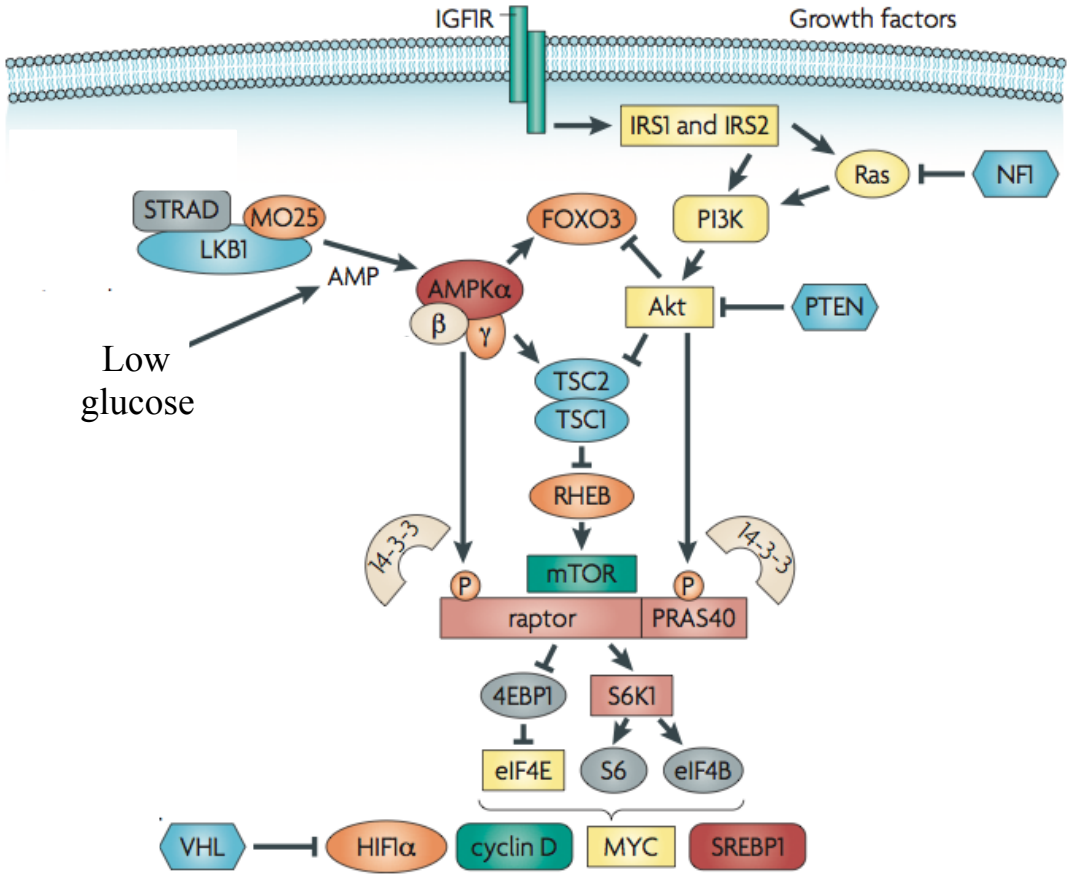


Figure 1.3 Schematic of LKB1 and AMPK interactions with mTOR pathway. LKB1 exerts its inhibitory effects on the mTOR pathway by phosphorylating AMPK. This phosphorylation is promoted by conformational changes in AMPK when AMP levels are increased. AMPK phosphorylates raptor directly and also activates the TSC2 tumor suppressor which further inhibits mTOR activation. Akt phosphorylation has opposite effects, inhibiting TSC2 and phosphorylating mTOR complex member PRAS40 to induce activation of the pathway (modified from Shackelford et al, 2009).

LKB1 and the mTOR Pathway

The mechanistic target of rapamycin, or mTOR, pathway is an important regulator of cellular and organismal growth and homeostasis. mTOR is a serine and threonine kinase in the PI3-kinase family, and its effects are carried out as part of two distinct large multi-protein complexes, TORC1 and TORC2 (Laplante and Sabatini, 2012). Both of these complexes share the mTOR kinase subunit, and associated proteins DEPTOR, Tti1, and Tle2. In TORC1 complexes, these associate with regulatory-

associated protein of mTOR (raptor) and proline rich Akt substrate 40 (PRAS40), whereas in TORC2 the shared complexes associate with rapamycin insensitive companion of mTOR (rictor), mSin1/2 and proto1/2. These two complexes carry out very different roles within the cell, with each phosphorylating different substrates to induce distinct downstream effects and each being regulated by different upstream mechanisms. TORC1 is known to be regulated by upstream signals including growth factors, cellular energy state, amino acid availability, hypoxia, and stress. Downstream effects include induction of protein synthesis by activating ribosomal S6 kinase and inhibition of 4E-BP1, activation of lipid synthesis, increase in energy metabolism by induction of HIF1a and increases in mitochondrial biogenesis through PPARGC1A (Cunningham et al., 2007), and inhibition of autophagy. TORC2, on the other hand, phosphorylates protein kinase C, Rho and Rac proteins to regulate cytoskeletal organization and motility, and activates the Akt and SGK1 kinases, which induce cellular survival mechanisms and can attenuate metabolic transcription by inhibiting FOXO transcription factors.

Activation of TORC1 is regulated by direct phosphorylation of mTOR or PRAS40 and also by a variety of processes that converge on the tuberous sclerosis complex TSC1 and TSC2. These proteins are tumor suppressors that inhibit mTOR activation by decreasing the activation of a small RAS-like protein, Rheb, through their GTPase activating activity. Thus, activators of TSC1/2 lead to TORC1 inhibition, while inhibitory effects lead to TORC1 activation. In addition to directly inhibitory phosphorylation of the mTOR complex, AMPK also directly activates TSC2 by phosphorylation, which further attenuates the mTOR pathway. In contrast, TSC2 is inhibited by AKT, ERK, and RSK kinases, causing the opposite effects.

Feedback effects are also important in the regulation of these pathways. AKT is a key activator of the TORC1 pathway, both by the inhibitory phosphorylation of TSC2 and by direct phosphorylation of the TORC1 component PRAS40 (Shaw and Cantley, 2006a). Correspondingly, activation of TORC1 is constrained by feedback inhibition of the PI3K/AKT pathway, which is thought to primarily involve inhibitory phosphorylation of the insulin receptor substrates IRS1 and IRS2, which transduce signals from upstream growth factor receptors to induce activation of PI3K signaling, resulting in AKT phosphorylation. On the other hand, activation of TSC2 has opposite effects on TORC1 and TORC2; TORC1 is inhibited, whereas TORC2 is stimulated by TSC2 (Yang et al, 2006; Huang and Manning, 2009). This leads to activation of AKT, resulting in feedback inhibition of TSC2 and restoration of balanced TORC1 activity. Thus, in TSC2 deficient tumors and cells, this feedback mechanism is largely abrogated, and TORC1 is constitutively active, while TORC2 is attenuated (Huang and Manning, 2009; Yang et al., 2006), as is downstream AKT signaling (Gan et al, 2010). Because mTOR and AKT are both oncogenic pathways whose activities are linked by multiple feedback mechanisms it can be difficult to predict how they will interact in the context of a given cancer, or they might respond to perturbations, such as targeted inhibition of a pathway. Understanding these interactions will be important for the rational selection of targeted agents for particular patients, and for determining which combined treatments may be particular beneficial.

Regulation of Other Processes by LKB1

AMPK-alpha belongs to the AMPK family of protein kinases, which includes AMPK-alpha and -beta subunits as well as 12 other related kinases: NUAK1, BRSK, SIK1, MARK1, SNRK, and their paralogs (Shackelford and Shaw, 2009). This family of kinases exhibits significant homology in the activation loop region, and it was subsequently shown that all 14 proteins are phosphorylated by LKB1. Of note, because the gamma subunit of AMPK is thought to be responsible for sensing metabolic stress, the activation of other LKB1 targets does not seem to be dependent on cellular nutrient state. AMPK-family kinases perform a variety of regulatory roles within the cell, including interaction with various transcription factors and chromatin remodeling proteins (Marignani, 2001; Walkinshaw et al., 2013), regulation of the cytoskeleton, and interactions with signaling pathways such as TGF-beta (Katajisto et al., 2008; Londesborough et al., 2008; Vaahtomeri et al., 2008). Through these interactions, in addition to its metabolic effects, LKB1 serves as a 'master regulator' of a number of phenotypes, including embryonic development (Lo et al., 2012; Ossipova et al., 2003), cellular polarity and motility (Alessi et al., 2006; Shackelford and Shaw, 2009). LKB1 has been shown to regulate cell morphology and epithelial integrity through AMPK, SIK1 and likely other kinases (Amin et al., 2009; Eneling et al., 2012; Lee et al., 2007; Mirouse et al., 2007; Partanen et al., 2007; 2012). Effects on polarity and epithelial organization that are mediated through AMPK are affected by energy stress (Lee et al., 2007; Mirouse et al., 2007). LKB1-induced effects on polarity can be pronounced; restoration of LKB1 activity has been shown to induce polarization within hours to previously depolarized cells even in single cell suspension (Baas et al., 2004). This was

also associated with the formation of brush border-like membrane protrusions in this colon cancer line, which was subsequently found to involve Mst4 and Ezrin (Klooster et al., 2009). Furthermore, the activity of oncogenic pathways such as c-Myc (Partanen et al., 2007), and the epithelial-mesenchymal transition (Roy et al., 2010) are altered by these effects as well.

Somatic LKB1 Loss in Human Tumors

As mentioned previously, germline mutations in LKB1 give rise to Peutz-Jeghers syndrome, which is characterized by mucocutaneous pigmentation changes and the development of hamartomous polyps throughout the gastrointestinal tract. Hamartomas are polyclonal proliferations of both epidermal and mesenchymal cells that are generally considered benign, although in some cases adenocarcinomas have developed within these growths. Moreover, the lifetime risk of cancer among individuals affected by Peutz-Jeghers syndrome has been shown to be approximately 20-fold greater than that of the general population. Carcinomas arise most commonly in the gastrointestinal tract, but may also arise in the lung, breast, pancreas, uterus, ovary, cervix, and testes (Boardman et al., 1998; Giardiello et al., 1987; 2000; Spigelman et al., 1989).

After LKB1 was identified as the causative tumor suppressor responsible for this syndrome, efforts have been made to estimate the prevalence of somatic loss of LKB1 across many different tumor types. Many of these studies have been somewhat small, and a variety of methods have been used to ascertain LKB1 loss, including manual sequencing of each exon for mutations and determination of intragenic deletions (Matsumoto et al., 2007), single strand conformation polymorphism, determination of

loss of heterozygosity, and examination of LKB1 protein expression by immunohistochemistry or western blot. Recently, high throughput sequencing of LKB1 and other genes has been achieved across large numbers of tumors and cell lines using, for instance, exon capture and next generation sequencing.

Non-small cell lung cancer has been shown to have the highest prevalence of LKB1 loss among human cancers. Ji et al found alterations of LKB1, including both mutations and deletions, in 27 of 80 lung adenocarcinomas (34%), as well as in 19% of squamous cell and 9% of other lung tumors (Ji et al., 2007). Koivunen et al determined that somatic mutations in LKB1 affected 25 of 143 lung adenocarcinomas (17%) in an American population, but only 5% of tumors in an Asian cohort (Koivunen et al., 2008). Matsumoto et al reported a characterization of LKB1 mutations in another Asian cohort with a prevalence of 7%, and also documented LKB1 mutations and deletions in 20 of 51 (39%) of NSCLC cell lines (Matsumoto et al., 2007). Because of its high prevalence in NSCLC, much of the research, and particularly the clinical associations with LKB1, have focused on this tumor type. Although it was reported in Matsumoto et al that LKB1 mutations were significantly more likely to occur in poorly differentiated tumors, Koivunen et al showed that there was no significant association between LKB1 mutation and tumor stage, or patient outcome. The lower prevalence of LKB1 mutations among Asian cohorts was statistically significant in Koivuven et al, and this low prevalence has been observed in multiple studies (Gao et al., 2010; Okuda et al., 2010; Sun et al., 2010; Suzuki, 2012), suggesting that either environmental factors or genetic differences can affect the genetic makeup of these tumors. LKB1 loss is strongly associated with smoking history, with most tumors exhibiting LKB1 loss occurring in patients with

greater than ten pack-years of cigarette exposure, and LKB1 mutations occurring infrequently among never smokers (Koivunen et al., 2008; Matsumoto et al., 2007). Finally, LKB1 mutations in lung cancer tend to co-occur with KRAS mutations at a rate greater than would be expected by chance; on the other hand, LKB1 and EGFR are rarely mutated in the same tumors, suggesting either redundancy, or more likely antagonism between their effects.

LKB1 loss in other tumor types has been studied in less detail but is significantly less prevalent than in lung cancer. LKB1 mutations were identified in 15 of 74 (20%) of cervical cancer and these mutations conferred significantly worse prognosis (Wingo et al., 2009). In breast cancer two large studies have identified only rare LKB1 mutations, occurring in one of 687 and two of 951 patients (Cancer Genome Atlas Network, 2012; Loi et al., 2013). LKB1 mutations have been observed in cell lines derived a wide variety of primary tumor types, including lung, cervix, melanoma, colon, small cell lung cancer, head and neck squamous cell carcinoma, prostate cancer, cholangiocarcinoma, hepatocellular carcinoma, ovarian cancer, thyroid cancer, renal cell carcinoma, uterine cancer, pancreatic cancer, certain leukemias, and breast cancer. Limited reports of primary tumors for most of these sites show low prevalence of LKB1 mutations, which for simplicity sake we will estimate to be less than 5% for most cases. More definitive characterization of LKB1 loss in these settings will depend on larger and more systematic efforts, especially those currently underway through The Cancer Genome Atlas (TCGA). Of the tumor types that have been characterized thus far by the TCGA, it appears that LKB1 mutations are found in less than 2% of colon, breast, AML, clear cell renal, colorectal, melanoma, glioblastoma, head and neck, ovarian, thyroid, and uterine cancer.

However, as we will demonstrate, reliance solely on sequence mutations probably underestimates the actual prevalence of LKB1 loss by a factor of two or more because a number of additional mechanisms can lead to functionally equivalent loss of the tumor suppressor activity.

Mouse Models of LKB1 Loss

The interactions between LKB1, AMPK and mTOR provide a mechanistic link between the tumor suppressor and both metabolic and oncogenic phenotypes that may be important in the biology of LKB1-deficient tumors. However, as described above, LKB1 is an important regulator of many additional downstream mediators and has been shown to influence diverse phenotypes. Therefore mouse models of LKB1 loss have been developed to gain better understand the effects of LKB1 loss on tumorigenesis and other processes in vivo.

Global deletion of LKB1 results in midgestational death, resulting in embryos that show significant neural tube defects, vascular and placental abnormalities, and mesenchymal cell death (Ylikorkala, 2001). Mice with heterozygous deletion of LKB1, however, are viable and develop gastrointestinal polyposis that is consistent with the hamartomous polyposis seen in human Peutz Jehgers syndrome. Study of the polyps that occur in this model has demonstrated increased activity of mTOR, hypoxia inducible factor 1a (Shackelford et al., 2009), and cyclooxygenase 2 (Rossi et al., 2002). Treating mice with oral rapamycin substantially reduced the polyp burden in these mice, and this could be a therapeutic strategy for patients with Peutz Jeghers syndrome (Robinson et al., 2009; Shackelford et al., 2009; Wei et al., 2003). Increased COX2 staining has also been

observed in human Peutz Jeghers polyps, and both genetic and pharmacologic inhibition of COX2 decreased polyp proliferation in the mouse model (Udd et al., 2004). Interestingly, tissue specific knockout of LKB1 using a smooth muscle promoter resulted in LKB1 loss only in mesenchymal cells, leaving LKB1 function intact in the gut epithelium, which was nevertheless sufficient to produce hamartomous polyps. The LKB1 deficient mesenchymal cells produced less TGF-beta, resulting in decreased TGF-beta signaling within the epithelium and enhanced epithelial proliferation, suggesting a stromal paracrine effect on the neighboring epithelial component of these polyclonal tumors, rather than direct effects of LKB1 loss (Katajisto et al., 2008). This observation also suggests that the effects of LKB1 loss on the biology of human tumors may differ substantially from the phenotypes observed in PJS polyps.

The murine model for PJS polyposis also results in spontaneous formation of hepatocellular carcinomas in mice over one year old (Nakau et al., 2002), as well as development of uterine cancers (Contreras et al., 2008). The hepatocellular carcinomas were shown to exhibit loss of both LKB1 alleles (Nakau et al., 2002). However, other tumor types were not observed. To provide models for the phenotypes associated with LKB1 loss in human tumors, LKB1 has been combined with loss of other tumor suppressors or activation of oncogenes in a variety of experimental settings.

Phosphatase and tensin homolog (PTEN) is another tumor suppressor that results in a gastrointestinal polyposis disease called Cowden's syndrome; it acts as a phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase, which attenuates this secondary messenger of oncogenic phosphatidyl inositol 3 kinase (PIK3CA) signaling. When a heterozygous hypomorphic LKB1 allele is combined with PTEN heterozygous loss,

accelerated polyposis results with greater polyp burden (Huang et al., 2008), as well as the development of follicular B-cell lymphomas (García-Martínez et al., 2011). Both of these effects could be reversed by targeting the mTOR pathway. Similarly, when heterozygous LKB1 loss is studied in mice heterozygous for the tumor suppressor p53 – analogous to the human Li-Fraumeni syndrome – gastrointestinal polyp development is accelerated, and the formation of aggressive neoplasia is also accelerated and more diverse tumor types are observed compared to the p53 heterozygous background (Wei et al., 2005).

To give more control over the context of these genetic lesions, and to achieve complete loss of LKB1 rather than heterozygous loss, genetically engineered murine models have been developed in which conditional LKB1 loss is achieved using a cre-recombinase loxP system. LoxP sites are situated around LKB1 exons three through six, and allow deletion of this region when cre-recombinase is expressed (Bardeesy et al., 2002). This allows LKB1 loss or alteration of other targeted genes to be restricted to particular organ systems using tissue specific promoters or locally introduced viruses. Restricted LKB1 loss targeted to the pancreas (Hezel et al., 2008), prostate (Pearson et al., 2008), breast (McCarthy et al., 2009), and endometrium (Contreras et al., 2008) resulted in invasive carcinomas in breast and endometrial tissue, and lower grade noninvasive neoplasias in the prostate and pancreas. Disordered cellular polarity was observed in the prostate and pancreatic lesions, and the prostatic intraepithelial neoplasia exhibited increased WNT, AKT, and mTOR signaling (Pearson et al., 2008).

Targeted deletion of LKB1 in hematopoietic stem cells resulted in hematopoietic failure, but did not lead to leukemic transformation (Gan et al., 2010a; Gurumurthy et al.,

2010; Nakada et al., 2010). Upon losing LKB1 these cells were initially induced to proliferate and expand, but were subsequently depleted. Although mTOR activation was demonstrated and alterations in metabolism were also observed, the failure of hematopoiesis could not be reversed by pharmacologic activation of AMPK or by mTOR inhibition; AMPK knockout did not produce the same phenotype. Rather the effect was shown to be due to failure of normal chromosomal segregation, which resulted in aneuploidy (Nakada et al., 2010).

Murine Model of LKB1/KRAS Mutant Lung Cancer

Conditional deletion of LKB1 in the lung has been achieved using the murine system described in the previous section (Bardeesy et al., 2002), with lung targeting achieved using inhaled adenovirus or lentivirus that expresses Cre recombinase. This system had previously been engineered to allow specific knock-in of the oncogenic mutant G12D KRAS allele, which resulted in the development of many KRAS-driven lesions. These were largely minimally invasive and could best be classified as adenomas. When LKB1 was deleted in a similar manner, no lung tumors developed. However, KRAS and LKB1 are often simultaneously altered in human tumors, with roughly 50% of KRAS mutant tumors exhibiting LKB1 loss in our analyses; when the mice were crossed such that Cre-recombinase could induce alterations of both of these genes simultaneously, aggressive carcinomas developed with a high propensity for metastasis. They exhibited histological diversity, with some tumors having adenocarcinoma differentiation, while others showed squamous cell morphology or mixed adenosquamous phenotype (Ji et al., 2007). This was in contrast to the tumors that resulted when KRAS

was combined with either p53 or p16 loss, which demonstrated similar aggressiveness, multiplicity, and metastasis, but gave rise exclusively to adenocarcinomas (Ji et al., 2007).

This model for LKB1/KRAS mutant tumors has been studied in more detail to gain insight into the behavior of LKB1 deficient lung tumors, which could lead to novel treatment approaches. Gene expression analysis of resulting KRAS-driven tumors allowed characterization of gene expression seen in KRAS/LKB1 tumors to be compared to those associated with KRAS/p53, KRAS/p16, or KRAS alone. The initial publication focused on the role of NEDD9 as an LKB1-regulated gene that induced invasion and metastasis in these tumors (Feng et al., 2012; Ji et al., 2007). Subsequent work expanded on these findings, characterizing both gene expression and protein expression in more detail, and comparing the phenotypes of KRAS/LKB1 metastases to those of primary lesions. This work identified activation of TGF-beta and SRC pathways as being important for the development of metastasis (Carretero et al., 2010).

Because LKB1 plays important roles in regulating tumor metabolism, the resulting changes in metabolic processes, and especially the altered response to metabolic stress, may provide novel targeted strategies that could be used to treat LKB1 deficient tumors. LKB1 and AMPK play important roles in governing glucose homeostasis, and have been shown to be at least partially responsible for the therapeutic effects of metformin (Shaw, 2005). Furthermore, retrospective analysis of clinical trials in diabetes has shown that metformin reduces cancer incidence and mortality (Evans et al., 2005; Landman et al., 2010; Libby et al., 2009; Margel et al., 2013) and improves outcome after chemotherapy treatment (Jiralerspong et al., 2009; Tan et al., 2011). The murine model

demonstrated that tumors lacking LKB1 could not activate a protective AMPK response after phenformin inhibition, and this resulted in decreased tumor size, decreased Ki67 staining, and increased necrosis and apoptosis compared to KRAS alone or KRAS/p53 tumors. Defective autophagy was implicated in mediating this effect, and restoration of LKB1 in cell lines reversed the sensitivity (Shackelford et al., 2013). Therefore, this safe, cheap, and ubiquitous drug has potential as a metabolic inhibitor that could be used as a targeted therapy for LKB1 deficient tumors.

The KRAS/LKB1 mouse model has also been used to probe the susceptibility of these tumors to potential therapeutic approaches. There are currently no targeted agents that have been conclusively shown to be effective in either KRAS or LKB1 mutant tumors, and mutated KRAS is associated with resistance to EGFR-targeted agents; in fact, mutated KRAS and mutated EGFR are essentially mutually exclusive. However, mutant KRAS activates the RAF/MEK/ERK pathway, and MEK inhibition is therefore a potentially attractive targeted agent in these tumors. To determine whether LKB1 loss influenced the susceptibility of tumors to such treatment a 'mouse clinical trial' was conducted comparing response to docetaxel plus the MEK inhibitor selumetinib in murine tumors with mutant KRAS alone, KRAS/p53$^{-/-}$ and KRAS/LKB1$^{-/-}$ genotypes. This study showed that KRAS/LKB1$^{-/-}$ tumors had a decreased response to MEK inhibition, with less reduction in growth and lacking the induction of apoptosis that was seen in the other two genotypes (Chen et al., 2012).

Although these models provide an excellent means to study the function of LKB1 loss in tumors in an in vivo setting, they have not been compared directly to human tumors and it is unknown how well they reflect human disease phenotypes. An alternative

pre-clinical model system relies on the use of cell lines derived from cancers in human patients. These may better represent the heterogeneity of human tumors and likely better represent the complex evolutionary background that gives rise to human cancer, and have been shown to have similar genetic alterations as primary human lung tumors (Gazdar et al, 2010). However, they have been cultured over extended periods in permissive conditions with no similarity to the tissues in which they arose, and without the influence of cells in the microenviroment such as immune cells, fibroblasts, and vasculature.

Gene Expression Analysis

Many different phenotypes must be altered in the evolution of a tumor. These have been listed broadly in the classic review 'The Hallmarks of Cancer' and include generation of growth signals, loss of response to anti-growth signals, evasion of apoptosis, and tissue invasion and metastasis (Hanahan and Weinberg, 2000; 2011). These phenotypes are complex and result from the combined effects of many signaling pathways that can be distinct but can also interact and overlap. Although the concerted actions of proteins are the final executors of these phenotypes, the activity of a given protein is dependent on the level of its corresponding mRNA expression, and on modulators of its translation, activation, and degradation. Many signaling pathways within the cell ultimately affect the activation of transcription factors and other modulators of gene expression such as micro-RNAs. Thus, the study of gene expression patterns associated with a particular phenotype may identify genes that represent key downstream effectors of the phenotype, or may allow the inference of transcriptional programs that have been activated.

For the complex network of pathways and phenotypes that are affected by LKB1, predicting which potential effects will actually be observed in human tumors is challenging. The use of gene expression analysis may be particularly helpful in elucidating dysregulated pathways that can form the basis for further investigation. Indeed, this is a key rationale for the characterization of gene expression in the LKB1/KRAS murine model (Carretero et al., 2010; Ji et al., 2007). In our work, we make use of several gene expression datasets from human resected lung tumors in which LKB1 mutations have been determined by sequencing to determine LKB1 associated gene expression patterns and make inferences about the underlying biology and phenotypes of these tumors. Some introduction of gene expression analysis is therefore needed.

Proteins carry out most enzymatic, structural, and signaling functions within a cell. Translation of a protein is dependent on the quantity of its corresponding mRNA and on factors that influence its rate of translation. After a protein is translated, its activity is regulated by a large number of posttranslational factors. The expression of a protein is affected by its half-life but also by modifications such as ubiquitination that can specifically target it for degradation. Conformational changes, cellular localization, formation of complexes with other protein partners, proteolytic cleavage and maturation, and a variety of post-translational modifications such as phosphorylation, lipidation, oxidation, and acetylation all affect what a given protein does within a cell. Thus, in many cases, the expression of a given gene and the functional activation of its protein product may be only weakly correlated. On the other hand, the expression of a particular gene, or especially of a co-expressed set of genes, may be a quite accurate reflection of the activity of particular transcription factors or transcriptional programs within a cell.

The difficulty is in identifying which sets of genes are associated with which upstream phenotypes or transcription factors.

To elucidate these relationships it is useful to study the gene expression patterns of large collections of tumors. One of the first such studies to utilize this approach used DNA microarray applied to acute myeloid leukemias to demonstrate two distinct 'classes' of gene expression patterns within this disease (Golub, 1999). In lung cancer three early studies demonstrated that particular sets of genes were co-regulated – showing correlated increases or decreases in expression from patient to patient – that could define patient subsets with differential prognosis (Beer et al., 2002; Bhattacharjee et al., 2001; Garber et al., 2001). Some sets were linked with putative phenotypes, for instance a set of co-regulated genes were identified that were associated with neuroendocrine differentiation (Bhattacharjee et al., 2001). A larger study then characterized the gene expression patterns of 442 lung adenocarcinomas to determine gene sets that influenced a patient's prognosis (Shedden et al., 2008). An algorithm was used to define 100 'clusters' of genes with similar expression patterns, and each of these was then investigated as a single variable to determine its association with outcome. This served as a data reduction approach for the purposes of the study, but by parsing the resulting lists of genes it was clear that many clusters could be linked to particular phenotypes. By studying the genes comprising these clusters, one could identify epithelial to mesenchymal (EMT) associated, vasculature, interferon stimulated, and lymphocyte associated genes, as well as genes associated with proliferation and cell cycle progression, which had strong associations with prognosis.

To go further in understanding the underlying processes driving different gene sets, more sophisticated analyses are required. My work has taken advantage of a great many publicly available resources studying gene expression patterns in collections of tumors and tools for interpreting these findings. Some of the most useful sources of information are sets of tumors in which gene expression is characterized alongside additional molecular characterization such as copy number alterations, somatic mutations, or protein expression. Such studies can employ Sanger sequencing to determine mutations in a defined number of genes and array-based technology to examine gene expression and copy number (Chitale et al., 2009; Ding et al., 2008; Hayes et al., 2006; Selamat et al., 2012; Wilkerson et al., 2012). Recently, the TCGA has published large scale, systematic, molecular characterizations of many tumor types including lung squamous cell carcinoma (Cancer Genome Atlas Research Network et al., 2012) and adenocarcinoma (Cancer Genome Atlas Network, 2013). These employ exome capture next-generation DNA sequencing of more than 10000 genes, RNA sequencing for determination of gene expression, along with analysis of microRNA expression, copy number changes, and proteomic analysis of expression and phosphorylation of important proteins. These resources can allow particular gene expression changes to be linked to functional molecular alterations within a cancer – for instance the mutation of a known oncogene, or the amplification of a key transcription factor.

Another important aspect of interpreting gene expression patterns is in comparing results obtained through a given analysis with gene expression changes seen in other experiments. If two studies examine similar phenotypes, then that phenotype may be associated with the expression of a set of genes that would show up in both studies.

Examining the amount of overlap between the two studies determines whether this is the case. Some overlap by chance is expected; if the number of overlapping genes significantly exceeds what is expected by chance then the two studies may share a common effect. As an example, a set of 200 genes could be found to be upregulated in tumors expressing mutated KRAS, and a different set of, say, 100 genes might be downregulated after treatment of a cell line with a MEK inhibitor. If only three genes would be expected to overlap by chance, but instead 30 are observed then this could indicate that the two studies are similar in some way.

Sources of data abound with which to compare a gene set of interest. One useful collection of many diverse genesets has been compiled in the Molecular Signatures database (Liberzon et al., 2011). A user-friendly tool enables statistical comparisons with these gene lists and reports a ranked list of the most significant overlaps. Another source is the Connectivity Map, which characterizes gene expression changes in MCF7, PC3, and HL60 cell lines after four hours of treatment with one of over 1000 small molecules (Lamb et al., 2006). Additionally, directed searches to find studies addressing phenotypes of interest can be made using either the GEO or Array Express databases. These studies can then be analyzed and one can query personalized, user-defined gene sets.

Statistics based on the Mann-Whitney test can determine the significance of the overlap, or 'Gene Set Enrichment Analysis' can be employed for the significance of overlap with a ranked list (Subramanian et al., 2005). It should be noted that many nonspecific effects could lead to significant overlap between studies. For instance, if two inhibitors both decrease proliferation, then they will likely both affect the expression of cell-cycle genes, which could cause a statistically significant overlap in gene expression

changes. Even with a very significant P-value, one could not conclude from this that the inhibitors shared a common mechanism of action. Thus, results should be interpreted cautiously and should most often be construed as hypothesis generating. When possible, interesting relationships derived from gene expression analysis should be confirmed by appropriate experiments.

CHAPTER II

GENE EXPRESSION SIGNATURE OF LKB1 LOSS

Introduction

As we elaborated in the general introduction, LKB1 exerts complex roles within the cell. Its interactions with AMPK-alpha, through which it regulates metabolism and the mTOR pathway, are most familiar. However, LKB1 affects a variety of additional functions, including development, cell polarity and motility, chromatin and transcriptional regulation, and cell growth, by phosphorylation of 12 other members of the AMPK family (Alessi et al., 2006; Lizcano et al., 2004; Shackelford and Shaw, 2009), and these perform a variety of functions. Thus, LKB1 also affects development, cell polarity and motility, chromatin and transcriptional regulation, and cell growth through its effects on these downstream kinases. LKB1 is one of the most frequently altered genes in lung adenocarcinomas – our work shows that approximately 30-35% of these tumors exhibit loss. Thus, identifying targeted treatments to which LKB1 deficient tumors are susceptible would be valuable to patients suffering from this disease. Understanding the complex interactions between LKB1 and the various downstream pathways that it influences may help identify such targeted strategies and could determine feedback and resistance mechanisms that may differ between LKB1 wild-type and mutant tumors.

To study these processes, and the biology of LKB1 deficient tumors in general, genetically engineered murine models of LKB1/KRAS mutant lung cancer have been

developed. The resulting tumors are aggressive, metastasize readily, and exhibit diverse histological differentiation similar to that observed in human non-small cell lung cancer (Carretero et al., 2010; Ji et al., 2007). This model implicates upregulation of TGF-beta and SRC pathways in the biology of these tumors and particularly in the progression to metastasis (Carretero et al., 2010). In vivo testing of treatment regimens demonstrates that these murine tumors exhibit sensitivity to metabolic stress induced by phenformin (Shackelford et al., 2013), but are resistant to MEK inhibition (Chen et al., 2012). Although mouse models may provide a powerful tool to study tumor biology, the validity of the LKB1/KRAS lung tumor model in predicting human disease phenotypes has not been evaluated.

In this study we perform a comprehensive analysis of the gene expression changes associated with LKB1 loss in human tumors. We show that LKB1 loss is associated with a consistent pattern of gene expression across resected human NSCLC tumors and cell lines. Importantly, this pattern is not recapitulated in the murine model. A predictive signature derived from this pattern accurately classifies mutational and non-mutational loss of LKB1 in multiple validation sets. We give evidence that intragenic deletion of one or more exons, which is observed in NSCLC cell lines, is also a common mechanism for loss of LKB1. The LKB1 signature is also significantly associated with LKB1 loss in tumors and cell lines representing non-lung primaries. Finally we show our initial work in developing a clinical assay for determining this signature in patient samples.

Methods and Materials

Analysis of Publicly Available Gene Expression Data

Publicly available datasets were downloaded from the Gene Expression Omnibus (Barrett et al., 2012; Edgar et al., 2002) and ArrayExpress (Parkinson et al., 2009) or from individual websites, as listed in Appendix A. Processed and normalized gene expression data generated from RNA sequencing of tumors characterized by TCGA were downloaded from the TCGA Genome Data Analysis Center (GDAC) website, as were data from RPPA analysis, copy number changes, and somatic mutations. For many tumor types, including lung adenocarcinoma, tumor characterization by TCGA is ongoing. All data used in this thesis were from the data update of September 15, 2013. For RNA sequencing expression data, the log2 RSEM normalized files were used. Processed data uploaded to GEO by their original contributors were downloaded as 'series matrix.txt' files. In cases where data were presented as linear expression values, log2 transformed values were used. For analyses in which gene expression data from several studies were pooled, probeset expression values were standardized within each dataset by subtracting the mean value and dividing by the standard deviation. To collapse gene lists such that each gene was represented only once in our analyses, standardized scores from multiple probesets representing the same gene were averaged to give a single value.

Determining Exon-Imbalance Score

Exon-level RNAseq data for TCGA-characterized lung adenocarcinomas were downloaded from https://confluence.broadinstitute.org/display/GDAC/Home. The reads

per kilobase per million (RPKM) expression data for the ten exons in LKB1 were extracted manually for each of 446 lung adenocarcinomas. For each patient, the measured expression of each exon was divided by the total expression for all ten LKB1 exons to give the fraction of reads corresponding to that exon. We then looked at the distribution of these fractions across all 446 tumors and calculated an exon loss score by subtracting the mean exon expression fraction from the observed exon expression fraction for each tumor and dividing by the standard deviation for all 446 tumors. Thus, if a patient had a score of negative two for the third LKB1 exon, this indicates that the fraction of reads corresponding to exon 3 is two standard deviations below the average fraction observed across all tumors. Finally, because exon loss could affect any exon in the gene we used the minimum score across all ten exons as a single exon-loss score for each tumor.

Analysis of Gene Expression Associated with LKB1 Mutations

For clinical and cell line datasets in which LKB1 status was known, a Student's t-test was used to determine statistically significant in gene expression between LKB1 mutant and wild-type tumors. For cell line data, LKB1 mutation status was annotated using the Catalog of Somatic Mutations in Cancer (COSMIC) database (Forbes et al., 2001; 2010), the Cancer Cell Line Encyclopedia (CCLE) resource (Barretina et al., 2012), and individual publications. For data from the Directors Challenge Lung consortium (Shedden et al., 2008), LKB1 mutation status was unknown, and associations with LKB1 expression were determined using P-values derived from linear regression modeling, fitting the expression of each probeset to the expression of each of the two probesets corresponding to LKB1: 204292_x_at and 41657_at. These analyses were

performed using the R Bioconductor software platform, with the lm () function in the limma package.

Development of LKB1-Deficient Gene Signature

We used a training and testing approach to develop and test a gene signature capable of classifying LKB1-deficient tumors. There was no use of testing set samples at any point during the training process. We generated three gene lists using statistical comparisons from two training sets: the Wash U (Ding et al., 2008) set with comparisons to documented LKB1 mutations and the Michigan samples from the Director's Challenge Consortium (Shedden et al., 2008) with comparisons to LKB1 expression. The LKB1 classifier was taken as the intersection of these three lists.

ListA in Wash U:
All probesets '*x*' such that raw P-value < 0.01 for Student's t-test comparing LKB1 mut (n=7) vs LKB1 WT (n=34), resulting in 601 selected probesets.

ListB in Mich:
All probesets '*x*' such that raw P-value < 0.01 for linear regression model of 178 tumors, resulting in 3679 probesets:
*expr (204292_x_at) ~ a * expr (x) + b*

ListC in Mich:
All probesets '*x*' such that raw P-value < 0.01 for linear regression model of 178 tumors, resulting in 3467 probesets:
*expr (41657_at) ~ a * expr (x) + b*

Classifier = (ListA) ∩ (ListB) ∩ (ListC)

Lists B and C show a high degree of overlap, sharing more than half their genes, as they are derived from the same source and represent association with the two distinct LKB1 probesets. The intersection of the three lists results in a classifier of 167 probesets, a significantly larger intersection than expected by chance (P-value = 6.8e-38 by

hypergeometric test). Standardized values were then averaged across probesets representing the same gene to give a final set of 129 standardized gene expression values. Using different P-value cutoffs gave similar classification results of unknown lung cancer samples (classification concordance greater than 90%).

Linear Regression Analysis to Classify Clinical Samples

To provide a scheme of patient classification that was not dependent on hierarchical clustering analysis, a linear regression model was used to determine the association between LKB1 loss and each of the four transcriptional nodes observed. Using the expression data from the Michigan training cohort, the cluster scores identified in the previous section were used as four variables in a linear regression model to determine the best fit for LKB1 mRNA expression, as measured by the 41657_at probeset:

$$expr\ (41657\_at) \sim a * expr\ (LKB1\_loss) + b * expr\ (Mito) + c * expr\ (NRF2) + d * expr\ (Down)$$

The 16-gene LKB1-loss score was found to have the strongest association with LKB1 loss and inclusion of additional variables in the classification model did not substantially affect its accuracy, with concordance in sample classification greater than 90% and equivalent performance in detecting LKB1 mutations (22 of 26 using LKB1-loss score alone versus 23 of 26 for a combined model). Thus, the 16 gene LKB1-loss score was used to classify the LKB1 loss status of samples in the remainder of this study. A cutoff of 0.2 was used to delineate LKB1 loss from LKB1 wild-type, resulting in the

classification of approximately 30-35% of lung adenocarcinomas as having loss of LKB1, similar to the fraction observed by hierarchical clustering.

Determining Association with LKB1 Loss

Our LKB1 classification score was used to predict LKB1-loss status for unknown samples from eight collections of resected lung adenocarcinomas (total n=851). The accuracy of predicting LKB1 mutations was assessed in resected LUAD using both the TCGA dataset (Cancer Genome Atlas Network, 2013) and the pooled MSKCC2 (Chitale et al., 2009), UNC (Wilkerson et al., 2012), and USC (Selamat et al., 2012) datasets, while predictions of LKB1 mutations in cell lines were assessed in the pooled Sanger and CCLE datasets. In tumors with known mutation status, three groups of tumors were considered: tumors with identified mutations in LKB1; tumors without observed mutations but predicted by the gene expression signature to have loss of LKB1; and tumors without mutations predicted to be LKB1 wild-type. Expression of LKB1 mRNA and of pAMPK-T172 was compared between these groups using a Student's t-test.

LKB1 mutation data was also available for the MSKCC1 (Chitale et al., 2009) dataset and these samples represented another potential test set. However, for unknown reasons univariate analysis comparing reported LKB1 mutant and wild-type tumors in this dataset yielded fewer significant gene associations than would be expected by chance. In this dataset only five probesets out of 22000 passed a P-value cutoff of 0.001; in contrast for the Wash U (Ding et al., 2008) and MSKCC2 (Chitale et al., 2009) cohorts, 118 probesets and 162 probesets passed this cutoff, respectively. Furthermore, the top ranked genes associated with LKB1 mutations in this dataset showed no

significant overlap with the consistent pattern of gene expression observed in each of the other clinical and cell line datasets (Fig. 2.1a). Based on these findings we considered this dataset an outlier and excluded these data from our validation.

Results

LKB1 Loss Results in Consistent Gene Expression Changes in Human Tumors

The effects of signaling pathways are mediated in part by activation of transcription factors affecting the expression of downstream genes. Inferences drawn from the analysis of the dysregulated genes may disclose novel links between pathways and phenotypes that would otherwise be difficult to predict. However, a variety of genes are functionally mutated in cancer, and not all of these will necessarily produce distinct patterns of gene expression. We first became interested in the effects of LKB1 loss on gene expression changes when we observed that many LKB1 mutant cell lines were grouped together after we looked at similarities in their overall gene expression using unsupervised hierarchical clustering. This effect appeared unlikely to be due to chance, and after more careful inspection, was being driven primarily by a number of genes that were over-expressed by the cell lines with LKB1 mutations. Indeed, we confirmed the association in a second cell line dataset as well as the limited clinical datasets available at the time. This project grew from that initial observation, as we wanted to further characterize these genes to determine what molecular phenotypes might cause their dysregulation and especially to determine targetable pathways that could be tested for benefit in the clinic.

To determine whether LKB1 loss was more broadly associated with a consistent pattern of gene expression changes, we first identified all lung cancer datasets in which LKB1 mutation status was known, which included six set of resected lung adenocarcinomas referred to as MSKCC1 (Chitale et al., 2009), MSKCC2 (Chitale et al., 2009), Wash U (Ding et al., 2008), UNC (Wilkerson et al., 2012), USC (Selamat et al., 2012), and TCGA (Cancer Genome Atlas Network, 2013) as well as two large collections of NSCLC cell lines – GDSC (Garnett et al., 2012) and CCLE (Barretina et al., 2012) and data from two studies using the LKB1/KRAS murine model (Carretero et al., 2010; Ji et al., 2007). Differential gene expression between LKB1 mutant and LKB1 wild-type samples was assessed by a Student's t-test, and genes were then ranked by statistical significance for each dataset. Additionally, a lung adenocarcinoma dataset in which LKB1 status was unknown – samples from the University of Michigan characterized in the Director's Challenge lung cohort (Shedden et al., 2008) – was included, and associations with LKB1 were made based linear regression to look for correlations with the mRNA expression of the LKB1. These LKB1-associated gene lists were then compared pairwise across all datasets, and the statistical significance of gene overlap shown visually (Fig. 2.1a). This reveals a consistent pattern of gene expression associated with LKB1 loss across human datasets (median P-value = 4.0e-22 for 45 pair-wise comparisons). Murine LKB1 loss also resulted in a consistent gene expression signature across the two studies, but without significant overlap with the human studies, suggesting important differences in tumor biology between mouse models and human tumors.

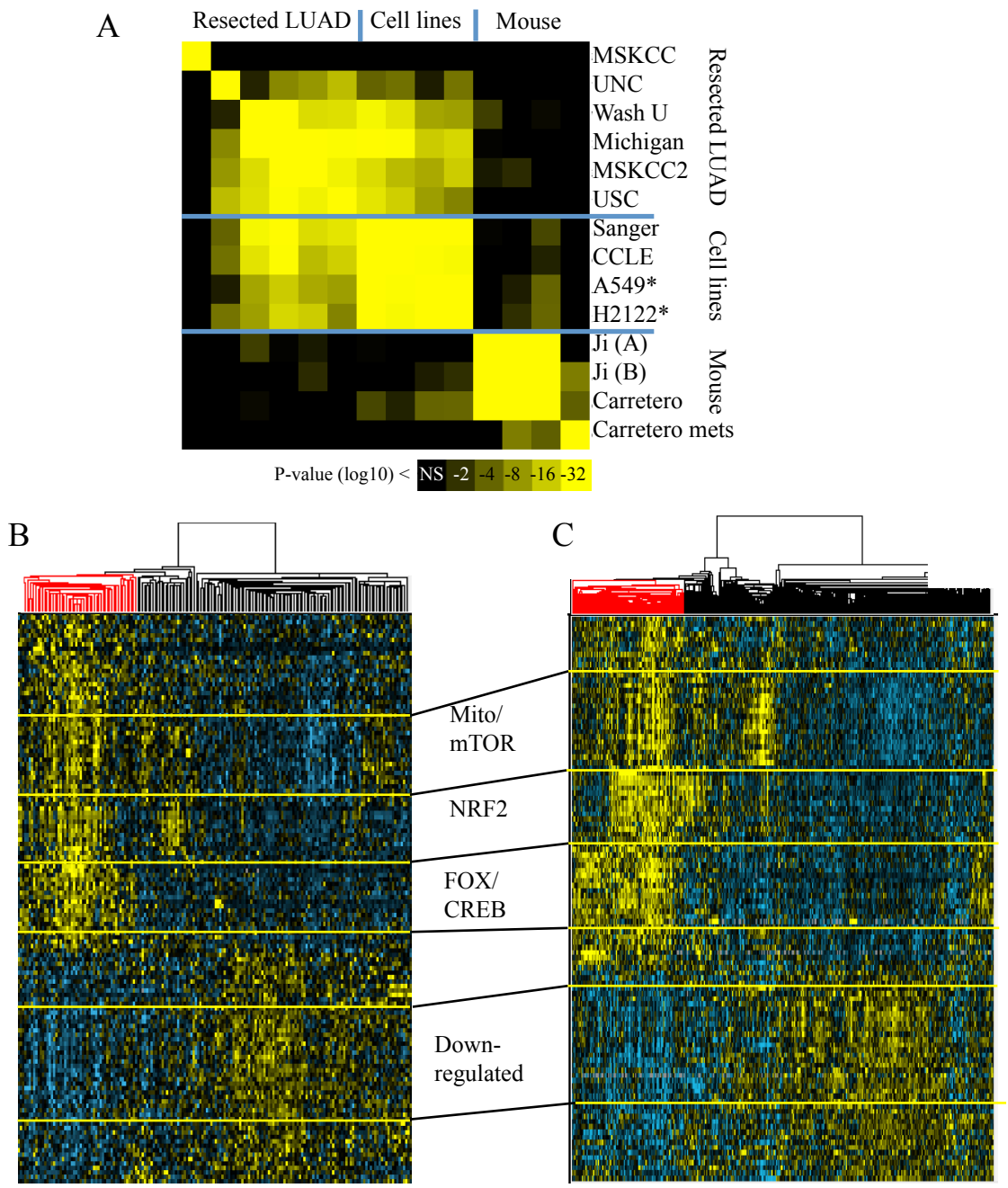


Figure 2.1. LKB1 loss produces a characteristic pattern of gene expression. A, The significance of gene overlap is shown for pairwise comparisons of the top 200 genes over-expressed in tumors with LKB1 loss in 15 studies of lung adenocarcinomas. Asterisks indicate comparisons between cell lines expressing vector control and those expressing wild-type LKB1. *P*-values from a hypergeometric test are color coded according to the legend. **B,C,** Unsupervised hierarchical clustering resected lung adenocarcinomas from the Michigan cohort of the Director's Challenge study (**B,** n=178) or the TCGA (**C,** n=446) using a 129 gene signature of LKB1 loss. Tumors are shown on the horizontal axis, with loss of LKB1 highlighted in red; genes are shown on the vertical axis, with four clusters of gene expression highlighted in red; names given to these clusters correspond with pathways implicated by our statistical analyses.

In addition to this visual depiction of the P-values for these associations, we also performed Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) between the LKB1-associated genes in the TCGA and those derived from each of the other 14 studies (Fig. 2.2). We used this analysis tool to rank 11933 genes in the TCGA according to their statistical association to LKB1 mutations. Then for each other study, the top 200 LKB1-associated genes were plotted as 'hits' along this ranked distribution. An 'enrichment score' is calculated based on this distribution, with higher values indicating a greater degree of similarity. In Fig. 2.2, we present the enrichment plots, along with the enrichment scores, the number of overlapping genes when comparing the two sets of 200 genes, and the P-value for this overlap, which corresponds to the color-coded P-value shown in Fig. 2.1a. This gives more detailed evidence for the similarities seen in these datasets.

We next wanted to determine whether the genes associated with LKB1 loss could be used to generate a consistent signature that could classify tumors that had lost LKB1, and also to determine the correlation patterns of these genes, which could reflect underlying phenotypes within the tumors. Although our initial observation of LKB1-associated gene expression was made in cell lines we decided to focus our subsequent analysis on resected lung cancer, where a large number of datasets were available, comprising over 1000 lung adenocarcinomas in total. As mentioned previously we identified six collections of resected lung adenocarcinomas in which LKB1 had undergone sequencing for somatic mutations (9-12). To preserve some of these datasets for validation purposes we used LKB1 associations from two studies as a training cohort. We used the Wash U (Ding et al., 2008) dataset of 41 tumors that contained 7 LKB1

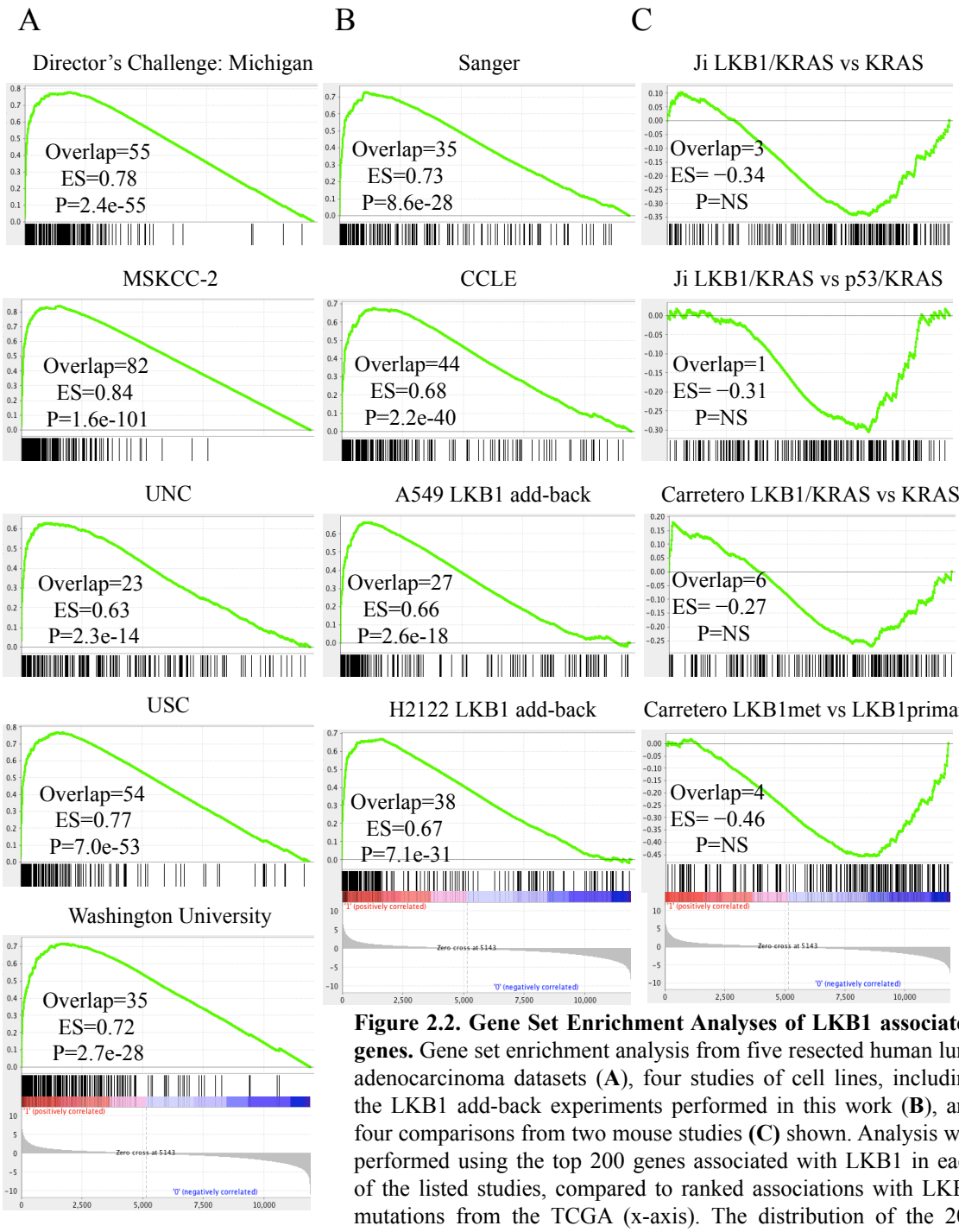


Figure 2.2. Gene Set Enrichment Analyses of LKB1 associated genes. Gene set enrichment analysis from five resected human lung adenocarcinoma datasets (**A**), four studies of cell lines, including the LKB1 add-back experiments performed in this work (**B**), and four comparisons from two mouse studies **(C)** shown. Analysis was performed using the top 200 genes associated with LKB1 in each of the listed studies, compared to ranked associations with LKB1 mutations from the TCGA (x-axis). The distribution of the 200 genes is shown as black hash marks along the x-axis. The enrichment score is plotted on the y-axis, and the maximum enrichment score is given for each plot. Also shown are the number of overlapping genes with the top 200 genes in the TCGA study, and the p-value for the significance of this overlap, calculated using the hypergeometic test. The number of overlapping genes expected by chance is four.

mutations to identify differentially expressed genes using a Student's t-test, and the Michigan set of 178 lung adenocarcinomas with unknown LKB1 mutation status, from which differentially expressed genes were identified by performing linear regression with the mRNA expression of LKB1. The overlapping genes from these two approaches defined a set of 129 genes associated with LKB1 loss. Unsupervised clustering of these genes identified a subset of 30-35% of lung adenocarcinomas that express an LKB1-deficient signature (Fig. 2.1b). We identified four transcriptional nodes of genes that showed high correlation in expression, which have been marked in this figure and named by key pathways implicated in our subsequent statistical analysis: 'mito/mTOR', 'NRF2', 'FOX/CREB' and 'Down-regulated'. We can see that these are observed across two independent sets of lung adenocarcinomas – the 178 tumors in the Michigan training set (Shedden et al., 2008) and the 446 tumors characterized by the TCGA (Cancer Genome Atlas Network, 2013). Thus, they represent consistent and reproducible genesets that may be driven by common underlying transcription factors or phenotypes. The characterization of these genes and interpretation of the potential phenotypes associated with them is developed in the following chapter of this work.

A 16-Gene LKB1-loss Classifier Accurately Predicts Mutational and Non-mutational Loss of LKB1 in Resected Human Lung Adenocarcinomas

Using unsupervised hierarchical clustering resulted in two visually distinct groups of tumors, which we found to be strongly associated with LKB1 loss (Fig. 2.1b). However, because each transcriptional node represents an independent phenotype, not every tumor with LKB1 loss will exhibit high expression of each node, and conversely,

activation of some these nodes can occur without LKB1 loss. Because of this, a simplistic clustering approach that attempts to convert the expression of multiple phenotypes into a single binary classification can be unreliable and inconsistent between datasets. For instance, in every dataset there is a subset of tumors that have activation of NRF2-driven genes without expression of the remainder of the LKB1-associated genes. In certain datasets, these samples will be clustered with the LKB1 wild-type group, while in other sets they will cluster with the LKB1-deficient group. A more accurate way to approach the data is to consider each of the observed transcriptional nodes to be an independent phenotype reflected by the expression of a set of genes. Thus, for each individual node one could make a binary classification, but when more than one node is considered this results in all possible combinations of the two phenotypes. We returned to the training set to determine which of the four transcriptional nodes had the strongest association with LKB1 loss, which revealed that 16 genes with high correlation of expression – labeled as the FOX/CREB cluster in Fig. 2.1b – had the strongest association to both LKB1 mutations and LKB1 mRNA expression. These genes comprising this cluster are AVPI1, BAG1, CPS1, DUSP4, FGA, GLCE, HAL, IRS2, MUC5AC, PDE4D, PTP4A1, RFK, SIK1, TACC2, TFF1, and TESC.

The expression of these genes can be combined to give a single 'LKB1-loss score', as detailed in the methods section. Higher scores indicate a greater likelihood of LKB1 loss. Among resected lung adenocarcinomas, these scores are found to occur in a bimodal distribution, as shown in (Fig. 2.2a) for the TCGA lung adenocarcinomas. By solving for the parameters of the two underlying normal distributions, we can see that there is a high score population of tumors that have lost LKB1, with a mean score of 0.79

and a standard deviation of 0.43, and a low score population of tumors with WT LKB1,

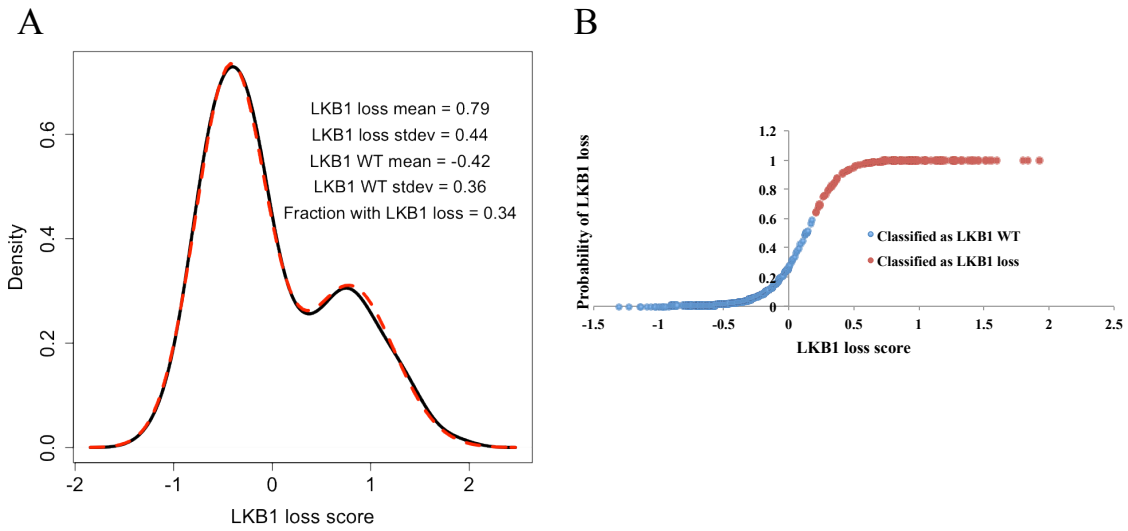


Figure 2.3. LKB1 loss scores exhibit a bimodal distribution. A, Population density graph showing the distribution of LKB1 loss scores seen in TCGA lung adenocarcinomas (n=446). Solid curve in black indicates the actual distribution; dashed red curve shows the results of a bimodal curve fit to the experimental data. Parameters for this fit are listed. **B,** Plot of probability of LKB1 loss at different values of the LKB1 loss score, as calculated by the best fit parameters. A score cutoff of 0.2 was used throughout the work to discriminate LKB1-loss from LKB1 wild-type; the classifications associated with scores is colored red for LKB1 loss or blue for LKB1 wild-type.

with a mean score of -0.42 and a standard deviation of 0.36. The proportion of tumors that belong to the LKB1 loss group is 0.34. From this calculation we can also derive the probability of LKB1 loss for any given score, which is plotted for the TCGA cohort of tumors. Throughout our work we have used an LKB1-loss score of 0.2 to delineate the classification of LKB1 loss from LKB1 WT, which corresponds to a probability of LKB1 loss of 63%. Because a certain percentage of tumors fall into an area of overlap between the two normal distributions, there is an inherent rate of misclassification when converting from a continuous probability score to a binary classification. Our bimodal parameters also allow estimation of this misclassification rate, which we expect to be an

overall misclassification rate of 6.5%, with 4.5% false positives among the tumors called LKB1 loss, and 7.3% false negatives among tumors called LKB1 WT (Fig.2.2b).

We next tested the accuracy of this 16-gene signature in predicting LKB1 mutations in two clinical validation sets: a pooled analysis of previously published resected lung adenocarcinoma and lung adenocarcinomas characterized by the TCGA. LKB1 mutations were accurately predicted in each of these validation cohorts, detecting 22 of 26 somatic LKB1 mutations in the pooled cohort and 65 of 67 mutations in the TCGA cohort (sensitivity 0.85 and 0.97; P-value = 2.8e-9 and 9.4e-32 by Fisher's exact test; Fig. 2.4a; Fig. 2.5a,b). To ensure that our results were not influenced by our choice of training set, we also confirmed our findings using a second classifier derived from an independent training cohort (Fig. 2.6). We used the same approach described above but started with the MSKCC2 (Chitale et al., 2009) dataset with known LKB1 mutations and the Director's Challenge cohort (Shedden et al., 2008) – excluding the University of Michigan samples – for association with LKB1 expression. The concordance between this alternate training set LKB1-loss score and the original LKB1-loss score was 94% among the TCGA lung adenocarcinomas, and this correctly predicted 64 of 67 LKB1 mutations (sensitivity 0.96, P-value = 2.6e-35, by Fisher's exact test).

Twenty three percent of tumors without known LKB1 mutations were classified as having LKB1 loss. These tumors could be considered misclassified tumors, which would result in a specificity of the test of 0.77 for predicting LKB1 mutations. However, some mutations may have been unrecognized, and there are multiple mechanisms by which tumor suppressors can be inactivated in addition to somatic mutation. Indeed, these LKB1 wild-type tumors demonstrate unequivocal evidence of LKB1 loss. They exhibit

A

Dataset	LKB1 status (n)	Sensitivity (%)	Specificity (%)	P-value
Lung Adeno Pooled	Loss: 26 WT: 213	85	76	2.8e-09
Lung Adeno TCGA	Loss: 67 WT: 339	97	78	2.5e-33
NSCLC Cell lines	Loss: 39 WT: 46	93	91	1.2e-16
Non-NSCLC Cell lines	Loss: 49 WT: 348	65	83	1.2e-11

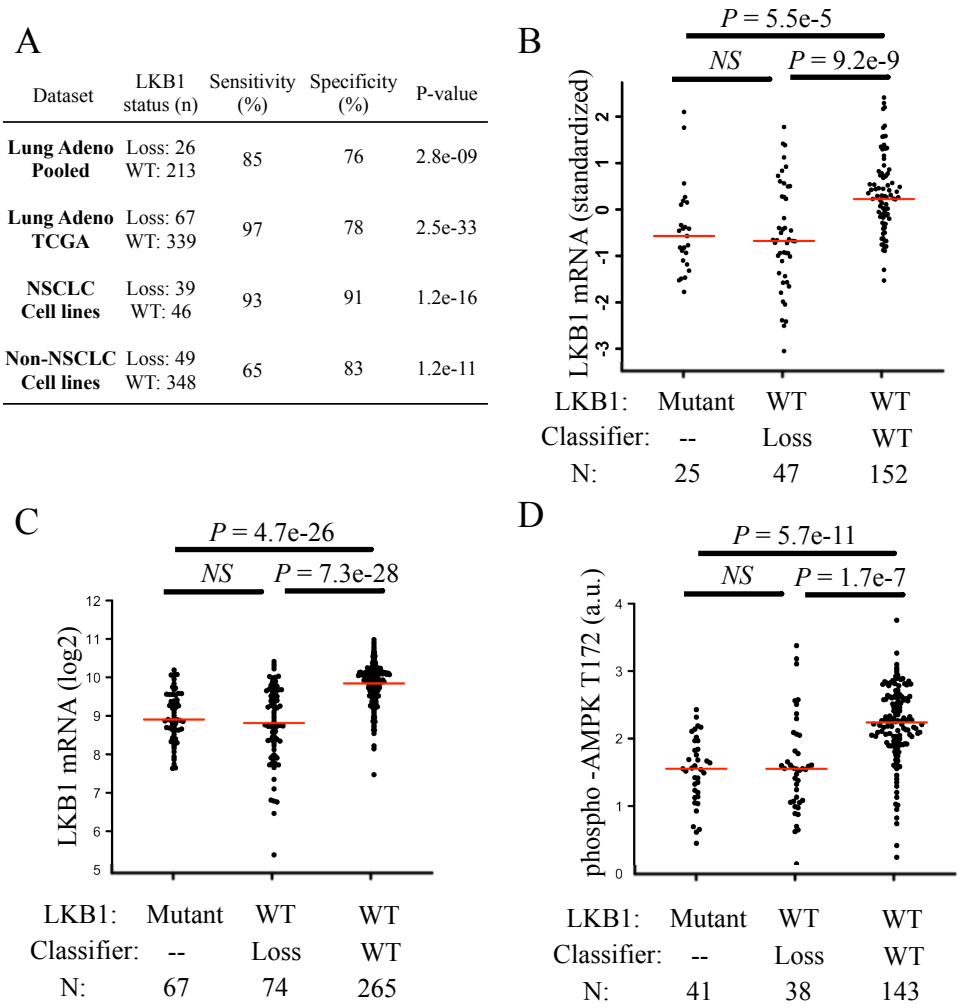


Figure 2.4. LKB1-loss signature is predictive of mutations and non-mutational loss of LKB1. A, Sensitivity and specificity of the LKB1 classifier for prediction of LKB1 mutations across independent testing sets; p-value represents the result of the Fisher's exact test. **B, C,** Expression of LKB1 mRNA is shown for tumors grouped by LKB1 mutation and classification status among a pooled analysis of resected lung adenocarcinomas (**B**), or among lung adenocarcinomas characterized by TCGA (**C**). **D,** RPPA values for expression of phospho-AMPK T172 are shown for lung adenocarcinomas characterized by TCGA and grouped by LKB1 mutation and classification status. For **B-D**, each dot represents one tumor, with red bars indicating the median expression. P-values are derived from the student's t-test comparing indicated groups.

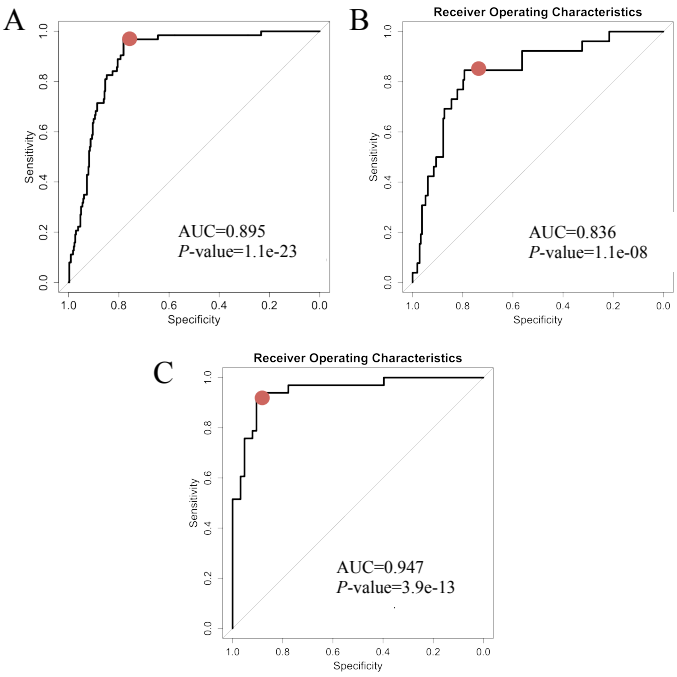


Figure 2.5. Receiver operating characteristics for LKB1-loss score in resected lung adenocarcinomas and NSCLC cell lines. Receiver operating curves to show the relationship of sensitivity and specificity of the LKB1 loss score for detecting LKB1 mutations in the TCGA lung adenocarcinomas (**A**), pooled analysis of other resected lung adenocarcinomas (**B**), or NSCLC cell lines (**C**). AUC refers to area under the curve, and the P-value reflects the results of the Mann-Whitney U test. The red dot indicates the location of the cutoff score of 0.2 used in this work.

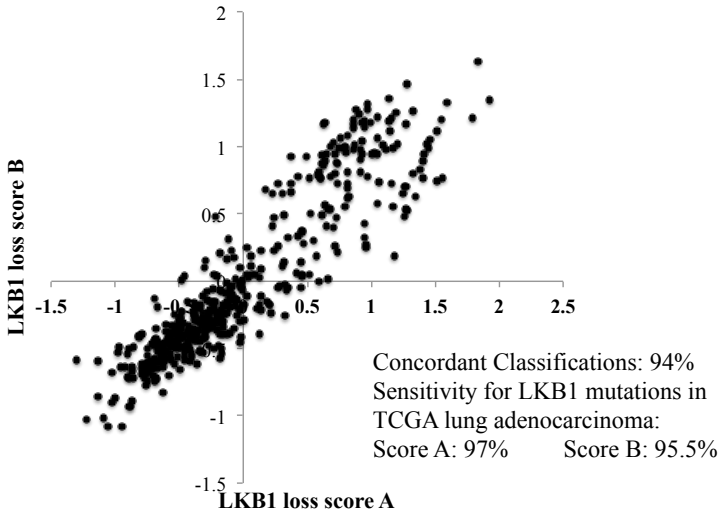


Figure 2.6. Comparison of LKB1 loss scores derived from two different training sets. LKB1-loss scores are plotted for two distinct classifiers: Score A is the primary 16-gene LKB1-loss classifier used throughout this study and Score B results from an independent classifier derived by applying the same statistical approach to a different training cohort. The concordance is the percentage of tumors that are given the same LKB1-loss classification score by each of the two classifiers.

low expression of LKB1 mRNA in both the pooled analysis (P-value = 9.2e-9 by Student's t-test; Fig. 2.3b) and among the TCGA samples (P-value = 5.5e-28 by Student's t-test; Fig. 2.3c). Moreover, loss of LKB1 kinase activity is attested by the significantly attenuated phosphorylation of AMPK at the threonine 172 position (P-value = 4.6e-8 by Student's t-test; Fig. 2.3d). No difference is observed in LKB1 mRNA or AMPK phosphorylation between known LKB1 mutants and wild-type tumors predicted to have LKB1 loss. Furthermore, other genes and proteins differentially expressed by tumors with known LKB1 mutations are concordantly dysregulated among the tumors with predicted loss (Fig. 2.7). This shows that in addition to predicting somatic mutations in LKB1, this signature detects LKB1 loss by other mechanisms, doubling the number of tumors identified.

To explore additional mechanisms of LKB1 loss we used the TCGA lung adenocarcinoma cohort to examine copy number changes affecting the LKB1 locus on chromosome 19, and also exon-level expression data of LKB1 from RNA sequencing. Although the actual copy number of LKB1 in any tumor is a discrete integer value, measured copy number data are reported as a single continuous variable for all tumors; thus, interpreting the significance of a particular value is not straightforward. More extreme values indicate a greater likelihood of copy number loss and are more associated with homozygous loss than heterozygous loss, but no single cutoff crisply discriminates two copies from one heterozygous copy number loss, and no value clearly distinguishes heterozygous from homozygous loss. Furthermore, some tumors may have single copy number loss affecting LKB1, but may have a second functional copy that renders their

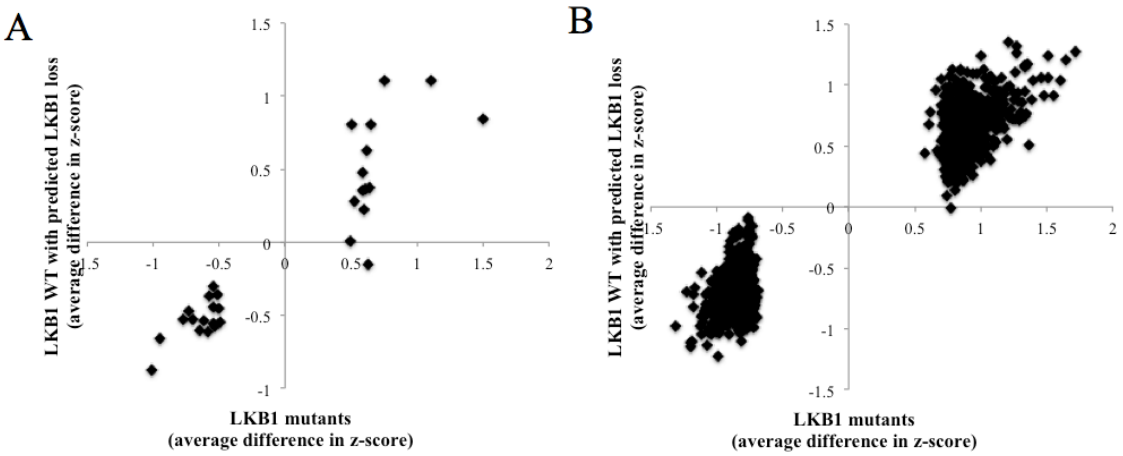


Figure 2.7. Comparison of protein and gene expression differences associated with known LKB1 mutations or associated with predicted LKB1 loss among LKB1 WT tumors. A, Differences in protein expression determined by TCGA RPPA are shown for proteins that have significant association with LKB1 mutations (with $p<0.01$). **B,** Differences in mRNA expression determined by TCGA RNAseq analysis are shown for genes that have significant association with LKB1 mutations (with $p<1e-6$). Dots represent individual proteins or genes. The x-axis shows the difference in average expression between LKB1 mutant tumors and LKB1-WT tumors with WT classification score. The y-axis shows the difference in average expression between LKB1-WT tumors with a LKB1-loss classification score compared to LKB1-WT tumors with a WT classification score.

tumor LKB1 wild-type, while other tumors may have a second event that inactivates the other copy of LKB1, giving an LKB1-deficient tumor.

We also calculated exon loss scores for each tumor, as described in the methods section of this chapter, by measuring the observed distribution of exon expression for each individual tumor and comparing it to the distribution observed across all 446 tumors in the dataset. For both exon loss and chromosomal deletion, then, we are not able to accurately classify the presence or absence of these lesions in a single tumor, nor can we use these to determine the functional status of LKB1 within the tumor. However, when measured across groups of tumors these variables can give useful evidence for the relative prevalence of these alterations and their contribution to LKB1 loss in lung cancer.

Both of these measures showed a strong statistical association with the LKB1 loss gene expression score. Using a stringent cutoff of the lowest 5% of scores to define chromosomal or exon loss, 70-75% of tumors with these lesions had positive LKB1-loss scores. This specificity of approximately 0.70 was observed as the cutoff threshold was increased, until an inflection point was reached, after which the specificity dropped rapidly. For LKB1 copy number loss, 10% of tumors had scores corresponding to loss of a single copy of LKB1; 68% of these were LKB1-loss signature positive (P-value = 6.3e-7 by Fisher's exact test). For exon loss, 17% of tumors had evidence of exon loss by this measure, of which 71% were LKB1-loss signature positive (P-value = 8.3e-13 by Fisher's exact test). The tumors that met the cutoff threshold for copy number or exon loss, but were classified as LKB1 wild-type by signature score could potentially represent misclassified cases of LKB1 inactivation. If this were the case we would expect to see decreased phosphorylation of AMPK, as we see with known somatic mutations and with LKB1 loss detected by our gene signature. To test this, we compared the pAMPK levels for these tumors with those LKB1 wild-type tumors that did not meet the cutoff threshold for exon or chromosomal loss. There was no significant decrease in AMPK phosphorylation in either case, indicating that the signature accurately classified LKB1 activity.

We next examined the distribution of these lesions among tumors with LKB1 loss. Evidence of single copy-number chromosomal loss was present in 22% of tumors with known somatic mutations in LKB1, and in 18% of LKB1-loss signature positive tumors without a detected mutation. However, the characteristic of having exon loss was significantly more common among tumors without detectable LKB1 mutations (44% vs.

23%, P-value = 0.011). The presence of either exon loss or copy number loss was found in 51% of the LKB1-loss signature positive wild-type tumors. This gives evidence for two additional mechanisms of LKB1 loss that are quite prevalent, but undetected by the TCGA exon sequencing efforts; furthermore, intragenic deletion of one or more exons has been reported to be a common in cell lines (Matsumoto et al., 2007).

The LKB1-loss Classifier Accurately Predicts Mutational and Non-mutational Loss of LKB1 in NSCLC Cell Lines

Our initial comparison of LKB1-associated gene expression showed that two collections of NSCLC cell lines showed similarities in gene dysregulation compared to datasets of resected lung adenocarcinomas. Having shown that our LKB1-loss signature predicts LKB1 loss in resected lung adenocarcinoma datasets, we next wanted to test its performance among NSCLC cell lines. We calculated LKB1-loss scores as described previously for all cell lines included in two large collections of gene expression data, one characterized by the Genomics of Drug Sensitivity in Cancer study (GDSC) (Garnett et al., 2012) and one characterized by the Cancer Cell Line Encyclopedia (CCLE) (Barretina et al., 2012). Determining the presence or absence of LKB1 in the cell lines is not straightforward, as multiple studies have characterized mutations and protein expression of LKB1 in lung cancer cell lines, and identical results are not found across studies. For instance, the CCLE performed targeted exome-capture of LKB1 followed by next generation sequencing of enriched target DNA and identified LKB1 mutations in only eight of 113 NSCLC cell lines studies. However, of the 105 cell lines reported to be LKB1 wild-type, 24 have been shown to have LKB1 loss by other studies, including two

well studied cell lines – A549 and H460 – that have nonsense mutations that were missed by the exon-capture approach but were detected by the CCLE using a different methodology (Oncomap). Therefore, any given study may have a low sensitivity, but high specificity, for detecting LKB1 loss. Furthermore, different groups have employed a variety of approaches for detecting LKB1 loss including examination of LKB1 protein expression by western blot (Spoerke et al., 2012), and determining chromosomal loss or intragenic deletions of one or more exons within the LKB1 gene (Matsumoto et al., 2007). We identified nine sources in which the LKB1 status of some set of NSCLC cell lines had been investigated and considered any mutation or loss of LKB1 described therein to be an instance of LKB1 loss. For determining LKB1 wild-type status we considered a cell line to have LKB1 WT status if no mutations were reported in either the CCLE or GDSC database. This approach resulted in 39 NSCLC cell lines with reported LKB1 loss, and 46 LKB1 WT cell lines. Of these, our LKB1-loss signature correctly classified 36 of 39 cell lines with LKB1 loss, while four of 46 LKB1 wild-type cell lines were classified as having LKB1 loss (Sensitivity 0.93, Specificity 0.91, P-value=1.2e-16 by Fisher's exact test; Fig. 2.4a; 2.5).

LKB1-loss Classifier is Associated with LKB1 Mutations in Cell Lines Derived from Non-lung Primary Cancers

LKB1 loss has been identified in other tumor types, but these have not been studied as extensively as in non-small cell lung cancer. It has been reported that LKB1 is lost in approximately 20% of cervical cancers (Wingo et al., 2009), as well as in cervical cancer cell lines. A very low rate – less than 1% – of LKB1 mutations has been observed

in breast cancer, and sporadic, low prevalence loss of LKB1 have been reported in other tumor types, but in many cases these characterizations have been incomplete. Using individual publications compiled in the Sanger Institute's catalog of mutations in cancer (COSMIC) database (Forbes et al., 2001; 2010), as well as the sequencing efforts of the GDSC (Garnett et al., 2012) and CCLE (Barretina et al., 2012) we identified 49 LKB1 mutations in cell lines that were not derived from NSCLC. Applying the same datasets and LKB1 classification score used above for NSCLC cell lines, we found a significant association between these mutations and our LKB1-loss score. 32 of 49 LKB1 mutant cell lines expressed a positive LKB1-loss score, compared to 182 of 1054 LKB1 wild-type or unknown cell lines (sensitivity 0.66, specificity 0.83, P-value = 6.7e-13 by Fisher's exact test; Fig. 2.4a). The presence and functional significance of many of these LKB1 mutations have not been validated, and there are likely some instances of LKB1 loss among cell lines found to be LKB1 WT or of unknown LKB1 status. However, we can conclude that LKB1 loss is significantly associated with similar gene expression changes in these, non-lung cell lines, although additional factors may affect the expression of the 16 genes in the signature, which may differ between NSCLC and non-lung cancers.

Although small sample sizes prevented a more thorough characterization of performance, the LKB1 classifier reached statistical significance for predicting mutations in the gastrointestinal tract, melanoma, cervical cancer, and central nervous system cell lines (Table 2.1). LKB1 mutations in some types of cell lines were not associated with signature expression. Seven LKB1 mutations have been identified in hematopoietic cell lines, and none of these displayed a positive LKB1-loss signature. The spectrum of

mutations seen in these hematopoietic cell lines also appeared to be different from that seen in NSCLC, with only one nonsense mutation, zero frame-shift mutations, and several missense mutations affecting C-terminal residues that have not been observed in

Table 2.1. Association of LKB1 mutations with LKB1 loss score in cell lines excluding NSCLC.

Histology[a]	Number of cell lines	LKB1 mutations	Correctly classified	Sensitivity	Specificity	P-value	
Biliary tract	11	1 (0.09)	1	1	0.7	0.36	
Breast	64	3 (0.05)	2	0.67	0.82	0.1	
Cervical	12	5 (0.42)	4	0.80	0.71	0.24	
CNS	122	1 (0.01)	1	1	0.96	0.049	*
GI tract	179	8 (0.04)	7	0.88	0.61	0.0083	**
Hematopoietic	228	8 (0.04)	0	0	0.97	1	
Hepatocellular	27	2 (0.07)	2	1	0.56	0.22	
Melanoma	73	5 (0.07)	5	1	0.63	0.0095	**
Ovarian	57	3 (0.05)	2	0.67	0.8	0.13	
Pancreas	45	2 (0.04)	2	1	0.51	0.49	
Prostate	8	1 (0.13)	1	1	0.71	0.38	
Renal	32	1 (0.03)	1	1	0.97	0.063	
SCLC	73	6 (0.08)	2	0.33	0.94	0.073	
Thyroid	16	1 (0.06)	1	1	0.8	0.25	
Uterus	31	2 (0.06)	1	0.50	0.9	0.25	
Total	1103[a]	49 (0.05)	32	0.66	0.83	6.7e-13	***

[a] No LKB1 mutations were observed in bladder (n=26), meosthelioma (n=10), sarcoma (n=76), or other (n=13). However, these are included in the total.
*p<0.05
**p<0.01
***p<0.001

NSCLC or Peutz-Jeghers syndrome (Alessi et al., 2006). It is not known whether these mutations have functional consequences. Interestingly, other C-terminal mutations have been tested for phenotypic effect and do not inhibit tumor growth or induce AMPK phosphorylation; however, there was a significant effect on restoring cell polarity. Only five hematopoietic cell lines out of 200 expressed a weakly positive LKB1-loss signature,

so the specificity was correspondingly high in hematopoietic cells. In contrast, among all cell lines from the gastrointestinal tract, seven of eight LKB1 mutant cell lines were associated with a positive LKB1-loss signature, but the specificity was lower among these cell lines as well (sensitivity 0.87, specificity 0.61, P-value=0.008 by Fisher's exact test). It may be possible to improve the accuracy of the LKB1-loss signature by modifying which genes are used for non-lung primaries.

Association of LKB1-Loss Signature with Loss of LKB1 in Resected, Non-lung Primary Tumors

Our results from the analysis of non-lung cell lines shows that LKB1 loss has similar effects on gene expression across multiple histological cancer types. This may represent consistent LKB1-induced alterations in pathway activation. We next wanted to test whether LKB1 loss in primary tumors of non-lung origin showed evidence of similar expression of the LKB1-loss gene expression signature. We downloaded all gene expression, protein expression, and mutation data currently available from the TCGA characterization of 23 primary tumor types in addition to lung adenocarcinomas. The mutation rate was lower across these tumor types than observed in lung adenocarcinomas; these data are summarized in Table 2.2. Considering all non-lung adenocarcinomas together, somatic mutations in LKB1 were observed in only 24 out of 4187 tumors. Of these, 21 tumors also had gene expression data available from RNA sequencing. We used a higher expression cutoff for discriminating LKB1 loss from wild-type classification, which was necessary because the low prevalence of LKB1 loss affects both the specificity of the test and the actual distribution of expression scores, which are

Table 2.2. Association of LKB1 mutations with LKB1 loss score in TCGA cancer cohorts other than lung adenocarcinoma.

Histology[a]	Tumors	LKB1 mutations	Correctly classified	Sensitivity	Specificity	P-value	
Breast	759	2 (0.003)	2	1	0.91	0.066	
Cervical	38	1 (0.026)	1	1	0.95	0.13	
Colon	185	2 (0.007)	0	0	0.9	1	
Glioblastoma	31	1 (0.006)	0	0	0.94	1	
Head and Neck	299	1 (0.003)	1	1	0.93	0.24	
Renal Clear Cell	405	1 (0.002)	0	0	0.97	1	
Renal Papillary	102	2 (0.018)	0	0	0.89	1	
Lung Squamous	178	3 (0.017)	3	1	0.96	0.015	*
Pancreas	32	1 (0.018)	0	0	0.91	1	
Melanoma	307	5 (0.016)	2	0.4	0.95	0.029	*
Stomach	210	2 (0.011)	2	1	0.9	0.083	
Total	3602	21 (0.05)	11	0.55	0.92	7.2e-08	***

[a] No LKB1 mutations were observed in bladder (n=26), renal chromophobe (n=66), low grade glioma (n=217), ovarian (n=59), prostate (n=82), rectal (n=68), or thyroid (n=391). However, these are included in the total.

*p<0.05
**p<0.01
***p<0.001

based on the distribution of expression across the dataset. At this higher score cutoff, ten of twenty tumors with LKB1 mutations expressed a positive LKB1 signature, compared to 8% of tumors without LKB1 mutations (sensitivity 0.55, specificity 0.92, P-value = 7.2e-08 by Fisher's exact test). Due to the limited sample sizes, the only tumor type to reach individual significance was lung squamous cell carcinoma, in which three of three mutations were correctly classified. Breast adenocarcinomas and stomach

adenocarcinomas were close to attaining significance, each with two of two mutations correctly classified. Half of the LKB1 mutations did not express the LKB1-loss signature. Using the original score cutoff only increases the sensitivity to 0.62, while worsening the specificity to 0.75, although this was still statistically significant (P-value = 0.001). From this analysis we conclude that LKB1 loss produces similar gene expression changes in both NSCLC and in non-lung primaries, as was the case for cell lines. However, a significant number of LKB1 mutations do not induce these changes.

We also take advantage of the fact that a significant number of LKB1 lesions reduce the mRNA expression of LKB1 to low levels that are obvious outliers compared to the overall distribution of LKB1 expression in a dataset. Using the TCGA RNAseq expression for all primary tumor types we used very low mRNA expression (two to nine standard deviations below the mean) as a surrogate for LKB1 loss. Similar results were observed using this approach to detect LKB1 loss, with a strong statistical association between LKB1 loss and signature expression, but with lower sensitivity than seen in lung adenocarcinomas (sensitivity 0.53, specificity 0.92, P-value = 2.9e-12 by Fisher's exact test). Statistically significant associations were observed for breast, cervical, lung squamous, and melanoma. Differences in the prevalence of LKB1 loss, and the sensitivity for detecting this loss were observed across the multiple tumor types assessed. These results are presented in Table 2.3.

For tumor types other than lung adenocarcinomas, the only datasets with mutational analysis and gene expression analysis are the TCGA dataset. However, we can use outlier analysis of LKB1 mRNA expression to determine LKB1 loss in other gene expression datasets. We identified two additional large cervical cancer datasets, and one

Table 2.3. Association of very low LKB1 expression with LKB1 loss score in TCGA cancer cohorts other than lung adenocarcinoma.

Histology[a]	Number of Tumors	Low LKB1	Correctly Classified	Sensitivity	Specificity	P-value	
Bladder	163	3 (0.018)	3	1	0.98	4.9E-05	***
Breast	914	13 (0.014)	6	0.46	0.91	7.5e-4	***
Cervical	116	7 (0.06)	5	0.71	0.91	3.4e-4	***
Head and Neck	303	1 (0.003)	1	1	0.93	0.073	
Renal Clear Cell	480	3 (0.006)	0	0	0.95	1	
Lung Squamous	408	4 (0.01)	2	0.5	0.92	0.0364	*
Melanoma	310	5 (0.016)	2	0.4	0.95	0.02875	*
total	4984	36 (0.007)	19	0.53	0.91	2.9e-12	***

[a] No loss of LKB1 was observed in colon (217), glioblastoma (160), renal chromophobe (66), renal papillary (106), low grade glioma (271), ovarian (265), pancreas (40), pancreas (176), rectal (78), thyroid (482), or endometrial (104) cancer. However, these are included in the total.
*p<0.05
**p<0.01
***p<0.001

dataset from squamous cell lung carcinoma. Plotting LKB1 expression compared to LKB1-loss signature score demonstrates a strong statistical association between these phenotypes in each of these datasets (Fig. 2.8). Based on the expression score and LKB1 mRNA expression we can also estimate the prevalence of LKB1 loss in cervical cancer to be around 15-20%, and the prevalence in lung squamous cell carcinomas to be about 5-10%, while the prevalence of LKB1 loss in breast cancer is approximately 1%. These are consistent with previously reported rates of LKB1 loss in these cancers (Cancer Genome Atlas Network, 2012; Cancer Genome Atlas Research Network et al., 2012; Wingo et al., 2009).

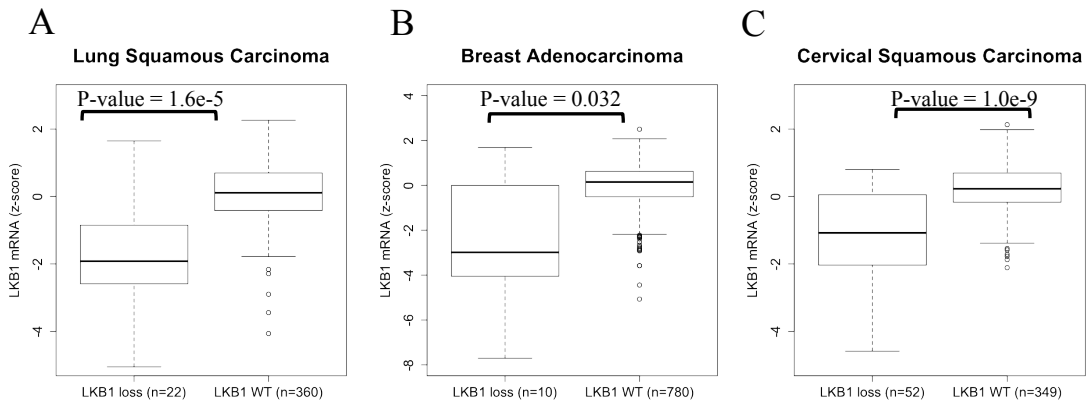


Figure 2.8. Decreased LKB1 mRNA is associated with LKB1-loss signature in resected breast cancer, lung squamous cell carcinoma, and cervical cancer. A, Expression of LKB1 mRNA is shown for tumors classified as LKB1 loss or LKB1 WT in a pooled cohort of lung squamous cell carcinomas (GSE4573 and TCGA). **B,** Association between LKB1 loss score with low LKB1 expression in breast cancer (TCGA). **C**, Association between LKB1 loss score and low LKB1 expression in pooled analysis of three cohorts of cervical squamous cell carcinoma (GSE38964, GSE20167, TCGA). The y-axis represents the standard deviations from the mean LKB1 expression; p-values represent the results of a Student's t-test.

Development of an Assay for LKB1-Loss Score Suitable for Analysis of Clinical Samples Using the nanoString n-Counter Assay

We have demonstrated that our signature accurately predicts LKB1 loss in lung adenocarcinoma and has significant associations with LKB1 loss in other primary tumor types. Because our signature detects LKB1 loss that occurs both through somatic point mutations and other mechanisms of loss, it is a more sensitive test for LKB1 loss than LKB1 sequencing alone. Currently there are no anti-cancer drugs known to preferentially benefit tumors with LKB1 loss. However, therapeutic strategies to exploit the metabolic differences have shown interesting results in preclinical in vivo studies (Shackelford et al., 2013). Moreover, strategies that target the mTOR pathway may be attractive approaches as well. Furthermore, our work implicates LKB1 in regulating the apoptotic

response to MEK inhibition, and MEK inhibitors may also prove effective in treating LKB1-deficient tumors. Thus, accurately determining LKB1 loss in clinical specimens will be useful both in testing associations with patient outcome in clinical trials, and ultimately in stratifying patients for appropriate targeted therapies.

We chose to develop an assay for our signature using a commercial testing platform, the nanoString nCounter analysis platform, which has been used successfully to develop a gene expression diagnostic test for breast cancer (Reis et al., 2011) (Prosigna, which is approved for use in Europe). After selecting a gene list, molecular probes are designed to specifically bind the mRNA of each gene of interest. Each probe is tagged with a unique 'molecular barcode' of different fluorescent colors, which serves as the method of detection. RNA is extracted from clinical samples, immobilized on a test slide, and the probes are hybridized to the sample mRNA. After washing probes that are not bound to the appropriate targets the color-coded tags are counted, from which we can determine the original mRNA expression for each target gene in the sample.

We performed an initial test of this platform, using our 16-gene signature as well as additional genes for normalization, in a set of 24 lung adenocarcinoma cell lines and 24 resected lung adenocarcinomas. After extracting mRNA and performing quality control measures on these samples, we obtained signature gene expression data from these cell lines using the nanoString nCounter® analysis system (nanoString, Seattle, WA). Gene expression was first normalized using the standard n-Counter positive and negative controls. Expression values were then log2 transformed, and adjusted based on the observed expression of nine control genes. This approach allows each patient or cell line to be analyzed as a single sample without having to compare the expression to a

range of samples or to a reference samples. The expression measured for the 16 genes was converted to a single numeric score using weighted averages based on the TCGA data. The resulting LKB1-loss scores successfully discriminated six of the seven LKB1 mutant from the 15 LKB1 wild-type cell lines, with one misclassified LKB1 mutant cell line; two cell lines were of unknown LKB1 mutation status (P-value = 9.3e-5; Fig. 2.8a).

The 24 resected lung adenocarcinoma samples were analyzed in the same manner as described for the cell lines. These samples had also been analyzed by RNA sequencing

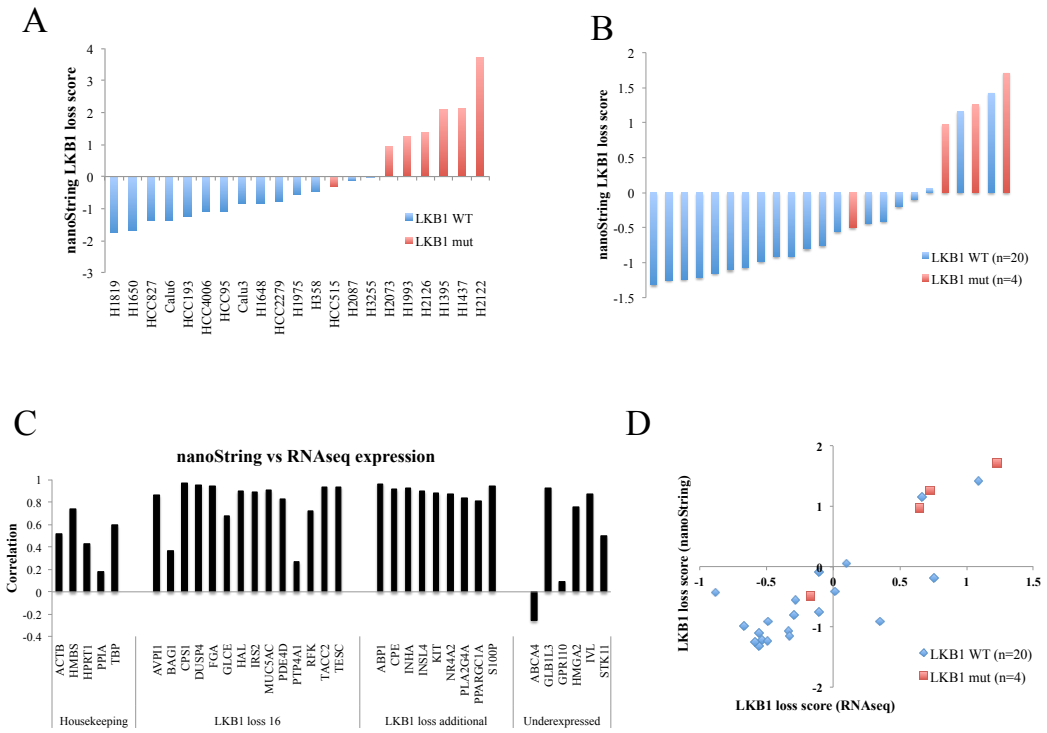


Figure 2.9. Performance of nanoString platform to assess LKB1 loss signature in cell lines and clinical samples. A, Association between LKB1 mutations and LKB1-loss score in 22 cell lines, using nanoString nCounter platform to measure gene expression for LKB1 loss signature genes. **B,** Association between LKB1 mutations and LKB1-loss score in 24 resected lung adenocarcinomas, using nanoString nCounter platform to measure gene expression for LKB1 loss signature genes. **C**, Correlation coefficients between gene expression measured by nanoString and those measured in the same samples by RNA sequencing for each gene included in the analysis. **D**, Plot of resulting LKB1-loss scores for the 24 resected lung adenocarcinomas, with scores resulting from nanoString plotted on the y-axis compared to scores from RNAseq on the x-axis.

(unpublished data); four LKB1 mutations had been detected through that analysis. Three of the four LKB1 mutations strongly expressed the LKB1-loss signature, while one did not, and two putatively wild-type tumors were classified as LKB1-deficient (P-value = 0.015; Fig. 2.8b). We also compared the correlation of gene expression scores calculated using the nanoString data to the score derived from the RNA sequencing data, which was determined using the same approach that was applied to all datasets in this work. The two scores were significantly correlated, and the same classification of the four LKB1 mutations was observed by both methods. However, two samples that had moderate expression of the LKB1-loss score by RNAseq and would have been classified as LKB1-loss tumors had lower scores using the nanoString platform, This resulted in an overall concordance of 0.875 between the two methods. Further analysis of these samples is planned, which will include analysis of additional samples, better testing for LKB1 mutations, and analysis of LKB1 protein loss by immunohistochemistry. This will allow us to make more firm conclusions about the performance of this test in clinical samples.

Discussion

Accurately determining loss of a tumor suppressor such as LKB1 can be challenging and may require multiple approaches. Sequencing such genes to determine the presence of somatic mutations can provide a specific test for such lesions, but this approach cannot detect other functional inactivation of a tumor suppressor, for instance through chromosomal deletion. Mutation testing could be coupled with immunohistochemistry to allow detection of loss of expression, but reproducibility of this assay may be a problem, and determining the appropriate cutoff of expression may be

difficult in the presence of background signal. Although no targeted agents are currently known to preferentially benefit LKB1-deficient tumors, classifying the functional status of this tumor suppressor will be important for clinically testing such drugs and ultimately for prescribing such agents to patients.

We show that LKB1 loss in lung cancer is strongly associated with activation of a particular gene expression signature, and that this can be used to accurately predict LKB1 loss in multiple validation sets. A gene expression pattern termed the 'magnoid' subtype of lung adenocarcinomas has been previously associated with LKB1 loss (Wilkerson et al, 2012). Although this subtype shows significant gene overlap with our signature, when used as a predictor of LKB1 loss it detected only 64% of LKB1 mutations in the TCGA cohort (unpublished data). In contrast, the sensitivity of LKB1-loss classifier in detecting LKB1 mutations observed in the TCGA lung adenocarcinoma cohort is 0.97 (65 of 67). Furthermore, we show that our test identifies an additional 74 tumors with LKB1 loss in which LKB1 mutations were not detected bringing the overall prevalence of LKB1 loss to 30-35% in this collection of tumors. These tumors showed the same degree of pAMPK attenuation as the tumors with known LKB1 mutations (P-value = 1.7e-7 vs. signature-negative tumors), and had low LKB1 mRNA (P-value = 7.3e-28). Furthermore, many of these tumors had identifiable lesions in LKB1, with evidence of either chromosomal loss or loss of one or more exons, which accounted for half of the additional tumors found. Based on the bimodal distribution of the scores seen in the TCGA dataset we observe that this test has an inherent 6.5% misclassification rate because the two tumor populations do not show complete separation of expression scores. Thus, based on this and the performance in validation sets, we conclude that this test has greater than 0.90 sensitivity

and specificity for detecting LKB1 loss in resected lung adenocarcinomas and in NSCLC cell lines. We also show significant association between the activation of the LKB1-loss signature and instances of LKB1 loss in other tumor types (P-value = 7.2e-08), where it is less prevalent. This shows that the association between these genes and LKB1 loss is not specific to lung cancer. However, only about 50-60% of LKB1 mutations in non-lung primaries express the LKB1-loss signature. These rates varied by tumor type, so there may be tumor-specific differences in the phenotypes induced by LKB1 loss. More work focused on LKB1 loss in specific other tumor types will be required to give adequate explanation for these differences, but it is nevertheless important to show that the associations we observe are applicable outside of lung cancer.

Our work may also have important implications for model systems used to study LKB1 loss in lung cancer. There are three main experimental systems used by cancer researchers: (1) Studies of human cancer, including large-scale molecular characterizations such as those carried out by the TGA as well as human experimentation in the form of clinical trials. (2) Studies of cell lines derived from human tumors and grown over extended periods of time in cell culture. (3) Studies of mouse models in which genetic or environmental factors can be manipulated in various ways that can give rise to cancer. However, it is unclear which model is most reliable for studying clinically relevant disease phenotypes. The initial genomic (Ji et al, 2007) and proteomic (Carretero et al, 2010) characterizations of the mouse model include limited data on the effects of LKB1 in human cell lines. Our work reanalyzes the gene expression data from these studies and compares them to LKB1-associated genes from seven studies of resected human lung adenocarcinomas and NSCLC cell lines. One of the most interesting findings

of our analysis is that, while LKB1 loss induces similar effects on gene expression in resected human tumors and NSCLC cell lines, these differ substantially from the LKB1-associated genes induced in the murine model. This may suggest that key phenotypes driving the expression of these genes are not recapitulated in the mouse model.

CHAPTER III

INTERPRETING THE BIOLOGICAL SIGNIFICANCE OF THE LKB1 SIGNATURE

Introduction

Our signature is an accurate classifier of LKB1 loss, correctly predicting LKB1 mutations in the largest and most extensively characterized test set from the TCGA, in which 65 of 67 LKB1 mutant tumors express our signature (sensitivity 0.96, P-value < 1e-16). Our signature also outperforms mutational sequencing for determining LKB1 loss, as it is associated not only with mutational loss but chromosomal and exon loss of LKB1, and most probably other potential mechanisms as well. Thus, this signature can be applied to accurately classify the LKB1 loss status of hundreds of lung cancer specimens in which functional loss of LKB1 is unknown or incompletely characterized by mutational profiling. Moreover, the existence of a pattern of gene expression strongly associated with LKB1 loss suggests that there may be underlying differences in pathway activation and/or transcription factor regulation in these cancers. In this chapter we use this classification as a powerful tool to characterize molecular and clinical associations with LKB1 that reveal important aspects of the phenotypes associated with LKB1 loss. Characterizing these dysregulated pathways is important, as they may lead to novel observations about the biology of these tumors and could identify drug targets that could benefit patients with LKB1-deficient lung cancer. Indeed this is a major impetus for the study of preclinical models, and has formed the basis of the several papers characterizing

the murine model of LKB1 loss in lung cancer (Carretero et al., 2010; Chen et al., 2012; Ji et al., 2007; Shackelford et al., 2013). However, because the murine model does not recapitulate the gene expression patterns seen in human tumors, it is likely that the study of these genes in the human tumors will be more germane to understanding the biology and treatment of LKB1-deficient lung cancer.

To characterize the tumor phenotypes associated with LKB1 loss we first apply statistical tests to a variety of types of molecular data to determine which proteins, genes, or mutations, show differential activation between LKB1-deficient and LKB1 wild-type tumors. Because somatic mutations in LKB1 account for only about half the prevalence of LKB1 loss, we use our LKB1-loss gene expression signature to separate tumors into these groups for the purpose of statistical comparison. Although no assay is 100% accurate, the predictive accuracy of our expression signature is superior to determinations based only on the presence of LKB1 mutations. Because the molecular characterization is more extensive in the TCGA collection of lung adenocarcinomas, we were able to use this dataset to determine associations with mutations, copy number alterations, microRNA expression, and protein expression and phosphorylation. Associations with mutations could then be tested in additional validation cohorts, in which either LKB1 mutations were determined by sequencing, or LKB1-loss signature expression could be measured using gene expression profiling data. We also characterized the expression patterns of the dysregulated genes to identify potential underlying phenotypes that could drive their expression. We found that differentially expressed genes represented multiple distinct phenotypes, which could be seen as separate ‘clusters’ of genes whose expression was tightly correlated (Fig. 2.1b).

Interpreting which phenotypic differences are responsible for causing altered gene expression can be challenging. Our work contributes significantly to understanding these phenotypes in LKB1-deficient lung cancer. We identify three distinct subsets of upregulated genes that we ascribe to phenotypes associated with LKB1 loss and link to the activity of multiple transcription factors. We also show that restoration of LKB1 affects the expression of the CREB/FOXO subset of genes, demonstrating that this phenotype is directly responsive to LKB1 activity. Many other phenotypes are also represented within the gene signature that could not be covered in this thesis. To give a few additional examples: around 80% of adenocarcinomas expressing neuroendocrine markers exhibit loss of LKB1; evidence of WNT dysregulation is evident in LKB1-deficient tumors; and tumors with LKB1 loss have low expression of a large number of immune markers, cytokines, and stromal genes. Future work to elucidate the causes and effects of these phenotypes, in addition to the ones we characterize in this paper, will be a rich source of hypotheses regarding the biology of LKB1 loss in lung cancer.

Methods and Materials

Characterization of Four Transcriptional Nodes Comprising the LKB1 Signature

Expression data for each gene was mean centered and normalized, and unsupervised hierarchical clustering was performed with Gene Cluster 3.0 (de Hoon et al., 2004) utilizing uncentered Pearson's correlations and the centroid linkage method. Similar results were obtained for clustering of genes and tumors when using Spearman's rank correlation as the similarity metric (data not shown). Resulting heat maps were visualized using Java TreeView application (Saldanha, 2004). Unsupervised clustering

revealed correlation patterns within the genes of the LKB1-deficient signature that were reproducible across multiple resected LUAD datasets. Four distinct sets of co-regulated genes were identified from resulting dendrograms as gene clusters with internal centroid correlation values greater than 0.5, including three transcriptional nodes showing increased expression among LKB1-deficient tumors and one with decreased expression. The genes comprising these transcriptional nodes are given in Appendix B. The same four nodes could be observed in classifiers independently derived from either the Michigan or non-Michigan patients of the Director's Challenge consortium (Shedden et al., 2008), demonstrating the reproducibility of these clustering patterns across multiple datasets. A numeric score for each of these transcriptional nodes was calculated by taking the average of the standardized expression values for the genes comprising the node.

We hypothesize that the expression of these transcriptional nodes were driven by different underlying phenotypes. For subsequent analyses to characterize the biological pathways reflected by these nodes we required larger gene lists corresponding to each node. Thus, we further characterized gene expression correlations with these four gene clusters using a generalized linear model applied to gene expression data from the Director's Challenge Consortium (Shedden et al., 2008) dataset. The lm () function in the limma package of R bioconductor platform was used to determine the best fitting parameters to relate the expression of each probeset to the scores of the four LKB1-asociated gene clusters; interaction terms were not included in the model:

$$expr(x) \sim a * expr(LKB1\_loss) + b * expr(Mito) + c * expr(NRF2) + d * expr(Down)$$

Gene lists for each of the four nodes were then constructed taking the top 200 most significantly upregulated genes for that node as determined by the corresponding P-value from this model.

Determination of Clinical and Molecular Associations with LKB1 loss

Our LKB1 classification score was used to predict LKB1-loss status for unknown samples from ten collections of resected lung adenocarcinomas (total n=1297). Statistical associations were made based on these classifications. We used the TCGA dataset as a discovery cohort to determine associations with clinical phenotypes, mutations in other genes, copy number alterations, protein expression and phosphorylation, and microRNA expression. Many of these molecular phenotypes, in particular the microRNA, copy number, and proteomic data are unique to the characterization conducted by TCGA and cannot be replicated in other test sets. However, association between LKB1 loss and clinical phenotypes as well as associations with p53, EGFR, and KRAS mutations could be assessed using a pooled analysis of nine lung adenocarcinoma datasets (total n=851) in which gene expression had been analyzed and could be correlated with some or all of these phenotypes. We were also able to test associations with other mutations in a second recently published characterization of mutations in lung adenocarcinoma that employed exon capture enrichment (Imielinski et al., 2012).

All samples were classified as LKB1 wild-type or LKB1-deficient based on the expression of the LKB1-loss score. Statistical associations were made using the Fisher's exact test for categorical data (smoking status, tumor stage, mutations in other genes) or the Student's t-test for continuous variables (protein expression, copy number data,

microRNA expression). Kaplan-Meier analysis of overall survival and progression free survival was performed between these two groups for patients in the Director's Challenge Consortium (Shedden et al., 2008), using the R statistical platform with the survival package. Statistical significance was determined using the log rank test. Additionally, a Cox proportional hazards model was used to determine the association of LKB1 loss with outcomes after accounting for initial tumor stage.

Analyses to Identify Molecular Phenotypes Underlying the Activation of Transcriptional Clusters

We used several approaches to identify candidate pathways that could potentially drive the expression of the four transcriptional nodes observed in our analysis. Gene set enrichment analysis was performed using the molecular signatures database (http://www.broadinstitute.org/gsea/msigdb/) (Liberzon et al., 2011) to determine enrichment of transcription factor consensus sequences in the promoter regions of these gene lists. This tool was also used to compare our eight gene lists to previously characterized perturbation and cancer-derived signatures.

The connectivity map was used to determine significant similarities between our eight gene lists and gene perturbations induced in the cell lines MCF7, HL60, and PC3 by six hours of treatment with 1309 different small molecules. We uploaded our gene lists onto the connectivity map online analysis tool to rank compounds that were significantly associated with the gene expression phenotypes we observed. For significant hits we then downloaded ranked perturbation lists from the Connectivity Map (Lamb et al., 2006) website to generate our own enrichment P-values.

Finally, we generated an association matrix using searches of GEO (Barrett et al., 2012; Edgar et al., 2002) and ArrayExpress (Parkinson et al., 2009) to obtain perturbations of interest to our study. Because the connectivity map did not employ a lung cancer derived cell line, we searched for all perturbations made to A549, a commonly studied lung adenocarcinoma cell line with a mutation in LKB1. We next performed targeted queries for perturbations related to the hypotheses suggested by our GSEA and connectivity map analyses; specifically we searched for perturbations involving pharmacologic or genetic modulations of the CREB and FOXO3 pathways, the NRF2 transcription factor, mitochondria, and protein translation. Also, for the connectivity map, associations highlighted in Tables 3.1, 3.2, and 3.3 were made using ranked perturbation lists from the Connectivity Map website to generate our own enrichment P-values. We eliminated redundant probesets to reduce our association matrix to a single probeset per gene, and then determined the top 200 over-expressed and under-expressed genes associated with each perturbation (roughly the top and bottom 2% of changes). Numeric overlap was then determined with each of the eight cluster scores and statistical significance calculated using a hypergeometric distribution by the phyper () function in the Bioconductor limma package.

Cell Culture and Gene Transduction

A549, H2122, and H460 cell lines were generously shared with us by John Minna and Luc Girard (University of Texas, Southwestern). They were tested to ensure that they were mycoplasma negative, and were cultured in RPMI1640 containing 5% FBS, without antibiotics. Empty pBABE viral plasmids, pBABE-LKB1 and pBABE-LKB1-K78I were

obtained from AddGene. Phoenix cells were transfected with viral plasmids and retroviral particles were harvested from media supernatant 48 hours after transfection. Viruses were added to target cells with polybrene, and selection with 1.0 ug/ml puromycin was begun 48-72 hours after infection. Cells were selected under puromycin for one to two weeks before subsequent experiments, with experiments being completed within two months from the time of transfection.

CRE-Luciferase Reporter

We designed a dual-luciferase reporter driven by a 3x CRE consensus binding sequence in the promoter region in addition to a TATA box, which was inserted into an FG12 lentiviral construct. Luciferase activity from this reporter was compared to a control reporter that was identical but with mutated CRE sites. Cells were stably transduced to express CRE wild-type or mutant reporters and ratios between the two were compared after subsequent perturbations.

Immunoblots

Cell lysates were harvested while cells were in exponential growth phase in RIPA lysis buffer containing phosphatase and protease inhibitors. Lysates were homogenized and run on pre-cast SDS-PAGE gels (BioRad). Phospho-ACC (s79), ACC, and LKB1 antibodies were obtained from Cell Signaling Technology (Danvers, MA). Resulting western blots were quantitated using ImageJ software to measure pixel density and area, with results compared to those of ACC as a loading control.

Analysis of Gene Expression After LKB1 Expression in H2122 and A549 Cells

For our own LKB1 perturbation analyses, mRNA was isolated from three biological replicates of A549 and H2122 after stable expression of pBABE, LKB1 or LKB1 K78I using a qiagen mRNA isolation kit, with trizol extraction reagent. RNA concentrations were measured, and the RNA integrity number and 28s:18s ratio were calculated for quality control purposes. Amplification of 130ng total RNA was performed using Ambion WT Expression kit, and in vitro transcription was carried out overnight. cRNA was subsequently cleaned using Ambion-WT bead cleanup kit. 10.5ug of cRNA was used for second cycle cDNA synthesis and resulting cDNA was cleaned using Ambion-WT bead cleanup kit. 5.5ug of purified cDNA products were used in fragmentation and labeling reactions. Samples were hybridized overnight to a HT Human Gene 1.1 ST PM16 array plate utilizing a GeneTitan instrument. They were then scanned on the Affymetrix Gene Titan AGCC v. 3.2.3 and then analyzed on Affymetrix Expression Console v. 1.1 using a RMA normalization algorithm producing log base 2 results.

Results

Genes Associated with LKB1 Loss Correspond to Particular Tumor phenotypes and Transcription Factors

The study of differentially expressed genes can elucidate upstream pathways that drive their expression; such pathways may represent novel biological features of these tumors. Our first important observation is that these differentially expressed genes do not all represent the same phenotype. They cluster together in different groups that are

reproducibly observed across multiple datasets. This suggests that each group of genes is associated with a different phenotype and is independently regulated. These phenotypes show significant levels of overlap among tumors that have lost LKB1 but the fact that they are independent is important not only to the understanding of the biology of these tumors but also in our approach to the subsequent analysis of the gene expression patterns. The most instructive illustration of this point is a set of eight genes – AKR1C1, AKR1C2, CBR, G6PD, ME1, PGD, PIR, and SLC7A11 – that we have determined to be regulated by activation of the NRF2 pathway.

Activation of the NRF2 Pathway is Common Among LKB1-Deficient Tumors

These eight genes – AKR1C1, AKR1C2, CBR, G6PD, ME1, PGD, PIR, and SLC7A11 – consistently form a cluster of correlated genes in unsupervised hierarchical analysis. They were selected among 129 genes associated with LKB1 based on their differential expression according to LKB1 status in two training cohorts. Indeed, their expression is significantly increased in LKB1-deficient tumors in each independent dataset that we have examined. However, not every LKB1-deficient tumor expresses this set of genes; conversely, there is also a small subset of LKB1 wild-type tumors that express these genes as well. We therefore suspected that these genes could be regulated independently of LKB1 status and endeavored to discover a more direct association.

We first focused only on tumors that had lost LKB1. We separated these tumors into groups with low or high expression of the NRF2 cluster and then determined the genes that were differentially expressed between these two groups. This expanded our list from the initial eight genes to a continuous ranking of association with all genes in the

dataset. Different statistical cutoffs applied to these rankings would then yield longer lists of genes associated with this phenotype. This is important, because statistical approaches used to study gene expression patterns can be limited in power when a list of genes is too short, but can be prone to yield spurious associations when gene lists are excessively long. Examination of the resulting gene lists revealed a large number of genes with metabolic functions, and in fact three of the original eight genes are enzymes in the pentose-phosphate pathway. Thus, our first hypothesis was that this set of genes represented a metabolic phenotype of unknown significance.

This idea was intriguing, given the prominent role of LKB1 in regulating metabolism. However, exactly what we were looking at was unclear at the time, and it was challenging to plan functional experiments to further explore this unknown phenotype. Thus, we performed additional statistical tests that ultimately firmly established the underlying mechanism behind the expression of these genes. Our first specific clue came by looking at predictions of transcription factor binding sites in the promoter regions of our genes. We used predictions from the TRANSFAC database (Matys, 2006) for 368 transcription factors and compared the likelihood of a hit within genes of the signature to the likelihood of control genes that did not have association with this phenotype. This revealed a highly significant (P-value less than 10e-10) association with the transcription factor NRF2, which induces the expression of proteins that detoxify reactive oxygen species (ROS) (Taguchi et al., 2011). This transcription factor is ubiquitinated by a binding partner, KEAP1 and is subsequently degraded. Elevated ROS levels disrupt this binding, leading to NRF2 activation. Somatic mutations in NRF2 or KEAP1 can also disrupt the interaction of these two proteins, leading to constitutive

NRF2 activation (Hayes and McMahon, 2009; Shibata et al., 2008; Singh et al., 2006). Mutation of both of these genes is prevalent in lung squamous cell carcinoma (Cancer Genome Atlas Research Network et al., 2012) but in lung adenocarcinoma only mutations in KEAP1 are common (Solis et al., 2010).

It is important to note that if this transcription factor analysis is applied to unselected genes associated with LKB1 loss then this association is 'washed out' by the inclusion of many additional genes that are unrelated to this phenotype. The association with NRF2 drops from being the second most significant hit, with a P-value of 4.3e-11 to the seventeenth hit with a P-value of 8.2e-5; it would be difficult to hone in on this phenotype based on this level of data. Thus, statistical analysis generated specific gene lists corresponding to different phenotypes, and was crucial to the successful unraveling of the complex interactions among multiple phenotypes and the observed differentially expressed genes.

After the success of this initial analysis we applied our approach in a more generalized way. We first visually examined hierarchical clustering of our 129 gene LKB1-loss signature to determine clusters of genes with strong expression correlation. The standardized expression of the genes comprising each cluster were averaged within each tumor to give a single numeric score for each cluster. A multivariable general linear model was then used to determine the two hundred genes most strongly associated with each of the four transcriptional nodes shown in Fig. 2.1b; these genes are listed in Appendix C. We generated hypotheses regarding the pathways or phenotypes that drive the expression of these clusters by mining public data sources, including predicted promoter transcription factor binding sites (Liberzon et al., 2011; Subramanian et al.,

2005) and drug-induced perturbations characterized by the connectivity map project (Lamb et al., 2006).

For the NRF2 cluster, the connectivity map analysis revealed a highly significant interaction with genes induced by a potent activator of NRF2 – 15-deoxy-**Δ**12,14-prostaglandin J2. Targeted searches of gene expression data from experiments deposited in the Gene Expression Omnibus (GEO) database also showed that expression of the NRF2 cluster could be attenuated by knocking down NRF2 using siRNA (Mitsuishi et al., 2012). Comparison with expression patterns from lung squamous cell carcinoma revealed nearly identical patterns of genes expressed by tumors with NRF2 or KEAP1 mutations. These associations are presented in Table 3.1. Finally, with the availability of mutational data from the TCGA characterization of lung adenocarcinomas we show that KEAP1 mutations are present in 62% of tumors with high expression of the NRF2

Table 3.1 Results from Gene Set Enrichment Analysis of NRF2-associated gene cluster.

	Tissue/Cell line	p-value
AP1Transcription Factor Predicted promoter elements (msiDB)		6.2e-12
NRF2 Transcription Factor Predicted promoter elements (msiDB)		5.2e-11
dmPGJ2 induction	HL60	1.2e-19
	PC3	5.5e-22
	MCF7	1.0e-39
KEAP1 -/-	Mouse liver	8.8e-8
KEAP1 mutant	Lung SqCC	1.7e-62
NRF2 mutant	Lung SqCC	9.6e-69
siNRF2 repressed	A549	2.9e-24

transcriptional cluster, but only 2% of tumors in which these genes are not expressed (Fig. 3.1a). This conclusively establishes this transcriptional node as being driven by a separate somatic mutation, which has a high degree of overlap with LKB1 mutations but represents a distinct phenotype. It also shows how our analytical approach can make statistical inferences from gene expression data to arrive at meaningful associations with tumor biology.

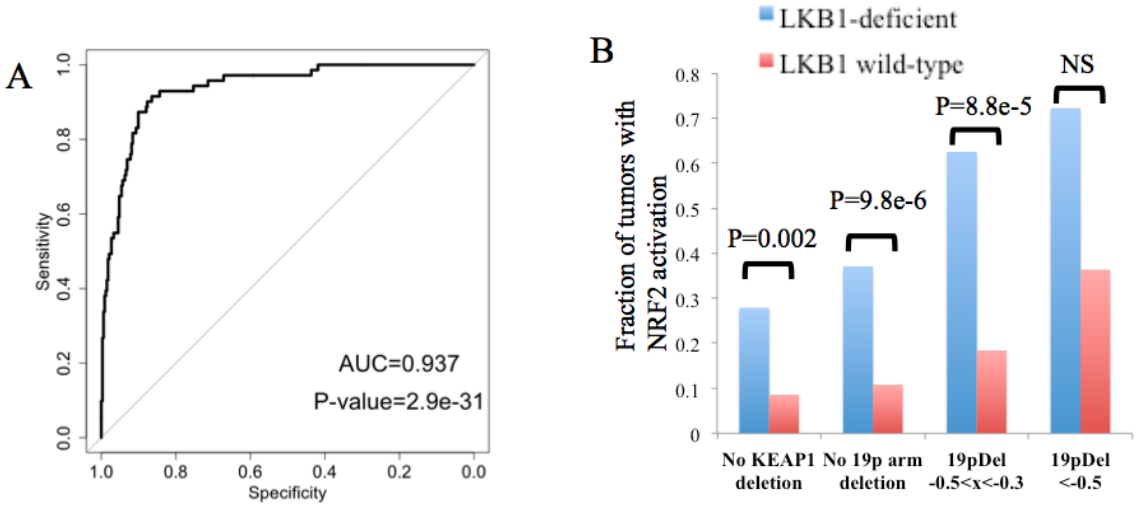


Figure 3.1. Association of NRF2 activation cluster with KEAP1 mutations and large deletions of chromosome 19. A, Receiver operating curves to show the relationship of sensitivity and specificity of the NRF2 activation score in detecting KEAP1 mutations among the TCGA lung adenocarcinomas **B,** Fraction of tumors showing high NRF2 activation score in TCGA lung adenocarcinomas groups according to LKB1 status and the presence of reported 19p chromosomal arm deletions. We considered tumors with no evidence of KEAP1 deletion, no evidence of 19p arm deletion, and tumors with low and high scores for 19p arm deletions. P-values show the significance determined by Fisher's exact test.

NRF2 Association with LKB1 is Partially Explained by Deletion Events Affecting LKB1 and KEAP1

We propose that the increase in prevalence of NRF2 activation can be attributed to selective pressures that must have existed during the evolution of these tumors.

However, both LKB1 and KEAP1 reside on chromosome 19p, with LKB1 located approximately 1MB from the telomere in cytogenetic band 19p13.3, and KEAP1 located 10MB from the telomere in adjacent cytogenetic band 19p13.2. Thus, large deletions of 19p could inactivate one copy of both LKB1 and KEAP1 and increase the likelihood of both genes being lost in the same tumor. Evidence for such lesions are observed in about 15% to 25% of TCGA lung adenocarcinomas. Thus, we must consider this as an alternative explanation to explain this overlap. If the observed phenotype were only due to the proximity of these two genes on the same chromosome, then this would have different implications for understanding the biology of these tumors and the selective pressures to which they were subjected during their evolution. To address this issue, we examined the distribution of these three phenotypes: LKB1-loss, chromosome 19p arm deletions, and NRF2 activation across TCGA lung adenocarcinomas. Indeed, the presence of a broad deletion of chromosomal arm 19p significantly increased the likelihood that either an LKB1 wild-type or LKB1-deficient tumor would have activation of the NRF2 pathway. However, after controlling for the level of chromosome 19p arm deletion we found that LKB1 loss was always significantly associated with a several-fold increased likelihood of NRF2 activation. This was true even after excluding tumors with any evidence of KEAP1 copy number loss (Fig. 3.1b). Thus, although chromosomal deletion is a mechanism that likely predisposes tumors to go on to lose either or both of these tumor suppressors, the evidence supports the notion that LKB1-deficient tumors experience significant pressure to activate NRF2.

Gene Set Enrichment Analysis of the Mitochondria/mTOR Cluster and the Downregulated Cluster

Analysis of the genes comprising the NRF2-driven cluster allowed us to concretely identify a distinct molecular pathway dysregulated in LKB1-deficient lung tumors, which is linked to somatic mutations in a different gene. This novel association between these two pathways informs us of the biology of LKB1-deficient tumors. We expect that similar analysis of genes comprising other clusters may give insight into other phenotypes relevant to LKB1-deficient lung cancer. The second cluster we consider is the 'mTOR/Mitochondria' cluster. This cluster had high expression of oxidative phosphorylation and mitochondria-associated genes as well as genes involved in protein translation. Promoter analysis identifies several transcription factors that may contribute to the induction of these genes, including ELK1, MYC, NRF1, and splicing factor 1. Furthermore, inhibition of the mTOR pathway, either by sirolimus or by the PI3K/Akt inhibitor Ly-294002 caused significant downregulation of genes in this pathway, suggesting that these genes may be reflective of mTOR activation (Table 3.2).

Downregulated genes of the signature represented contributions from multiple phenotypes, including TGF-beta (Fig. 3.2a) and NF-kB signaling as well as stroma-related genes. We were unable to identify recurrent patterns of gene expression within the downregulated genes that were consistent among datasets. This may be due to the fact that fewer downregulated genes were identified in our training cohort and clusters with fewer genes may be more prone to random differences that could vary among datasets. Also, whereas the LKB1-deficient tumors share a common mutational influence on their behavior and evolution, LKB1-wild-type tumors may be more diverse and heterogeneous,

Table 3.2 Results from Gene Set Enrichment Analysis of Mitochondria/mTOR-associated gene cluster.

	Tissue/Cell line	p-value
Mitochondrial Localization (MitoCarta)		3.6e-22
ELK1 Transcription Factor Predicted promoter elements (msigDB)		2.2e-13
Splicing Factor 1 Transcription Factor Predicted promoter elements (msigDB)		1.1e-9
NRF1 Transcription Factor Predicted promoter elements (msigDB)		3.9e-7
MYC Transcription Factor Predicted promoter elements (msigDB)		4.2e-7
LY-294002 Repression	HL60	1.1e-8
	MCF7	2.2e-17
	PC3	3.0e-12
Sirolimus Repression	HL60	3.0e-12
	MCF7	1.2e-19
	PC3	1.2e-19
PD0325901 Repression	Multiple	1.9e-24
PGC1A Induction	C2C12	1.3e-10

which would tend to make the gene expression patterns less coherent, and could potentially 'wash out' statistical associations with any of the underlying phenotypes.

We observed that in the A549 cell line TGF-beta induced a significant subset of these downregulated genes, while simultaneously attenuating both the NRF2 and LKB1-loss transcriptional components; conversely, c-SRC inhibition induced activation of the LKB1-loss transcriptional node, suggesting that TGF-beta and SRC signaling can antagonize the activity of transcription factors upstream of this node (Fig. 3.2b).

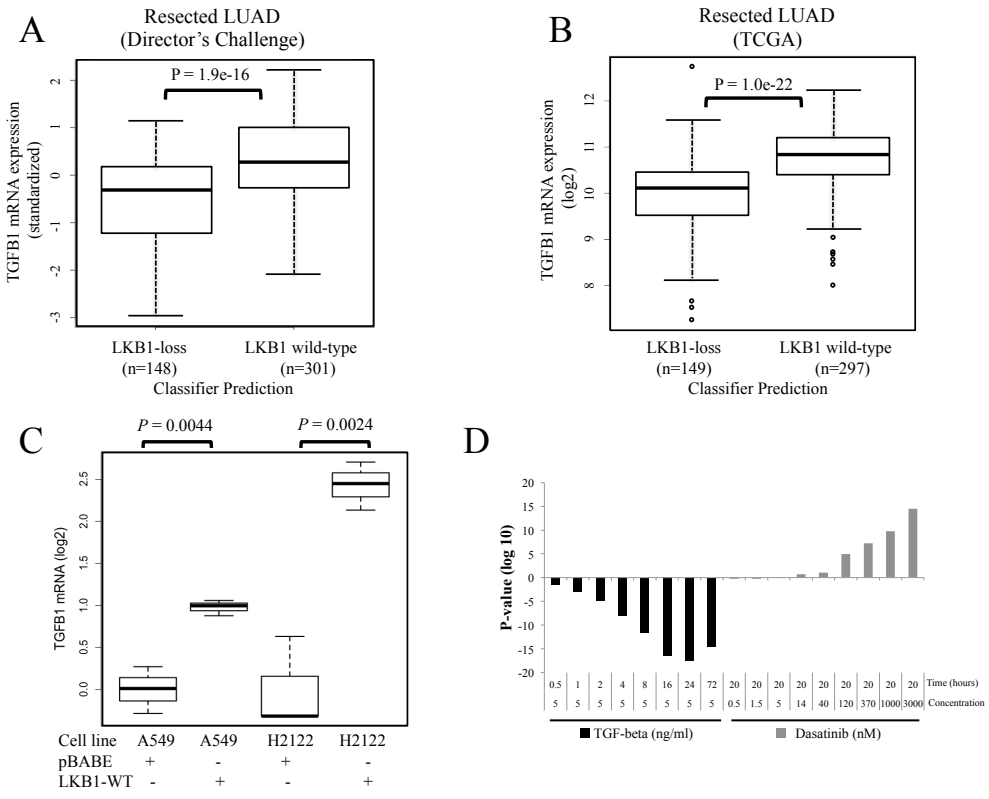


Figure 3.2. TGF-beta mRNA expression is decreased in tumors with LKB1 loss, and TGF-beta can cause down-regulation of the CREB/FOXO3 transcriptional cluster. The distribution of TGF-beta mRNA expression is shown for resected lung adenocarcinomas that are predicted by the LKB1-loss classifier to exhibit LKB1 loss or wild-type LKB1 in **A,** the Director's Challenge Consortium (n=449) or **B,** the TCGA cohort (n=446). The P-value represents the result of a student's t-test comparing these groups. **C,** Induction of TGF-beta mRNA relative to pBABE vector control is shown for A549 and H2122 cell lines after stable expression of wild-type LKB1. The range is plotted and p-values represent the result of student's t-test of the indicated comparisons. **D,** The significance of gene overlap is shown for comparisons of the FOX/CREB signature to the genes perturbed by TGF-beta or dasatinib treatment of the LKB1-mutant cell line A549 at the various time points or concentrations shown. P-values from a hypergeometric test are shown on the ordinate axis with positive values indicating an induction of FOX/CREB-associated genes and negative values indicating repression.

Gene Set Enrichment Analysis of the LKB1-loss Cluster Identifies Association with FOXO3, FOXA2, and CREB Transcription Factors

In the previous chapter we showed that of the four transcriptional nodes we identified for further characterization, one node was most strongly associated with LKB1 mutations and decreased LKB1 expression in multivariate analysis, while the other three nodes showed very little independent association with LKB1 loss. Therefore, the 16 genes comprising this node were used throughout this work as the LKB1-loss classifier. The accuracy of this classifier and its ability to detect LKB1 loss by multiple mechanisms make this an attractive approach to classifying LKB1 status in clinical specimens and in in vitro studies. Because this is the node with the strongest association to LKB1 status, the inferences drawn from these genes may also have the most direct importance to the understanding the biology of LKB1 deficient tumors.

Promoter analysis of the top 200 genes associated with the 16-gene LKB1-loss signature implicated CREB, FOXO3, and FOXA2 (or HNF3-beta) transcription factors (Table 3.3). Analysis of perturbed genes from the connectivity map also revealed induction of this cluster by colforsin, an adenylate cyclase stimulator that activates CREB, and by the typical antipsychotics thioridazine, prochlorperazine, and trifluoperazine, which have been identified as stimulators of FOXO3 transcription factors that block AKT-induced nuclear export of FOXO3 (Kau et al., 2003). We then searched the GEO and Array Express data repositories and found corroborating evidence for CREB (Zhang et al., 2005) and FOXO3 activation within this cluster (Eijkelenboom et al., 2013; Gan et al., 2010b; Tenbaum et al., 2012). Moreover, HNF3-beta promoter

occupancy is significantly increased among these genes both in the LKB1 mutant A549 lung cell line and in HEPG2 liver cell line and in human liver tissue (Rosenbloom et al., 2012). These associations are presented in Table 3.3. In addition to FOXO3, HNF3, and CREB, other transcription factors showed significant enrichment, but further corroborating evidence was lacking. These factors may also be of functional significance in LKB1-deficient tumors, especially LEF1, since increased WNT signaling has been previously implicated in these tumors (Lin-Marq et al., 2005; Liu et al., 2012b; Ossipova et al., 2003). Our analysis shows that this gene cluster represents the effects of a specific set of transcription factors that are dysregulated downstream of LKB1.

Table 3.3 Results from Gene Set Enrichment Analysis of CREB/FOXO-associated gene cluster.

	Tissue/Cell Line	p-value
CREB Transcription Factor Predicted promoter elements (msigDB)		3.1E-04
Colforsin induction	MCF7	2.2E-17
	PC3	2.6E-28
	PC12	1.6E-05
CREB regulated	Islet Cells	2.0E-14
	MIN6	4.5E-09
	HEK293T	6.5E-06
FOXO1/3/4 Transcription Factor Predicted promoter elements (msigDB)		9.5E-05
Induction by CA-FOXO3	DLD1	5.4E-11
	HuVEC	1.4E-05
	RCC4	4.2E-07
	UMRC2	3.3E-08
Prochlorperazine induction	HL60	1.6E-03
	MCF7	2.0E-06
	PC3	9.2E-03
Thioridazine induction	HL60	7.2E-05
	MCF7	1.5E-17
	PC3	7.2E-13
Trifluoperazine induction	HL60	6.0E-14
	MCF7	4.5E-10
	PC3	2.0E-06
FOXA2 Transcription Factor (HNF3B) Predicted promoter element (msigDB)		2.0E-04
Promoter occupancy (ChIP-Seq)	A549	7.1E-09
	HEPG2	4.4E-05
	Human Liver	1.1E-04

Association Between LKB1 Loss and Prevalence of Other Mutations

In addition to gene expression analysis, other types of molecular data, primarily characterized by TCGA, offer the opportunity to study functional alterations such as mutations or copy number changes in oncogenes or other tumor suppressors, and indicators of pathway activation such as phosphorylation. Co-occurrence or mutual exclusivity of mutations can reveal information about how particular pathways are activated or interact with each other within cancer cells. Mutual exclusivity – a statistically significant decrease in the co-occurrence of two somatic mutations, can occur because of functional redundancy, i.e. that both mutations could result in the same effects in the cell, or because of negative interactions between pathways downstream of the mutations, for instance if the activity of one mutation leads to attenuation of the other mutated gene due to feedback inhibitory effects. On the other hand, increased likelihood of two mutations occurring in the same tumor suggests that the effects are synergistic, perhaps by significantly amplifying oncogenic signaling through a single pathway, or by activating distinct pathways that have cooperative effects. There could be alternative explanations for co-occurrence of different mutations independent of the function of the genes. A single phenotype might increase the likelihood of the two genes undergoing mutations; for instance, this could result if a tumor had disrupted mismatch repair genes, or if a large chromosomal gain or loss led to increased chances of mutation in closely positioned genes.

To determine the association of LKB1 loss with other mutations we used 403 lung adenocarcinomas characterized by the TCGA that had both RNAseq gene expression data and determination of somatic mutations by exon capture DNAseq. Mutations were found

in 16302 genes, with a total of 102425 somatic alterations observed. Because of the high levels of exposure to carcinogens in cigarette smoke, non-small cell lung cancer has one of the highest rates of DNA mutations in cancer, with a median of five to ten nucleotide changes per megabase of DNA (Lawrence et al., 2013). Thus, most of the observed alterations are 'noise' mutations that occur at random throughout the genome having no functional significance. There are sophisticated algorithms to rank the genes affected by mutation, taking into account effects such as gene length, the context of the specific nucleotides surrounding an observed mutation, and the pattern and rate of specific base changes seen in a given tumor, in order to determine which genes are most likely to have been mutated in a nonrandom way (Imielinski et al., 2012). Non-random mutations are more likely to have been selected during the evolution of a cancer because of their functional effects. For our analysis, we narrowed our search to the top 200 most frequently mutated genes without a priori elimination of probable random mutations. We then subjected these 200 genes to a statistical test – the Fisher's exact test – to determine if their prevalence is significantly different between LKB1 mutant and LKB1-deficient tumors. Any genes included in the 200 that are affected only by chance mutations would be unlikely to segregate significantly with LKB1 status. Thus, the inclusion of such genes is unlikely to affect the results of our analysis, except to increase the likelihood of a false positive result, necessitating correction of raw P-values by multiple hypothesis testing.

Using this approach we identified five genes with significantly different mutation rates in LKB1-deficient and wild-type lung adenocarcinomas. KRAS, KEAP1, and ATM each showed significantly higher mutation rates among LKB1-deficient tumors, while EGFR and p53 had a lower rate of mutation. These results are presented in Table 3.4.

Table 3.4. Association of LKB1 loss with other somatic mutations in lung adenocarcinoma.

	LKB1-loss classifier		Fisher test		
	Number of Samples (LKB1 loss / total)	Fraction LKB1 loss (%)	Odds Ratio (95% C.I.)	P-value	FDR
TCGA Lung Adenocarcinomas					
KRAS mutant	57 / 117	48.7	2.5 (1.6, 4.0)	7.28E-05	1.46E-02
KRAS wild-type	79 / 286	27.6			
EGFR mutant	2 / 58	3.4	0.056 (0.0066, 0.22)	5.93E-09	1.19E-06
EGFR wild-type	134 / 345	38.8			
KEAP1 mutant	50 / 71	70.4	6.8 (3.7, 12.6)	3.15E-12	6.30E-10
KEAP1 wild-type	86 / 332	25.9			
ATM mutant	22 / 36	61.1	3.5 (1.6, 7.6)	6.49E-04	1.30E-01
ATM wild-type	114 / 367	31.1			
Pooled analysis Lung Adenocarcinomas					
KRAS mutant	50 / 111	45	2.3 (1.4, 3.7)	0.00035	
KRAS wild-type	85 / 322	26.4			
EGFR mutant	4 / 76	5.3	0.086 (0.02, 0.24)	4.3E-10	
EGFR wild-type	116 / 293	39.6			
NRF2 active	118 / 198	59.6	5.1 (3.6 7.3)	7.40E-22	
NRF2 low	146 / 653	22.4			
Imielinski Lung Adenocarcinomas[a]					
KEAP1 mutant	7 / 22	31.8	3.3 (1.0, 9.8)	2.50E-02	
KEAP1 wild-type	20 / 161	12.4			
ATM mutant	8 / 20	40	5.0 (1.6, 15)	3.00E-03	
ATM wild-type	19 / 163	11.7			

[a] Gene expression was not available for the Imielinski study, so these comparisons are made on the basis of somatic mutations in LKB1.

We then looked for additional datasets of somatic mutations to see if these observations could be made in independent tumor sets. We used our previously characterized pooled analysis of 851 lung adenocarcinomas to test associations with EGFR, KRAS, and p53, because a subset of these tumors had undergone selected mutational profiling. For these associations we used the LKB1-signature score to classify the LKB1 status of tumors and made statistical comparisons based on these classifications. This analysis confirmed that LKB1 has a significant association with EGFR and KRAS prevalence, but no association was seen with p53. For KEAP1 and ATM we made use of mutations profiled in 183 lung adenocarcinomas characterized by Imielinski et al (Imielinski et al., 2012). Because gene expression data were not available for this sample set we made comparisons with LKB1 mutations observed in the study. Although this is a smaller dataset and limited by the lack of gene expression data, KEAP1 and ATM both showed significantly increased rate of loss among LKB1-mutant tumors, confirming our initial analysis. While LKB1 has been previously associated with differences in KRAS and EGFR mutation rate (Koivunen et al., 2008; Matsumoto et al., 2007), the increased prevalence of KEAP1 and ATM mutations is novel and may give insight into the biology of these tumors.

Clinical Phenotypes of LKB1-Deficient Lung Adenocarcinomas

Loss of LKB1 in the murine model results in aggressive, metastatic tumors, and a LKB1-metastasis gene signature derived from this study predicts worse prognosis in patients with lung cancer (Carretero et al., 2010; Ji et al., 2007). However, a large study of LKB1 mutations in resected human tumors showed that LKB1 loss was not associated

with worse prognosis (Koivunen et al., 2008). We used both the TCGA dataset and our pooled analysis to determine associations between LKB1 loss and clinical variables. We see that there is no difference in the prevalence of LKB1 loss between tumors of different initial stage (Table 3.5), and that LKB1 loss does not affect prognosis (Fig. 3.3). Results shown in our Kaplan-Meier analysis were derived considering all stages of tumors. Separately we performed a Cox proportional hazards model to assess the association of LKB1 loss with inclusion of clinical stage as a covariate; this also reported no association (P-value = 0.37). We did observe a significant association with smoking history, with ever-smokers having a higher prevalence of LKB1 loss than lifelong never-smokers (P-value = 0.03 for TCGA, 6.5e-09 for pooled analysis; Table 3.5). This association has also been previously reported (Koivunen et al., 2008).

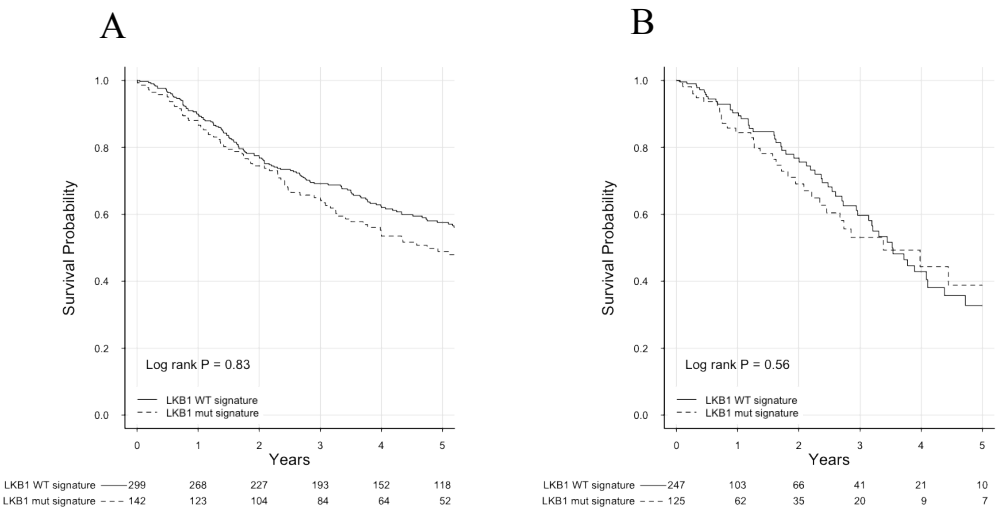


Figure 3.3. Association between patient outcome and LKB1 loss.
Resected LUAD tumors from the Director's Challenge Consortium (**A,** n=441), or from TCGA (**B,** n=373) were classified as LKB1-loss or LKB1 WT using the LKB1-classifier score. Kaplan-Meier curves were used to plot cumulative events for these two groups for overall survival, P-values represent the results of the log-rank test; the number of evaluable tumors remaining are given at yearly intervals below each plot.

Table 3.5. Association of LKB1 loss with clinical variables in lung adenocarcinoma.

	LKB1-loss classifier		Fisher test	
	Number of Samples (LKB1 loss / total)	Fraction LKB1 loss (%)	Odds Ratio (95% C.I.)	P-value
TCGA Lung Adenocarcinomas				
Stage I	71 / 217	32.7	NA	NA
Stage II	32 / 88	36.4	1.2 (0.67, 2.0)	0.59
Stage III	22 / 64	34.4	1.1 (0.57, 2.0)	0.88
Stage IV	9 / 20	45	1.7 (0.59 4.7)	0.32
Never smoker	13 / 55	23.6	0.48 (0.22, 0.97)	0.031
Ever smoker	96 / 245	39.2		
Pooled Analysis Lung Adenocarcinomas				
Stage I	153 / 476	32.1	NA	NA
Stage II	51 / 148	34.5	1.1 (0.68 1.7)	0.62
Stage III	31 / 118	26.3	0.73 (0.43, 1.2)	0.27
Stage IV	2 / 10	20	0.52 (0.052, 2.7)	0.51
Never smoker	12 / 116	10.3	0.20 (0.010, 0.38)	6.50E-09
Ever smoker	200 / 553	36.2		

Protein, MicroRNA, and Copy Number Alterations Associated with LKB1 Status

We also performed statistical comparisons to determine differences in expression of proteins, phosphorylated proteins, and micro-RNA and to characterize the differences in prevalence of copy-number alterations between LKB1-deficient and LKB1 wild-type lung adenocarcinomas. For proteomic data generated from reverse phase protein arrays and for microRNAseq expression data we used a Student's t-test to calculate P-values for observed differences. For copy number alterations we used processed data from TCGA corresponding to 79 amplification or deletion regions that had been selected as

statistically significant recurrent lesions using GISTIC analysis. Continuous variable copy number data had been converted into discrete values to indicate low and high level amplifications, and the presence or absence of deletions for these peaks, and Fisher's exact test was used to calculate P-values. Copy number associations are presented in Table 3.6; microRNA associations are presented in Table 3.7; proteomic associations are presented in Table 3.8.

Table 3.6. Association of LKB1 loss with copy number alterations in lung adenocarcinoma.

Chromosomal region	Alteration	Cancer genes in region	P-value
Increased prevalence in LKB1-deficient lung cancer			
19p13.2	Deletion	STK11	1.2E-09
19p13.3	Deletion	KEAP1	1.4E-07
3q29	Deletion	TP63	2.0E-05
10p15.1	Amplification	AKR1C2	5.8E-05
Decreased prevalence in LKB1-deficient lung cancer			
5p15.33	Amplification	TERT	3.8E-08
15q11.2	Deletion		1.4E-07
10q26.3	Deletion		1.3E-06
7p21.1	Amplification		5.4E-05
17q12	Amplification	ERBB2	0.00013
7p11.2	Amplification	EGFR	0.00024
5p13.1	Amplification	AMPK	0.0007
7q31.2	Amplification	MET	0.0011

Table 3.7. Association of LKB1 loss with differences in microRNA expression in lung adenocarcinoma.

	P-value	microRNA	P-value
Increased expression in LKB1-deficient cancer		**Decreased expression in LKB1-deficient cancer**	
hsa-mir-582	1.4E-21	hsa-mir-146b	1.6E-16
hsa-mir-148b	8.4E-14	hsa-mir-222	2.1E-15
hsa-mir-141	2.1E-13	hsa-mir-221	2.5E-13
hsa-mir-203	2.5E-13	hsa-mir-452	5.0E-11
hsa-mir-192	1.8E-11	hsa-mir-205	6.3E-10
hsa-mir-194-1	2.2E-11	hsa-mir-542	6.4E-10
hsa-mir-194-2	4.1E-11	hsa-mir-146a	1.0E-08
hsa-mir-375	7.8E-11	hsa-mir-500a	2.5E-08
hsa-mir-200c	1.3E-09	hsa-mir-378	5.5E-08
hsa-mir-338	2.9E-06	hsa-mir-342	3.4E-06
		hsa-mir-155	4.7E-06
		hsa-mir-103-1	7.6E-06
		hsa-mir-15a	2.1E-05
		hsa-mir-181b-1	3.4E-05
		hsa-mir-589	3.9E-05
		hsa-mir-181a-1	6.6E-05

Table 3.8. Association of LKB1 loss with differences in protein expression and phosphorylation in lung adenocarcinoma.

Protein or phosphorylated protein	P-value	Protein or phosphorylated protein	P-value
Increased expression in LKB1-deficient cancer		**Decreased expression in LKB1-deficient cancer**	
Claudin-7	2.5E-19	AMPK pT172	4.3E-14
c-Kit	1.6E-17	PKC-alpha pS657	3.4E-10
AMPK alpha	1.6E-07	Annexin I	2.0E-07
TIGAR	1.3E-06	Axl	1.8E-06
HER3	1.7E-06	PKC-alpha	2.1E-06
IGFBP2	8.6E-06	KEAP1	3.2E-06
Rab11	9.3E-06	STAT5-alpha	4.1E-06
AR	1.2E-05	NF-kB-p65 pS536	8.3E-06
CDK1	5.5E-05	P38 MAPK	2.6E-05
Nrf2	1.7E-04	NF2	3.5E-05
LCN2a	2.1E-04	Dvl3	4.2E-05
Bim	2.4E-04	PDK1 pS241	1.2E-04
CD31	5.6E-04	ERK2	2.1E-04
		TSC2	3.1E-04
		p70S6K	4.4E-04
		PI3K-p110-alpha	4.7E-04
		PI3K-p85	6.1E-04
		Syk	1.1E-03

Wild-type LKB1 Decreases the Expression of the LKB1-associated Signature Genes

To test the direct effects of LKB1 on the regulation of the observed gene expression patterns we stably expressed LKB1 or mutated LKB1 K78I in three NSCLC cell lines – H2122, A549 and H460 – lacking functional tumor suppressor. Expression of LKB1 was confirmed by western blot for both wild-type and K78I LKB1. However, only wild-type LKB1 significantly induced phosphorylation of acetyl-CoA-carboxylase, a well-recognized downstream target of AMPK (Fig. 3.4a). The CREB transcription factor

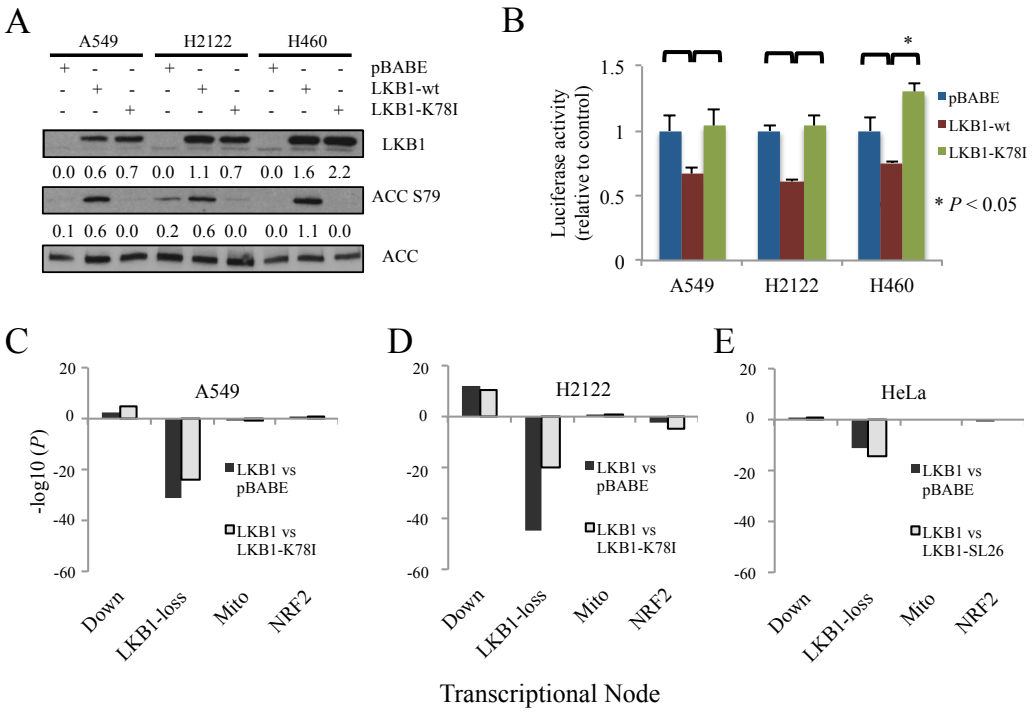


Figure 3.4. Restoring wild-type LKB1 in cell lines harboring mutations slows growth and attenuates the expression of the LKB1-deficient gene signature.
A, Immunoblots of whole-cell lysates from A549, H2122, and H460 stably expressing emtpy pBABE vector, LKB1 or K78I LKB1. Quantitation of western blot signal is shown relative to ACC using ImageJ software. A single replicate was performed. **B,** Activity of CRE-luciferase is shown for A549, H2122, and H460 cell line after stable expression of LKB1 or K78I LKB1. Reporter activations were determined relative to a control luciferase with mutated CRE sites, and are shown relative to the pBABE control. P-values show the significance of unpaired student's t-tests. **C-E,** Changes in gene expression of A549, H2122, or HeLa cell lines after re-expressing wild-type or mutant LKB1 were compared to the gene lists for each of the four LKB1-associated clusters using a hypergeometric test. Log10 P-values are indicated on the y-axis, with positive values indicating induction of expression and negative values indicating repression.

was identified in our previous analysis as a putative driver of a significant fraction of the LKB1-loss associated genes and has been previously shown to be regulated by LKB1 (Feng et al., 2012). To give further evidence that our in silico findings translate to actual biological changes we used a luciferase reporter driven by the CRE-consensus sequence to assess activation of this transcription factor. Attenuation of CREB activation was confirmed in A549, H2122, and H460 cell lines yielding a reduction in reporter activity

of 30-40% (Fig. 3.4b; P-value less than 0.05 for each cell line). Microarray analysis of gene expression changes in A549 and H2122 shows that LKB1 significantly

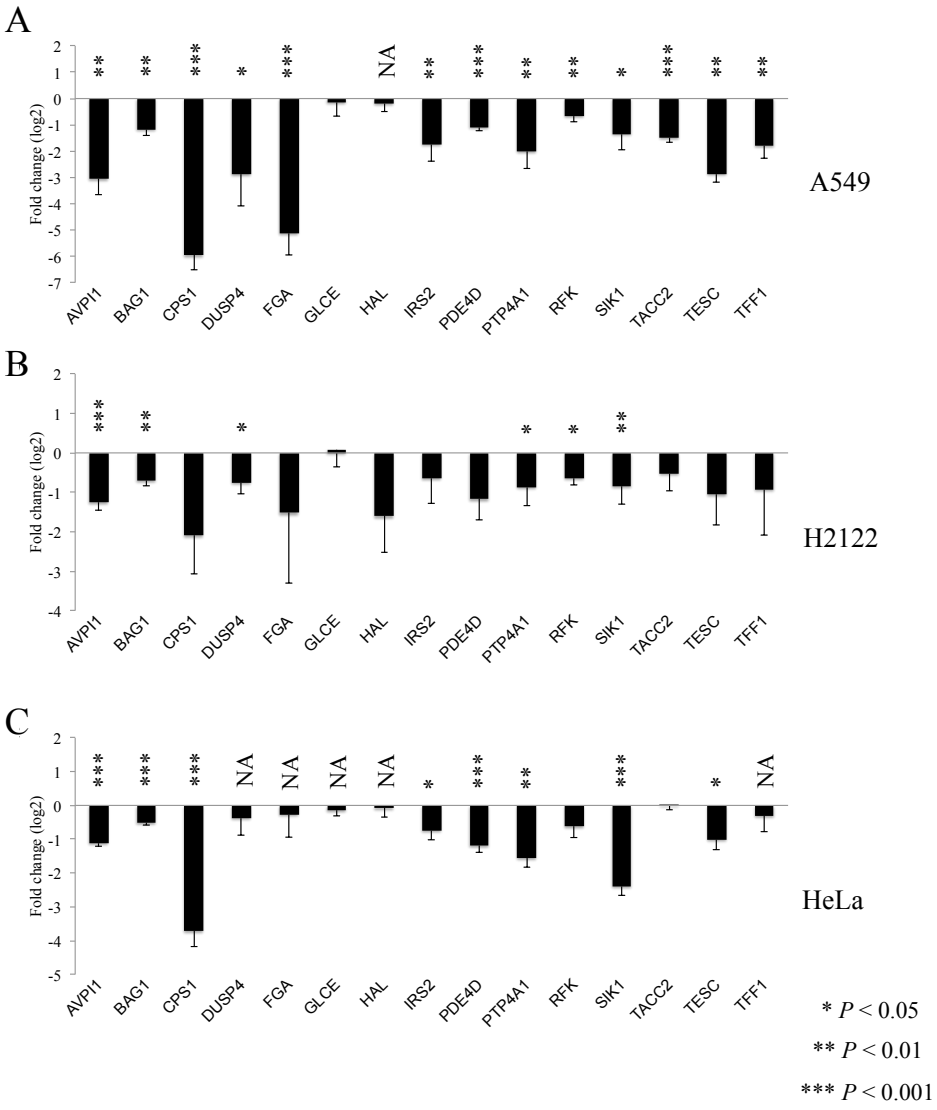


Figure 3.5. Expression of wild-type LKB1 in A549, H2122, or HeLa cell lines decreases the expression of the genes in the CREB transcriptional node.
Microarray gene expression was measured in triplicate after stable expression of pBABE vector or LKB1-wild-type in A549, **A,** H2122, **B.** or HeLa, **C,** For each gene comprising the CREB transcriptional node the average change in gene expression (log base 2) is plotted comparing LKB1 wild-type to pBABE control. Error bars represent standard deviations; P- values represent the results of student's t-test comparing these groups. NA indicates minimal basal expression.

downregulates the over-expressed genes associated with LKB1 in each of these cell lines (Fig. 2.1a, and Fig. 2.2). We then used the genesets associated with each of the four transcriptional clusters do determine whether LKB1 affected the activation of these clusters to a similar degree or whether some clusters were more affected than others. This analysis revealed that restoration of LKB1 significantly (P-value < 1.0e-30 by hypergeometric test) downregulates the LKB1-loss gene cluster, while increasing the expression of a subset of the downregulated genes. MTOR/mitochondria and NRF2 associated clusters were unaffected (Fig. 3.4c,d). Similarly, analysis of gene expression changes after restoring LKB1 in HeLa cells (data shared by Dr. Lin-Marq (Lin-Marq et al., 2005)) also showed attenuation only of the LKB1-loss cluster (Fig. 3.4e; *P-value* = 5.0e-15 by hypergeometric test), demonstrating that LKB1 induces consistent effects across different cancer types. In addition to this statistical approach, we also show the effect of LKB1 on the expression of the individual genes in the 16-gene signature (Fig. 3.5).

Discussion

In this chapter we have used statistical methods to give a detailed analysis of a variety of phenotypes associated with LKB1 loss in lung cancer. The main validation of the LKB1-loss classifier is the demonstration that it accurately predicts LKB1 mutations and non-mutational loss in independent validation sets. Here we show that this accuracy makes it a powerful tool for discovering novel associations with this tumor suppressor. Before going into some detail about the novel findings uncovered with this approach, it is useful to consider the associations that corroborate previously established observations in

LKB1-deficient lung cancer. Koivunen et al performed sequencing of LKB1, BRAF, KRAS, and EGFR in 310 non-small cell lung cancers and reported associations with mutations and clinical variables (Koivunen et al., 2008). This analysis showed that LKB1-mutant lung tumors had a higher prevalence of KRAS mutations and a very low prevalence of EGFR mutations. There was also a significant association with smoking status, adenocarcinoma histology, and Caucasian race compared to tumors from Asian patients. No difference in tumor stage or outcome was observed. The associations between LKB1 and EGFR (Ding et al., 2008; Matsumoto et al., 2007), KRAS (Mahoney et al., 2009; Matsumoto et al., 2007), and smoking (Matsumoto et al., 2007) have also been observed in other datasets. Thus, our analysis is able to recapitulate the previous molecular and clinical associations made by multiple studies, and to increase the statistical significance of these observations substantially by virtue of the increased sample size available to us.

With regard to our statistical inferences from gene expression data, we identify the CREB transcription factor as being activated in LKB1-deficient lung cancer. This association is also well established in the literature both in cancer (Feng et al., 2012; Gu et al., 2012; Komiya et al., 2009; Shackelford and Shaw, 2009) and in the regulation of glucose homeostasis (He et al., 2009; Koo et al., 2005a; Screaton et al., 2004; Shaw, 2005). LKB1 controls the activation of CREB through the phosphorylation of the CREB transcriptional coactivators CRTC1, CRTC2, and CRTC3, which is carried out by the salt-inducible kinases and AMPK, which are in turn directly activated by LKB1. Phosphorylation by these kinases results in sequestration of CRTC family members in the cytoplasm, leading to attenuation of CREB activity homeostasis (He et al., 2009; Koo et

al., 2005b; Screaton et al., 2004; Shaw, 2005)). CREB has been shown to have oncogenic roles in lung cancer (Aggarwal et al., 2008; Seo et al., 2008), and CRTC1 and CREB activity have been shown to be important for the growth and migration of LKB1-deficient lung cancers, at least in part due to their role in upregulating a key mediator, NEDD9 (Feng et al., 2012; Ji et al., 2007). Thus, the connection we see between LKB1 loss and CREB activation is not a novel observation, but shows that our statistical approaches yield reliable conclusions in accord with previous findings.

This ability to independently demonstrate previously established findings using our statistical methods lends credence to the many novel findings we observe. We see increased expression of two clusters of genes that have important roles in tumor metabolism – one related to oxidative phosphorylation and mitochondrial biogenesis, the other – the NRF2 pathway – related to detoxification of reactive oxygen species. Increased expression of mitochondrial genes in these tumors may represent compensation for defective mitochondrial function, which has been demonstrated in LKB1-deficient cells due to loss of AMPK-induced mitophagy (Egan et al., 2011; Shackelford et al., 2013). Alternatively, there could be activation of transcription factors that govern mitochondrial biogenesis, either downstream of mTOR or other pathways. PPAR-gamma coactivator 1A (PGC-1A) is one such regulator that links mTOR to mitochondrial biogenesis (Cunningham et al., 2007; Fernandez-Marcos and Auwerx, 2011); it is also a target of CREB and one of the most highly over-expressed genes in LKB1-deficient tumors. Understanding how these pathways interact to regulate tumor metabolism in the absence of LKB1 may be important in the design of therapeutic strategies that

incorporate metabolic inhibitors such as metformin, which may show promise in the treatment of these tumors (Shackelford et al., 2013).

Dysregulated production of potentially defective mitochondria could also be related to the activation of NRF2 that occurs preferentially among lung tumors lacking LKB1, since these mitochondria may produce increased reactive oxygen species (Shackelford et al., 2013). NRF2 is a key activator of the oxidative stress response and also plays a role in metabolic reprogramming of cancer cells (DeNicola et al., 2011; Mitsuishi et al., 2012). Although the ROS detoxification induced by constitutive NRF2 activation is beneficial in some contexts (DeNicola et al., 2011; Homma et al., 2009; Ohta et al., 2008; Shibata et al., 2008; Singh et al., 2008), ROS can alter signal transduction within the cell by inactivating phosphatases – leading to PI3K, EGFR (Chen, 2006; Hirota, 2001), Src, and TGF-beta activation (Murillo et al., 2007) – and can be important for tumorigenesis (De Raedt et al., 2011; Weinberg et al., 2010). LKB1-deficient tumors have been shown to be susceptible to oxidative stress, as they are unable to make the appropriate adaptive responses in metabolism and biosynthesis (Jeon et al., 2012). Furthermore, NRF2 activation has been shown to confer resistance to chemotherapy, and thus may be an important clinical phenotype (Homma et al., 2009; Singh et al., 2008; Solis et al., 2010). NRF2 is frequently activated by somatic mutations in KEAP1 in NSCLC (Singh et al., 2006; Solis et al., 2010), and our analysis of the TCGA lung adenocarcinomas shows that roughly two thirds of tumors with expression of the NRF2 transcriptional signature harbor KEAP1 mutations. We demonstrated that this increase in prevalence is still seen after controlling for different levels of chromosome 19p deletion, suggesting that selective pressure exists for NRF2 as a secondary protective

mechanism either in response to increased levels of ROS production, for instance by defective mitochondria, or decreased ability to respond to normal levels of stress appropriately in the absence of LKB1. It is likely that both alternatives could play a role. Understanding the interactions between these metabolic effects, pathway activation, and drug sensitivity presents an interesting focus for future research.

ATM is another tumor suppressor that is mutated at a higher rate among LKB1 deficient cells. ATM is a serine and threonine kinase that is involved in the DNA damage recognition checkpoint. When it is activated, by DNA damage or other cellular stress, it phosphorylates a number of downstream targets, such as p53, BRCA1, CHK1 and CHK2, to inhibit cell cycle progression and either repair the DNA damage or send the cell into apoptosis (Shiloh, 2003). Interestingly, ATM is known to phosphorylate both LKB1 (Sapkota et al., 2002) and AMPK, which can be induced by etoposide (Luo et al, 2013), reactive oxygen species (Alexander et al., 2010) a pharmacologic activator of AMPK (Sun et al., 2007), or through IGF-1 induced oncogenic signaling (Suzuki et al., 2004). Thus, ATM can be placed in the same AMPK regulating pathway as LKB1, and this may underlie the increased frequency of ATM loss among LKB1-deficient cancers. Furthermore, because AMPK activation can be induced by reactive oxygen species in the absence of LKB1, this may be another driving force behind the selection for NRF2 activation in these tumors.

The gene cluster with the strongest association to LKB1 loss also provided interesting insights into pathway activation in LKB1-deficient tumors. We identified three transcription factors that had strong associations with the genes in this cluster, CREB, FOXO3, and FOXA2 (HNF3). The association between LKB1 and CREB is well

established in the literature and was discussed above. FOXA2 is a forkhead box transcription factor also known as hepatic nuclear factor 3-beta. It plays roles in lung and gastrointestinal development (Snyder et al., 2013; Wan, 2004) FOXA2 affects the differentiation state of tumors (Gupta et al., 2012; Qi et al., 2010; Snyder et al., 2013), and has been reported to exert tumor suppressive effects (Basseres et al., 2012; Liu et al., 2012a; Tang et al., 2010). AMPK activation has been shown to lead to downregulation of FOXA2 activation after phosphorylation by AKT (Yokoyama et al., 2011). Chromatin immunoprecipitation of FOXA2 in A549 cells or HEPG2 cells shows binding to several of the genes most strongly associated with LKB1-deficient lung cancer, including AVPI1, DUSP4, FGA, ID1, NR4A2, RFK, S100P, and TFF1. Some of these genes have also been shown to have CREB or FOXO3 binding sequences in their promoter, raising the possibility of cooperativity between the different transcription factors we identified. More research, and into the general effects of this in the context of LKB1-deficient lung cancer could be enlightening.

The final transcription factor we found to be associated with the LKB1-loss gene cluster is FOXO3. FOXO3 has tumor suppressive roles in cancer, and is known to induce the expression of the pro-apoptotic factors BIM and PUMA. Its transcriptional activity is controlled by phosphorylation from ERK and AKT kinases, which cause it to maintain cytoplasmic localization and induce little gene expression (Calnan and Brunet, 2008; Zhang et al., 2011). It has been previously reported that restoring LKB1 in LKB1-mutant lung cancer cell lines led to activation of AKT and inhibition of FOXO3 and other apoptotic regulators (Zhong et al., 2008). Our gene expression analysis suggests that activation of this transcription factor may be a phenotype generally associated with LKB1

loss, with important implications for therapy. We go into more detail regarding the potential for AKT to play a role in this gene signature as well, and the consequences of FOXO3 activation in the following chapter.

These phenotypes that are associated with LKB1 loss in humans are distinctly different from those observed in the murine model of LKB1/KRAS mutant lung cancer. In contrast to human tumors with LKB1 loss, murine tumors did not express genes characteristic of FOXO3, CREB or NRF2 activation. There are a number of potential effects that could explain the differences observed between the mouse and human data. For instance, the mouse model only examines the role of LKB1 loss in KRAS-mutant lung tumors, while comparisons of human tumors encompass heterogeneous genetic backgrounds. Furthermore, mouse tumor models of mutant KRAS in the absence of additional genetic alterations produce relatively benign lesions (Ji et al, 2007) that may be more similar to adenomas than adenocarcinomas and thus may be a poor representation of LKB1 wild-type lung cancer. For the Carretero study, all gene expression data from LKB1 wild-type tumors were derived from this KRAS mutant model. However, in the Ji study we could include a comparison of murine LKB1/KRAS tumors with p53/KRAS tumors, which both have similarly aggressive phenotypes (Fig. 2.1). To control for the greatest number of potential confounding variables, we also determined differentially expressed genes in the TCGA lung adenocarcinomas for comparison of LKB1/KRAS mutant tumors (n=29) to KRAS/p53 mutant tumors (n=41). This gene list was essentially unchanged from the overall analysis of LKB1-associated genes in the TCGA study (116 overlapping genes among the top 200 from each study) but still showed only three overlapping genes when compared to the mouse study of the same genetic groups.

Another possibility is that key features of the biology of LKB1-deficient human tumors may not be recapitulated in the murine model. As mentioned above, these murine tumors do not express genes that we associate with several putatively activated transcription factors in LKB1-deficient tumors. Additionally, TGF-beta and related signaling pathways, which are either unchanged or downregulated in LKB1-deficient human tumors, show increased expression in the murine model (Carretero et al., 2010; Ji et al., 2007). TGF-beta has many effects on cell signaling, but we show that TGF-beta treatment of the A549 cell line significantly downregulates both the LKB1-loss and NRF2 cluster, and thus may antagonize the activation of the pathways that are characteristically activated in human tumors with LKB1 loss. Although we do not have an explanation of the clear differences between human tumors and the manipulated murine model it is possible that manipulating one or more of these pathways in the mouse model could result in tumors that more closely reflect the phenotype of LKB1 loss in human lung cancer.

We argue that each of the four gene clusters represents a distinct phenotype that may be regulated independently of each other. To determine whether this regulation could be directly affected by LKB1 expression we used an in vitro cell line model that has been used in several other studies (Feng et al., 2012; Ji et al., 2007; Zhong et al., 2008) of restoring LKB1 expression in NSCLC cell lines that harbor LKB1 mutations. Gene expression analysis of the perturbations induced by expressing LKB1 in A549 and H2122 cell lines showed that LKB1 significantly ($p<1e-30$) downregulated the gene cluster associated with the LKB1-loss signature and linked to FOXO3, FOXA2, and CREB transcription factors. This shows that LKB1 has direct effects on the activation of

these genes, potentially through regulation of the transcription factors we identified. For the CREB transcription factor we confirmed this attenuation using a luciferase assay for CREB activity.

Thus, we have shown that a specific gene expression pattern is consistently associated with LKB1 loss in multiple datasets and can be used to predict loss of LKB1 in clinical samples. We have demonstrated that many molecular phenotypes are associated with loss of this tumor suppressor. Some of these have been previously characterized but many are novel and give new insight into the biology of these tumors. Furthermore, we show that the subset of genes most strongly associated with LKB1 loss – including the 16 genes used to classify LKB1 mutational status – is directly regulated downstream of LKB1. We next wanted to determine the mechanism by which LKB1 regulates these genes and transcription factors and to determine whether any of the implicated pathways represent potential targets for therapy in lung cancer. Because we have shown that restoring LKB1 expression in vitro produces changes in expression of the FOX/CREB transcriptional cluster, this isogenic LKB1-addback model is an ideal system for testing hypotheses regarding interactions of LKB1, FOXO3, and clinically relevant phenotypes such as drug sensitivity and apoptosis, which we explore in the following chapter.

CHAPTER IV

LKB1 LOSS IS ASSOCIATED WITH SENSITIVITY TO MEK INHIBITION AND ALTERATIONS IN PI3K-AKT-FOXO3 SIGNALING

Introduction

The goal of this work is improved treatment of patients with lung cancer by identification of novel aspects of tumor biology that could lead to new or better therapies. Tumors that are driven by mutated constitutively active oncogenes, so called 'oncogene addicted' tumors, are responsive to targeted pharmacologic inhibition of the oncogene. Both monoclonal antibodies and small molecule inhibitors have proved to be marvelously effective in several tumor types. For example, tyrosine kinase inhibitors (TKIs) have changed the natural history of chronic myelogenous leukemia (Druker et al., 2001), gastrointestinal stromal tumor (Demetri et al., 2002), and mutated oncogene-driven subsets of lung cancer (Lynch et al., 2004; Paez et al., 2004; Shaw et al., 2013) and melanoma (Chapman et al., 2011; Flaherty et al., 2012; 2010; Sosman et al., 2012); similarly, monoclonal antibody inhibitors of HER2/neu are dramatically effective in HER2 amplified breast cancer (Slamon et al., 2001) and gastric cancer. Research to improve on these treatments has focused on understanding innate and acquired resistance mechanisms and developing strategies to overcome resistance, discovery of new drugs targeting different epitopes on the oncogene or combining two inhibitors with different sites of action. On the other hand, reversing the tumorigenic effects of inactivated tumor suppressors genes such as LKB1 is considerably more challenging. Restoration of lost

tumor suppressor activity in patients has not proved feasible so far. Thus, we and others are focusing efforts on identifying potentially targetable downstream 'driver' pathways activated as a consequence of suppressor loss.

One such pathway in LKB1-deficient tumors is mTOR activation. There is a direct link from LKB1 loss, which results in inhibition of AMPK phosphorylation, thence to mTOR activation. Clinical trials have not yet assessed mTOR inhibitors in the subset of LKB1-deficient tumors, so it is so far unknown whether these approaches will have efficacy. Our analysis of gene expression patterns has revealed several upregulated transcription factors, in addition to the mTOR pathway, specifically CREB, NRF2, and FOXO3, that could influence drug sensitivity or lead to the identification of drug targets in LKB1 deficient lung cancers. The CREB transcription factor has oncogenic effects (Aggarwal et al., 2008; Feng et al., 2012; Seo et al., 2008), and is activated downstream of several pathways through which oncogenic signals can be induce, such as G-coupled protein receptors, the MEK/ERK pathway, and intracellular calcium signaling. NRF2 activation has been shown to influence drug response, likely by suppressing the cytotoxic effects of free radicals induced by treatment.

Finally, the FOXO3 transcription factor has been shown to be an important determinant of the apoptotic response by induction of the expression of pro-apoptotic factors BIM, PUMA, and FAS ligand, as well as cell cycle inhibitors p21 and p27. The activity of FOXO3 is tightly regulated by post-translational modification including acetylation, phosphorylation, and ubiquitination (Calnan and Brunet, 2008; Fu and Tindall, 2008; Zhang et al., 2011). Phosphorylation of FOXO3 at threonine 32, serine 253, and serine 315 is induced by AKT and SGK kinases, which results in inhibition of

FOXO3 activity by promoting its cytoplasmic sequestration by 14-3-3 proteins. Acetylation by CBP and p300 inhibits FOXO3 as well, also primarily by affecting localization (Calnan and Brunet, 2008). Inhibition of FOXO3 by ERK phosphorylation also occurs, and this has been shown to cause FOXO3 degradation in an MDM2 dependent manner (Yang et al., 2008). Conversely, other amino acids of FOXO3 are phosphorylated by AMPK, and this has been shown to have an activating effect on the transcription factor (Greer et al., 2007). Thus, there are a number of pathways that interact to regulate the activity of this transcription factor, the dysregulation of which may result in the activation of downstream targets identified in our gene expression data. Identifying and understanding these dysregulated pathways, as well as the effects of FOXO3 itself, may inform our understanding of LKB1 influence on drug sensitivity for particular targeted agents.

In addition to identification of new targets in LKB1 deficient tumors, we have also used empiric drug sensitivity data to identify drugs with specific activity in LKB1 tumors, utilizing our 16-gene signature to classify the LKB1 deficient cell lines. Extensive characterizations of in vitro drug response to a variety of targeted inhibitors were recently reported (Barretina et al., 2012; Garnett et al., 2012). In the Cancer Cell Line Encyclopedia (CCLE) study (Barretina et al., 2012) 505 cell lines were tested for sensitivity to 24 compounds, and in the Genomics of Drug Sensitivity in Cancer (GDSC) study (Garnett et al., 2012) 715 cell lines were tested with 138 drugs (note: some drugs were not tested in all cell lines). Because our signature is associated with LKB1 loss in cell lines of multiple histological types, we used the signature as a tool to look for statistical associations with drug sensitivity in these studies to identify candidate drugs

that may have more efficacy in LKB1-deficient tumors. It is also possible that the signature will serve as a measure of activity for specific transcription factors, such as FOXO3, that could be activated in other contexts, independent of LKB1 activity. In either case, we these associations could lead to the discovery of novel treatment options for patients with tumors that have lost LKB1.

Methods and Materials

Statistical Analysis of Drug Sensitivity Associations

Data for drug sensitivity across two large multi-histology collections of cell lines were obtained from the GDSC (Garnett et al., 2012) and CCLE (Barretina et al., 2012) studies. The same studies provided microarray analysis of gene expression for these cell lines. From these data LKB1 classifier scores were derived, as detailed in Chapter II. The gene expression scores for these collections were merged to give a single set of 1244 independent cell lines. In cases where cell lines were included in both studies, the average LKB1 classifier score was used.

To identify inhibitors that may show differential sensitivity in tumors lacking LKB1, we performed univariate linear regression analysis to determine the association between the LKB1 classifier score and the IC50 values for 131 different compounds included in the GDSC study. To allow for training and testing analyses the CCLE study was split into two cell line groups. The set of cell lines from the CCLE that were also included in the GDSC study was used as a training set confirmation, while samples that were not included among GDSC cell lines were used as an independent validation set. Linear regression was used to determine associations between cluster scores and the IC50

values and maximum inhibitory effects seen for each of the 24 inhibitors included in the CCLE study. Distributions were also compared for groups of cell lines given a binary classification as high or low LKB1-loss score and Student's t-tests were used to compare drug sensitivity between the two groups.

To demonstrate that the association between the LKB1-loss signature and MEK sensitivity was a novel observation not accounted for by previous findings, we used a multivariable general linear model relating maximum inhibition by the MEK inhibitor selumetinib to the LKB1-loss score and each of three previously published MEK sensitivity signatures (Dry et al., 2010; Garnett et al., 2012; Loboda et al., 2010), as well as additional variables representing mutations in KRAS, NRAS, HRAS, BRAF, and LKB1. The published gene signatures were used to calculate sensitivity scores for each cell line by averaging standardized expression for each of the published probesets. The correlations between the genes comprising these predictive gene signatures were examined visually in heat maps to ensure they were strongly correlated with one another, such that each signature could be justifiably represented as a single numeric value. Mutations were determined based on data from COSMIC and the CCLE. Linear regression modeling was performed using the R statistical platform with the Limma package.

Proliferation and Drug Sensitivity Assays

In vitro proliferation assays were performed in 96-well plates after seeding 1000 cells in each well. Quantitation of relative cell growth was made using the Alamar Blue (Invitrogen) colorimetric assay. Similarly, for drug sensitivity assays, 1000 cells per well

were added to 96-well plates. Inhibitors were added at the specified concentrations 24 hours after seeding, and relative cell viability was quantified 72 hours after adding inhibitors using Alamar Blue. Trametinib, selumetinib, PD0325901, BEZ235, dasatinib, and paclitaxel were purchased from Chemitek.

Immunoblots

Cell lysates were harvested while cells were in exponential growth phase in RIPA lysis buffer containing phosphatase and protease inhibitors. Lysates were homogenized and electrophoreses were performed on pre-cast SDS-PAGE gels (BioRad). Phospho-ACC (s79), ACC, LKB1, pAKT, pFOXO1/3/4 (T32), pERK1/2 (T202/Y204), pMEK (S217/S221), ERK and AKT antibodies were obtained from Cell Signaling Technology (Danvers, MA). Resulting western blots were quantitated using ImageJ software to measure pixel density and area, with results compared to those of Akt as a loading control.

Results

LKB1-deficient Cell Lines Show Increased Susceptibility to MEK Inhibition

To identify potential candidates for targeted therapy among patients with LKB1-deficient tumors, we investigated drug sensitivity associations using data from the Genomics of Drug Sensitivity in Cancer (GDSC) study, which measured in vitro susceptibility of 715 cell lines to 138 diverse pharmacologic inhibitors (Garnett et al., 2012). Cells with high expression of the LKB1-loss signature were significantly more sensitive to each of the four different MEK inhibitors included in the study – PD-

0325901 (P-value=3.2e-6), selumetinib (AZD6244; P-value=0.0056), Cl-1040 (P-value=0.0073), and RDEA119 (P-value=1.5e-5) (Fig. 4.1a). This novel association with MEK inhibition was confirmed using an independent testing set of cell lines from the Cancer Cell Line Encyclopedia (CCLE), a second large-scale analysis of in vitro drug susceptibility that included data on both selumetinib and PD-0325901 (Barretina et al., 2012) (Fig. 4.1b,c).

Mutations in the RAS/RAF pathway lead to activation of MEK and ERK, and thus, tumors and cell lines with such mutations are more sensitive to inhibitors of these pathways. We observed that the magnitude of the association of the LKB1-loss signature with selumetinib sensitivity was similar to the effect of mutational RAS/RAF activation (Fig. 4.1d). Because LKB1 loss has significant overlap with these mutations, we next wanted to determine whether LKB1 loss was an independent determinant of MEK sensitivity after controlling for these mutations. In addition, we used gene expression signatures of MEK sensitivity from three previously published studies to determine whether our signature is associated with a novel phenotype or is accounted for by these previous analyses. We used a multivariable general linear model to account for the following variables: mutations in KRAS, NRAS, HRAS, and BRAF, as well as previously reported gene signatures from Loboda et al, Garnett et al, and Dry et al (Barretina et al., 2012; Dry et al., 2010; Garnett et al., 2012; Loboda et al., 2010). Multivariate analysis demonstrated that our signature of LKB1 loss was independently associated with sensitivity in a model that included all of these variables (P-value 2.8e-8). To show that these effects are independent of mutational status, we present the distributions of drug sensitivity data for signature-positive and signature-negative cell

lines separated according to mutation status (Fig. 4.1e-g). Significant associations are observed for each mutational class; the statistical significance is higher in the multivariate analysis because the sample size is much larger, and potentially also because LKB1-loss score was used as a continuous variable rather than a binary classification.

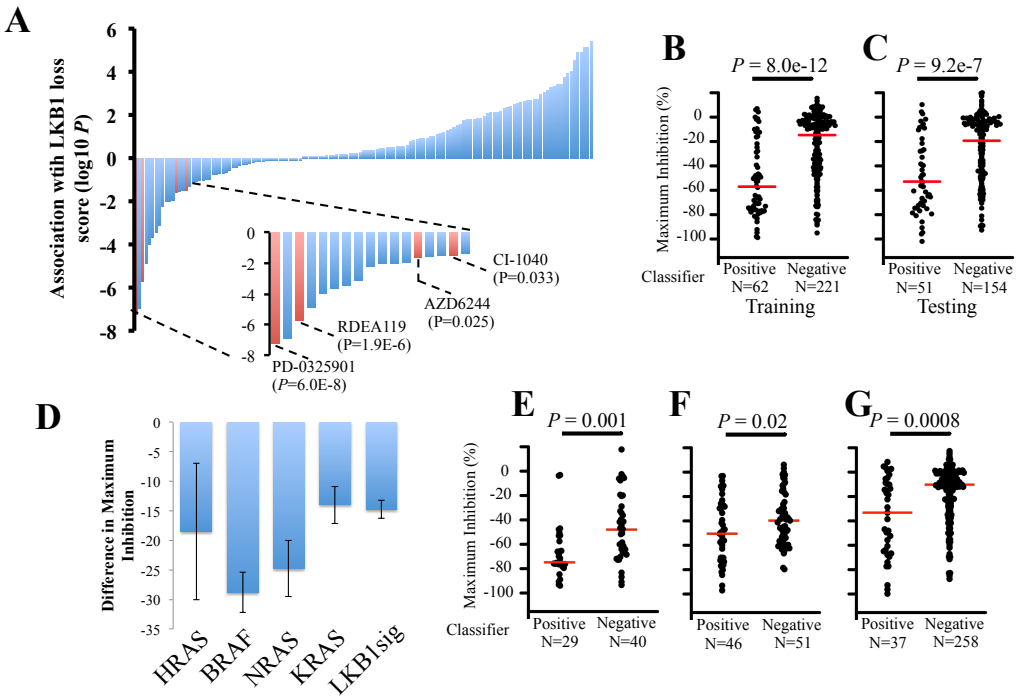


Figure 4.1. LKB1 loss score is associated with sensitivity to MEK inhibitors.
A, Inhibitors tested in the GDSC study (n=138) are ranked according to their statistical association with the LKB1-loss signature. Seventeen inhibitors that have p-values less than 0.05 are shown in the expanded inset, which includes the four MEK inhibitors included in this study. P-values represent the results of linear regression analysis between the LKB1 loss signature expression and the IC50 values for a given drug. Log10 transformed p-values are depicted on the y-axis, with positive values indicating resistance and negative values indicating sensitive phenotypes. **B-C,** Maximum inhibitory effect of selumetinib is shown for cell lines with high expression of the LKB1 classifier compared to those with low expression in both (**B**) training and (**C**) testing cohorts from the CCLE dataset. **D,** Univariate linear regression coefficients associated with mutations in RAS/RAF family members or the LKB1 signature are shown corresponding to differences in maximum selumetinib inhibition in the CCLE study. Bars indicate standard error of these regression estimates. **E-G** Association between selumetinib and LKB1-loss signature is shown for cell lines with mutations in (**E**) BRAF, (**F**) KRAS and (**G**) wild-type for BRAF, KRAS, NRAS, and HRAS. Cell lines classified as LKB1-loss are marked as 'classifier positive' while those with a wild-type signature are marked as 'classifier negative'. Distributions and medians are plotted; P-values represent the significance of the Student's t-test.

Restoring LKB1 Expression in LKB1-mutant Cell Lines Induces Resistance to MEK Inhibitors

Our statistical comparison used a gene expression signature that is strongly associated with LKB1 loss in lung cancer, and significantly associated with LKB1 loss in other types of tumors and cell lines, albeit to a lesser degree. LKB1 loss has not been thoroughly studied in non-lung cancer cell lines, so some of the signature-positive cell lines may have LKB1 loss that has not been described; however, it is likely that the signature is not as specifically associated with LKB1 loss in these non-lung cell lines, especially those that express these genes at a lower level close to the score cut-off for the binary classification. Thus, there are several ways that a falsely or misleadingly positive statistically significant association could be observed with the signature without being the phenotype being truly dependent on LKB1 loss. Furthermore, many factors influence sensitivity to any targeted agent, including activation of the pathway, ability to shift oncogenic signaling to a redundant pathway, or activation of other noninhibited pathways that could induce growth or activate resistance mechanisms to cell death. There could be a significant association between LKB1 loss and alteration of one or more such pathways that was not directly dependent on LKB1 signaling.

To determine whether LKB1 loss directly affects in vitro sensitivity to MEK inhibition, we turned to our isogenic LKB1 add-back model previously described. We derived LKB1 add-back lines, along with vector only and kinase dead control lines for A549, H2122, and H460 and tested their sensitivity to several inhibitors. We found significant induction of resistance to the MEK inhibitors selumetinib, trametinib, and PD0329501, by wild-type LKB1 in A549 and H2122, but not H460 (Fig. 4.2a). On the

other hand, induction of resistance with LKB1 re-expression was not observed for dasatinib (multi-targeted kinase inhibitor, including c-Src), BEZ235 (dual PI3K and mTOR inhibitor), or paclitaxel (microtubule inhibitor), demonstrating that this is specific for inhibition of the MEK pathway, not a general drug resistance effect (Fig. 4.2b). These findings are consistent with the MEK-specific effect observed in our in silico analysis of drug sensitivity, and support a direct role of LKB1 in determining MEKi sensitivity.

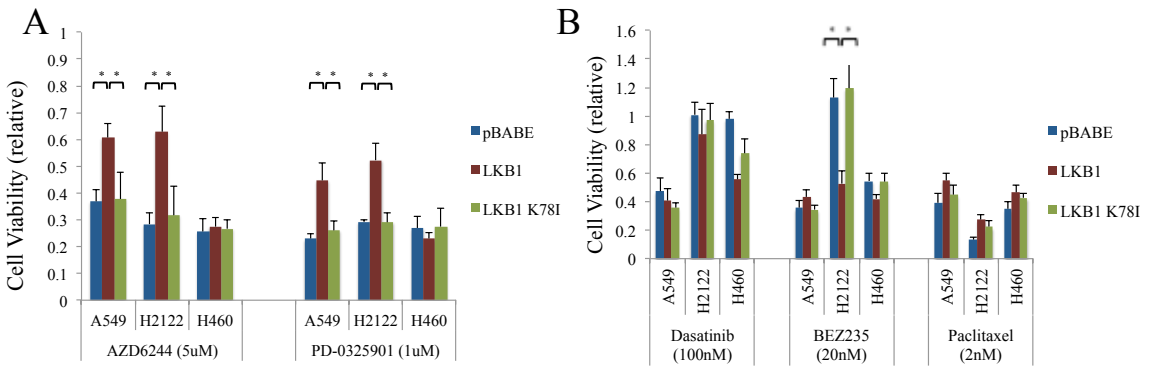


Figure 4.2. Restoring LKB1 confers resistance to MEK inhibition .
A-B, Cell viability is shown for A549, H2122, and H460 cell lines stably transduced with pBABE, LKB1, or LKB1 K78I and treated for 72 hours with the indicated inhibitors. Mean viability, determined by colorimetric Alamar Blue assay, is shown relative to DMSO treated controls with error bars representing standard deviation. P-values represent the significance of unpaired student's t-tests. Full concentration curves were not performed for these inhibitors and IC50 values were not calculated.

The Expression of the 16-gene Signature is Independent of MEK/ERK Signaling.

To give insight into mechanisms that might explain the association of LKB1 loss with MEK sensitivity we examined the effects of MEK inhibition on the expression of genes comprising the LKB1-loss signature. If the expression of genes in our signature is downregulated by MEK inhibition, then the signature might be indicative of MEK pathway activation rather than an independent phenotype. To answer this question we identified two publicly available datasets in which changes in gene expression of a total

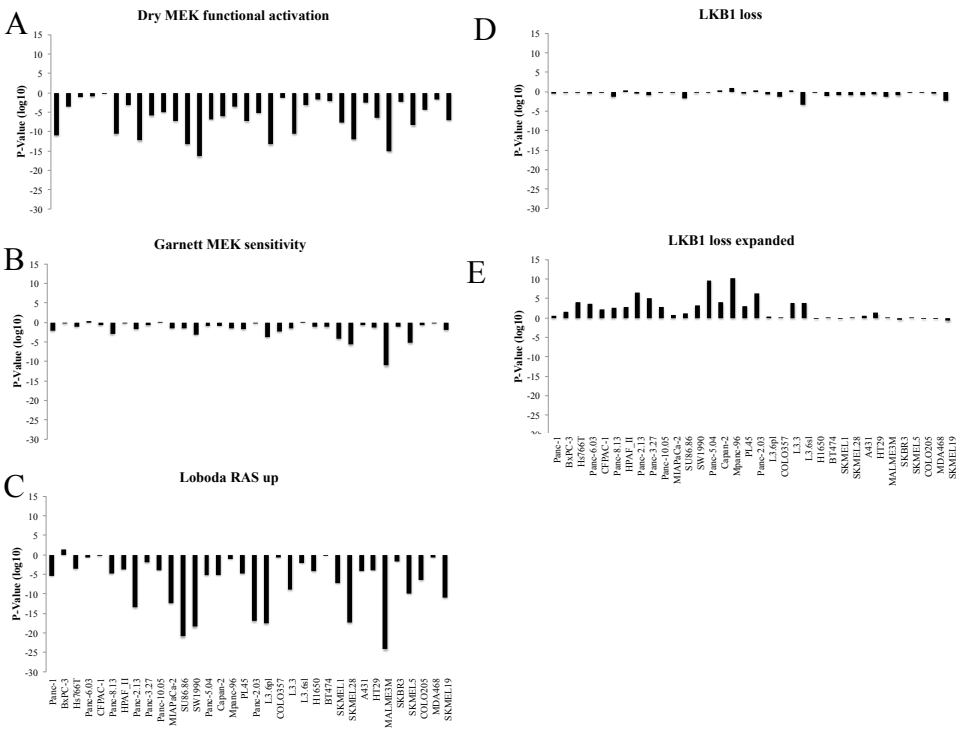


Figure 4.3. Influence of MEK inhibition on gene expression for three previously published MEK signatures and the LKB1 loss signature.
Changes in gene expression corresponding to different signatures of MEK sensitivity are shown after treatment with MEK inhibitor. **a,** Analysis of 15 gene 'MEK functional activation' signature reported in Dry, et al, 2010. **b,** Analysis of 58 genes reported to be correlated with MEK sensitivity in Garnett, et al, 2012. **c,** Analysis of 90 gene signature reported in Loboda, et al, 2010. **d,** Analysis of 16 gene signature of LKB1 loss reported here. **e,** Analysis of top 200 genes associated with the 16 gene LKB1 loss signature. The y-axis represents the log- transformed p-values for hypergeometric test of overlap significance between each signature and the genes perturbed by MEK inhibition in 34 cell lines (22 pancreatic, six skin, three breast, two colon, one lung) from two studies. Pancreatic cell lines were treated with 2mM CI-1040 for 24 hours, while other cell lines were treated with 50nM PD0325901 for eight hours.

of 34 cell lines were analyzed after treatment with MEK inhibitors in vitro or in vivo (Gysin et al., 2012; Pratilas et al., 2008). The overall expression of the 16-gene signature was not affected by MEK inhibition; the only significant interaction observed was the downregulation of DUSP4, a known component of the MAPK pathway. When the 200 genes of the FOX/CREB cluster were studied, gene expression was induced for a subset of the cell lines, but in no cases was signature expression significantly attenuated.

Conversely, each of the three previously published signatures of MEK sensitivity (Dry et al., 2010; Garnett et al., 2012; Loboda et al., 2010) that we analyzed showed significant downregulation after MEKi treatment, indicating that these signatures reflect the level of RAS/RAF/MAPK activation within a cell, whereas, the LKB1-loss signature is determined by an independent phenotype (Fig. 4.3).

Restoring LKB1 to Cell Lines in Vitro Induces Phosphorylation of AKT and FOXO3.

Because FOXO transcription factors were implicated by our analysis of the LKB1-associated gene signature and AKT/FOXO3 has been shown to directly affect MEK sensitivity (Catalanotti et al., 2013; Gopal et al., 2010; Meng et al., 2010), we examined the effects of LKB1 on activation of this pathway. In A549 and H2122, two cell lines in which LKB1 restoration induced resistance to selumetinib, we observed increased phosphorylation of AKT at serine 473 and increased phosphorylation of its downstream target FOXO3 at threonine 32 (Fig. 4.4), which causes FOXO3 nuclear export and downregulation of target genes (Calnan and Brunet, 2008). Activation of AKT and downregulation of FOXO3 are both consistent with our interpretation of the LKB1-loss gene expression signature and with the finding of increased MEK sensitivity among these cell lines, and represents a mechanistic link between these phenotypes.

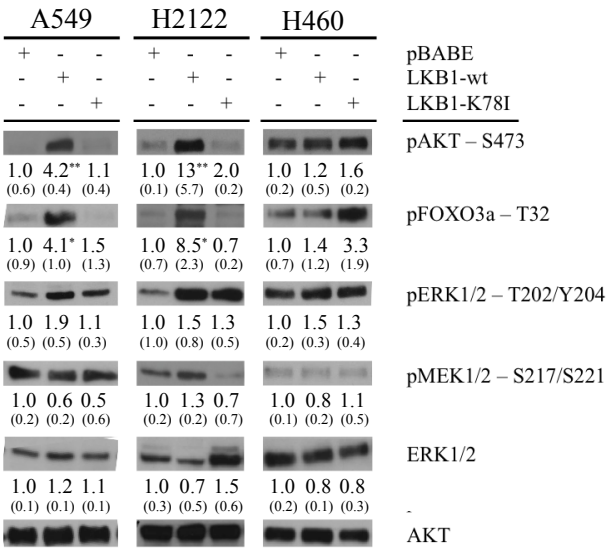


Figure 4.4. Restoring LKB1 alters AKT and FOXO3 phosphorylation.
Immunoblots of whole-cell lysates from A549, H2122, and H460 stably expressing pBABE vector, LKB1 or LKB1 K78I after overnight incubation in serum free medium. Blot density was quantitated using ImageJ software in reference to total AKT. Mean density is indicated below each band, relative to pBABE control. Standard deviations across replicates are given in parentheses. * P<0.05; **P<0.01 for Student's T-test.

PI3K/AKT Signaling is Attenuated in Resected Human Tumors with LKB1 Loss

To determine whether PI3K and AKT signaling was also affected by LKB1 loss in human tumors we analyzed proteomic RPPA data from TCGA-characterized lung adenocarcinomas. The LKB1-loss signature was used to distinguish tumors with LKB1 loss from LKB1 wild-type tumors and expression levels for 174 proteins and phosphorylation sites were compared between these groups using the Student's t-test. In tumors with LKB1 loss, both the p85 and p110 subunits of PI3K showed significant decrease in expression (P = 0.00061 and 0.00047 respectively), as well as decreased phosphorylation of PDK1 at serine 241 (P = 0.00012), and decreased total AKT (P = 0.02) and phospho-S473 AKT (P = 0.018). Surprisingly, proteomic evidence did not

suggest significant mTOR activation, showing only modest increase in eIF4E expression and decrease in 4E-BP1 (P = 0.037, 0.032 respectively), with no significant differences in other components of mTOR signaling. We also did not see significant differences in most proteins associated with the MEK/ERK pathway, with a modest increase in phosphorylated MEK (P = 0.036) and a decrease in total ERK (P = 2.1e-04). We present a simplified model based on established pathway interactions, in which both PI3K/AKT and MEK/ERK pathways play important roles in regulating mTOR activity (Fig. 4.5). The downregulation of PI3K/AKT and LKB1/AMPK seen in LKB1-deficient tumors is shown on this schema using a color code to represent the statistical significance seen in our analysis of RPPA data.

Attenuation of PI3K/AKT signaling would also be expected to have significant effects on apoptotic regulation (Fig. 4.6). Loss of AKT-mediated repression of FOXO3 could induce expression of pro-apoptotic factors such as BIM, which shows elevated protein levels in LKB1 deficient tumors (*P-value* = 0.0002, Fig. 4.6). We further stratified LKB1-deficient tumors by the level of phosphorylated AMPK that was detected, to determine if the level of activation of this pathway affected expression of BIM. We found that BIM protein expression was significantly higher in the tumors with the most complete attenuation of pAMPK (P-value = 0.01), whereas, both groups of LKB1-deficient tumors still expressed significantly higher levels than LKB1 wild-type tumors (P-value = 0.008 for high AMPK group and 7.7e-5 for low AMPK groups vs. LKB1-wild type). While downregulation of AKT signaling can induce pro-apoptotic signaling, these effects can be counteracted on several levels by the MEK/ERK pathway, which phosphorylates FOXO3, BIM, and BAD to inhibit apoptosis (Fig. 4.6). The

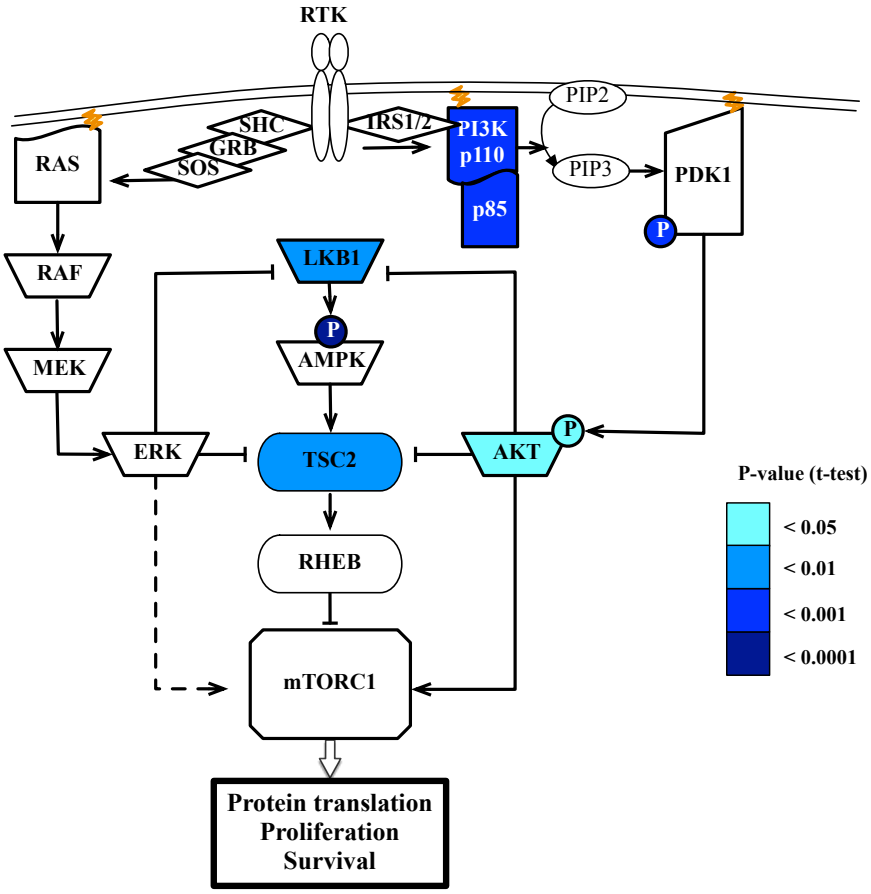


Figure 4.5. Signaling through MEK/ERK and PI3K/AKT pathways regulates mTOR activation.
A, Schema representing interactions of MEK/ERK, PI3K/AKT, and LKB1/AMPK signaling in the regulation of mTOR activity. For altered proteins or phosphorylated proteins that were mentioned in the text, we depict statistical differences in expression between LKB1 wild-type and LKB1-deficient lung adenocarcinomas using RPPA data characterized by the TCGA. The statistical significance of a Student's t-test is represented for various proteins using colors that correspond to p-values indicated in the legend. Blue indicates decreased expression among LKB1-deficient lung adenocarcinomas.

implications of these models are that in the absence of LKB1, downregulation of PI3K/AKT signaling renders tumors more reliant on MEK/ERK signaling for proliferation and inhibition of apoptosis, and thus more susceptible to targeted inhibition of this pathway.

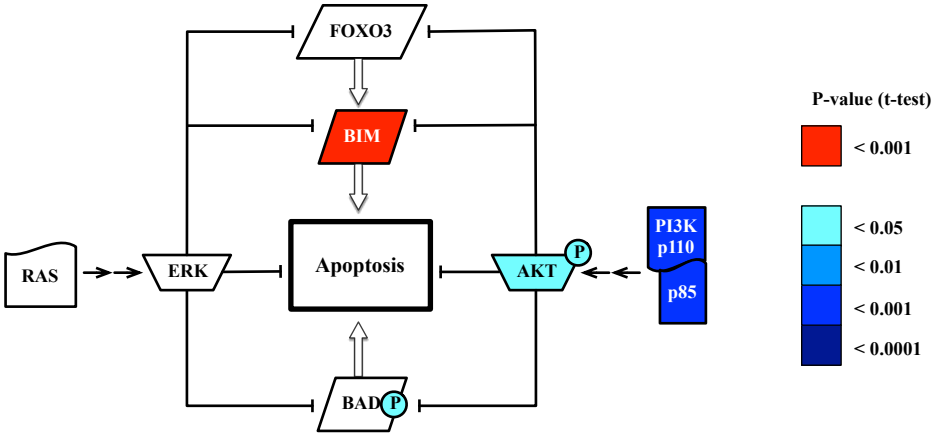


Figure 4.6. Signaling through MEK/ERK and PI3K/AKT pathways regulates apoptosis.
A, Schema representing interactions of MEK/ERK and PI3K/AKT signaling in the regulation of apoptosis. For altered proteins or phosphorylated proteins that were mentioned in the text, we depict statistical differences in expression between LKB1 wild-type and LKB1-deficient lung adenocarcinomas using RPPA data characterized by the TCGA. The statistical significance of a Student's t-test is represented for various proteins using colors that correspond to p-values indicated in the legend. Red indicates increased expression among LKB1-deficient lung adenocarcinomas, while blue indicates decreased expression.

Discussion

In this chapter we have used our 16-gene LKB1 signature as a tool to investigate empiric associations with drug sensitivity. The top candidate identified in this analysis was inhibitors of the MEK pathway, which has recently shown efficacy in the treatment of BRAF mutant melanoma (Catalanotti et al., 2013; Flaherty et al., 2012) and promising results in a phase II clinical trial in KRAS mutant advanced stage NSCLC (Jänne et al., 2013). Four MEK inhibitors were represented in the training set, and sensitivity to each was significantly associated with the LKB1-loss signature. Two of these compounds were also represented in a second set of cell lines, which allowed independent confirmation of our findings with strong statistical significance (P-value = 4.1e-7 in training; 4.8e-8 in

testing set). An association between MEK inhibition and LKB1 mutations has been previously posited by a small study of ten cell lines (Mahoney et al., 2009). However, this study is limited in scope, and an apparent post-hoc grouping of cell lines may invalidate the reported statistical association. Thus, our study supports the previous observation, but greatly expands its scope and significance.

Although this is an empiric observation, we have endeavored in the previous chapters to ascribe biological significance to the genes comprising each of the clusters we identified. This information has now proved useful in generating testable hypotheses to explain the association with MEK sensitivity. In particular, we know that our signature is predictive of LKB1 loss, and the expression of its constituent genes are directly affected by restoring LKB1 in cell lines; therefore, the pathways driving the expression of these genes must have been altered as well. Our analysis of these genes led us to implicate three transcription factors that are likely active in LKB1-deficient tumors. The FOXO3 transcription factor was particularly interesting in the context of MEK sensitivity, since FOXO3 and AKT have been shown to be key determinants of response to MEK inhibition, with significant attenuation of apoptosis seen after expression of constitutively active AKT or after siRNA knockdown of FOXO3 or BIM (Meng et al., 2010). LKB1 has previously been linked to AKT and FOXO3 activation. Zhong et al showed that depletion of LKB1 in an EGFR mutant cell line led to decreased AKT activation and increased apoptosis due to loss of AKT phosphorylation of several anti-apoptotic proteins, while expressing LKB1 in mutant cell lines induced AKT and FOXO3 phosphorylation (Zhong et al., 2008).

Our work replicates this effect showing increased AKT and FOXO3 phosphorylation after restoring LKB1 in cell lines. This, taken together with our gene expression and proteomic analysis, suggests that FOXO3 is generally upregulated among lung tumors with LKB1 loss. Furthermore, we show that restoring LKB1 induces resistance to MEK inhibition in a subset of cell lines, demonstrating that the association between MEK sensitivity and LKB1 suggested in our statistical approach can be conclusively confirmed experimentally.

Our analysis of proteomic data from resected human tumors further supports the hypothesis that AKT and BIM are dysregulated in LKB1-deficient tumors. Furthermore, the literature-based model that we present linking LKB1, mTOR, and AKT activity may provide mechanistic insight into the biology of these tumors. These pathways are intimately connected through a network of protein interactions coordinating the balance between mTOR and PI3K/AKT activation (Shaw and Cantley, 2006b; Vivanco and Sawyers, 2002). Activation of mTOR is induced by AKT phosphorylation of the mTOR complex itself and by inhibitory phosphorylation of the upstream TSC2 inhibitor (Huang and Manning, 2009). To limit this signal mTOR phosphorylates IRS1/2 substrates, resulting in their degradation and blocking PI3K activation by receptor tyrosine kinase signaling. A separate feedback component involves TSC2 induction of mTORC2 activation in concert with its downregulation of mTORC1, resulting in AKT activation and restoration of pathway balance. Perturbations of this network result in shifts in the balance of pathway activation. For instance, pharmacologic inhibition of mTOR relieves feedback inhibition, resulting in activation of PI3K and AKT (Carracedo et al., 2008; Rodrik-Outmezguine et al., 2011). Conversely, loss of TSC2 causes strong mTOR

activation and results in abrogation of PI3K and AKT signaling (Manning, 2005; Shaw et al., 2004a), with resultant activation of FOXO3, similar to the effects seen here due to LKB1 loss. Thus, although LKB1 acts as a tumor suppressor, its loss leads to the downregulation of important oncogenic pathways; this feedback inhibition may explain why LKB1 and EGFR mutations rarely occur in the same tumors, since EGFR signals prominently through PI3K, mTORC2, and NF-kB (Tanaka et al., 2011).

Interestingly, MEK sensitivity has been tested in the murine model of LKB1/KRAS mutant lung cancer, but showed the opposite association (Chen et al., 2012). Thus, the differences in LKB1-associated gene expression between human and mouse tumors may also be reflected in clinically relevant phenotypes produced by these models. Whereas, our work implicates activation of the FOXO3 transcription factor as a potential mediator of this effect, the murine model showed activation of TGF-beta and SRC pathways in LKB1-deficient tumors, which are known to induce resistance to MEK inhibitors (Dry et al., 2010; Ferguson et al., 2012; Girotti et al., 2013; Huang et al., 2012). Indeed, the addition of dasatinib to a combination of BEZ235 and selumetinib induced tumor shrinkage in the LKB1-deficient murine tumors, suggesting that this phenotypic difference could explain the discrepancy in MEK sensitivity between human tumor cell lines and murine models (Carretero et al., 2010). Cell lines and genetically modified mouse models may be useful pre-clinical models for studying LKB1 loss and other scientific questions. In this case, however, opposite conclusions are drawn from the two different models. Testing the hypothesis in human clinical trial samples will be required to conclusively answer this question.

Our work shows that LKB1 directly affects both FOXO3 and MEK sensitivity. In lung cancer our signature is strongly associated with LKB1 loss; however in other tumor types, this association may be less specific. We have linked this gene expression signature to FOXO3 activation, which likely represents the mechanism underlying MEK sensitivity. However, in other tumor types where LKB1 loss is less prevalent, the gene signature could be induced by activation of FOXO3 for other reasons. If this is the case then the signature might be effective at predicting response to MEK inhibition even in tumor types where LKB1 loss is rarely seen. Additional study would be required to determine if this is the case, and ultimately clinical trials would be needed to see whether expression of our signature is associated with improved response to MEK inhibition. Understanding the mechanisms of resistance associated with this class of inhibitors will be crucial for the rational design of targeted combination therapies, as well as selection of patients most likely to receive clinical benefit. Thus, analysis of LKB1 loss in clinical trials of MEK inhibition will be informative as to whether this phenotype is predictive of patient outcome and whether it could be used prospectively to guide treatment decisions.

CHAPTER V

CONCLUSION

Our work yields several novel insights into the biology and potential treatment of tumors with loss of LKB1. We present four major findings: First, LKB1 loss is associated with a characteristic pattern of gene expression changes that is observed in multiple types of human cancers and in both lung and non-lung carcinoma cell lines. Second, the differentially expressed genes associated with LKB1-deficient human tumors are substantially different from the pattern of genes observed in the mouse model of LKB1/KRAS mutant lung cancer. Third, we demonstrate a link between LKB1 loss and MEK sensitivity that may have important clinical implications. Fourth, we identify dysregulation of the PI3K/AKT/FOXO3 pathway as an important characteristic of LKB1-deficient tumors, which provides a potential mechanistic link between our genomic and proteomic analyses and the observed MEK sensitivity phenotype.

All of the observations made to reach these conclusions stem from our first finding, that LKB1 loss in lung cancer has a very strong association with a particular gene expression signature. This is clearly evident when we compare the similarity in gene associations among eleven studies, which we presented visually in (Fig. 2.1a). The median P-value for these comparisons was 10e-22, a highly significant similarity. We took advantage of these consistent gene expression changes to identify a 16-gene signature predictive of LKB1 loss, which we then tested in multiple validation cohorts. Gene expression data from a total of 1297 lung adenocarcinomas were analyzed in this

work. LKB1 sequencing results were available for approximately half, with mutations reported in 93. Of these, 87 were correctly predicted by our LKB1-loss signature (sensitivity 94.5%, P-value = 2.9e-40 by Fisher's exact test). The extensive testing of our gene expression signature places this work among the best-validated genomics studies. The excellent performance of our signature suggests that it could be a useful tool for classifying functional LKB1 status in patients, especially since we show that tumors with identified sequence mutations in LKB1 comprise only about half the total of LKB1 loss. We have included our preliminary results to show that a targeted gene expression detection assay using nanoString nCounter technology can adequately determine this signature in clinical specimens.

Our primary interest, however, has been to discover new insights into the biology of LKB1 deficient tumors and ultimately to improve therapy of these human cancers. LKB1 is known to regulate cellular metabolism and the mTOR pathway by its effects on AMPK, but many other interesting phenotypes are also affected by LKB1, some through AMPK, and others by AMPK-independent effects carried out by other kinases regulated by LKB1. Predicting how these complex pathways interact to influence clinically relevant phenotypes is difficult and appropriate model systems must be studied to make these connections. This was, of course, the main impetus for the development and characterization of the genetically engineered murine model of LKB1/KRAS mutant lung cancer, which has been used to directly test hypotheses related to LKB1 biology in an in vivo setting. Our studies, on the other hand, focus entirely on the effect of LKB1 loss in human tumors. The characterization of hundreds of lung tumors by the TCGA has provided a valuable resource for identifying such associations. Of the 403 tumors

currently with mutation data, 67 have somatic mutations in LKB1, which provides a large sample size that can provide great statistical power for making comparisons. Furthermore, these tumors are thoroughly characterized molecularly, including data on gene mutations, copy number changes, proteomic analysis, microRNA expression, mRNA expression, and methylation allowing an integrated approach to answering many questions regarding tumor biology.

Our work was begun several years prior to the availability of these data, however, so we initially employed a meta-analytic approach to combine several smaller studies. We saw that our LKB1 signature was predictive of both LKB1 mutations and low LKB1 mRNA expression in these cell line and patient datasets, and consistently identified between 30 and 35% of resected lung adenocarcinomas. Thus, we could use these classifications to identify LKB1 wild-type and LKB1-deficient tumors in smaller datasets and combine them into a large pooled sample set that could be used to determine associations between LKB1 status and clinical variables or other mutations characterized by these studies. This allowed us to show significant associations with smoking status, EGFR, and KRAS mutations prior to the availability of the TCGA cohort. These associations fit with previously reported results in the literature and gave further assurance that our model was effective in classifying the LKB1 status of tumors.

Based on our initial analyses we could identify ranked lists of genes that were statistically associated with LKB1 loss in lung cancer. However, transforming these lists into useful hypotheses regarding the biology of these tumors was challenging. We first examined the genes individually, and searched the literature to determine whether any had been previously associated with LKB1 loss or other phenotypes that might be worth

pursuing. We also employed commercially and publicly available tools such as the Molecular Signatures Database and Ingenuity Pathway Analysis to look for significant features of the entire gene set that might give clues regarding the phenotypes driving the expression of the genes in our signature. We ultimately found it more informative to study their patterns of activation and expression in more detail. Here we saw that the expression patterns of these genes clustered consistently across multiple studies. Thus, we came to view these genes as being associated with several discrete phenotypes that could be investigated independently. This led us to employ a novel statistical approach in which general linear models were used to determine genes with significant associations with each observed phenotype, bringing us from one gene list to four. This approach was effective in allowing us to identify specific phenotypes underlying these transcriptional clusters. Analyzing the new gene lists with statistical comparison tools resulted in high confidence hits to transcription factors, drug perturbations, etc., which proved informative in understanding the biology of these tumors.

The effectiveness of this approach is highlighted by our identification of the NRF2 transcription factor as a driver of one cluster. The gene expression signature showed a strong statistical association with NRF2 and with and known perturbations of NRF2. This statistical association was subsequently confirmed in the TCGA dataset, in that the NRF2 activation signature was strongly associated with somatic mutations in KEAP1, low KEAP1 protein expression, and high NRF2 protein expression.

Thus, the NRF2 cluster is independently regulated phenotype that does not appear to be under direct control by LKB1; indeed, re-expression of LKB1 caused no change in the level of expression of the genes in the NRF2 cluster. On the other hand, we had

expected that restoration of LKB1 activity would be associated with downregulation of the mTOR pathway, but were surprised to find no change in expression of the genes we had putatively linked to mTOR activation. However, we did observe significant downregulation of LKB1-associated genes, primarily affecting the genes in the FOX/CREB cluster that was most strongly associated with LKB1. Thus, this cluster is directly regulated downstream of LKB1. The associations we made with this cluster were particularly informative in our understanding of LKB1-deficient cancers. We identified CREB and FOXA2 as transcription factors that are activated in these LKB1 deficient tumors. Although FOXA2 has been reported to have tumor suppressive properties (Liu et al., 2012a; Tang et al., 2010), the significance of its upregulation in LKB1-deficient tumors is unknown. CREB, and its CRTC activators, on the other hand, have been shown to have oncogenic effects in lung cancer generally (Aggarwal et al., 2008), as well as specifically in LKB1-deficient lung tumors (Feng et al., 2012). We also saw activation of the FOXO3 transcription factor, which has important effects in regulating apoptosis, and this proved to be of significant interest later in our work when we identified an association between LKB1 loss and sensitivity to MEK inhibitors.

In addition to our gene expression associations, we employed our LKB1-loss signature to perform a comprehensive analysis of mutations, microRNA, copy number alterations, and proteomic changes associated with LKB1 loss in the TCGA cohort. We have incorporated the proteomic data into a proposed model of an LKB1-regulated pathway network. However, the significance of many of the other associations remains unclear. These will be fruitful hypotheses-generating bases for further research.

The mechanistic model we propose ties together several aspects of our work. Dysregulation of PI3K/AKT/FOXO3 pathway is shown in our in vitro work and is understood to directly affect MEK sensitivity in cancer. In fact, at least nine clinical trials are ongoing to determine the benefit of combining PI3K or AKT inhibitors with MEK inhibition (Britten, 2013). Dysregulation of these pathways is evident in the proteomic analysis of human tumors, and FOXO3 drives the expression of many upregulated genes in our LKB1-loss signature. AKT is a potent activator of the mTOR pathway, which then induces feedback inhibition of AKT by attenuating signal transduction through PI3K. Thus, this represents a likely mechanistic link between our findings and the loss of mTOR inhibition that occurs in the absence of LKB1. This is analogous to effects seen in cancers induced by TSC1 or TSC2 loss, in which mTOR activation leads to feedback inhibition of AKT signaling and activation of the FOXO3 pathway, which constrains tumor growth (Gan et al., 2010b; Manning, 2005). Although these effects have been well established in other tumor types, and have been shown previously to be associated with LKB1 loss in a small in vitro study (Zhong et al., 2008), our work draws on proteomic and genomic data to show that these effects are prominent and seen across human lung tumors with LKB1 loss. Furthermore, our in vitro work and analysis of drug sensitivity data show that this suppression of PI3K/AKT signaling is directly affected by LKB1 and can be exploited pharmacologically by targeting the MEK pathway.

The next important conclusion drawn from our work is that our model differs substantially from the model derived from the study of LKB1/KRAS mutant tumors in mice. Observations of proteomic and genomic changes in these mice led the authors to implicate increased c-Src and TGF-beta signaling as being particularly important in the

biology of these tumors, features that were not associated with LKB1 loss in human tumors. On the other hand, none of the specific pathways we identified in our work showed differential expression in the mouse data. We did not perform any experimentation with the mouse model ourselves, so we can only theorize as to why these tumors show such different phenotypes. It is possible that the activation of TGF-beta signaling may underpin some of the differences. Genes implicated with TGF-beta signaling were increased in LKB1-deficient mouse tumors, and increased secretion of TGF-beta was measured (Carretero et al., 2010; Ji et al., 2007). TGF-beta signaling has been shown to be downregulated in the context of LKB1 loss in other settings (Londesborough et al., 2008; Vaahtomeri et al., 2008), though not in cancer, and we show that TGF-beta is decreased in human tumors with LKB1 loss and increased after re-expressing LKB1 (Fig. 3.2). Mesenchymal cells lacking LKB1 have been shown to secrete significantly less TGF-beta; this effect may contribute to the formation of polyps in Peutz-Jeghers syndrome (Katajisto et al., 2008). Furthermore, A549 cells treated with TGF-beta showed rapid attenuation of the NRF2 and FOX/CREB gene expression clusters, showing that these pathways can be antagonized by TGF-beta. Thus, while it is unclear why this pathway becomes activated in murine tumors, it is possible that blocking TGF-beta activation with further genetic modifications could allow formation of a mouse tumor model that more closely resembles human LKB1 loss.

Our results may have important ramifications on the use of mouse model systems in general. Preclinical models are invaluable to test hypotheses and generate new understanding of tumor biology that can ultimately lead to testing in clinical trials and improvement in patient care. Cell lines and mouse models are complementary and each

has certain experimental applications for which it is better suited. Both are indispensable for the study of cancer. Both have substantial advantages and drawbacks. Cell lines are relatively cheap and easy to manipulate, can be worked with rapidly, and several hundred cell lines are available for study, theoretically encompassing the heterogeneity found in human tumors. However, they have been grown for years in plastic dishes, bathed in nutrient and growth-factor-rich media in the absence of stromal cells such as fibroblasts, blood vessels, and immune cells. In contrast, mouse models take time to develop and characterize, and are expensive and cumbersome to maintain. It is difficult to do large-scale studies with more than, say, several dozen mice. Furthermore, resulting tumors are much more genetically homogeneous than the human counterparts, lacking the many mutations and chromosomal alterations accumulated with time in human lung cancers as they evolve progressively over many years during exposure to the carcinogens in the 350,000 cigarettes accumulated in a patient with 50 pack-years of smoking – the average exposure among tumors with LKB1 loss. (Tyler Jacks, AACR 2013 "Genomic characterization of mouse models of lung cancer"). However, they can provide a depth of knowledge about the cancers that arise, including tumorigenesis, histology, patterns of progression, pathway activation, gene and protein expression, propensity to metastasize, and response to therapeutics in 'mouse clinical trials'. These tumors are also studied in the setting of living animals with fully functioning immune systems, stromal cells, and vasculature, which offers a significant advantage over in vitro studies.

For disease phenotypes such as the association with MEK inhibition identified in this study, the only way to adequately assess the accuracy of preclinical models is through analysis of clinical trials. However, our study shows that for LKB1-associated

gene expression, human-derived cell lines show patterns that resemble those from resected human lung adenocarcinomas, while the mouse model produces tumors that appear substantially different. This may have significant implications for the use of the LKB1/KRAS mouse model to predict human disease, although further correlation is necessary.

Study of these model systems can lead to improved understanding of the biology of LKB1 loss, and may ultimately alter the way these tumors are treated, the primary rationale for our study. We have identified one specific target in LKB1 deficient tumors, namely MEK. (This finding highlights the differences between mouse and cell line preclinical models, as LKB1 loss in the mouse model induced significant resistance to such inhibitors.) MEK can be inhibited by pharmaceutical compounds such as selumetinib and trametinib, which are FDA approved for the treatment of BRAF mutant melanoma and are currently undergoing testing in non-small cell lung cancer. The statistical strength of the association with MEK sensitivity is strong. We see concordant findings for all four MEK inhibitors included in two large characterizations of cell line drug sensitivity published last year in Nature (Barretina et al, 2012; Garnett et al, 2012). The P-value for the association in our training set is 4.8e-07, and for our testing set it is 4.1e-08. Together almost 500 cell lines are included to draw these conclusions. We show that the association with the LKB1 signature is independent of mutations in the RAS/RAF pathway that are known to confer sensitivity to MEKi, and also to previously published gene signatures of MEK response. Loss of LKB1 is certainly not the only determinant of MEK sensitivity, but the magnitude of the effect is similar to what is seen for mutational activation of the pathway. Moreover, we demonstrate that restoration of

LKB1 in vitro can induce resistance to MEK inhibitors, but does not similarly affect sensitivity to paclitaxel or targeted inhibitors of other pathways. This is a significant piece of evidence supporting our claim that LKB1 status directly influences the response to MEK inhibition.

Thus, we have identified a novel association between sensitivity to targeted MEK inhibition and LKB1 loss in cell lines, which correlates well with in vivo drug sensitivity phenotype, and is reversed in vitro with restoration LKB1 activity. These results warrant further testing in the clinical trial setting. Indeed, we are actively pursuing this goal by developing a clinical assay using a custom-based assay for our signature using the commercial nanoString platform that will be capable of determining the LKB1 status of an unknown lung cancer given 100ng of RNA. We have been in contact with pharmaceutical companies and academic medical centers involved in testing MEK inhibitors to find potential sources of clinical trial material, as well as commercial interests that may wish to develop our test further in CLIA certified labs for testing in clinical settings.

Appendix A. Data Sources and Statistical Comparisons

Name	Tissue	Source	Comparison
		Fig. 2.1	
MSKCC	lung adeno	Chitale et al, 2009	t-test: LKB1 mut (16) vs LKB1 WT (75)
UNC	lung adeno	GSE26939	t-test: LKB1 mut (n=6) vs LKB1 WT (n=75)
Wash U	lung adeno	GSE12667	t-test: LKB1 mut (n=7) vs LKB1 WT (n=34)
Michigan	lung adeno	Shedden et al, 2008	Linear regression with LKB1 probeset 41657_at (n=178)
TCGA	lung adeno	https://confluence.broadinstitute.org/display/GDAC/Home (LUAD RNAseqv2 Level3 RSEM downloaded 2013/07/15)	t-test: LKB1 mut (n=67) vs LKB1 WT (n=339)
MSKCC2	lung adeno	Chitale et al, 2009	t-test: LKB1 mut (n=12) vs LKB1 WT (n=90)
USC	lung adeno	GSE32861	t-test: LKB1 mut (n=8) vs LKB1 WT (n=48)
Sanger	NSCLC cell lines	www.broadinstitute.org/cgi-bin/cancer/datasets.cgi (Sanger_Cell_Line_Project_Affymetrix_QCed_Data_n798.gct)	t-test: LKB1 mut (n=25) vs LKB1 WT (n=44)
CCLE	NSCLC cell lines	www.broadinstitute.org/ccle/home (CCLE_Expression_2012-09-29.res)	t-test: LKB1 mut (n=34) vs LKB1 WT (n=46)
A549	NSCLC cell line	GSE51266	avg diff: LKB1 WT (n=3) vs pBABE vector (n=3)
H2122	NSCLC cell line	GSE51266	avg diff: LKB1 WT (n=3) vs pBABE vector (n=2)
Ji (A)	Mouse lung adeno	GSE6135	t-test: LKB1/KRAS primary adeno (n=5) vs KRAS primary adeno (n=5)
Ji (B)	Mouse lung adeno	GSE6135	t-test: LKB1/KRAS primary adeno (n=5) vs KRAS/p53 primary adeno (n=5)
Carretero	Mouse lung adeno	GSE21581	t-test: LKB1/KRAS primary adeno (n=9) vs KRAS primary adeno (n=9)
Carretero Mets	Mouse lung adeno	GSE21581	t-test: LKB1/KRAS metastases (n=17) vs LKB1/KRAS primary (n=9)
		Table 3.3	
CREB		http://www.broadinstitute.org/gsea/msigdb	V$CREB_01
Colforsin	MCF7	www.broadinstitute.org/cmap/#	avg diff: MCF7; 0.5uM (n=1) or 50uM (n=1) vs DMSO
Colforsin	PC3	www.broadinstitute.org/cmap/#	avg diff: PC3; 0.5uM (n=2) vs DMSO
Forskolin	PC12	GSE2071	avg diff: PC12; 10uM forskolin (n=4) vs DMSO
CREB regulated	Human Islet cells	natural.salk.edu/CREB/	Table S5: Islet

Appendix A. (cont.)

Name	Tissue	Source	Comparison
Table 3.3 (cont)			
CREB regulated	MIN6	natural.salk.edu/CREB/	Table S5: MIN6
CREB regulated	HEK293T	natural.salk.edu/CREB/	Table S5: HEK293T
FOXO1/3/4		www.broadinstitute.org/gsea/msigdb	TTGTTT_V$FOXO4_01
induction by CA-FOXO3	DLD1	E-MEXP-3262	With constitutively active, tamoxifen inducible FOXO3: avg diff: 24h tamoxifen induction (n=3) vs control (n=3)
induction by CA-FOXO3	HuVEC	GSE16573	With constitutively active, tamoxifen inducible FOXO3: avg diff: 12h tamoxifen induction (n=3) vs control (n=3)
induction by CA-FOXO3	RCC4	GSE23926	With constitutively active, tamoxifen inducible FOXO3: avg diff: 12h tamoxifen induction (n=1) vs control (n=1)
induction by CA-FOXO3	UMRC2	GSE23926	With constitutively active, tamoxifen inducible FOXO3: avg diff: 12h tamoxifen induction (n=1) vs control (n=1)
Prochlorperazine induction	HL60	www.broadinstitute.org/cmap/#	avg diff: 10uM Prochlorperazine (n=4) vs DMSO
Prochlorperazine induction	MCF7	www.broadinstitute.org/cmap/#	avg diff: 10uM Prochlorperazine (n=9) vs DMSO
Prochlorperazine induction	PC3	www.broadinstitute.org/cmap/#	avg diff: 10uM Prochlorperazine (n=3) vs DMSO
Thioridazine induction	HL60	www.broadinstitute.org/cmap/#	avg diff: 10uM Thioridazine (n=4) vs DMSO
Thioridazine induction	MCF7	www.broadinstitute.org/cmap/#	avg diff: 10uM Thioridazine (n=11) vs DMSO
Thioridazine induction	PC3	www.broadinstitute.org/cmap/#	avg diff: 10uM Thioridazine (n=5) vs DMSO
Trifluoperazine induction	HL60	www.broadinstitute.org/cmap/#	avg diff: 10uM Trifluoperazine (n=4) vs DMSO
Trifluoperazine induction	MCF7	www.broadinstitute.org/cmap/#	avg diff: 10uM Trifluoperazine (n=9) vs DMSO
Trifluoperazine induction	PC3	www.broadinstitute.org/cmap/#	avg diff: 10uM Trifluoperazine (n=3) vs DMSO
FOXA2/HNF3		www.broadinstitute.org/gsea/msigdb	TGTTTGY_V$HNF3_Q6
Promoter occupancy	A549	http://genome.ucsc.edu/ENCODE/	'broadPeak' file used to determine ChIP-seq peaks within 1000 bp of gene start codon
Promoter occupancy	HEPG2	http://genome.ucsc.edu/ENCODE/	'broadPeak' file used to determine ChIP-seq peaks within 1000 bp of gene start codon
Promoter occupancy	Human Liver	GSE25836	'Bed' file used to determine ChIP-seq peaks within 1000 bp of gene start codon

Appendix A. (cont.)

Name	Tissue	Source	Comparison or gene set
		Table 3.1	
AP1		http://www.broadinstitute.org/gsea/msigdb	TGANTCA_V$AP1_C
NRF2		http://www.broadinstitute.org/gsea/msigdb	V$NRF2_Q4
15dPGJ2	MCF7	www.broadinstitute.org/cmap/#	avg diff: 10uM 15-delta prostaglandin J2 (n=5) vs DMSO
15dPGJ2	HL60	www.broadinstitute.org/cmap/#	avg diff: 10uM 15-delta prostaglandin J2 (n=3) vs DMSO
15dPGJ2	PC3	www.broadinstitute.org/cmap/#	avg diff: 10uM 15-delta prostaglandin J2 (n=2) vs DMSO
keap1-/-	mouse liver	GSE11287	avg diff: KEAP1-/- liver (n=3) vs control
KEAP1 mut	LUSQ	https://tcga-data.nci.nih.gov/tcga/	t-test: KEAP1 mut (n=22) vs KEAP1/NRF2 WT (n=171)
NRF2 mut	LUSQ	https://tcga-data.nci.nih.gov/tcga/	t-test: NRF2 mut (n=24) vs KEAP1/NRF2 WT (n=171)
siNRF2	A549	GSE28230	avg diff: A549; siNRF2 (n=3) vs control siRNA
		Table 3.2	
Mitochondrial localization		http://www.broadinstitute.org/pubs/MitoCarta/	MITOCARTA_LIST
ELK1		http://www.broadinstitute.org/gsea/msigdb	SCGGAAGY_V$ELK1_02
SF1		http://www.broadinstitute.org/gsea/msigdb	V$SF1_Q6
NRF1		http://www.broadinstitute.org/gsea/msigdb	RCGCANGCGY_V$NRF1_Q6
MYC		http://www.broadinstitute.org/gsea/msigdb	CACGTG_V$MYC_Q2
LY-294002	HL60	www.broadinstitute.org/cmap/#	avg diff: 10uM LY-294002 (n=9) vs DMSO
LY-294002	MCF7	www.broadinstitute.org/cmap/#	avg diff: 10uM LY-294002 (n=18) vs DMSO
LY-294002	PC3	www.broadinstitute.org/cmap/#	avg diff: 10uM LY-294002 (n=6) vs DMSO
Sirolimus	HL60	www.broadinstitute.org/cmap/#	avg diff: 100nM sirolimus (n=9) vs DMSO
Sirolimus	MCF7	www.broadinstitute.org/cmap/#	avg diff: 100nM sirolimus (n=19) vs DMSO
Sirolimus	PC3	www.broadinstitute.org/cmap/#	avg diff: 100nM sirolimus (n=6) vs DMSO
PD0325901	multiple	GSE10087	paired t-test: 12 cell lines treated 8hr with 50nM PD-0325901 (n=1 rep each) vs DMSO

Appendix A. (cont.)

Name	Tissue	Source	Comparison or gene set
		Fig. 3.2	
TGF-beta	A549	GSE17708	avg diff: A549; 5ng/ml TGF-beta at various times (n=3 reps each) vs control
Dasatinib	A549	E-TAMB-585	avg diff: variable concentrations Dasatinib (n=1 rep each) vs DMSO
		Fig. 3.4, 3.5	
LKB1-add-back	HeLa	PMID: 15731909	avg diff: pLOX-LKB1-YFP vs pLOX-YFP or pLOX-LKB1-YFP vs pLOX-LKB1-SL26-YFP (2 replicates each)
		Fig. 4.3	
MEKi perturbation	multiple	GSE10087	avg diff: PD0325901 50nM for 12 h vs control for 12 cell lines of various histology (1 replicate each)
MEKi perturbation	Pancreatic cell lines	GSE45765	avg diff: CI-1040 2mM for 24h vs control for 22 pancreatic adenocarcinoma cell lines (3 replicates each)

Appendix B. Initial genes in four transcriptional clusters

Gene Symbol	Transcriptional Node	Gene Symbol	Transcriptional Node	Gene Symbol	Transcriptional Node	Gene Symbol	Transcriptional Node
AVPI1	FOX/CREB	ACSL3	mito	AKR1C1	NRF2	ANKRD25	Down
BAG1	FOX/CREB	ATP5B	mito	AKR1C2	NRF2	C1orf139	Down
CPS1	FOX/CREB	C1QBP	mito	CBR1	NRF2	COL8A2	Down
DUSP4	FOX/CREB	C20orf24	mito	G6PD	NRF2	DOC1	Down
FGA	FOX/CREB	COX5A	mito	ME1	NRF2	EDNRA	Down
GLCE	FOX/CREB	CYC1	mito	PGD	NRF2	EVC	Down
HAL	FOX/CREB	DLAT	mito	PIR	NRF2	FBLN1	Down
IRS2	FOX/CREB	DLD	mito	SLC7A11	NRF2	GAS7	Down
MUC5AC	FOX/CREB	ECHS1	mito			GSN	Down
PDE4D	FOX/CREB	FDX1	mito			KIAA1641	Down
PTP4A1	FOX/CREB	FLJ22555	mito			KIRREL	Down
RFK	FOX/CREB	GHITM	mito			MACF1	Down
SIK1	FOX/CREB	LOC92482	mito			MFGE8	Down
TACC2	FOX/CREB	MDH2	mito			NOTCH2	Down
TFF1	FOX/CREB	MRPL46	mito			PCF11	Down
TESC	FOX/CREB	MRPS11	mito			PTK7	Down
		MRPS16	mito			RGL1	Down
		MRPS33	mito			RNF38	Down
		NDUFAB1	mito			SLC34A2	Down
		NDUFS1	mito			SLC9A6	Down
		NDUFV2	mito			TIAM1	Down
		THEM2	mito			TIMP3	Down
		TXNL4A	mito			TNC	Down
		VDAC2	mito			TXNIP	Down
						ZFP36L1	Down
						ZNF161	Down

Appendix C. Top 200 genes associated with each cluster

Rank	FOX/CREB	P-value	Mito	P-value	NRF2	P-value	Down-regulated	P-value
1	PDE4D	8.5E-33	MRPS16	1.8E-34	G6PD	4.2E-49	ZFPM2	3.9E-30
2	IRS2	3.3E-29	TSR1	1.1E-31	AKR1C1	7.6E-47	LTBP2	5.2E-30
3	TESC	1.1E-28	VDAC2	1.7E-31	AKR1C2	1.3E-42	SGCD	8.9E-30
4	SIK1	1.3E-27	TFAM	2.4E-28	PGD	1.2E-39	PRKG1	1.3E-29
5	DUSP4	2.4E-27	SSBP1	2.0E-27	ME1	1.4E-33	TIMP3	1.6E-28
6	BAG1	1.6E-25	PFDN4	4.9E-26	TALDO1	9.3E-31	KANK2	1.9E-28
7	RHOB	2.2E-21	MRTO4	9.7E-26	NQO1	2.1E-30	HSPG2	4.6E-28
8	AVPI1	1.7E-20	SAC3D1	3.1E-25	PIR	1.0E-29	COX7A1	1.1E-27
9	HAL	2.3E-19	CIAO1	4.0E-25	GCLM	2.4E-29	MACF1	1.6E-27
10	CPS1	9.8E-19	NDUFS1	4.8E-25	AKR1C3	7.0E-29	MYLK	1.6E-26
11	NR4A2	5.2E-17	ATP5G3	7.0E-25	SLC7A11	3.1E-28	TMEM204	7.7E-26
12	TACC2	7.1E-17	NDUFB8	1.1E-24	OSGIN1	5.5E-26	ADARB1	1.2E-25
13	CHMP1B	1.7E-16	C2orf47	2.2E-24	AKR1B10	2.2E-25	ECM2	1.3E-25
14	FGA	2.4E-16	TIMM23	4.9E-24	CYP4F11	6.1E-24	AEBP1	2.3E-25
15	C8orf4	1.5E-15	MRPL15	6.2E-24	CBR1	8.6E-23	DKK3	3.0E-25
16	RFK	2.9E-15	DDX50	1.1E-23	GCLC	4.0E-22	COL8A2	3.2E-25
17	ETS2	3.4E-15	KIAA1279	1.0E-22	CBR3	4.5E-22	FILIP1L	4.2E-25
18	NR4A1	2.1E-14	BUD31	1.6E-22	UGDH	3.4E-21	TGFB1I1	8.8E-25
19	MUC5AC	1.1E-13	COX5A	1.6E-22	GPX2	3.9E-20	LAMA2	1.1E-24
20	MSLN	1.8E-13	TMEM189	1.9E-22	ALDH3A1	5.9E-20	LRP1	1.0E-23
21	PPARGC1A	3.5E-13	DNAJC9	2.4E-22	PRDX1	6.5E-20	PKD2	2.0E-23
22	CHL1	5.2E-13	PWP1	2.9E-22	FASLG	6.7E-20	SFXN3	2.1E-23
23	EPAS1	8.9E-13	RAN	3.6E-22	KIAA0319	1.7E-19	CDH11	3.6E-23
24	MUC5B	1.1E-12	PPIF	7.0E-22	CABYR	1.7E-18	COL10A1	3.8E-23
25	GABARAPL1	1.4E-12	SSB	9.9E-22	TSPAN7	2.2E-18	FLNA	5.1E-23
26	PER2	1.7E-12	MRPL13	1.4E-21	SQSTM1	3.9E-18	SPON1	6.0E-23
27	SORBS2	1.7E-12	MDH2	1.8E-21	FTL	1.4E-17	ITGBL1	7.2E-23
28	FAM46A	1.9E-12	GLRX3	2.1E-21	CES1	1.4E-17	ZEB1	1.2E-22
29	PLA2G10	2.0E-12	MRPS11	2.7E-21	FTH1	1.8E-17	ZCCHC24	1.4E-22
30	DUSP1	2.5E-12	CCDC86	2.7E-21	TRIM16	1.8E-16	AOC3	1.5E-22
31	CYP2C18	2.8E-12	C1QBP	2.7E-21	HTATIP2	1.9E-16	MYL9	2.0E-22
32	MTUS1	5.4E-12	EEF1E1	2.8E-21	MEGF9	3.7E-16	BACE1	2.0E-22
33	PTP4A1	6.5E-12	C16orf61	3.3E-21	NR0B1	3.7E-16	LMOD1	2.8E-22
34	TOB1	7.4E-12	ATP5B	3.6E-21	IDH1	4.7E-16	SEPT11	2.8E-22
35	SLC16A4	8.9E-12	PSMA7	3.9E-21	TXNRD1	6.6E-16	DCN	3.2E-22
36	NEDD9	9.4E-12	SNRPA1	4.0E-21	ABCC1	2.5E-15	RUNX1	7.0E-22
37	CATSPERB	1.1E-11	CACYBP	4.1E-21	GSR	6.3E-14	PEA15	1.5E-21
38	TSPAN8	1.3E-11	MRPL42	4.2E-21	UGT1A1	7.0E-14	EDNRA	2.3E-21
39	PTPRM	1.3E-11	ZC3H15	6.1E-21	UGT1A1	2.0E-13	LOXL1	3.0E-21
40	PLA2G4A	1.4E-11	TMEM93	1.2E-20	RIT1	3.2E-13	PRELP	3.2E-21
41	AQP3	1.9E-11	POP7	1.9E-20	ABCB6	8.9E-13	GLT8D2	3.3E-21
42	FURIN	2.4E-11	CHCHD3	3.2E-20	EPHX1	9.2E-13	MEOX2	4.8E-21
43	GEM	3.8E-11	PSMD14	3.7E-20	TXN	9.6E-13	ELN	6.2E-21
44	KIT	8.6E-11	ADRM1	6.5E-20	EGF	1.1E-12	EHD2	8.1E-21
45	KCNQ1	9.2E-11	COX4NB	7.8E-20	CYP4F3	2.7E-12	ZFP36L1	8.7E-21
46	TFCP2L1	9.9E-11	MTIF2	8.2E-20	SOD1	4.2E-12	MFAP4	1.1E-20
47	CD55	1.0E-10	WDR12	1.0E-19	TSKU	4.4E-12	ITGB5	1.3E-20
48	GABARAPL1	1.0E-10	DDX21	1.3E-19	DMPK	4.5E-12	ILK	1.4E-20
49	FOS	1.1E-10	PA2G4	1.5E-19	HGD	6.1E-11	FBN1	1.4E-20
50	MSMB	1.1E-10	GEMIN6	1.7E-19	TKT	1.0E-10	PHLDB1	1.9E-20
51	SMAD2	1.2E-10	PSMD9	2.7E-19	ADH1C	1.4E-10	C18orf1	2.2E-20
52	SPDEF	1.5E-10	SNRNP27	2.7E-19	P2RX5	1.6E-10	PDGFRB	3.3E-20
53	NNMT	1.5E-10	SMNDC1	3.8E-19	MAP2	2.6E-10	PDZRN3	3.9E-20
54	INSL4	1.6E-10	GHITM	4.1E-19	FTH1P5	3.1E-10	ERG	4.3E-20
55	ATP1B1	2.1E-10	GRPEL1	4.4E-19	TFE3	3.1E-10	MYH10	5.0E-20
56	RPL13A	2.2E-10	MRPL46	4.8E-19	SLC48A1	3.3E-10	CYR61	5.5E-20
57	ALDH3A2	2.2E-10	NOP16	4.9E-19	GLA	9.5E-10	MFGE8	6.7E-20
58	TFF1	2.4E-10	MRPL35	6.1E-19	PLAC1	1.1E-09	RECK	9.9E-20
59	GLCE	2.4E-10	WAPAL	6.7E-19	C20orf24	1.5E-09	TNFSF12	1.1E-19
60	PTPRB	2.5E-10	HCCS	6.7E-19	AKR1C4	2.3E-09	ZNF423	1.3E-19
61	MAP2K3	2.8E-10	MRPS30	6.8E-19	MAFG	3.0E-09	LEPROT	1.4E-19
62	ASAH1	3.7E-10	NUP37	7.1E-19	ABCC3	4.8E-09	DPYSL2	1.5E-19
63	AHCYL2	3.7E-10	RFC2	7.3E-19	RAP1GAP	4.8E-09	ENG	1.5E-19
64	GPRC5C	4.7E-10	SDHB	8.6E-19	IGHA1	6.2E-09	GSN	1.6E-19
65	KIF13B	5.2E-10	CCDC59	9.8E-19	TMOD1	6.3E-09	HEG1	2.5E-19
66	RHOBTB2	6.3E-10	ECD	1.4E-18	NMB	2.1E-08	TAGLN	2.7E-19
67	MUC13	6.3E-10	CDK1	1.5E-18	TBC1D2	2.2E-08	VGLL3	2.8E-19
68	ODC1	6.9E-10	EIF2S1	1.6E-18	HIGD1B	2.9E-08	ASPN	2.9E-19
69	CD46	9.8E-10	RAB22A	1.7E-18	DNAI2	3.7E-08	SMAD7	3.1E-19

Appendix C. (cont.)

Rank	FOX/CREB	P-value	Mito	P-value	NRF2	P-value	Down-regulated	P-value
70	TSC22D1	1.1E-09	EXOSC4	1.8E-18	SLC39A8	3.7E-08	ACTA2	3.2E-19
71	ENTPD4	1.4E-09	NUP88	1.9E-18	CLN5	3.8E-08	TNS1	3.4E-19
72	AGR2	1.5E-09	EIF5B	2.0E-18	SLN	4.4E-08	PTRF	3.4E-19
73	FLRT3	1.6E-09	FASTKD2	2.4E-18	UCHL1	6.4E-08	DAB2	4.6E-19
74	F3	1.8E-09	MRPL22	2.4E-18	SLC3A2	7.4E-08	STX12	4.6E-19
75	ULK2	2.1E-09	MRPL19	2.5E-18	ABHD4	8.0E-08	PLSCR4	5.2E-19
76	KLF2	3.3E-09	CYC1	2.6E-18	ASF1A	8.8E-08	RGS3	6.9E-19
77	HGSNAT	3.4E-09	PDCD11	4.0E-18	ABHD2	1.0E-07	DIXDC1	7.9E-19
78	SMPDL3B	4.2E-09	GOT2	4.4E-18	MYH14	1.0E-07	PARVA	1.0E-18
79	ERG	4.4E-09	ASCC1	4.6E-18	LAMP1	1.1E-07	DCHS1	1.1E-18
80	SIK2	6.1E-09	CCNJ	4.7E-18	CTSD	1.6E-07	SPARC	1.3E-18
81	CRLF1	6.1E-09	COX5B	5.3E-18	STAB1	1.7E-07	KIAA0754	1.4E-18
82	LPAR1	6.3E-09	FIP1L1	5.5E-18	BHMT2	1.8E-07	HEPH	1.6E-18
83	MAST4	6.3E-09	NIP7	5.7E-18	TDP2	2.0E-07	CALD1	1.8E-18
84	CARKD	8.3E-09	ZWINT	6.4E-18	LRP8	2.2E-07	PRRX1	1.9E-18
85	ELN	8.7E-09	NUP93	7.6E-18	FECH	3.1E-07	SFRP4	2.0E-18
86	NAAA	9.0E-09	MRPS34	7.7E-18	LTBP2	3.5E-07	RGL1	3.2E-18
87	PDZD2	1.0E-08	AURKA	7.9E-18	CREG1	4.0E-07	SLC34A2	3.6E-18
88	GALNT2	1.4E-08	UBE2N	8.0E-18	CYP4F2	5.2E-07	KLC1	3.6E-18
89	POGZ	1.7E-08	IMP4	8.3E-18	MAOA	5.7E-07	HTRA1	3.8E-18
90	AKAP12	1.7E-08	CSTF1	8.3E-18	ACSL1	5.9E-07	A2M	4.6E-18
91	MECOM	2.2E-08	EIF4E2	9.2E-18	BLVRB	6.8E-07	NRP1	4.6E-18
92	DNAJC12	2.6E-08	COPS3	1.1E-17	ATP7A	6.9E-07	CSGALNACT2	4.7E-18
93	ABLIM1	2.6E-08	C12orf11	1.2E-17	SLC6A6	7.3E-07	NOTCH2	5.2E-18
94	ATG12	3.2E-08	MTCH2	1.3E-17	NQO2	9.6E-07	SSPN	5.6E-18
95	RND1	3.3E-08	EIF2S2	1.3E-17	CLDN8	9.7E-07	SNED1	7.0E-18
96	MEIS3P1	3.3E-08	FAM64A	1.4E-17	CLDN15	1.0E-06	GAS7	7.1E-18
97	BMP2	3.4E-08	MRPL17	1.5E-17	ABCA4	1.0E-06	TCF21	7.9E-18
98	BARX1	3.4E-08	CCNB1	1.5E-17	DZIP3	1.1E-06	HLX	8.6E-18
99	FBLN5	3.5E-08	C14orf156	1.5E-17	LRP4	1.2E-06	TGFBR2	1.0E-17
100	RPL15	3.6E-08	KCTD5	1.8E-17	SEPX1	1.2E-06	APBB2	1.2E-17
101	CREB3L1	3.7E-08	UBE2I	1.9E-17	ROD1	1.3E-06	LHFP	1.5E-17
102	SNED1	3.7E-08	ZWILCH	2.3E-17	PHKB	1.3E-06	CNN1	1.6E-17
103	ALG9	4.0E-08	PSMB5	2.9E-17	SULT1A1	1.5E-06	MYH11	1.7E-17
104	ARSE	4.0E-08	MRPL12	3.5E-17	KYNU	1.5E-06	MEF2A	2.1E-17
105	WIF1	4.4E-08	ETFA	3.7E-17	TNS1	1.5E-06	OMD	2.4E-17
106	LIMCH1	4.5E-08	NDUFAB1	4.2E-17	SFN	1.6E-06	KIAA1462	3.0E-17
107	EPHA5	4.7E-08	EIF4A1	4.3E-17	SULT1A2	1.6E-06	PMP22	3.0E-17
108	SPRY1	5.3E-08	PPM1G	4.8E-17	AGA	1.7E-06	ATXN1	3.3E-17
109	URB1	5.7E-08	LRRC42	5.0E-17	TMED1	1.8E-06	TBX3	3.6E-17
110	SFTPB	5.9E-08	MRPS33	5.5E-17	KIAA0232	1.9E-06	CTGF	3.7E-17
111	PER1	6.0E-08	FEN1	5.6E-17	AKR1B1	2.1E-06	FHOD1	4.8E-17
112	HSPA12A	6.2E-08	DLAT	5.8E-17	MLPH	2.5E-06	PICALM	5.2E-17
113	TRAK1	6.2E-08	PFDN2	5.8E-17	SLC46A3	2.5E-06	EPS15	7.3E-17
114	KLF5	7.3E-08	ATP5J2	6.4E-17	GALNS	3.2E-06	TCF4	7.6E-17
115	CEBPD	7.5E-08	GAR1	6.5E-17	SULT1A3	3.8E-06	FXYD6	7.9E-17
116	ID1	7.8E-08	FAM149B1	7.3E-17	PACSIN2	4.9E-06	VIM	8.1E-17
117	ACACB	8.1E-08	RSL24D1	7.3E-17	ADCY7	5.3E-06	GAS6	8.8E-17
118	IGF1R	9.4E-08	GMFB	7.7E-17	CLCN4	5.3E-06	COLEC12	9.1E-17
119	PDE3A	9.6E-08	VPS26A	8.5E-17	NOL3	6.2E-06	CTSO	9.1E-17
120	BCAS1	1.0E-07	NOLC1	9.2E-17	CCND3	6.3E-06	GALNT10	1.4E-16
121	PHF17	1.0E-07	MRPL11	1.0E-16	ACE2	7.3E-06	PALLD	1.4E-16
122	PRKAB1	1.1E-07	SSSCA1	1.0E-16	FAH	7.5E-06	TGFB1	1.4E-16
123	GRAMD1B	1.1E-07	HSPD1	1.1E-16	PPFIBP2	7.8E-06	IDS	1.5E-16
124	BMP6	1.2E-07	ACP1	1.1E-16	ACOT13	8.3E-06	RIN2	1.5E-16
125	SGPP1	1.2E-07	EBNA1BP2	1.2E-16	CEACAM6	1.3E-05	PDLIM2	1.7E-16
126	BACE2	1.3E-07	IMMT	1.2E-16	SEMG2	1.3E-05	ARHGDIB	1.8E-16
127	MCF2L	1.4E-07	UBE2G1	1.3E-16	RNF24	1.4E-05	CD93	1.8E-16
128	C2orf67	1.5E-07	SNRPF	1.3E-16	PCOLCE2	1.5E-05	NEK1	1.9E-16
129	FNDC3A	1.8E-07	MCM4	1.3E-16	ATP6V1A	1.5E-05	ACVR1	2.1E-16
130	HYAL1	1.8E-07	METTL5	1.4E-16	MDFI	1.5E-05	SMPD1	2.1E-16
131	ATP2C2	2.0E-07	KIF4A	1.5E-16	IL9R	1.9E-05	ANKH	2.4E-16
132	SPINK1	2.0E-07	BOLA2	1.5E-16	GLB1	2.1E-05	UNC5B	2.4E-16
133	CHP	2.0E-07	MRPL16	1.8E-16	PSG3	2.1E-05	ITSN1	2.5E-16
134	C4BPA	2.1E-07	PGAM1	2.5E-16	NOS3	2.3E-05	SH3GLB1	2.9E-16
135	HGD	2.1E-07	AURKB	2.5E-16	RTN4	2.5E-05	OLFML1	3.0E-16
136	DAPK1	2.4E-07	PAICS	2.8E-16	DSTN	2.5E-05	EMP1	3.2E-16
137	GOLPH3L	2.5E-07	ACTR1A	3.1E-16	SEL1L3	2.5E-05	MMP2	3.3E-16
138	POU5F1P3	2.5E-07	ADSL	3.3E-16	ZBTB20	2.7E-05	CSRP1	3.5E-16

Appendix C. (cont.)

Rank	FOX/CREB	P-value	Mito	P-value	NRF2	P-value	Down-regulated	P-value
139	CDK2AP2	2.6E-07	NUP205	3.4E-16	HEXB	2.7E-05	LMCD1	4.0E-16
140	ATP9B	2.6E-07	DNPEP	3.6E-16	CHODL	2.9E-05	MOXD1	5.0E-16
141	TNFSF11	2.6E-07	NCAPG	4.2E-16	CEACAM1	3.6E-05	FAM134A	5.2E-16
142	FAM63A	2.7E-07	CHUK	4.2E-16	C11orf49	3.6E-05	NOX4	5.3E-16
143	OBSL1	2.7E-07	DBF4	4.3E-16	TBCD	3.7E-05	GPR124	6.2E-16
144	C14orf147	2.8E-07	AGPS	4.5E-16	SLC38A6	3.9E-05	CNN2	7.8E-16
145	SELENBP1	2.9E-07	PHB	4.7E-16	AHCYL1	4.0E-05	ARHGAP1	8.3E-16
146	FAT1	3.2E-07	RANBP1	5.0E-16	ZDHHC7	4.4E-05	TAX1BP3	9.0E-16
147	MOSC2	3.2E-07	PSMD12	5.1E-16	CLU	4.8E-05	MXRA7	1.1E-15
148	FAM83E	3.2E-07	DTYMK	5.3E-16	SYBU	4.8E-05	BGN	1.1E-15
149	CPE	3.3E-07	BUB3	5.5E-16	GPX3	4.9E-05	H2AFY	1.1E-15
150	COBL	3.3E-07	NOL7	5.5E-16	SERF2	5.1E-05	FBLN1	1.2E-15
151	NEDD4L	3.5E-07	FAM35A	6.2E-16	TCN2	5.5E-05	RIN3	1.3E-15
152	FGL1	3.7E-07	CIAPIN1	6.2E-16	GSTP1	5.5E-05	TLR2	1.3E-15
153	FGG	3.7E-07	ANXA7	6.3E-16	GULP1	6.2E-05	NAV3	1.4E-15
154	TNS1	4.5E-07	DENR	6.3E-16	LYVE1	6.4E-05	PAFAH1B1	1.4E-15
155	KCNK1	4.7E-07	MRPS17	6.8E-16	PTCH2	6.5E-05	SYNC	1.6E-15
156	C19orf21	4.9E-07	PPP1CC	7.1E-16	ELP4	6.7E-05	CTSD	1.6E-15
157	PDE8A	5.1E-07	TXNDC9	7.3E-16	ABCC5	6.9E-05	EPB41L2	1.6E-15
158	ARNT2	5.2E-07	C12orf10	8.1E-16	MFAP1	7.6E-05	VCL	2.1E-15
159	FOXO1	5.3E-07	SF3B3	8.1E-16	POU2F1	7.7E-05	FBXL5	2.3E-15
160	FZD3	5.3E-07	DRAP1	8.8E-16	LMNA	7.8E-05	MAP1LC3B	2.5E-15
161	CITED2	5.8E-07	NAE1	9.8E-16	CD63	8.1E-05	MICAL2	3.8E-15
162	PGC	7.0E-07	CEBPZ	1.1E-15	TBXAS1	8.3E-05	WDFY3	4.5E-15
163	FRAT1	7.0E-07	SERBP1	1.1E-15	MYO1D	8.8E-05	PLA2G15	4.6E-15
164	CSGALNACT1	7.4E-07	RPP40	1.1E-15	S100P	8.8E-05	MEF2C	4.8E-15
165	HYAL2	7.7E-07	SNRPD1	1.2E-15	SERINC5	9.5E-05	APLP2	5.0E-15
166	MLPH	7.8E-07	ATIC	1.3E-15	NEIL3	1.0E-04	DOCK4	6.1E-15
167	NEO1	8.0E-07	MDH1	1.3E-15	NAMPT	1.1E-04	ZEB2	6.4E-15
168	ALPL	8.1E-07	TMEM185B	1.4E-15	GNA15	1.1E-04	VDR	6.6E-15
169	UFC1	8.2E-07	COMMD4	1.5E-15	GAA	1.1E-04	EVC	6.7E-15
170	GYG2	1.0E-06	TIMM17B	1.5E-15	NUPR1	1.1E-04	ABCA6	7.4E-15
171	RPH3AL	1.0E-06	CEP55	1.5E-15	ZNF323	1.3E-04	COL6A2	7.6E-15
172	EDNRB	1.1E-06	ZNF259	1.5E-15	UBL3	1.3E-04	SNX19	7.6E-15
173	COL14A1	1.1E-06	MRPS7	1.6E-15	DAPK2	1.4E-04	TXNIP	9.1E-15
174	GALNT4	1.1E-06	CPSF6	1.7E-15	DUSP22	1.4E-04	SORBS1	9.1E-15
175	LIMD1	1.2E-06	BUB1B	1.7E-15	CD81	1.4E-04	FBLN5	9.5E-15
176	ABCA8	1.2E-06	HSPE1	1.7E-15	SLC7A8	1.4E-04	ISLR	1.0E-14
177	CHAT	1.2E-06	SMC3	1.7E-15	SNX19	1.4E-04	NBL1	1.0E-14
178	FZD10	1.3E-06	RNASEH1	1.9E-15	SC4MOL	1.4E-04	PGCP	1.0E-14
179	EFNA1	1.3E-06	MRPS22	2.1E-15	SLC22A18AS	1.5E-04	FAM178A	1.0E-14
180	PBXIP1	1.4E-06	SPC25	2.2E-15	TTC9	1.5E-04	MMP23A	1.0E-14
181	ERN2	1.4E-06	C20orf24	2.3E-15	BCL2L13	1.6E-04	MYO1B	1.1E-14
182	ZFP36	1.4E-06	SLMO2	2.3E-15	ALDH1A1	1.6E-04	IGFBP4	1.1E-14
183	JTB	1.5E-06	RARS	2.4E-15	ATP10D	1.7E-04	TNFSF12	1.1E-14
184	ADARB1	1.5E-06	BRAP	2.4E-15	TRMT61A	1.7E-04	LOH3CR2A	1.2E-14
185	RER1	1.5E-06	LRPPRC	2.5E-15	ATP6V0A1	1.7E-04	MSN	1.2E-14
186	ORAI2	1.5E-06	MARCH5	2.6E-15	C17orf108	1.7E-04	COL8A1	1.2E-14
187	RBPMS	1.6E-06	MSH2	2.7E-15	RNASE1	1.8E-04	MXRA8	1.3E-14
188	EIF4B	1.7E-06	TCEB1	2.7E-15	VGLL1	1.8E-04	TNFSF13	1.4E-14
189	PDE10A	1.7E-06	PAK1IP1	2.8E-15	ALCAM	1.8E-04	ACTG2	1.4E-14
190	TRIM31	1.7E-06	MOBKL3	2.8E-15	HPSE	1.9E-04	SEC24B	1.5E-14
191	ALDH2	1.8E-06	TUBA3C	2.8E-15	FMO2	1.9E-04	FN1	1.6E-14
192	TNFRSF10B	1.9E-06	PSMB6	2.9E-15	ALDOA	1.9E-04	TBX2	1.7E-14
193	FAM8A1	2.0E-06	AHSA1	2.9E-15	SH3BGRL	2.1E-04	TACC1	1.7E-14
194	PDK4	2.0E-06	OGFOD1	3.1E-15	MYOT	2.1E-04	LYST	1.7E-14
195	PARM1	2.0E-06	BUB1	3.1E-15	C5orf30	2.2E-04	CAPN2	1.7E-14
196	CDH15	2.1E-06	CDC20	3.1E-15	PCYT2	2.2E-04	HNMT	1.9E-14
197	RORC	2.1E-06	MRPS12	3.4E-15	TTC39A	2.3E-04	CHD9	1.9E-14
198	ATP11A	2.4E-06	TOMM40	3.6E-15	ANXA4	2.3E-04	JAM3	1.9E-14
199	SUCLG2	2.5E-06	NUDT21	3.7E-15	MPP2	2.3E-04	CAPN3	1.9E-14
200	CYP2C9	2.5E-06	CDT1	3.9E-15	HHIPL2	2.4E-04	MCOLN1	2.2E-14

REFERENCES

Aggarwal, S., Kim, S.W., Ryu, S.H., Chung, W.C., and Koo, J.S. (2008). Growth Suppression of Lung Cancer Cells by Targeting Cyclic AMP Response Element-Binding Protein. Cancer Res. *68*, 981–988.

Alessi, D.R., Sakamoto, K., and Bayascas, J.R. (2006). LKB1-dependent signaling pathways. Annu. Rev. Biochem. *75*, 137–163.

Alexander, A., Cai, S.L., Kim, J., Nanez, A., Sahin, M., MacLean, K.H., Inoki, K., Guan, K.L., Shen, J., Person, M.D., et al. (2010). Cozzarelli Prize Winner: ATM signals to TSC2 in the cytoplasm to regulate mTORC1 in response to ROS. Proc. Nat. Acad. Sci. U.S.A. *107*, 4153–4158.

Amin, N., Khan, A., St Johnston, D., Tomlinson, I., Martin, S., Brenman, J., and McNeill, H. (2009). LKB1 regulates polarity remodeling and adherens junction formation in the Drosophila eye. Proc. Nat. Acad. Sci. U.S.A. *106*, 8941–8946.

Baas, A.F., Kuipers, J., van der Wel, N.N., Batlle, E., Koerten, H.K., Peters, P.J., and Clevers, H.C. (2004). Complete polarization of single intestinal epithelial cells upon activation of LKB1 by STRAD. Cell *116*, 457–466.

Bardeesy, N., Sinha, M., Hezel, A.F., Signoretti, S., Hathaway, N.A., Sharpless, N.E., Loda, M., Carrasco, D.R., and DePinho, R.A. (2002). Loss of the Lkb1 tumour suppressor provokes intestinal polyposis but resistance to transformation. Nature *419*, 162–167.

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603–307.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., et al. (2012). NCBI GEO: archive for functional genomics data sets--update. Nucl. Acids Res. *41*, D991–D995.

Basseres, D.S., D'Alò, F., Yeap, B.Y., Löwenberg, E.C., Gonzalez, D.A., Yasuda, H., Dayaram, T., Kocher, O.N., Godleski, J.J., Richards, W.G., et al. (2012). Lung Cancer. Lung Cancer *77*, 31–37.

Beer, D.G., Kardia, S.L.R., Huang, C.-C., Giordano, T.J., Levin, A.M., Misek, D.E., Lin, L., Chen, G., Gharib, T.G., Thomas, D.G., et al. (2002). Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med.

Bhattacharjee, A., Richards, W.G., Staunton, J., Li, C., Monti, S., Vasa, P., Ladd, C., Beheshti, J., Bueno, R., Gillette, M., et al. (2001). Classification of human lung

carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc. Nat. Acad. Sci. U.S.A. *98*, 13790–13795.

Bin Zheng, Jeong, J.H., Asara, J.M., Yuan, Y.-Y., Granter, S.R., Chin, L., and Cantley, L.C. (2009). Oncogenic B-RAF Negatively Regulates the Tumor Suppressor LKB1 to Promote Melanoma Cell Proliferation. Mol. Cell *33*, 237–247.

Boardman, L.A., Thibodeau, S.N., Schaid, D.J., Lindor, N.M., McDonnell, S.K., Burgart, L.J., Ahlquist, D.A., Podratz, K.C., Pittelkow, M., and Hartmann, L.C. (1998). Increased risk for cancer in patients with the Peutz-Jeghers syndrome. Ann. Intern. Med. *128*, 896–899.

Britten, C.D. (2013). PI3K and MEK inhibitor combinations: examining the evidence in selected tumor types. Cancer Chemother. Pharmacol. *71*, 1395–1409.

Calnan, D.R., and Brunet, A. (2008). The FoxO code. Oncogene *27*, 2276–2288.

Cancer Genome Atlas Network (2012). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61–70.

Cancer Genome Atlas Network (2013). Diversity of Lung Adenocarcinoma Revealed by Integrative Molecular Profiling. Nature (Submitted).

Cancer Genome Atlas Research Network, Hammerman, P.S., Hayes, D.N., Wilkerson, M.D., Schultz, N., Bose, R., Chu, A., Collisson, E.A., Cope, L., Creighton, C.J., et al. (2012). Comprehensive genomic characterization of squamous cell lung cancers. Nature *489*, 519–525.

Carracedo, A., Ma, L., Teruya-Feldstein, J., Rojo, F., Salmena, L., Alimonti, A., Egia, A., Sasaki, A.T., Thomas, G., Kozma, S.C., et al. (2008). Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer. J. Clin. Invest.

Carretero, J., Shimamura, T., Rikova, K., Jackson, A.L., Wilkerson, M.D., Borgman, C.L., Buttarazzi, M.S., Sanofsky, B.A., McNamara, K.L., Brandstetter, K.A., et al. (2010). Integrative Genomic and Proteomic Analyses Identify Targets for Lkb1-Deficient Metastatic Lung Tumors. Cancer Cell *17*, 547–559.

Catalanotti, F., Solit, D.B., Pulitzer, M.P., Berger, M.F., Scott, S.N., Iyriboz, T., Lacouture, M.E., Panageas, K.S., Wolchok, J.D., Carvajal, R.D., et al. (2013). Phase II Trial of MEK Inhibitor Selumetinib (AZD6244, ARRY-142886) in Patients with BRAFV600E/K-Mutated Melanoma. Clin. Cancer Res. *19*, 2257–2264.

Chapman, P.B., Hauschild, A., Robert, C., Haanen, J.B., Ascierto, P., Larkin, J., Dummer, R., Garbe, C., Testori, A., Maio, M., et al. (2011). Improved Survival with Vemurafenib in Melanoma with BRAF V600E Mutation. N. Engl. J. Med. *364*, 2507–2516.

Chen, C.H. (2006). Reactive Oxygen Species Generation Is Involved in Epidermal Growth Factor Receptor Transactivation through the Transient Oxidization of Src Homology 2-Containing Tyrosine Phosphatase in Endothelin-1 Signaling Pathway in Rat Cardiac Fibroblasts. Mol. Pharmacol. *69*, 1347–1355.

Chen, Z., Cheng, K., Walton, Z., Wang, Y., Ebi, H., Shimamura, T., Liu, Y., Tupper, T., Ouyang, J., Li, J., et al. (2012). A murine lung cancer co-clinical trial identifies genetic modifiers of therapeutic response. Nature *483*, 613–617.

Chitale, D., Gong, Y., Taylor, B.S., Broderick, S., Brennan, C., Somwar, R., Golas, B., Wang, L., Motoi, N., Szoke, J., et al. (2009). An integrated genomic analysis of lung cancer reveals loss of DUSP4 in EGFR-mutant tumors. Oncogene *28*, 2773–2783.

Contreras, C.M., Gurumurthy, S., Haynie, J.M., Shirley, L.J., Akbay, E.A., Wingo, S.N., Schorge, J.O., Broaddus, R.R., Wong, K.-K., Bardeesy, N., et al. (2008). Loss of Lkb1 Provokes Highly Invasive Endometrial Adenocarcinomas. Cancer Res. *68*, 759–766.

Cunningham, J.T., Rodgers, J.T., Arlow, D.H., Vazquez, F., Mootha, V.K., and Puigserver, P. (2007). mTOR controls mitochondrial oxidative function through a YY1–PGC-1α transcriptional complex. Nature *450*, 736–740.

de Hoon, M.J.L., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. Bioinformatics *20*, 1453–1454.

De Raedt, T., Walton, Z., Yecies, J.L., Li, D., Chen, Y., Malone, C.F., Maertens, O., Jeong, S.M., Bronson, R.T., Lebleu, V., et al. (2011). Exploiting Cancer Cell Vulnerabilities to Develop a Combination Therapy for Ras-Driven Tumors. Cancer Cell *20*, 400–413.

Demetri, G.D., Mehren, von, M., Blanke, C.D., Van den Abbeele, A.D., Eisenberg, B., Roberts, P.J., Heinrich, M.C., Tuveson, D.A., Singer, S., Janicek, M., et al. (2002). Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors. N Engl J Med *347*, 472–480.

DeNicola, G.M., Karreth, F.A., Humpton, T.J., Gopinathan, A., Wei, C., Frese, K., Mangal, D., Yu, K.H., Yeo, C.J., Calhoun, E.S., et al. (2011). Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. Nature *475*, 106–109.

Ding, L., Getz, G., Wheeler, D.A., Mardis, E.R., McLellan, M.D., Cibulskis, K., Sougnez, C., Greulich, H., Muzny, D.M., Morgan, M.B., et al. (2008). Somatic mutations affect key pathways in lung adenocarcinoma. Nature *455*, 1069–1075.

Druker, B.J., Talpaz, M., Resta, D.J., Peng, B., Buchdunger, E., Ford, J.M., Lydon, N.B., Kantarjian, H., Capdeville, R., Ohno-Jones, S., et al. (2001). Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N. Engl. J. Med. *344*, 1031–1037.

Dry, J.R., Pavey, S., Pratilas, C.A., Harbron, C., Runswick, S., Hodgson, D., Chresta, C.,

McCormack, R., Byrne, N., Cockerill, M., et al. (2010). Transcriptional Pathway Signatures Predict MEK Addiction and Response to Selumetinib (AZD6244). Cancer Res. *70*, 2264–2273.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucl. Acids Res. *30*, 207–210.

Egan, D.F., Shackelford, D.B., Mihaylova, M.M., Gelino, S., Kohnz, R.A., Mair, W., Vasquez, D.S., Joshi, A., Gwinn, D.M., Taylor, R., et al. (2011). Phosphorylation of ULK1 (hATG1) by AMP-Activated Protein Kinase Connects Energy Sensing to Mitophagy. Science *331*, 456–461.

Eijkelenboom, A., Mokry, M., de Wit, E., Smits, L.M., Polderman, P.E., van Triest, M.H., van Boxtel, R., Schulze, A., de Laat, W., Cuppen, E., et al. (2013). Genome-wide analysis of FOXO3 mediated transcription regulation through RNA polymerase II profiling. Mol. Sys. Biol. *9*, 1–15.

Eneling, K., Brion, L., Pinto, V., Pinho, M.J., Sznajder, J.I., Mochizuki, N., Emoto, K., Soares-da-Silva, P., and Bertorello, A.M. (2012). Salt-inducible kinase 1 regulates E-cadherin expression and intercellular junction stability. The FASEB J. *26*, 3230–3239.

Esteller, M., Avizienyte, E., Corn, P.G., Lothe, R.A., Baylin, S.B., Aaltonen, L.A., and Herman, J.G. (2000). Epigenetic inactivation of LKB1 in primary tumors associated with the Peutz-Jeghers syndrome. Oncogene *19*, 164–168.

Esteve-Puig, R., Canals, F., Colomé, N., Merlino, G., and Recio, J.Á. (2009). Uncoupling of the LKB1-AMPKα Energy Sensor Pathway by Growth Factors and Oncogenic BRAFV600E. PLoS ONE *4*, e4771.

Evans, J.M.M., Donnelly, L.A., Emslie-Smith, A.M., Alessi, D.R., and Morris, A.D. (2005). Metformin and reduced risk of cancer in diabetic patients. Bmj *330*, 1304–1305.

Feng, Y., Wang, Y., Wang, Z., Fang, Z., Li, F., Gao, Y., Liu, H., Xiao, T., Li, F., Zhou, Y., et al. (2012). The CRTC1-NEDD9 signaling axis mediates lung cancer progression caused by LKB1 loss. Cancer Res.

Ferguson, J., Arozarena, I., Ehrhardt, M., and Wellbrock, C. (2012). Combination of MEK and SRC inhibition suppresses melanoma cell growth and invasion. Oncogene *32*, 86–96.

Fernandez-Marcos, P.J., and Auwerx, J. (2011). Regulation of PGC-1 , a nodal regulator of mitochondrial biogenesis. Am. J. Clin. Nutr. *93*, 884S–890S.

Flaherty, K.T., Infante, J.R., Daud, A., Gonzalez, R., Kefford, R.F., Sosman, J., Hamid, O., Schuchter, L., Cebon, J., Ibrahim, N., et al. (2012). Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations. N. Engl. J. Med. *367*, 1694–1703.

Flaherty, K.T., Puzanov, I., Kim, K.B., Ribas, A., McArthur, G.A., Sosman, J.A.,

O'Dwyer, P.J., Lee, R.J., Grippo, J.F., Nolop, K., et al. (2010). Inhibition of mutated, activated BRAF in metastatic melanoma. N. Engl. J. Med. *363*, 809–819.

Forbes, S.A., Bhamra, G., Bamford, S., Dawson, E., Kok, C., Clements, J., Menzies, A., Teague, J.W., Futreal, P.A., and Stratton, M.R. (2001). The Catalogue of Somatic Mutations in Cancer (COSMIC) (Hoboken, NJ, USA: John Wiley & Sons, Inc.).

Forbes, S.A., Bindal, N., Bamford, S., Cole, C., Kok, C.Y., Beare, D., Jia, M., Shepherd, R., Leung, K., Menzies, A., et al. (2010). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucl. Acids Res. *39*, D945–D950.

Fu, Z., and Tindall, D.J. (2008). FOXOs, cancer and regulation of apoptosis. Oncogene *27*, 2312–2319.

Gan, B., Hu, J., Jiang, S., Liu, Y., Sahin, E., Zhuang, L., Fletcher-Sananikone, E., Colla, S., Wang, Y.A., Chin, L., et al. (2010a). Lkb1 regulates quiescence and metabolic homeostasis of haematopoietic stem cells. Nature *468*, 701–704.

Gan, B., Lim, C., Chu, G., Hua, S., Ding, Z., Collins, M., Hu, J., Jiang, S., Fletcher-Sananikone, E., Zhuang, L., et al. (2010b). FoxOs Enforce a Progression Checkpoint to Constrain mTORC1-Activated Renal Tumorigenesis. Cancer Cell *18*, 472–484.

Gao, B., Sun, Y., Zhang, J., Ren, Y., Fang, R., Han, X., Shen, L., Liu, X.-Y., Pao, W., Chen, H., et al. (2010). Spectrum of LKB1, EGFR, and KRAS mutations in chinese lung adenocarcinomas. J. Thorac. Oncol. *5*, 1130–1135.

Garber, M.E., Troyanskaya, O.G., Schluens, K., Petersen, S., Thaesler, Z., Pacyna-Gengelbach, M., van de Rijn, M., Rosen, G.D., Perou, C.M., Whyte, R.I., et al. (2001). Diversity of gene expression in adenocarcinoma of the lung. Proc. Natl. Acad. Sci. U.S.A. *98*, 13784–13789.

García-Martínez, J.M., Wullschleger, S., Preston, G., Guichard, S., Fleming, S., Alessi, D.R., and Duce, S.L. (2011). Effect of PI3K- and mTOR-specific inhibitors on spontaneous B-cell follicular lymphomas in PTEN/LKB1-deficient mice. Br. J. Cancer *104*, 1116–1125.

Garnett, M.J., Edelman, E.J., Heidorn, S.J., Greenman, C.D., Dastur, A., Lau, K.W., Greninger, P., Thompson, I.R., Luo, X., Soares, J., et al. (2012). Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature *483*, 570–575.

Gazdar AF, Girard L, Lockwood WW, Lam WL, Minna JD. (2010). Lung cancer cell lines as tools for biomedical discovery and research. J. Natl. Cancer Inst. *102*, 1310-21

Giardiello, F.M., Welsh, S.B., Hamilton, S.R., Offerhaus, G.J., Gittelsohn, A.M., Booker, S.V., Krush, A.J., Yardley, J.H., and Luk, G.D. (1987). Increased risk of cancer in the Peutz-Jeghers syndrome. N. Engl. J. Med. *316*, 1511–1514.

Giardiello, F.M., Brensinger, J.D., Tersmette, A.C., Goodman, S.N., Petersen, G.M.,

Booker, S.V., Cruz-Correa, M., and Offerhaus, J.A. (2000). Very High Risk of Cancer in Familial Peutz-Jeghers Syndrome. Gastroenterology *119*, 1447–1453.

Gill, R.K., Yang, S.-H., Meerzaman, D., Mechanic, L.E., Bowman, E.D., Jeon, H.-S., Chowdhuri, S.R., Shakoori, A., Dracheva, T., Hong, K.-M., et al. (2011). Frequent homozygous deletion of the LKB1/STK11 gene in non-small cell lung cancer. Oncogene *30*, 3784–3791.

Girotti, M.R., Pedersen, M., Sanchez-Laorden, B., Viros, A., Turajlic, S., Niculescu-Duvaz, D., Zambon, A., Sinclair, J., Hayes, A., Gore, M., et al. (2013). Inhibiting EGF Receptor or SRC Family Kinase Signaling Overcomes BRAF Inhibitor Resistance in Melanoma. Cancer Disc. *3*, 158–167.

Godlewski, J., Nowicki, M.O., Bronisz, A., Nuovo, G., Palatini, J., De Lay, M., Van Brocklyn, J., Ostrowski, M.C., Chiocca, E.A., and Lawler, S.E. (2010). MicroRNA-451 Regulates LKB1/AMPK Signaling and Allows Adaptation to Metabolic Stress in Glioma Cells. Mol. Cell *37*, 620–632.

Golub, T.R. (1999). Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. Science *286*, 531–537.

Gopal, Y.N.V., Deng, W., Woodman, S.E., Komurov, K., Ram, P., Smith, P.D., and Davies, M.A. (2010). Basal and treatment-induced activation of AKT mediates resistance to cell death by AZD6244 (ARRY-142886) in Braf-mutant human cutaneous melanoma cells. Cancer Res. *70*, 8736–8747.

Greer, E.L., Oskoui, P.R., Banko, M.R., Maniar, J.M., Gygi, M.P., Gygi, S.P., and Brunet, A. (2007). The energy sensor AMP-activated protein kinase directly regulates the mammalian FOXO3 transcription factor. J. Biol. Chem. *282*, 30107–30119.

Gu, Y., Lin, S., Li, J.-L., Nakagawa, H., Chen, Z., Jin, B., Tian, L., Ucar, D.A., Shen, H., Lu, J., et al. (2012). Altered LKB1/CREB-regulated transcription co-activator (CRTC) signaling axis promotes esophageal cancer cell migration and invasion. Oncogene *31*, 469–479.

Gupta, A., Yu, X., Case, T., Paul, M., Shen, M.M., Kaestner, K.H., and Matusik, R.J. (2012). Mash1 expression is induced in neuroendocrine prostate cancer upon the loss of Foxa2. Prostate *73*, 582–589.

Gurumurthy, S., Xie, S.Z., Alagesan, B., Kim, J., Yusuf, R.Z., Saez, B., Tzatsos, A., Ozsolak, F., Milos, P., Ferrari, F., et al. (2010). The Lkb1 metabolic sensor maintains haematopoietic stem cell survival. Nature *468*, 659–663.

Gwinn, D.M., Shackelford, D.B., Egan, D.F., Mihaylova, M.M., Mery, A., Vasquez, D.S., Turk, B.E., and Shaw, R.J. (2008). AMPK Phosphorylation of Raptor Mediates a Metabolic Checkpoint. Mol. Cell *30*, 214–226.

Gysin, S., Paquette, J., and McMahon, M. (2012). Analysis of mRNA Profiles after

MEK1/2 Inhibition in Human Pancreatic Cancer Cell Lines Reveals Pathways Involved in Drug Sensitivity. Mol Cancer Res. *10*, 1607–1619.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57–70.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. Cell *144*, 646–674.

Hardie, D., and Alessi, D.R. (2013). LKB1 and AMPK and the cancer-metabolism link - ten years after. BMC Biol. *11*, 36.

Hawley, S., Boudeau, J., Reid, J., Mustard, K., Udd, L., Mäkelä, T., Alessi, D., and Hardie, D.G. (2003). Complexes between the LKB1 tumor suppressor, STRADα/β and MO25α/β are upstream kinases in the AMP-activated protein kinase cascade. J. Biol. *2*, 28.

Hayes, D.N., Monti, S., Parmigiani, G., Gilks, C.B., Naoki, K., Bhattacharjee, A., Socinski, M.A., Perou, C., and Meyerson, M. (2006). Gene Expression Profiling Reveals Reproducible Human Lung Adenocarcinoma Subtypes in Multiple Independent Patient Cohorts. J. Clin. Oncol. *24*, 5079–5090.

Hayes, J.D., and McMahon, M. (2009). NRF2 and KEAP1 mutations: permanent activation of an adaptive response in cancer. Trends Biochem. Sci. *34*, 176–188.

He, L., Sabet, A., Djedjos, S., Miller, R., Sun, X., Hussain, M.A., Radovick, S., and Wondisford, F.E. (2009). Metformin and Insulin Suppress Hepatic Gluconeogenesis through Phosphorylation of CREB Binding Protein. Cell *137*, 635–646.

Hearle, N. (2006). Frequency and Spectrum of Cancers in the Peutz-Jeghers Syndrome. Clin. Cancer Res. *12*, 3209–3215.

Hemminki, A., Markie, D., Tomlinson, I., Avizienyte, E., Roth, S., Loukola, A., Bignell, G., Warren, W., Aminoff, M., and Höglund, P. (1998). A serine/threonine kinase gene defective in Peutz–Jeghers syndrome. Nature *391*, 184–187.

Hezel, A.F., Gurumurthy, S., Granot, Z., Swisa, A., Chu, G.C., Bailey, G., Dor, Y., Bardeesy, N., and DePinho, R.A. (2008). Pancreatic Lkb1 Deletion Leads to Acinar Polarity Defects and Cystic Neoplasms. Mol. Cell. Biol. *28*, 2414–2425.

Hirota, K. (2001). Redox-sensitive Transactivation of Epidermal Growth Factor Receptor by Tumor Necrosis Factor Confers the NF-kappa B Activation. J. Biol. Chem. *276*, 25953–25958.

Homma, S., Ishii, Y., Morishima, Y., Yamadori, T., Matsuno, Y., Haraguchi, N., Kikuchi, N., Satoh, H., Sakamoto, T., Hizawa, N., et al. (2009). Nrf2 Enhances Cell Proliferation and Resistance to Anticancer Drugs in Human Lung Cancer. Clin. Cancer Res. *15*, 3423–3432.

Huang, J., and Manning, B.D. (2009). A complex interplay between Akt, TSC2 and the two mTOR complexes. Biochem. Soc. Trans *37*, 217.

Huang, S., Hölzel, M., Knijnenburg, T., Schlicker, A., Roepman, P., McDermott, U., Garnett, M., Grernrum, W., Sun, C., Prahallad, A., et al. (2012). MED12 Controls the Response to Multiple Cancer Drugs through Regulation of TGF-β Receptor Signaling. Cell *151*, 937–950.

Huang, X., Wullschleger, S., Shpiro, N., McGuire, V.A., Sakamoto, K., Woods, Y.L., Mcburnie, W., Fleming, S., and Alessi, D.R. (2008). Important role of the LKB1–AMPK pathway in suppressing tumorigenesis in PTEN-deficient mice. Biochem. J. *412*, 211.

Imielinski, M., Berger, A.H., Hammerman, P.S., Hernandez, B., Pugh, T.J., Hodis, E., Cho, J., Suh, J., Capelletti, M., Sivachenko, A., et al. (2012). Mapping the Hallmarks of Lung Adenocarcinoma with Massively Parallel Sequencing. Cell *150*, 1107–1120.

Jänne, P.A., Shaw, A.T., Pereira, J.R., and Jeannin, G. (2013). Selumetinib plus docetaxel for KRAS-mutant advanced non-small-cell lung cancer: a randomised, multicentre, placebo-controlled, phase 2 study. Lancet Oncol. *14*, 38–47.

Jeghers, H., McKusick, V.A., and Katz, K.H. (1949). Generalized intestinal polyposis and melanin spots of the oral mucosa, lips and digits; a syndrome of diagnostic significance. N. Engl. J. Med. *241*, 1031–1036.

Jeon, S.-M., Chandel, N.S., and Hay, N. (2012). AMPK regulates NADPH homeostasis to promote tumour cell survival during energy stress. Nature *485*, 661–665.

Ji, H., Ramsey, M.R., Hayes, D.N., Fan, C., McNamara, K., Kozlowski, P., Torrice, C., Wu, M.C., Shimamura, T., Perera, S.A., et al. (2007). LKB1 modulates lung cancer differentiation and metastasis. Nature *448*, 807–810.

Jiralerspong, S., Palla, S.L., Giordano, S.H., Meric-Bernstam, F., Liedtke, C., Barnett, C.M., Hsu, L., Hung, M.C., Hortobagyi, G.N., and Gonzalez-Angulo, A.M. (2009). Metformin and Pathologic Complete Responses to Neoadjuvant Chemotherapy in Diabetic Patients With Breast Cancer. J. Clin. Oncol. *27*, 3297–3302.

Katajisto, P., Vaahtomeri, K., Ekman, N., Ventelä, E., Ristimäki, A., Bardeesy, N., Feil, R., DePinho, R.A., and Mäkelä, T.P. (2008). LKB1 signaling in mesenchymal cells required for suppression of gastrointestinal polyposis. Nat. Genet. *40*, 455–459.

Kau, T.R., Schroeder, F., Ramaswamy, S., Wojciechowski, C.L., Zhao, J.J., Roberts, T.M., Clardy, J., Sellers, W.R., and Silver, P.A. (2003). A chemical genetic screen identifies inhibitors of regulated nuclear export of a Forkhead transcription factor in PTEN-deficient tumor cells. Cancer Cell *4*, 463–476.

Klooster, ten, J.P., Jansen, M., Yuan, J., Oorschot, V., Begthel, H., Di Giacomo, V., Colland, F., de Koning, J., Maurice, M.M., Hornbeck, P., et al. (2009). Mst4 and Ezrin Induce Brush Borders Downstream of the Lkb1/Strad/Mo25 Polarization Complex.

Developmental Cell *16*, 551–562.

Koivunen, J.P., Kim, J., Lee, J., Rogers, A.M., Park, J.O., Zhao, X., Naoki, K., Okamoto, I., Nakagawa, K., Yeap, B.Y., et al. (2008). Mutations in the LKB1 tumour suppressor are frequently detected in tumours from Caucasian but not Asian lung cancer patients. Br. J. Cancer *99*, 245–252.

Komiya, T., Coxon, A., Park, Y., Chen, W.-D., Zajac-Kaye, M., Meltzer, P., Karpova, T., and Kaye, F.J. (2009). Enhanced activity of the CREB co-activator Crtc1 in LKB1 null lung cancer. Oncogene 1–9.

Koo, S.-H., Flechner, L., Qi, L., Zhang, X., Screaton, R.A., Jeffries, S., Hedrick, S., Xu, W., Boussouar, F., Brindle, P., et al. (2005a). The CREB coactivator TORC2 is a key regulator of fasting glucose metabolism. Nat. Cell Biol. *437*, 1109–1111.

Koo, S.-H., Flechner, L., Qi, L., Zhang, X., Screaton, R.A., Jeffries, S., Hedrick, S., Xu, W., Boussouar, F., Brindle, P., et al. (2005b). The CREB coactivator TORC2 is a key regulator of fasting glucose metabolism. Nature *437*, 1109–1111.

Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.-P., Subramanian, A., Ross, K.N., et al. (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science *313*, 1929–1935.

Landman, G.W.D., Kleefstra, N., van Hateren, K.J.J., Groenier, K.H., Gans, R.O.B., and Bilo, H.J.G. (2010). Metformin Associated With Lower Cancer Mortality in Type 2 Diabetes: ZODIAC-16. Diabetes Care *33*, 322–326.

Laplante, M., and Sabatini, D.M. (2012). mTOR Signaling in Growth Control and Disease. Cell *149*, 274–293.

Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature *499*, 214–218.

Lee, J.H., Koh, H., Kim, M., Kim, Y., Lee, S.Y., Karess, R.E., Lee, S.-H., Shong, M., Kim, J.-M., Kim, J., et al. (2007). Energy-dependent regulation of cell structure by AMP-activated protein kinase. Nature *447*, 1017–1020.

Libby, G., Donnelly, L.A., Donnan, P.T., Alessi, D.R., Morris, A.D., and Evans, J.M.M. (2009). New Users of Metformin Are at Low Risk of Incident Cancer: A cohort study among people with type 2 diabetes. Diabetes Care *32*, 1620–1625.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739–1740.

Lin-Marq, N., Borel, C., and Antonarakis, S.E. (2005). Peutz-Jeghers LKB1 mutants fail to activate GSK-3beta, preventing it from inhibiting Wnt signaling. Mol. Genet. Genomics *273*, 184–196.

Liu, M., Lee, D.-F., Chen, C.-T., Yen, C.-J., Li, L.-Y., Lee, H.-J., Chang, C.-J., Chang, W.-C., Hsu, J.-M., Kuo, H.-P., et al. (2012a). IKKa Activation of NOTCHLinks Tumorigenesis via FOXA2 Suppression. Mol. Cell *45*, 171–184.

Liu, W., Monahan, K.B., Pfefferle, A.D., Shimamura, T., Sorrentino, J., Chan, K.T., Roadcap, D.W., Ollila, D.W., Thomas, N.E., Castrillon, D.H., et al. (2012b). LKB1/STK11 Inactivation Leads to Expansion of a Prometastatic Tumor Subpopulation in Melanoma. Cancer Cell *21*, 751–764.

Lizcano, J.M., Göransson, O., Toth, R., Deak, M., Morrice, N.A., Boudeau, J., Hawley, S.A., Udd, L., Mäkelä, T.P., Hardie, D.G., et al. (2004). LKB1 is a master kinase that activates 13 kinases of the AMPK subfamily, including MARK/PAR-1. EMBO J. *23*, 833–843.

Lo, B., Strasser, G., Sagolla, M., Austin, C.D., Junttila, M., and Mellman, I. (2012). Lkb1 regulates organogenesis and early oncogenesis along AMPK-dependent and -independent pathways. J. Cell Biol. *199*, 1117–1130.

Loboda, A., Nebozhyn, M., Klinghoffer, R., Frazier, J., Chastain, M., Arthur, W., Roberts, B., Zhang, T., Chenard, M., Haines, B., et al. (2010). A gene expression signature of RAS pathway dependence predicts response to PI3K and RAS pathway inhibitors and expands the population of RAS pathway activated tumors. BMC Med. Genomics *3*, 26.

Loi, S., Michiels, S., Lambrechts, D., Fumagalli, D., Claes, B., Kellokumpu-Lehtinen, P.L., Bono, P., Kataja, V., Piccart, M.J., Joensuu, H., et al. (2013). Somatic Mutation Profiling and Associations With Prognosis and Trastuzumab Benefit in Early Breast Cancer. J. Natl. Cancer Inst. *105*, 960–967.

Londesborough, A., Vaahtomeri, K., Tiainen, M., Katajisto, P., Ekman, N., Vallenius, T., and Makela, T.P. (2008). LKB1 in endothelial cells is required for angiogenesis and TGF -mediated vascular smooth muscle cell recruitment. Development *135*, 2331–2338.

Luo, L., Huang, W., Tao, R., Hu, N., Xiao, Z.-X., and Luo, Z. (2013). ATM and LKB1 dependent activation of AMPK sensitizes cancer cells to etoposide-induced apoptosis. Cancer Letters *328*, 114–119.

Lützner, N., Arce, J.D.-C., and Rösl, F. (2012a). Gene Expression of the Tumour Suppressor LKB1 Is Mediated by Sp1, NF-Y and FOXO Transcription Factors. PLoS ONE *7*, e32590.

Lützner, N., Kalbacher, H., Krones-Herzig, A., and Rösl, F. (2012b). FOXO3 is a glucocorticoid receptor target and regulates LKB1 and its own expression based on cellular AMP levels via a positive autoregulatory loop. PLoS ONE *7*, e42166.

Lynch, T.J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G., Haluska, F.G., et al. (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N. Engl. J. Med. *350*, 2129–2139.

Mahoney, C.L., Choudhury, B., Davies, H., Edkins, S., Greenman, C., Haaften, G.V., Mironenko, T., Santarius, T., Stevens, C., Stratton, M.R., et al. (2009). LKB1/KRAS mutant lung cancers constitute a genetic subset of NSCLC with increased sensitivity to MAPK and mTOR signalling inhibition. Br. J. Cancer *100*, 370–375.

Manning, B.D. (2005). Feedback inhibition of Akt signaling limits the growth of tumors lacking Tsc2. Genes & Dev. *19*, 1773–1778.

Margel, D., Urbach, D.R., Lipscombe, L.L., Bell, C.M., Kulkarni, G., Austin, P.C., and Fleshner, N. (2013). Metformin Use and All-Cause and Prostate Cancer-Specific Mortality Among Men With Diabetes. J. of Clin. Oncol.

Marignani, P.A. (2001). LKB1 Associates with Brg1 and Is Necessary for Brg1-induced Growth Arrest. J. Biol. Chem. *276*, 32415–32418.

Matsumoto, S., Iwakawa, R., Takahashi, K., Kohno, T., Nakanishi, Y., Matsuno, Y., Suzuki, K., Nakamoto, M., Shimizu, E., Minna, J.D., et al. (2007). Prevalence and specificity of LKB1 genetic alterations in lung cancers. Oncogene *26*, 5911–5918.

Matys, V. (2006). TRANSFAC (R) and its module TRANSCompel (R): transcriptional gene regulation in eukaryotes. Nucl. Acids Res. *34*, D108–D110.

McCarthy, A., Lord, C.J., Savage, K., Grigoriadis, A., Smith, D.P., Weigelt, B., Reis-Filho, J.S., and Ashworth, A. (2009). Conditional deletion of the Lkb1gene in the mouse mammary gland induces tumour formation. J. Pathol. *219*, 306–316.

Meng, J., Fang, B., Liao, Y., Chresta, C.M., Smith, P.D., and Roth, J.A. (2010). Apoptosis Induction by MEK Inhibition in Human Lung Cancer Cells Is Mediated by Bim. PLoS ONE *5*, e13026.

Mirouse, V., Swick, L.L., Kazgan, N., St Johnston, D., and Brenman, J.E. (2007). LKB1 and AMPK maintain epithelial cell polarity under energetic stress. J. Cell Biol. *177*, 387–392.

Mitsuishi, Y., Taguchi, K., Kawatani, Y., Shibata, T., Nukiwa, T., Aburatani, H., Yamamoto, M., and Motohashi, H. (2012). Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming. Cancer Cell *22*, 66–79.

Murillo, M.M., Carmona-Cuenca, I., del Castillo, G., Ortiz, C., Roncero, C., Sánchez, A., Fernández, M., and Fabregat, I. (2007). Activation of NADPH oxidase by transforming growth factor-β in hepatocytes mediates up-regulation of epidermal growth factor receptor ligands through a nuclear factor-κB-dependent mechanism. Biochem. J. *405*, 251.

Nakada, D., Saunders, T.L., and Morrison, S.J. (2010). Lkb1 regulates cell cycle and energy metabolism in haematopoietic stem cells. Nature *468*, 653–658.

Nakau, M., Miyoshi, H., Seldin, M.F., Imamura, M., Oshima, M., and Taketo, M.M. (2002). Hepatocellular carcinoma caused by loss of heterozygosity in Lkb1 gene knockout mice. Cancer Res. *62*, 4549–4553.

Ohta, T., Iijima, K., Miyamoto, M., Nakahara, I., Tanaka, H., Ohtsuji, M., Suzuki, T., Kobayashi, A., Yokota, J., Sakiyama, T., et al. (2008). Loss of Keap1 Function Activates Nrf2 and Provides Advantages for Lung Cancer Cell Growth. Cancer Res. *68*, 1303–1309.

Okuda, K., Sasaki, H., Hikosaka, Y., Kawano, O., Moriyama, S., Yano, M., and Fujii, Y. (2010). LKB1 gene alterations in surgically resectable adenocarcinoma of the lung. Surg Today *41*, 107–110.

Orlova, K.A., Parker, W.E., Heuer, G.G., Tsai, V., Yoon, J., Baybis, M., Fenning, R.S., Strauss, K., and Crino, P.B. (2010). STRADα deficiency results in aberrant mTORC1 signaling during corticogenesis in humans and mice. J. Clin. Invest. *120*, 1591–1602.

Ossipova, O., Bardeesy, N., DePinho, R.A., and Green, J.B.A. (2003). LKB1 (XEEK1) regulates Wnt signalling in vertebrate development. Nat. Cell Biol. *5*, 889–894.

Paez, J.G., Jänne, P.A., Lee, J.C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., et al. (2004). EGFR Mutations in Lung Cancer: Correlation with Clinical Response to Gefitinib Therapy. Science *304*, 1497–1500.

Parkinson, H., Kapushesky, M., Kolesnikov, N., Rustici, G., Shojatalab, M., Abeygunawardena, N., Berube, H., Dylag, M., Emam, I., Farne, A., et al. (2009). ArrayExpress update--from an archive of functional genomics experiments to the atlas of gene expression. Nucl. Acids Res. *37*, D868–D872.

Partanen, J.I., Nieminen, A.I., Mäkelä, T.P., and Klefström, J. (2007). Suppression of oncogenic properties of c-Myc by LKB1-controlled epithelial organization. Proc. Natl. Acad. Sci. U.S.A. *104*, 14694–14699.

Partanen, J.I., Tervonen, T.A., Myllynen, M., Lind, E., Imai, M., Katajisto, P., Dijkgraaf, G.J.P., Kovanen, P.E., Mäkelä, T.P., Werb, Z., et al. (2012). Tumor suppressor function of Liver kinase B1 (Lkb1) is linked to regulation of epithelial integrity. Proc. Natl. Acad. Sci. U.S.A *109*, E388–E397.

Pearson, H.B., McCarthy, A., Collins, C.M.P., Ashworth, A., and Clarke, A.R. (2008). Lkb1 Deficiency Causes Prostate Neoplasia in the Mouse. Cancer Res. *68*, 2223–2232.

Peutz, J. (1921). Very remarkable case of familial case of polyposis of mucous membrane of intestinal tract and nasopharynx accompanied by peculiar pigmentations of skin and mucous membrane. Ned Maandschr Geneeskd *10*, 134–146.

Pratilas, C.A., Hanrahan, A.J., Halilovic, E., Persaud, Y., Soh, J., Chitale, D., Shigematsu, H., Yamamoto, H., Sawai, A., Janakiraman, M., et al. (2008). Genetic Predictors of MEK Dependence in Non-Small Cell Lung Cancer. Cancer Res. *68*, 9375–9383.

Puffenberger, E.G., Strauss, K.A., Ramsey, K.E., Craig, D.W., Stephan, D.A., Robinson, D.L., Hendrickson, C.L., Gottlieb, S., Ramsay, D.A., Siu, V.M., et al. (2007). Polyhydramnios, megalencephaly and symptomatic epilepsy caused by a homozygous 7-kilobase deletion in LYK5. Brain *130*, 1929–1941.

Qi, J., Nakayama, K., Cardiff, R.D., Borowsky, A.D., Kaul, K., Williams, R., Krajewski, S., Mercola, D., Carpenter, P.M., Bowtell, D., et al. (2010). Siah2-Dependent Concerted Activity of HIFand FoxA2 Regulates Formation of Neuroendocrine Phenotype and Neuroendocrine Prostate Tumors. Cancer Cell *18*, 23–38.

Reis, P.P., Waldron, L., Goswami, R.S., Xu, W., Xuan, Y., Perez-Ordonez, B., Gullane, P., Irish, J., Jurisica, I., and Kamel-Reid, S. (2011). mRNA transcript quantification in archival samples using multiplexed, color-coded probes. BMC Biotech. *11*, 46.

Robinson, J., Lai, C., Martin, A., Nye, E., Tomlinson, I., and Silver, A. (2009). Oral rapamycin reduces tumour burden and vascularization in Lkb1+/−mice. J. Pathol. *219*, 35–40.

Rodrik-Outmezguine, V.S., Chandarlapaty, S., Pagano, N.C., Poulikakos, P.I., Scaltriti, M., Moskatel, E., Baselga, J., Guichard, S., and Rosen, N. (2011). mTOR Kinase Inhibition Causes Feedback-Dependent Biphasic Regulation of AKT Signaling. Cancer Discov. *1*, 248–259.

Rosenbloom, K.R., Sloan, C.A., Malladi, V.S., Dreszer, T.R., Learned, K., Kirkup, V.M., Wong, M.C., Maddren, M., Fang, R., Heitner, S.G., et al. (2012). ENCODE Data in the UCSC Genome Browser: year 5 update. Nucl. Acids Res. *41*, D56–D63.

Rossi, D.J., Ylikorkala, A., Korsisaari, N., Salovaara, R., Luukko, K., Launonen, V., Henkemeyer, M., Ristimäki, A., Aaltonen, L.A., and Mäkelä, T.P. (2002). Induction of cyclooxygenase-2 in a mouse model of Peutz-Jeghers polyposis. Proc. Natl. Acad. Sci. U.S.A. *99*, 12327–12332.

Roy, B.C., Kohno, T., Iwakawa, R., Moriguchi, T., Kiyono, T., Morishita, K., Sanchez-Cespedes, M., Akiyama, T., and Yokota, J. (2010). Involvement of LKB1 in epithelialâ€“mesenchymal transition (EMT) of human lung cancer cells. Lung Cancer *70*, 136–145.

Saldanha, A.J. (2004). Java Treeview--extensible visualization of microarray data. Bioinformatics *20*, 3246–3248.

Sanchez-Cespedes, M., Parrella, P., Esteller, M., Nomoto, S., Trink, B., Engles, J.M., Westra, W.H., Herman, J.G., and Sidransky, D. (2002). Inactivation of LKB1/STK11 is a common event in adenocarcinomas of the lung. Cancer Res. *62*, 3659–3662.

Sapkota, G.P. (2001). Phosphorylation of the Protein Kinase Mutated in Peutz-Jeghers Cancer Syndrome, LKB1/STK11, at Ser431 by p90RSK and cAMP-dependent Protein Kinase, but Not Its Farnesylation at Cys433, Is Essential for LKB1 to Suppress Cell Growth. J. Biol. Chem. *276*, 19469–19482.

Sapkota, G.P., Boudeau, J., Deak, M., Kieloch, A., Morrice, N., and Alessi, D.R. (2002). Identification and characterization of four novel phosphorylation sites (Ser31, Ser325, Thr336 and Thr366) on LKB1/STK11, the protein kinase mutated in Peutz-Jeghers cancer syndrome. Biochem. J. *362*, 481–490.

Screaton, R.A., Conkright, M.D., Katoh, Y., Best, J.L., Canettieri, G., Jeffries, S., Guzman, E., Niessen, S., Yates, J.R., Takemori, H., et al. (2004). The CREB coactivator TORC2 functions as a calcium- and cAMP-sensitive coincidence detector. Cell *119*, 61–74.

Selamat, S.A., Chung, B.S., Girard, L., Zhang, W., Zhang, Y., Campan, M., Siegmund, K.D., Koss, M.N., Hagen, J.A., Lam, W.L., et al. (2012). Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. Genome Res. *22*, 1197–1211.

Seo, H.S., Liu, D.D., Bekele, B.N., Kim, M.K., Pisters, K., Lippman, S.M., Wistuba, I.I., and Koo, J.S. (2008). Cyclic AMP Response Element-Binding Protein Overexpression: A Feature Associated with Negative Prognosis in Never Smokers with Non-Small Cell Lung Cancer. Cancer Res. *68*, 6065–6073.

Shackelford, D.B., and Shaw, R.J. (2009). The LKB1-AMPK pathway: metabolism and growth control in tumour suppression. Nat. Rev. Cancer *9*, 563–575.

Shackelford, D.B., Abt, E., Gerken, L., Vasquez, D.S., Seki, A., Leblanc, M., Wei, L., Fishbein, M.C., Czernin, J., Mischel, P.S., et al. (2013). LKB1 Inactivation Dictates Therapeutic Response of Non-Small Cell Lung Cancer to the Metabolism Drug Phenformin. Cancer Cell 1–16.

Shackelford, D.B., Vasquez, D.S., Corbeil, J., Wu, S., Leblanc, M., Wu, C.-L., Vera, D.R., and Shaw, R.J. (2009). mTOR and HIF-1alpha-mediated tumor metabolism in an LKB1 mouse model of Peutz-Jeghers syndrome. Proceedings of the National Academy of Sciences *106*, 11137–11142.

Shaw, A.T., Kim, D.-W., Nakagawa, K., Seto, T., Crinò, L., Ahn, M.-J., De Pas, T., Besse, B., Solomon, B.J., Blackhall, F., et al. (2013). Crizotinib versus Chemotherapy in Advanced ALK-Positive Lung Cancer. N. Engl. J. Med. *368*, 2385–2394.

Shaw, R.J. (2005). The Kinase LKB1 Mediates Glucose Homeostasis in Liver and Therapeutic Effects of Metformin. Science *310*, 1642–1646.

Shaw, R.J., and Cantley, L.C. (2006a). Ras, PI (3)K and mTOR signalling controls tumour cell growth. Nat. Cell. Biol. *441*, 424–430.

Shaw, R.J., and Cantley, L.C. (2006b). Ras, PI (3)K and mTOR signalling controls tumour cell growth. Nat. Cell. Biol. *441*, 424–430.

Shaw, R.J., Bardeesy, N., Manning, B.D., Lopez, L., Kosmatka, M., DePinho, R.A., and Cantley, L.C. (2004a). The LKB1 tumor suppressor negatively regulates mTOR signaling. Cancer Cell *6*, 91–99.

Shaw, R.J., Kosmatka, M., Bardeesy, N., Hurley, R.L., Witters, L.A., DePinho, R.A., and Cantley, L.C. (2004b). The tumor suppressor LKB1 kinase directly activates AMP-activated kinase and regulates apoptosis in response to energy stress. Proc. Natl. Acad. Sci. U.S.A. *101*, 3329–3335.

Shedden, K., Taylor, J.M.G., Enkemann, S.A., Tsao, M.-S., Yeatman, T.J., Gerald, W.L., Eschrich, S., Jurisica, I., Giordano, T.J., Misek, D.E., et al. (2008). Gene expression–based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. Nat. Med. *14*, 822–827.

Shibata, T., Ohta, T., Tong, K.I., Kokubu, A., Odogawa, R., Tsuta, K., Asamura, H., Yamamoto, M., and Hirohashi, S. (2008). Cancer related mutations in NRF2 impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy. Proc. Natl. Acad. Sci. U.S.A. *105*, 13568–13573.

Shiloh, Y. (2003). ATM and related protein kinases: safeguarding genome integrity. Nat. Rev. Cancer *3*, 155–168.

Singh, A., Boldin-Adamsky, S., Thimmulappa, R.K., Rath, S.K., Ashush, H., Coulter, J., Blackford, A., Goodman, S.N., Bunz, F., Watson, W.H., et al. (2008). RNAi-Mediated Silencing of Nuclear Factor Erythroid-2-Related Factor 2 Gene Expression in Non-Small Cell Lung Cancer Inhibits Tumor Growth and Increases Efficacy of Chemotherapy. Cancer Res. *68*, 7975–7984.

Singh, A., Misra, V., Thimmulappa, R.K., Lee, H., Ames, S., Hoque, M.O., Herman, J.G., Baylin, S.B., Sidransky, D., Gabrielson, E., et al. (2006). Dysfunctional KEAP1–NRF2 Interaction in Non-Small-Cell Lung Cancer. Plos Med. *3*, e420.

Slamon, D.J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M., et al. (2001). Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N. Engl. J. Med. *344*, 783–792.

Snyder, E.L., Watanabe, H., Magendantz, M., Hoersch, S., Chen, T.A., Wang, D.G., Crowley, D., Whittaker, C.A., Meyerson, M., Kimura, S., et al. (2013). Nkx2-1 Represses a Latent Gastric Differentiation Program in Lung Adenocarcinoma. Mol. Cell *50*, 185–199.

Solis, L.M., Behrens, C., Dong, W., Suraokar, M., Ozburn, N.C., Moran, C.A., Corvalan, A.H., Biswal, S., Swisher, S.G., Bekele, B.N., et al. (2010). Nrf2 and Keap1 Abnormalities in Non-Small Cell Lung Carcinoma and Association with

Clinicopathologic Features. Clin. Cancer Res. *16*, 3743–3753.

Sosman, J.A., Kim, K.B., Schuchter, L., Gonzalez, R., Pavlick, A.C., Weber, J.S., McArthur, G.A., Hutson, T.E., Moschos, S.J., Flaherty, K.T., et al. (2012). Survival in BRAF V600-mutant advanced melanoma treated with vemurafenib. N. Engl. J. Med. *366*, 707–714.

Spigelman, A.D., Murday, V., and Phillips, R.K. (1989). Cancer and the Peutz-Jeghers syndrome. Gut *30*, 1588–1590.

Spoerke, J.M., O'Brien, C., Huw, L., Koeppen, H., Fridlyand, J., Brachmann, R.K., Haverty, P.M., Pandita, A., Mohan, S., Sampath, D., et al. (2012). Phosphoinositide 3-Kinase (PI3K) Pathway Alterations Are Associated with Histologic Subtypes and Are Predictive of Sensitivity to PI3K Inhibitors in Lung Cancer Preclinical Models. Clin. Cancer Res. *18*, 6771–6783.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., and Lander, E.S. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U.S.A. *102*, 15545–15550.

Sun, Y., Ren, Y., Fang, Z., Li, C., Fang, R., Gao, B., Han, X., Tian, W., Pao, W., Chen, H., et al. (2010). Lung Adenocarcinoma From East Asian Never-Smokers Is a Disease Largely Defined by Targetable Oncogenic Mutant Kinases. J. Clin. Oncol. *28*, 4616–4620.

Sun, Y., Connors, K.E., and Yang, D.-Q. (2007). AICAR induces phosphorylation of AMPK in an ATM-dependent, LKB1-independent manner. Mol. Cell Biochem. *306*, 239–245.

Suzuki, A., Kusakai, G.-I., Kishimoto, A., Shimojo, Y., Ogura, T., Lavin, M.F., and Esumi, H. (2004). IGF-1 phosphorylates AMPK-α subunit in ATM-dependent and LKB1-independent manner. Biochem. and Biophys. Res. Comm. *324*, 986–992.

Suzuki, Y. (2012). LKB1, TP16, EGFR, and KRAS somatic mutations in lung adenocarcinomas from a Chiba Prefecture, Japan cohort. Drug Discov. Ther.

Taguchi, K., Motohashi, H., and Yamamoto, M. (2011). Molecular mechanisms of the Keap1-Nrf2 pathway in stress response and cancer evolution. Genes to Cells *16*, 123–140.

Tan, B.-X., Yao, W.-X., Ge, J., Peng, X.-C., Du, X.-B., Zhang, R., Yao, B., Xie, K., Li, L.-H., Dong, H., et al. (2011). Prognostic influence of metformin as first-line chemotherapy for advanced nonsmall cell lung cancer in patients with type 2 diabetes. Cancer *117*, 5103–5111.

Tanaka, K., Babic, I., Nathanson, D., Akhavan, D., Guo, D., Gini, B., Dang, J., Zhu, S., Yang, H., De Jesus, J., et al. (2011). Oncogenic EGFR Signaling Activates an mTORC2-

NF- B Pathway That Promotes Chemotherapy Resistance. Cancer Discov. *1*, 524–538.

Tang, Y., Shu, G., Yuan, X., Jing, N., and Song, J. (2010). FOXA2 functions as a suppressor of tumor metastasis by inhibition of epithelial-to-mesenchymal transition in human lung cancers. Cell Res. *21*, 316–326.

Tenbaum, S.P., Ordóñez-Morán, P., Puig, I., Chicote, I., Arqués, O., Landolfi, S., Fernández, Y., Herance, J.R., Gispert, J.D., Mendizabal, L., et al. (2012). β-catenin confers resistance to PI3K and AKT inhibitors and subverts FOXO3a to promote metastasis in colon cancer. Nat. Med. *18*, 892–901.

Udd, L., Katajisto, P., Rossi, D.J., Lepistö, A., Lahesmaa, A.M., Ylikorkala, A., Järvinen, H.J., Ristimäki, A.P., and Mäkelä, T.P. (2004). Suppression of Peutz—Jeghers polyposis by inhibition of cyclooxygenase-2. Gastroenterol. *127*, 1030–1037.

Vaahtomeri, K., Ventela, E., Laajanen, K., Katajisto, P., Wipff, P.J., Hinz, B., Vallenius, T., Tiainen, M., and Makela, T.P. (2008). Lkb1 is required for TGF -mediated myofibroblast differentiation. J. of Cell Sci. *121*, 3531–3540.

Vivanco, I., and Sawyers, C.L. (2002). The phosphatidylinositol 3-Kinase–AKT pathway in human cancer. Nat. Rev. Cancer *2*, 489–501.

Walkinshaw, D.R., Weist, R., Kim, G.W., You, L., Xiao, L., Nie, J., Li, C.S., Zhao, S., Xu, M., and Yang, X.J. (2013). The Tumor Suppressor Kinase LKB1 Activates the Downstream Kinases SIK2 and SIK3 to Stimulate Nuclear Export of Class IIa Histone Deacetylases. J. of Biol. Chem. *288*, 9345–9362.

Wan, H. (2004). Foxa2 regulates alveolarization and goblet cell hyperplasia. Development *131*, 953–964.

Wei, C., Amos, C.I., Rashid, A., Sabripour, M., Nations, L., McGarrity, T.J., and Frazier, M.L. (2003). Correlation of staining for LKB1 and COX-2 in hamartomatous polyps and carcinomas from patients with Peutz-Jeghers syndrome. J. Histochem. Cytochem. *51*, 1665–1672.

Wei, C., Amos, C.I., Stephens, L.C., Campos, I., Deng, J.M., Behringer, R.R., Rashid, A., and Frazier, M.L. (2005). Mutation of Lkb1 and p53 genes exert a cooperative effect on tumorigenesis. Cancer Res. *65*, 11297–11303.

Weinberg, F., Hamanaka, R., Wheaton, W.W., Weinberg, S., Joseph, J., Lopez, M., Kalyanaraman, B., Mutlu, G.M., Budinger, G.R.S., and Chandel, N.S. (2010). Mitochondrial metabolism and ROS generation are essential for Kras-mediated tumorigenicity. Proc. Natl. Acad. Sci. U.S.A. *107*, 8788–8793.

Wilkerson, M.D., Yin, X., Walter, V., Zhao, N., Cabanski, C.R., Hayward, M.C., Miller, C.R., Socinski, M.A., Parsons, A.M., Thorne, L.B., et al. (2012). Differential pathogenesis of lung adenocarcinoma subtypes involving sequence mutations, copy number, chromosomal instability, and methylation. PLoS ONE *7*, e36530.

Wingo, S.N., Gallardo, T.D., Akbay, E.A., Liang, M.-C., Contreras, C.M., Boren, T., Shimamura, T., Miller, D.S., Sharpless, N.E., Bardeesy, N., et al. (2009). Somatic LKB1 mutations promote cervical cancer progression. PLoS ONE *4*, e5137.

Woods, A., Johnstone, S.R., Dickerson, K., Leiper, F.C., Fryer, L.G.D., Neumann, D., Schlattner, U., Wallimann, T., Carlson, M., and Carling, D. (2003). LKB1 Is the Upstream Kinase in the AMP-Activated Protein Kinase Cascade. Current Biol. *13*, 2004–2008.

Xiao, B., Heath, R., Saiu, P., Leiper, F.C., Leone, P., Jing, C., Walker, P.A., Haire, L., Eccleston, J.F., Davis, C.T., et al. (2007). Structural basis for AMP binding to mammalian AMP-activated protein kinase. Nature *449*, 496–500.

Xiao, B., Sanders, M.J., Underwood, E., Heath, R., Mayer, F.V., Carmena, D., Jing, C., Walker, P.A., Eccleston, J.F., Haire, L.F., et al. (2011). Structure of mammalian AMPK and its regulation by ADP. Nature *472*, 230–233.

Xie, Z., Dong, Y., Scholz, R., Neumann, D., and Zou, M.H. (2008). Phosphorylation of LKB1 at Serine 428 by Protein Kinase C- Is Required for Metformin-Enhanced Activation of the AMP-Activated Protein Kinase in Endothelial Cells. Circulation *117*, 952–962.

Xie, Z., Dong, Y., Zhang, J., Scholz, R., Neumann, D., and Zou, M.H. (2009). Identification of the Serine 307 of LKB1 as a Novel Phosphorylation Site Essential for Its Nucleocytoplasmic Transport and Endothelial Cell Angiogenesis. Mol. Cell. Biol. *29*, 3582–3596.

Yang Q, Inoki K, Kim E, Guan KL. (2006). TSC1/TSC2 and Rheb have different effects on TORC1 and TORC2 activity. Proc. Natl. Acad. Sci. U. S. A. *103*, 6811-6.

Yang, J.-Y., Chang, C.-J., Xia, W., Wang, Y., Wong, K.-K., Engelman, J.A., Du, Y., Andreeff, M., Hortobagyi, G.N., and Hung, M.-C. (2010). Activation of FOXO3a is sufficient to reverse mitogen-activated protein/extracellular signal-regulated kinase kinase inhibitor chemoresistance in human cancer. Cancer Res. *70*, 4709–4718.

Yang, J.-Y., Zong, C.S., Xia, W., Yamaguchi, H., Ding, Q., Xie, X., Lang, J.-Y., Lai, C.-C., Chang, C.-J., Huang, W.-C., et al. (2008). ERK promotes tumorigenesis by inhibiting FOXO3a via MDM2-mediated degradation. Nat. Cell Biol. *10*, 138–148.

Ylikorkala, A. (2001). Vascular Abnormalities and Deregulation of VEGF in Lkb1-Deficient Mice. Science *293*, 1323–1326.

Yokoyama, Y., Iguchi, K., Usui, S., and Hirano, K. (2011). AMP-activated protein kinase modulates the gene expression of aquaporin 9 via forkhead box a2. Arch. Biochem. Biophys. *515*, 80–88.

Zeqiraj, E., Filippi, B.M., Deak, M., Alessi, D.R., and van Aalten, D.M.F. (2009). Structure of the LKB1-STRAD-MO25 Complex Reveals an Allosteric Mechanism of

Kinase Activation. Science *326*, 1707–1711.

Zhang, X., Tang, N., Hadden, T.J., and Rishi, A.K. (2011). Akt, FoxO and regulation of apoptosis. BBA - Mol. Cell Res. *1813*, 1978–1986.

Zhang, X., Odom, D.T., Koo, S.-H., Conkright, M.D., Canettieri, G., Best, J., Chen, H., Jenner, R., Herbolsheimer, E., Jacobsen, E., et al. (2005). Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. Proc. Natl. Acad. Sci. U.S.A. *102*, 4459–4464.

Zhong, D., Liu, X., Khuri, F.R., Sun, S.Y., Vertino, P.M., and Zhou, W. (2008). LKB1 Is Necessary for Akt-Mediated Phosphorylation of Proapoptotic Proteins. Cancer Res. *68*, 7270–7277.