Co-occurring Genomic Alterations Define Major Subsets of KRAS-Mutant Lung Adenocarcinoma with Distinct Biology, Immune Profiles, and Therapeutic Vulnerabilities

Ferdinandos Skoulidis1, Lauren A. Byers1, Lixia Diao2, Vassiliki A. Papadimitrakopoulou1, Pan Tong2, Julie Izzo3, Carmen Behrens1, Humam Kadara3, Edwin R. Parra3, Jaime Rodriguez Canales3, Jianjun Zhang4, Uma Giri4, Jayanthi Gudikote1, Maria A. Cortez5, Chao Yang1, Youhong Fan1, Michael Peyton5, Luc Girard6, Kevin R. Coombes7, Carlo Toniatti8, Timothy P. Heffernan8, Murim Choi9, Garrett M. Frampton10, Vincent Miller10, John N. Weinstein2, Roy S. Herbst11, Kwok-Kin Wong12, Jianhua Zhang8, Padmanee Sharma13, Gordon B. Mills14, Waun K. Hong15, John D. Minna6, James P. Allison16, Andrew Futreal4, Jing Wang2, Ignacio I. Wistuba3, and John V. Heymach1
ABSTRACT
The molecular underpinnings that drive the heterogeneity of KRAS-mutant lung adenocarcinoma are poorly characterized. We performed an integrative analysis of genomic, transcriptomic, and proteomic data from early-stage and chemorefractory lung adenocarcinoma and identified three robust subsets of KRAS-mutant lung adenocarcinoma dominated, respectively, by co-occurring genetic events in STK11/LKB1 (the KL subgroup), TP53 (KP), and CDKN2A/B inactivation coupled with low expression of the NKX2-1 (TTF1) transcription factor (KC). We further revealed biologically and therapeutically relevant differences between the subgroups. KC tumors frequently exhibited mucinous histology and suppressed mTORC1 signaling. KL tumors had high rates of KEAP1 mutational inactivation and expressed lower levels of immune markers, including PD-L1. KP tumors demonstrated higher levels of somatic mutations, inflammatory markers, immune checkpoint effector molecules, and improved relapse-free survival. Differences in drug sensitivity patterns were also observed; notably, KL cells showed increased vulnerability to HSP90-inhibitor therapy. This work provides evidence that co-occurring genomic alterations identify subgroups of KRAS-mutant lung adenocarcinoma with distinct biology and therapeutic vulnerabilities.

SIGNIFICANCE: Co-occurring genetic alterations in STK11/LKB1, TP53, and CDKN2A/B—–the latter coupled with low TTF1 expression—define three major subgroups of KRAS-mutant lung adenocarcinoma with distinct biology, patterns of immune-system engagement, and therapeutic vulnerabilities.

INTRODUCTION
The identification of subsets of lung adenocarcinoma with oncogenic drivers has transformed the treatment of non–small cell lung cancer (NSCLC), particularly for patients whose tumors harbor activating mutations in EGFR or oncogenic fusions involving the ALK, RET, and ROS1 kinases (1, 2). Despite these advances, the goal of developing specific therapeutic strategies for the 25% to 30% of lung adenocarcinomas that bear activating mutations in KRAS, the most common oncogenic driver in NSCLC, has thus far proven elusive. Both efforts to directly target RAS oncoproteins with small molecules as well as alternative approaches focused on inhibiting RAS posttranslational modifications or downstream signaling pathways have been employed with only limited success; recently, however, some promising results have been reported (3–5).

A formidable challenge to the development of effective therapies for KRAS-mutant lung adenocarcinomas is heterogeneity in their biology and therapeutic responsiveness. Unlike NSCLC patients with tumors bearing EGFR activating mutations or ALK fusions, diseases for which targeted agents typically achieve objective responses in 60% to 80% of cases (1, 2, 6), clinical testing of agents targeting downstream pathways, such as MEK and PI3K–AKT, in patients with KRAS-mutant tumors has typically yielded response rates of less than 20% (7). This suggests that there is greater molecular diversity in KRAS-mutant tumors compared with those initiated by other known driver events. Preclinical studies provide further support for this notion, and in fact markers designed to predict the varying degrees of KRAS dependence have been developed (8). However, the underlying mechanisms that drive this divergent biologic and clinical behavior are not well understood. Previously, we demonstrated that different KRAS amino acid substitutions—particularly common hydrophobic alterations such as G12C and G12V compared with hydrophilic alterations such as G12D—differ in their patterns of downstream signaling and response to targeted agents, suggesting that specific KRAS alleles may account for at least part of this heterogeneity (9). Other reports suggested that specific KRAS codon-12 alleles are not predictive...
of response to adjuvant therapy (10). Preclinical studies have indicated that 
Kras-mutant lung tumors bearing Tap53 or 
Stk11/Lkb1 co-mutations differ in their response to docetaxel 
with or without selumetinib, suggesting that co-mutations 
may also affect treatment responsiveness (11).

Here, in order to systematically address the heterogeneity 
exhibited by KRAS-mutant lung adenocarcinomas, we 
describe an integrative analysis that incorporates transcriptional, 
mutational, copy-number, and proteomic data from 
cohorts of both chemotherapy-naive and heavily pretreated 
KRAS-mutant tumors and identify three distinct and robust 
disease subsets. Significantly, we uncover co-occurring genetic 
events as major determinants of signaling diversification 
downstream of mutant KRAS and highlight subtype-selective 
dependencies that can be exploited therapeutically with 
agents currently undergoing clinical development.

RESULTS

Unsupervised NMF Clustering Identifies Three Robust and Reproducible Subsets of KRAS-Mutant Lung Adenocarcinoma

Expression profiling can capture the heterogeneous behavior of complex biologic systems and has been successfully applied for the molecular stratification of human tumors. We thus initially interrogated RNA sequencing (RNA-Seq) expression data from a training set of 68 KRAS-mutant lung adenocarcinomas from The Cancer Genome Atlas (TCGA; ref. 12). In order to identify naturally occurring biologic patterns, we employed non-negative matrix factorization (NMF)—an unsupervised approach—followed by consensus clustering, an iterative process that assesses the stability of partitioning over a number of clustering runs (13, 14). Application of consensus NMF revealed three to five robust clusters of KRAS-mutant lung adenocarcinoma with cluster stability peaking for k = 3, as evidenced by the maximal value of the cophenetic correlation coefficient, a quantitative measure of consensus matrix dispersion (Fig. 1A and B). A heatmap depicting expression levels of 384 genes selected for the NMF algorithm is shown in Fig. 1C. Potential functional modules within the 384 genes are highlighted in Supplementary Fig. S1.

We next sought to validate our expression-based classification using independent datasets. In order to enable class prediction, we first derived an 18-gene signature consisting of 6 genes whose expression correlated with each subtype in the TCGA cohort using a nearest centroid-based classifier (ref. 15; Fig. 1D; see also Supplementary Methods). For validation, we assembled two clinically distinct collections of KRAS-mutant lung adenocarcinomas with available mRNA expression data: (i) a combined surgical set of 88 chemotherapynaive, mostly early-stage or locally advanced tumors, consisting of 41 tumors prospectively collected at The University of Texas MD Anderson Cancer Center (PROSPECT dataset) and 47 previously reported lung adenocarcinomas (16), and (ii) 36 metastatic, platinum-refractory KRAS-mutant lung adenocarcinomas from the BATTLE-2 clinical trial with tumor re-biopsy performed at trial entry (17). Critically, the three expression clusters retained similar representation in the training and validation datasets (Fig. 1E). Cluster composition also remained stable as a factor of increasing disease stage (Fig. 1F).

Co-occurring Genetic Events Are Major Determinants of the Biologic Heterogeneity of KRAS-Mutant Lung Adenocarcinoma

We sought to determine whether alterations in the tumor genome could account for the robustness and reproducibility of cluster assignment. Previously, we have reported that individual KRAS codon 12 mutations that result in distinct amino acid substitutions differentially engage downstream effector pathways (9). We compared the distribution of the three most common mutant KRAS alleles [KRAS\(^{G12C}\), KRAS\(^{G12V}\), KRAS\(^{G12D}\) and all other grouped together] in the expression clusters but found no evidence for enrichment of specific amino acid substitutions (P = 0.3, Fisher exact test). This suggests that specific KRAS amino acid substitutions are not primary drivers of the molecular diversity of KRAS-mutant lung adenocarcinomas.

Next, we investigated whether somatic mutations or other genomic alterations outside of KRAS itself are dominant determinants of expression cluster membership. In order to formally address this possibility, we identified prevalent somatic mutations [present in \(\geq 14.7\%\) (10/68) of evaluated tumors] that were significantly enriched in the three cohorts using an FDR of 0.05 as statistical cutoff. This analysis yielded 11 genes with nonsilent somatic mutational events (Supplementary Fig. S2).

Co-mutations in TP53 (P = 3.8e\(^{-10}\)) and STK11/LKB1 (P = 1.03 e\(^{-10}\)) were the most significantly enriched genetic events in cluster 3 and cluster 2 (henceforth referred to as the KP and KL clusters), respectively (Fig. 2A) and were largely nonoverlapping in the context of chemotherapy-naive disease. We confirmed this in a second cohort of KRAS-mutant tumors from TCGA as well as in a larger merged dataset of 176 previously untreated (mostly early-stage) KRAS-mutant lung adenocarcinomas with available somatic mutation data, where only 4% of KRAS-mutant tumors harbored co-mutations in both TP53 and STK11 (Supplementary Fig. S2; refs. 16, 18–20).

The occurrence of triple-mutant tumors (KRAS;STK11;TP53, referred to as KPL\(^{+}\) in order to distinguish from the expression-based clusters) was significantly less frequent than expected by chance in both the original TCGA set (n = 68; P = 0.0018, permutation test based on 10,000 permutations) and the second set of 77 KRAS-mutant TCGA lung adenocarcinomas analyzed since (P = 0.01692, permutation test based on 10,000 permutations), and of borderline significance in the merged validation cohort (n = 176; P = 0.0693, permutation test based on 10,000 permutations). Other prominent genes with nonrandom mutation patterns across the three subsets included two that were enriched in the KL subgroup: ATM (P = 0.002), a gene encoding an apical kinase in the DNA damage response, and Kelch-Like ECH-Associated Protein 1 (KEAP1; P = 0.006), a gene encoding an adaptor protein that functions as a negative regulator of Nuclear Factor, Erythroid 2-Like 2 (NRF2; also known as NFE2L2)-mediated transcription (Fig. 2A). For a list of individual mutations in KRAS, STK11/LKB1, TP53, ATM, and KEAP1, see Supplementary Table S1.

Interestingly, the mean overall number of nonsilent somatic mutations per tumor also differed among the three subgroups, with KP tumors harboring a significantly higher overall mutational load (Fig. 2B, left) despite comparable exposure to smoking (measured in pack years; Fig. 2B, right), suggesting increased mutation tolerance or genomic instability in this subset.
Co-occurring Genomic Alterations Define KRAS Subgroups

Marked genomic copy-number aberrations can alter gene function and are considered functionally relevant genetic events. We thus computed the distribution of focal biallelic deletions and high copy-number gains across the genome in the three clusters using GISTIC2.0-derived deletion and amplification peaks generated from Affymetrix SNP 6.0 array hybridization data (21). This analysis revealed biallelic deletions of CDKN2A (encoding for the p16 tumor suppressor) and CDKN2B (encoding for the p15 tumor suppressor), both located at 9p21.3, as significantly enriched events in the first cluster (henceforth referred to as KC; $P = 0.004$ and $P = 0.002$, respectively, Fisher exact test; Fig. 2A; and Supplementary

Figure 1. Consensus NMF clustering identifies three robust and reproducible subsets of KRAS-mutant lung adenocarcinomas. **A**, consensus matrices of 68 KRAS-mutant lung adenocarcinomas from the TCGA dataset, computed for $k = 2$ to $k = 7$. **B**, cophenetic correlation coefficient plot reveals peak cluster stability for $k = 3$ ranks. **C**, heatmap depicting mRNA expression levels of 384 genes selected for the NMF algorithm. **D**, relative expression levels of individual genes that comprise the 18-gene cluster assignment signature in lung adenocarcinomas from the TCGA dataset. **E**, subgroup representation is unaffected by increasing disease stage. Stage 4 platinum-refractory tumors in this analysis represent the BATTLE-2 clinical cohort.

TCGA ($n = 68$)

PROSPECT/CHITALE ($n = 88$)

BATTLE-2 ($n = 36$)

Stage 1 ($n = 81$)

Stage 2/3 ($n = 65$)

Stage 4 Platinum-refractory ($n = 36$)
Fig. S3). It is worth noting, however, that many tumors in the KC cluster did not have observable copy-number alterations in these genes. When combined with somatic mutations, alterations in \textit{CDKN2A} were present in 7 of 15 tumors in the KC cluster compared with 1 of 23 and 5 of 30 tumors in the KL and KP clusters, respectively. Among other components of the G\textsubscript{1}–S transition machinery, somatic mutations in \textit{RB1} were more frequent in the KL cluster (\(P = 0.044\)), although this was based on a small number of events (\(n = 3\); Supplementary Fig. S4). Inclusion of both somatic mutations and biallelic deletions in \textit{STK11, TP53, KEAP1, ATM,} and \textit{CDKN2A} in the analysis of co-occurring genetic events did not alter the statistical conclusions (\textit{STK11}: \(P = 0.000049\), \textit{TP53}: \(P = 0.0000038\), \textit{KEAP1}: \(P = 0.01258\), \textit{ATM}: \(P = 0.0015\), \textit{CDKN2A}: \(P = 0.00648\); Fisher exact test).

A direct corollary of the conjecture that key co-mutations dictate signaling diversification downstream of mutant \textit{KRAS} is the prediction that these genetic events are likely to be clonal.

**Figure 2.** Co-occurring genetic events in pivotal tumor suppressor genes are differentially represented in the three \textit{KRAS}-mutant lung adenocarcinoma subgroups. \textbf{A}, co-mutation plot for 68, predominantly early-stage, \textit{KRAS}-mutant lung adenocarcinomas from the TCGA dataset. \textbf{B}, comparison of overall nonsynonymous somatic mutation rate (left) and cumulative exposure to smoking (expressed in pack years; right plot) in the KC, KL, and KP subgroups. ANOVA was used for the three-group comparison, and Tukey post-test was applied to all pair-wise comparisons. Asterisks denote statistical significance at \(P \leq 0.05\). NS, not significant. \textbf{C}, co-mutation plot for 36 metastatic, platinum-refractory, \textit{KRAS}-mutant lung adenocarcinomas with available somatic mutation data from the BATTLE-2 trial.
Co-occurring Genomic Alterations Define KRAS Subgroups

in nature and occur early during the process of branched tumor evolution. Indeed, clonality analysis utilizing an algorithm that incorporates variant allele counts with estimates of sample purity and tumor ploidy concluded that the overwhelming majority of reported somatic mutation events in KRAS (98.53%), STK11/LKB1 (100%), TP53 (95.24%), and KEAP1 (100%) in the TCGA dataset were likely to be clonal (Supplementary Fig. S5; ref. 22). In contrast, ATM mutations were clonal in only 81.82% of mutant lung adenocarcinomas. Overall, 67.64% of all identified somatic mutations in this analysis were reported to be clonal, in agreement with recent work from our group and others utilizing deep multiregion sequencing of localized, chemotherapy-naive lung adenocarcinomas (23, 24).

We then tested these findings using other datasets and confirmed that the expression cluster enrichment in STK11/LKB1 and TP53 co-mutations was recapitulated in the two independent validation cohorts. Among 66 KRAS-mutant lung adenocarcinomas with available somatic mutation data from the combined Chirale and PROSPECT datasets, the distribution of KRAS, LKB1, KRAS, TP53, and KRAS/LKB1/TP53 genotypes differed significantly between the three expression clusters (P = 0.047, Fisher exact test), with LKB1 somatic mutations accumulating, as predicted, in the KL cluster (P = 0.036, Fisher exact test; Supplementary Fig. S6). Lack of copy-number data precluded evaluation of CDKN2A/CDKN2B inactivation in this cohort. With TCGA data, the distribution of the KRAS$^{G12D}$, KRAS$^{G12V}$, and KRAS$^{G12C}$ alleles was not different across the three subgroups in this independent set of lung adenocarcinomas (P = 0.338, Fisher exact test).

We next examined whether similar patterns were seen among tumors from platinum-refractory, metastatic lung adenocarcinoma, assessed as part of the BATTLE-2 trial (Fig. 2C). Consistent with the analysis from early-stage, treatment-naïve tumors, highly significant association between co-occurring genetic events in STK11/LKB1 and TP53 (incorporating both somatic mutations and biallelic deletions) and subtype allocation was also evident among 36 metastatic, platinum-refractory TNBC patients harboring tumors from the combined Chirale and PROSPECT datasets, the distribution of KRAS, LKB1, TP53, and KRAS/LKB1/TP53 genotypes differed significantly between the three expression clusters (P = 0.047, Fisher exact test), with LKB1 somatic mutations accumulating, as predicted, in the KL cluster (P = 0.036, Fisher exact test; Supplementary Fig. S6). Lack of copy-number data precluded evaluation of CDKN2A/CDKN2B inactivation in this cohort. With TCGA data, the distribution of the KRAS$^{G12D}$, KRAS$^{G12V}$, and KRAS$^{G12C}$ alleles was not different across the three subgroups in this independent set of lung adenocarcinomas (P = 0.338, Fisher exact test).

In agreement with a recently reported role for TTF1/NKX2-1 in restraining a latent gastric differentiation program in genetically engineered murine models (GEMM) of hemizygous or homozygous Nkx2-1 deletion in the context of mutant Kra$^{G12D}$, 6 of 15 KC lung adenocarcinomas were classified as invasive mucinous carcinomas. This histologic subtype was not detected in the KL and KP subgroups (with the exception of a single colloid KL tumor, P = 0.000125, Fisher exact test). Furthermore, expression of HNF4A (P = 3.17e–05, unpaired t test) and PDX1 (P = 6.85e–05, unpaired t test), two master regulators of gastrointestinal (GI) cell fate, was significantly elevated in tumors from the KC cluster (refs. 27, 28; data not shown). As expected, markers of mucinous differentiation, including CK20, MUC5B, and AGR2, were also upregulated (data not shown), and several
Figure 3. Multiplatform profiling identifies low expression of TTF1 (NKX2-1) as a defining feature of KRAS-mutant lung adenocarcinomas in the KC cluster. 

A, supervised hierarchical clustering of RPPA expression data demonstrates consistently suppressed levels of TTF1 protein in the KC cluster. 

B, quantitative analysis of TTF1 protein (top plots) and TTF1 mRNA (bottom plots) expression in the TCGA, PROSPECT, and BATTLE-2 cohorts. Statistical comparison between the three groups is based on ANOVA. 

C, IHC analysis of TTF1 expression in KRAS-mutant lung adenocarcinomas from PROSPECT. Representative images from tumors in the three clusters are shown in the left plot. Scale bar, 200 μm. The Kruskal-Wallis test was used for statistical comparison. Error bars, SD of the mean. 

D, dot plot representation of PI3K proteomic (RPPA) score in the three KRAS-mutant lung adenocarcinoma subsets. 

E, GSEA shows enrichment of gene expression signatures reflecting both upper and lower GI neoplastic processes as well as wild-type p53 transcriptional activity in the KC cluster. 

F, supervised hierarchical clustering reveals distinct patterns of miRNA expression in the three KRAS-mutant lung adenocarcinoma subsets.
Co-occurring Genomic Alterations Define KRAS Subgroups

were infrequently observed among KP tumors in either the TCGA or BATTLE-2 cohorts, as well as in a separate collection of 49 surgically resected KRAS-mutant lung adenocarcinomas reported previously (18), unless the tumor also harbored an **STK11/LKB1** mutation (i.e., triple-mutant **KRAS**/**STK11/LKB1** tumors). KEAP1 mutations and/or somatic copy-number losses (monoallelic or biallelic) were also present in the majority of **STK11/LKB1**-wild type KL tumors (Fig. 4E). The significant co-occurrence of **KEAP1** and **STK11/LKB1** copy-number losses is not surprising, as the corresponding genomic loci reside on the short arm of chromosome 19 (at 19p13.2 for **KEAP1** and 19p13.3 for **STK11**).

Notably, 4 of 15 tumors in the KC subgroup also harbored mutations in **KEAP1** (Fig. 2A). In two cases, these co-occurred with mutations in **STK11/LKB1**, whereas one additional tumor with **KEAP1** mutation exhibited suppressed levels of phospho-AMPK Thr172.

In keeping with the function of **KEAP1** as a regulator of the proteasomal degradation of **NRF2**, a transcription factor with a pivotal role in the cellular defense against oxidative stress and xenobiotics, an **NRF2** gene expression signature was significantly enriched in the KL and KC clusters compared with KP (Fig. 4F). Indeed, several prototypical **NRF2** target genes were among the most differentially expressed transcripts in both the TCGA and PROSPECT datasets (Fig. 4G). Thus, activation of an **NRF2**-driven antioxidant and cytoprotective transcriptional program is a common feature and may be selected for in the evolution of tumors with functional inactivation of the **LKB1–AMPK** axis.

Additional mechanisms of **NRF2** upregulation and/or stabilization may account for the prominent upregulation of **NRF2** target genes in LC lung adenocarcinomas given the relatively low frequency of **KEAP1** somatic mutations in this subgroup. For example, expression levels of miR-200a, a microRNA that has been previously reported to suppress **KEAP1** (31), differed significantly between the three clusters (P = 0.0065, ANOVA), with higher levels detected in LC lung adenocarcinomas.

In addition to antioxidant genes, several molecular chaperones and components of the cellular proteostasis network were also expressed at higher levels in the KL cluster (Fig. 4H). These involved cytoplasmic and endoplasmic reticulum (ER) chaperones, several members of the protein disulfide isomerase family, as well as **XBP1** and **ATF4**. LC lung loss-dependent upregulation of nodal components of the unfolded protein response (UPR) was further demonstrated in pairs of isogenic **LKB1**-deficient/proficient NSCLC cell lines (Supplementary Fig. S8). Accordingly, **XBP1** and **ATF4** expression signatures were significantly enriched in the KL cluster (Fig. 4I), which also exhibited evidence of altered bio-energetics, in agreement with previous reports (32).

### Inflammation and Active Immunoediting in **KP** Lung Adenocarcinomas Suggest Potential Avenues for Therapeutic Intervention

**GSEA** and Ingenuity Pathway Analysis (IPA) highlighted inflammation and several immune-related pathways as prominent modules in the KP cluster. Both gene sets associated with activation of antitumor immunity and immune tolerance/escape were enriched in this subset of **KRAS**-driven lung adenocarcinomas (Fig. 5A and B). In contrast, KL tumors demonstrated a comparative lack of immune system engagement, whereas KC tumors demonstrated a mixed picture. Notably,
Figure 4. KL tumors with functional inactivation of the LKB1–AMPK pathway display evidence of adaptation to oxidative, proteotoxic, and energetic stress. **A**, functional inactivation of the LKB1–AMPK axis in KL tumors. Box plot representation of (from left to right) LKB1 mRNA/protein and phospho-AMPK (Thr172) protein expression. Comparison is based on ANOVA. **B**, suppressed levels of LKB1 mRNA and protein are noted even among STK11/LKB1 wild-type (WT) tumors in the KL subgroup. The unpaired t test is used for comparison with KP lung adenocarcinomas. **C**, frequent single copy-number (CN) loss at the STK11/LKB1 locus among LKB1 somatic mutation-negative KL tumors. **D**, GSEA reveals enrichment of two STK11/LKB1-related signatures in the KL subgroup. **E**, frequent genetic abrogation of the KEAP1 locus in LKB1 wild-type KL lung adenocarcinomas. **F**, significant enrichment of an NRF2 (NFE2L2) expression signature in KL and KC tumors. **G**, heatmap display of relative mRNA expression levels of several prototypical NRF2 target genes in the TCGA and combined PROSPECT/Chitale datasets. **H**, higher expression of several cytoplasmic and ER chaperone proteins and core unfolded protein response components among KL lung adenocarcinomas. **P** values are based on an unpaired t test. FC, fold change. **I**, GSEA identifies altered cellular bio-energetics, activation of the unfolded protein response, and HIF1α pathway upregulation as prominent modules among lung adenocarcinomas in the KL cluster.
Co-occurring Genomic Alterations Define KRAS Subgroups

**Figure 5.** KRAS-mutant lung adenocarcinoma subsets exhibit distinct patterns of immune system engagement. A, IPA identifies several immune-related modules among the top ten upregulated pathways in the KP cluster. B, GSEA reveals prominent enrichment of signatures relating to inflammation, antitumor immunity, and immune tolerance/escape in the KP subgroup. C, heatmap representation of relative mRNA expression levels of selected targetable immune checkpoint mediator/effecter molecules. D, expression of CD274 (PD-L1) mRNA in KRAS-mutant lung adenocarcinoma subsets. ANOVA is used for statistical comparison between the three groups. E, comparison of PD-L1 H-score between KL and KP tumors in a tissue microarray from PROSPECT. A single confirmed triple-mutant tumor (KRAS;TP53;LKB1) in this cohort is excluded from the analysis. Wilcoxon rank-sum test is used for statistical comparison. Asterisk denotes statistical significance at \( P \leq 0.05 \). Error bars, SD of the mean. H-score, histoscore. F, representative images of CD274 immunostaining in lung adenocarcinomas from the KL (KL1/KL2) and KP (KP1/KP2) clusters. Scale bar, 200 μm.
several targetable mediators of cell-intrinsic coinhibitory signals were expressed at higher levels in the KP cluster, including PD-L1 (CD274; \( P = 0.001147 \), unpaired \( t \) test), PD-L2 (CD279; \( P = 0.01278 \), unpaired \( t \) test), and CTLA-4 (CD152; \( P = 0.0269 \), unpaired \( t \) test; Fig. 5C and D), and there was a trend toward more dense infiltration with CD3\(^+\), CD8\(^+\), and CD45RO\(^+\) populations of lymphocytes (Supplementary Fig. S9). PD-L1 mRNA expression was also higher among KP lung adenocarcinomas in the BATALL-2 cohort (\( P = 0.007 \), unpaired \( t \) test; Fig. 5D). Furthermore, we confirmed the association of KP with elevated PD-L1 expression in the PROSPECT dataset, using an IHC (histo)score (H-score; Fig. 5E and F).

**Prognostic Relevance of KRAS-Mutant Lung Adenocarcinoma Subgroups**

Increased lymphocytic infiltration has been associated with improved clinical outcome after surgical resection in colorectal cancer, which also harbors frequent mutations in **KRAS** (33). We therefore explored the prognostic utility of the **KRAS**-mutant lung adenocarcinoma subgroups using relapse-free survival (RFS) data from PROSPECT. This analysis revealed that KP tumors had significantly longer RFS compared with KL and KC lung adenocarcinomas combined (\( P = 0.029 \), log-rank test; Fig. 6A, left) and compared with the KC subgroup alone (\( P = 0.0269 \), log-rank test; Fig. 6A, right). Cluster membership remained a significant independent predictor of RFS on multivariate analysis [incorporating adjuvant therapy and nodal status; \( P = 0.03 \) for the KP vs. (KL+KC) comparison and \( P = 0.03 \) for the KP vs. KC comparison]. In contrast, overall survival (OS) did not differ between the subgroups in a larger merged dataset of **KRAS**-mutant tumors with available OS data (\( P = 0.3 \); Supplementary Fig. S10). Due to the limited numbers of tumors with available molecular and clinical data reported here, these results require further validation in larger, prospectively collected datasets.

An integrated heatmap encompassing the key genomic, transcriptional, and proteomic features of the three **KRAS**-mutant lung adenocarcinoma subgroups, as well as their relationship to the previously described TCGA and iClusters (32, 35), is presented in Fig. 6B. Corresponding heatmaps for a higher number of **KRAS**-mutant lung adenocarcinoma subgroups (\( k = 4 \) and \( k = 5 \)) based on the NMF algorithm are included in Supplementary Fig. S11.

**KRAS-Mutant Cell Lines with LKB1 Loss Show Enhanced Sensitivity to HSP90 Inhibition**

We posited that differences in the underlying biology of **KRAS**-driven lung adenocarcinomas with co-mutations in **STK11/LKB1** or **TP53** would result in distinct therapeutic vulnerabilities. In support of this hypothesis, reanalysis of publicly available large-scale cell line drug sensitivity data (Genomics of Drug Sensitivity in Cancer; ref. 34) for **KRAS**-mutant NSCLC cell lines on the basis of co-mutations in **STK11/LKB1** (KL\(^m\)/KPL\(^m\)) and **TP53** (KP\(^m\)) yielded divergent patterns of drug sensitivity (Fig. 7A).

17-AAG, an ansamycin benzoquinone ATP-competitive HSP90 inhibitor, emerged as the top hit in this analysis with enhanced activity observed in **STK11/LKB1** co-mutant (KL\(^m\) and KPL\(^m\)) lines. We therefore further assessed whether HSP90 inhibition represented a novel therapeutic vulnerability for cells bearing **KRAS** mutation and **LKB1** loss.

We first validated and extended these findings by assessing the sensitivity of an extended panel of 22 **KRAS**-mutant NSCLC cell lines with different **LKB1** status (10 KL\(^m\)/KPL\(^m\) and 12 KP\(^m\)) to three distinct HSP90 inhibitors (17-AAG, ganetespib, AUY-922) using a 72-hour luminescent cell viability assay (Cell-Titer Glo; Promega). Significantly, KL\(^m\)/KPL\(^m\) cell lines were more sensitive to ganetespib and 17-AAG and showed a trend toward increased sensitivity to AUY922 compared with KP\(^m\) lines (17-AAG: \( P = 0.0237 \), ganetespib: \( P = 0.0044 \), AUY922: \( P = 0.0523 \), Mann–Whitney test for all comparisons; Fig. 7B).

**LKB1** status–dependent sensitization to HSP90 inhibition was further assessed in pairs of isogenic cell lines (Fig. 7C). Expression of wild-type **LKB1** in A549 and H460 (two naturally **LKB1**-deficient cell lines) and knockdown of **LKB1** in Calu-6 (a KP\(^m\) line that expresses wild-type **LKB1** protein) resulted in statistically significant increases in IC\(_{50}\) values to ganetespib, AUY-922, and 17-AAG (with the single exception of 17-AAG in the A549 isogenic pair; Fig. 7D). Thus, sensitivity to HSP90 inhibition in these lines is at least partially mediated by **LKB1** inactivation.

Hypersensitivity to HSP90 inhibitors is usually attributed to destabilization of oncopgenic kinases or other HSP90 client proteins that are required for the survival or proliferation of malignant cells or, alternatively, may result from the induction of irresolvable levels of proteotoxic stress. In order to explain the heightened responsiveness of **KRAS;LKB1** co-mutant lines to HSP90 inhibition, we focused our attention on nodal components of signaling networks that are known to regulate cellular fitness in the context of **LKB1** inactivation. Following treatment of a panel of 8 **KRAS**-mutant cell lines with 150 nmol/L of ganetespib for 24 hours, we observed reduction in levels of c-MYC as well as carbonyl anhydrase 9, a prototypical hypoxia-inducible factor-1\(\alpha\) (HIF1\(\alpha\)) target gene (Fig. 7E, top). Both c-MYC and HIF1\(\alpha\) have been previously reported to underpin the metabolic adaptation of **LKB1**-deficient cells (36). In addition, treatment with ganetespib simultaneously suppressed mTORC1 signaling output (evidenced by low levels of phospho-S6, phospho-p70S6K, and phospho-4EBP1 protein), MAPK pathway activation (measured by levels of phospho-ERK1/2), and phosphorylated SRC kinase levels (Fig. 7E, top), thus emulating the therapeutic combination that has previously been demonstrated to induce regression of **Kras;Lkb1** co-mutant tumors in a GEMM model of NSCLC (36). Furthermore, CHK1, a checkpoint protein that recently emerged in a functional genetic screen as synthetic lethal in the context of **LKB1** inactivation (37), was markedly depleted in response to HSP90 inhibition (Fig. 7E, top). Finally, levels of IRE1A and PERK, two nodal components of the proteostasis network, were potently suppressed following HSP90 inhibition in all tested cell lines, in agreement with previous reports (Fig. 7E, bottom; ref. 38).

For 17-AAG and other chemically related benzoquinone ansamycin HSP90 inhibitors that are subject to NAD(P)H Dehydrogenase, Quinone 1 (NQO1)–mediated bioactivation, we further reasoned that robust NRF2-driven upregulation of NQO1 in **LKB1**-mutant lines—which frequently, like primary
tumors, harbor mutations in KEAP1—would represent an additional mechanism accounting for their hypersensitivity to this chemical class of inhibitors (39). As predicted, LKB1-mutant cell lines—and most notably those with concurrent mutations in KEAP1—expressed higher levels of NQO1 protein (Fig. 7F). Furthermore, concurrent treatment with dicumarol (50 μmol/L), a known inhibitor of NQO1 enzymatic activity, conferred resistance to 17-AAG, supporting a role for NQO1-mediated bioactivation in these cells. Finally, KL tumors expressed higher levels of NQO1 transcripts than KP tumors in two independent datasets (Fig. 7G), in keeping with widespread inactivation of KEAP1 and elevated basal NRF2 transcriptional activity.

**DISCUSSION**

The development of more effective treatment strategies for lung adenocarcinoma bearing activating mutations in KRAS is hampered by the biologic and phenotypic heterogeneity of KRAS-mutant tumors. Here, we implemented an integrated approach to the discovery of biologically distinct subgroups of KRAS-mutant lung adenocarcinoma. Using NMF, an algorithm based on “decomposition by parts,” we identified three major expression-based subgroups that are highly reproducible across diverse clinical cohorts of both early-stage, chemotherapy-naïve and metastatic, platinum-refractory tumors. We subsequently explored the molecular
Figure 7. KRAS co-mutations are associated with distinct therapeutic vulnerabilities. A, reanalysis of drug sensitivity data (34) for 19 KRAS-mutant NSCLC cell lines based on the presence of co-mutations in STK11/LKB1 and TP53, KRAS/LKB1 (KL−) and KRAS/TP53/LKB1 (KPL−) triple-mutant lines were grouped together (n = 9) for this analysis and compared with KRAS/TP53 (KP) lines with wild-type STK11/LKB1 status (n = 10). Drugs with Wilcoxon rank-sum test-derived P values ≤ 0.1 for the KL−/KPL− versus KP+ comparison are displayed, and fold change in mean IC50 values is plotted on the x-axis. Asterisk denotes statistical significance at the P ≤ 0.05 level; **, significance at the P ≤ 0.01 level. B, KRAS-mutant NSCLC cell lines with LKB1 inactivation show increased sensitivity to HSP90 inhibition. Scatter plots of median IC50 values (nmol/L; from three to five independent experiments) of 10 KL−/KPL− and 12 KP+ cell lines for three chemically distinct HSP90 inhibitors. Log10 IC50 (nmol/L) values from five to six independent experiments for each isogenic pair were compared using the unpaired Wilcoxon rank-sum test. *, significance at the P ≤ 0.05 level; **, significance at the P ≤ 0.01 level. Error bars, SD of the mean. C, Western blot analysis of LKB1 expression in three isogenic pairs of LKB1-deficient/proficient KRAS-mutant NSCLC cell lines. D, LKB1 status-dependent sensitization of KRAS-mutant NSCLC cell lines to HSP90 inhibitors. Log10 IC50 (nmol/L) values from five to six independent experiments for each isogenic pair were compared using the unpaired t test. *, significance at the P ≤ 0.05 level; **, significance at the P ≤ 0.01 level. Error bars, SD of the mean. E, ganetespib (GAN) simultaneously destabilizes multiple proteins that support the fitness of KRAS-mutant NSCLC cell lines with LKB1 inactivation. DMSO, dimethyl sulfoxide. F, NQO1-mediated bioactivation contributes to the sensitivity of NSCLC cell lines to 17-AAG. Top, robust expression of NQO1 among KRAS/TP53/LKB1 lines, which frequently harbor co-occurring mutations in KEAP1/K(P)LKm. KP lines display variable NQO1 expression (CALU-6 expresses NQO1 after prolonged exposure). The 2NQO1-polyomorphic HS96 cell line is used as negative control. Bottom, cotreatment with dicumarol, an NQO1-inhibitor, renders KRAS-mutant NSCLC partially resistant to 17-AAG. Unpaired t test is used for all statistical comparisons. Error bars, SD of the mean from two independent experiments. G, differential expression of NQO1 mRNA in the three KRAS-mutant lung adenocarcinoma subgroups in the TCGA (left) and PROSPECT (right) cohorts. ANOVA is used for statistical comparison.
underpinnings and potential therapeutic vulnerabilities of these subgroups.

Previously, we and others have demonstrated variable effector pathway engagement downstream of different common mutant KRAS alleles \((KRAS^{G12C}, KRAS^{G12V}, KRAS^{G12D}; ref\ 9)\). Based on these observations, we initially explored the possibility that specific alleles may be associated with the three subgroups but found generally no consistent association between KRAS alleles and the three expression clusters. We then considered other molecular features and discovered that co-occurring genetic events (nonsilent mutations and/or biallelic deletions) in distinct tumor suppressor genes were significantly enriched in the three subgroups. Specifically, the second and third expression clusters were dominated, respectively, by genetic alterations in \(STK11/LKB1\) and \(TP53\). Somatic events involving these genes were clonal and largely nonoverlapping among chemotherapy-naive lung adenocarcinomas initiated by oncogenic \(KRAS\), indicating that they likely represent gatekeepers to divergent pathways of tumor evolution. This finding is consistent with prior work in GEMMs of lung adenocarcinoma initiated by somatic activation of an endogenous \(Kras^{G12C}\) allele, whereby concomitant genetic inactivation of either \(Stk11/Lkb1\) or \(Trp53\) sufficed to promote fully invasive and widely metastatic lung adenocarcinoma with shortened latency and complete penetrance, but variable phenotypic characteristics \(40\). In contrast, in the platinum-refractory BATTLE-2 cohort, increased frequency of triple-mutant tumors \((KRAS,STK11;TP53–KPL)\) was observed. It is unclear whether this represents expansion of a preexisting subclone or accumulation of \(de novo\) mutations arising during therapy. We are currently investigating this issue using multiregion sequencing as we have recently described \(24\). It is notable, however, that all \(KRAS,STK11;TP53\) triple-mutant tumors in BATTLE-2 still clustered together in the KL expression subgroup, suggesting that transcriptional fingerprints of LKB1 inactivation are retained in the context of co-occurring mutations in \(TP53\).

Two additional frequently co-mutated genes, \(ATM\) and \(KEAP1\), showed a predilection for the KL cluster. Interestingly, somatic mutations in \(ATM\) and \(KEAP1\) were also largely nonoverlapping in the TCGA, Imielski, and BATTLE-2 tumor collections. The frequent co-occurrence of \(KEAP1\) mutations in KL tumors implies the application of positive selection pressure for upregulation of an NRF2-mediated antioxidant, cytoprotective, and anti-inflammatory transcriptional program. Increased generation of reactive oxygen species \(\text{(ROS)}\) from cells undergoing energetic stress, coupled with defective \(H_{2}O_{2}\) detoxification due to impaired NADPH maintenance, may provide the impetus for this adaptive response \(41,42\). However, it should be noted that an NRF2 expression signature was also enriched in KC lung adenocarcinomas, despite lower rates of \(KEAP1\) somatic mutation, indicating potential alternative mechanisms for NRF2 upregulation and/or stabilization in this subgroup.

\(ATM\) mutations were most frequently encountered among KL lung adenocarcinomas that displayed evidence of functional LKB1–AMPK pathway inactivation but retained wild-type \(STK11/LKB1\) genomic sequence. Activation of LKB1 through ATM-dependent phosphorylation at Thr366 in response to genotoxic or oxidative stress has been recently reported and may account for this predilection \(43\). Therefore, inactivating mutations in \(ATM\), in addition to their effects on the DNA damage signaling cascade, may also promote a functional LKB1-deficient state.

The upregulation of several ER chaperone proteins including \(GRP78\) (glucose-regulated protein—otherwise known as BiP) in KL lung adenocarcinomas indicates that adaptation to ER stress may constitute an additional homeostatic mechanism that contributes to cellular fitness in the context of LKB1 deficiency. Elevated rates of protein synthesis driven by constitutively active mTORC1 and high levels of intracellular ROS, coupled with hypoxia and energy/nutrient deprivation that are characteristic of \textit{in vivo} tumor growth, may, under conditions of impaired LKB1 function, promote increased dependency on ER and cytoplasmic chaperone proteins in order to mitigate the consequences of a high proteotoxic load. Activation of the UPR downstream of active mTORC1 has in previous been demonstrated in GEMMs of constitutive TSC1/TSC2 depletion and contributes to several phenotypic characteristics of the tuberous sclerosis clinical syndrome \(43\). In addition, several nodal components of the UPR, including GRP78 and IRE1a, have been reported to exert context-dependent tumor-promoting functions in various experimental models of cancer, and ATF4 can promote cytoprotective autophagy independently of AMPK and mTORC1 \(44\); this latter function may be particularly relevant in the context of \(KRAS,LKB1\) co-mutant cancer cells, which rely on efficient lysosomal degradation of macromolecules for supply of mitochondrial TCA-cycle substrates \(45\).

Unexpectedly, KP and KL tumors displayed distinct patterns of immune system engagement. KP tumors were characterized by an inflammatory response, and their expression profiles showed enrichment for signatures of Jak–STAT pathway activation and interferon signaling. There was further evidence of active immunoeediting, as indicated by robust expression of several costimulatory \(\text{(e.g., CD28)}\) and coinhibitory molecules, including PD-L1, which was further validated immunohistochemically in the PROSPECT dataset. Thus, intrinsic regulation of T-cell function may be crucial for the establishment of functional tolerance toward KP lung adenocarcinomas. Accordingly, these tumors may be particularly amenable to therapeutic strategies that incorporate immune checkpoint blockade with inhibitors of PD-L1, PD-1, CTLA-4, or other checkpoint mediators \(46\). Recent preclinical studies from our group and others are consistent with this possibility \(47\). In keeping with the premise that a high somatic mutational load fosters tumor immunogenicity by generating a rich repertoire of neo-antigens, KP lung adenocarcinomas displayed higher global mutation rates than KL tumors in the TCGA cohort, despite similar cumulative exposure to smoking.

Contrary to tumors with somatic \(TP53\) mutations, KL lung adenocarcinomas appeared largely “immune-inert.” The mechanism underlying this phenotype is not clear, although a lower rate of somatic mutations may partly account for this relative lack of immune system engagement. It is also notable that secreted lactic acid, a by-product of HIF1\(\alpha\)-mediated metabolic reprogramming of LKB1-deficient cells toward glycolysis \(32\), is a potent lymphotxin and may thus actively promote establishment of loco-regional immune privilege. In addition, anti-inflammatory signals emanating from NRF2 may further oppose and prevent mounting of a significant inflammatory response to developing tumors with perturbations in the LKB1–AMPK pathway. Detailed characterization...
of the immune contexture of developing and established KL and KP tumors using GEMMs and large collections of human specimens will likely yield further insights into oncogene-type-specific immunoediting mechanisms.

Integrative analysis identified low expression of the NKX2-1 (TTF1), lineage-specific homeodomain transcription factor as a defining feature among tumors in the KC cluster. In striking resemblance to the lung cancer phenotype of GEMMs expressing oncogenic KRAS\(^{G12D}\) in the context of hemizygous or homozygous $\text{NKx2-1}$ inactivation, both transdifferentiation toward a GI-like phenotype—mediated by expression of the prototypical gastrointestinal transcription factors HNF4A and PDX1—and dedifferentiation toward a more primitive state through derepression of the normally embryonically restricted chromatin regulator HMGAl2 were evident among tumors in this group (27, 28). Notably, 6 of 15 tumors in this subgroup were classified histologically as invasive mucinous carcinomas, a phenotype that was not observed in the KL and KP subgroups. These results are concordant with a separate study that reported a strong association between negative TTF1 IHC staining and mucinous histology in KRAS-mutant lung adenocarcinomas (48).

The KC subgroup was significantly enriched in tumors bearing biallelic loss of CDKN2A and/or CDKN2B. Consistent with this finding, concomitant genetic ablation of Cldn2a and Cldn2b in a murine model of $\text{Kras}$-mutant lung cancer resulted in a high percentage of TTF1-negative tumors that frequently overexpressed HMGAl2, suggesting that loss of this locus may directly affect the differentiation state of at least some lung adenocarcinomas in the KC subgroup (25). It is also noteworthy—and potentially clinically relevant—that KC tumors displayed evidence of reduced mTORC1 signaling output and expressed high basal levels of wild-type p53-regulated transcripts.

It should be noted that the KRAS-focused analysis presented here affords additional and complementary biologic insights into the three expression and six integrated clusters proposed by TCGA based on an unselected pool of lung adenocarcinomas. For example, the KC subgroup integrates tumors from all three TCGA expression clusters and includes most iCluster6 lung adenocarcinomas, an integrative cluster for which no distinctive features had hitherto been identified. In addition, iCluster2 incorporates most KL lung adenocarcinomas with $\text{ATM}$ mutations, and the majority of CIMP-H (CpG-island methylator phenotype-high; iCluster3 and iCluster4) KRAS-mutant tumors are classified as KP in our analysis. In contrast, lower levels of $\text{p16}$ (CDKN2A) promoter methylation in KL lung adenocarcinomas are consistent with an intact CDKN2A/CDKN2B genomic locus in the majority of tumors in this subgroup.

Comparison of RFS in the three KRAS-mutant lung adenocarcinoma subgroups in PROSPECT demonstrated improved RFS in KP compared with either KC alone or non-KP (KL and KC) tumors. This result is intriguing in view of the immune phenotype of KP lung adenocarcinomas and previous reports linking lymphocytic infiltration in surgically resected colorectal tumors with reduced likelihood of relapse (33). However, due to the limited number of patients reported here, definitive assessment of the prognostic utility of the proposed classification will require additional, adequately powered, prospective studies.

Finally, we provide evidence that the proposed subclassification of KRAS-mutant lung adenocarcinomas is therapeutically relevant. By coupling reanalysis of publicly available large-scale drug sensitivity data with targeted screening of an extensive panel of comprehensively annotated KRAS;LKB1 (KL\(^{-}\)/KPL\(^{-}\)) and KRAS;TP53 (KP\(^{-}\)) NSCLC cell lines, we uncover differential sensitivity to a number of drugs, including enhanced sensitivity of LKB1-mutant (KL\(^{-}\) or KPL\(^{-}\)) cells to several chemically distinct inhibitors of the HSP90 molecular chaperone. We further demonstrate that in response to ganetespib, an HSP90 inhibitor that is currently undergoing phase III clinical evaluation in combination with docetaxel in patients with metastatic lung adenocarcinoma following failure of first-line therapy, several signaling pathways and adaptive biologic responses with an established role in maintenance of cellular fitness in the setting of LKB1 inactivation are concomitantly destabilized. This enhanced sensitivity is therefore likely to be multifactorial and may be at least partly attributable to (i) degradation of HIF1\(\alpha\) and c-MYC, which underpin metabolic adaptation to LKB1 loss by promoting a shift toward glycolysis; (ii) depletion of CHK1, which may be necessary for the processing of stalled replication forks in the setting of depleted nucleoside pools; and (iii) inactivation of HSP90-dependent cytotoxic protective branches of the UPR that facilitate tolerance of a high proteotoxic load. Regardless of the mechanism, however, we find that HSP90 sensitivity can be directly affected by modulating LKB1 status \textit{in vitro}, thus supporting a direct role for LKB1. Furthermore, this analysis identifies NRF2-mediated antioxidant responses and UPR pathways as potential therapeutic targets for LKB1-deficient KRAS-mutant lung adenocarcinomas.

Taken together, our findings have several important implications for the molecular stratification and precise therapeutic targeting of lung adenocarcinomas initiated by oncogenic KRAS. First, STK11/LKB1 and TP53 appear to control distinct tumorigenic pathways downstream of mutant KRAS. Therefore, clinical trials enrolling patients with KRAS-mutant lung adenocarcinomas should take into account the co-mutation status of individual tumors. Second, it is important to recognize that patterns of co-occurring or mutually exclusive genetic events may evolve under the selective pressures imposed by cancer therapy—as is the case for somatic mutations in STK11/ LKB1 and TP53 that rarely co-occur in KRAS-mutant chemo-naive tumors but frequently do so following development of secondary resistance to cytotoxic agents. Thus, tumors initiated by the same apical oncogenic event may warrant different therapeutic approaches depending on their treatment history. Third, it is evident that even in the absence of identifiable genomic aberrations, a substantial number of lung adenocarcinomas exhibit functional abrogation of the LKB1-AMPK pathway; development of validated assays for the identification of these cases based on IHC detection of LKB1 or phospho-AMPK Thr172 protein expression (49), serum biomarkers, mRNA-based expression signatures (50), or even “surrogate” somatic mutations in KEAP1 or ATM is therefore a critical step toward appropriate patient stratification. Finally, our analysis highlights dependence on the molecular chaperone machinery as a therapeutic vulnerability in cancer cells with co-mutations in KRAS;LKB1 that can be tackled with HSP90 inhibitors; on the other hand, tolerance of a higher mutation load in tumors with concurrent mutations in TP53 and establishment of a prominent inflammatory response may render this subset particularly susceptible to immune checkpoint blockade and other novel immunotherapy approaches. Thus, our work
advances current understanding of KRAS-driven lung carcinogenesis, suggests a framework for the molecular classification of KRAS-mutant lung adenocarcinomas, and facilitates the implementation of personalized therapy.

**METHODS**

**Clinical Cohorts**

The TCGA, PROSPECT, Imaielinski and colleagues, Chitale and colleagues, and JBR.10 clinical cohorts of untreated, mostly early-stage lung adenocarcinomas have been previously reported (12, 16, 18–20, 51). BATTLE-2 enrolled patients with advanced NSCLC, following failure of at least one first-line chemotherapy regimen, with fresh tumor biopsy mandated before adaptive randomization. Forty-one KRAS-mutant lung adenocarcinomas from stage 1 of BATTLE-2 were included in the current study, with mRNA expression data available for 36 tumors. In all cases, biospecimens were obtained following patient informed consent, under protocols approved by Institutional Review Boards at all participating institutions. All human studies were conducted in accordance with the Declaration of Helsinki.

**Molecular Profiling**

Level 3 somatic mutation, copy-number (GISTIC 2.0), RNA-Seq, miRNA-seq, and RPPA data for TCGA lung adenocarcinomas were accessed through the TCGA portal. Expression profiling of PROSPECT and BATTLE-2 tumors was performed using the Illumina Human WG-6 v3 beadChip and Affymetrix GeneChip Human Gene 1.0 ST Array, respectively. Massively parallel sequencing exome capture of 19 KRAS-mutant lung adenocarcinomas performed by Foundation Medicine, as previously described (56). RPPA was undertaken on the Illumina HiSeq2000 platform as previously described (52). Sample preparation, basic alignment, and downstream analyses were performed according to previously described methods (53–55). Detailed methodology can be found in Supplementary Methods. Targeted sequencing of coding exons and selected introns of cancer-related genes from 41 KRAS-mutant BATTLE-2 tumors and copy-number analysis based on Affymetrix SNP 6.0 array profiling was performed by Foundation Medicine, as previously described (56). RPPA analysis was performed as previously reported (12, 57). The ABSOLUTE algorithm was employed for clonality analyses (22). Unsupervised NMF consensus clustering was performed using NMF Version 0.5.06 (R package; ref. 14). Development of the cluster assignment signature was based on the CnANC algorithm (15). Further details can be found in Supplementary Methods.

**mRNA Expression Profiling**

For lung adenocarcinomas included in the TCGA cohort, experimental procedures regarding RNA extraction from tumors, mRNA library preparation, sequencing (on the Illumina HiSeq platform), quality control, and subsequent data processing for quantification of gene expression have been previously reported (12). Details regarding sample collection, storage, selection, and RNA extraction for PROSPECT tumors have also been previously published (20). Array-based expression profiling of PROSPECT tumors was performed using the Illumina Human WG-6 v3 BeadChip, according to the manufacturer’s protocol. BATTLE-2 lung adenocarcinomas were profiled using the Illumina Human Gene 1.0 ST Array from Affymetrix. The tumors reported by Chitale and colleagues were profiled using the HG-U133A and HG-U133A 2.0 arrays from Affymetrix, as previously reported (16). Gene expression data for the PROSPECT cohort have been previously deposited in the GEO repository (GSE42127). Gene expression data for the BATTLE-2 tumors have been deposited in the GEO repository (GSE51913). Tumors from the JBR.10 trial were profiled using the U133A oligonucleotide microarray from Affymetrix as previously described (51) and were accessed through the GEO repository (GSE14814). For cross-platform integration of mRNA expression data, expression levels for each gene were derived by averaging corresponding probesets and standardized within each dataset. Missing values were replaced with 0.

**Immunohistochemistry**

Tissue sections (4 µm thick) were stained with the indicated antibodies (see Supplementary Table S2) using an automated staining system (Leica Bond Max; Leica Microsystems), according to standard protocols. The Aperio Image Analysis Toolbox (Aperio; Leica Microsystems) was used for digital analysis of images obtained from scanned slides, as described in Supplementary Methods. The H-score for TTF1 was determined manually, jointly by two thoracic pathologists, as previously described (58).

**NSCLC Cell Line Propagation and Authentication**

NSCLC cell lines were established by Dr. J.D. Minna and Dr. A. Gazdar at the NCI and The University of Texas Southwestern Medical Center (UTSW; Dallas, TX) or were obtained from the ATCC. They were maintained in RPMI-1640 (R8758; Sigma Life Science), supplemented with 10% heat-inactivated FBS (Gibco), and 1% penicillin/streptomycin (Sigma Life Science) in a humidified chamber at 5% CO2. All cell lines were authenticated between 2009 and 2011 using short tandem repeat (STR) profiling (PowerPlex 1.2; Promega) for at least eight different loci, and results were compared with reference STR profiles available through the ATCC or provided by Dr. Minna at UTSW. Following authentication, cell line stocks were frozen and maintained in liquid nitrogen until they were used in the reported experiments. All cell lines were Mycoplasma-tested before experiments.

**Western Blotting and qRT-PCR**

qRT-PCR and Western analysis were performed following standard protocols, as detailed under Supplementary Methods. Total protein lysate (35 µg) was transferred to polyvinylidene fluoride membranes and blotted with the indicated antibodies before development using picoLUCENT PLUS-HRP (G-Bioscience) or enhanced chemiluminescence (Amersham) detection reagents.

**Cell Viability Assay**

Cell viability was determined following 72-hour exposure of NSCLC cells (in triplicate wells) to seven serial fold dilutions of the indicated drugs using the Cell-Titer Glo luminescent assay (Promega) according to the manufacturer’s protocol, with minor modifications. For IC50 determination, multiple models were fitted and the best one selected based on residual standard error. Details of the procedures can be found in Supplementary Methods.

**Statistical Analyses**

Statistical analyses were conducted using GraphPad Prism version 6.0 for Windows (GraphPad Software) or the R system for statistical computing. All reported P values are two-tailed, and for all analyses, P ≤ 0.05 is considered statistically significant, unless otherwise specified. Box and whisker plots indicate median and interquartile range. Kaplan-Meier RFS and OS curves were compared using the log-rank test. Multivariate analysis was based on the Cox proportional hazards model and included adjuvant therapy and nodal status as additional covariates. Further details can be found in Supplementary Methods.

**Disclosure of Potential Conflicts of Interest**

G.M. Frampton has ownership interest (including patents) in Foundation Medicine. V. Miller is CMO of Foundation Medicine and has ownership interest (including patents) in Foundation Medicine. V. Miller is CMO of Foundation Medicine.
has ownership interest (including patents) in the same. P. Sharma has ownership interest (including patents) in Jounce Therapeutics and is a consultant/advisory board member for Amgen, Bristol-Myers-Squibb, and GlaxoSmithKline. J.D. Minna has received cell line licensing fees from the NIH and Genentech. J.P. Allison has ownership interest (including patents) in Bristol-Myers-Squibb and Jounce Therapeutics, and is a consultant/advisory board member for Jounce Therapeutics. J.V. Heymach reports receiving commercial research support from AstraZeneca and is a consultant/advisory board member for AstraZeneca, Boehringer Ingelheim, Genentech, GlaxoSmithKline, and Synata. No potential conflicts of interest were disclosed by the other authors.

**Authors’ Contributions**

Conception and design: F. Skoulidis, R.S. Herbst, W.K. Hong, J.D. Minna, J. Wang, J.V. Heymach


Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc): F. Skoulidis, L.A. Byers, V.A. Papadimitriakopoulou, C. Behrens, H. Kadara, J.R. Canales, J. Zhang, U. Giri, J. Gudikote, M.A. Cortez, C. Yang, Y. Fan, M. Peyton, T.P. Heffernan, G.M. Frampton, V. Miller, R.S. Herbst, J.D. Minna, I.I. Wistuba, J.V. Heymach


Administrative, technical, or material support (i.e., reporting or organizing data, provided facilities, etc): L.A. Byers, C. Behrens, H. Kadara, Kadara, E.R. Parra, J.R. Canales, J. Zhang, J. Gudikote, M. Choi, G.M. Frampton, V. Miller, R.S. Herbst, J. Zhang, G.B. Mills, A. Futreal, J. Wang, J.V. Heymach

**Acknowledgments**

The authors thank Emily Roarty, Ph.D., for critical review of the manuscript and editorial assistance, and Ming-Sound Tsao, M.D., Frances A. Shepherd, M.D., and Chang-Qi Zhu, M.D., Ph.D., for providing information on the KRAS mutation status of lung adenocarcinomas from the JBR.10 clinical trial.

**Grant Support**

This work was supported by The University of Texas Southwestern Medical Center and The University of Texas MD Anderson Cancer Center Lung SPORE grant S P50 CA070907; DoD PROSPECT grant W81XWH-07-1-0306; 1 R01 CA155196-01A1 (to V.A. Papadimitriakopoulou and R.S. Herbst); Lung Cancer Moon Shot Program; M.D. Anderson Immunotherapy Platform; National Institute of Health Cancer Center Support Grant (CA016672); 1R01 CA168484-01 (to J.V. Heymach); V Foundation Grant (to J.V. Heymach); the David Brutton Jr. Endowed Chair (to J.V. Heymach); and a Ford Petrin donation. F. Skoulidis was supported by the Jeanine T. Rainbolt Advanced Scholar Endowment, the Eric and Pat Bodin Lung Cancer Research Endowment, AstraZeneca, and Synata. No potential conflicts of interest were disclosed by the other authors.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 16, 2014; revised March 15, 2015; accepted April 15, 2015; published online First June 11, 2015.

**REFERENCES**

Co-occurring Genomic Alterations Define KRAS Subgroups


Co-occurring Genomic Alterations Define Major Subsets of KRAS-Mutant Lung Adenocarcinoma with Distinct Biology, Immune Profiles, and Therapeutic Vulnerabilities

Ferdinandos Skoulidis, Lauren A. Byers, Lixia Diao, et al.


Updated version Access the most recent version of this article at: doi:10.1158/2159-8290.CD-14-1236

Supplementary Material Access the most recent supplemental material at: http://cancerdiscovery.aacrjournals.org/content/suppl/2015/06/11/2159-8290.CD-14-1236.DC1

Cited articles This article cites 58 articles, 22 of which you can access for free at: http://cancerdiscovery.aacrjournals.org/content/5/8/860.full#ref-list-1

Citing articles This article has been cited by 29 HighWire-hosted articles. Access the articles at: http://cancerdiscovery.aacrjournals.org/content/5/8/860.full#related-urls

E-mail alerts Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.