Published OnlineFirst October 7, 2014; DOI: 10.1158/1078-0432.CCR-14-0519

We searched for drugs that could potentially target the significantly altered genes in the discovery cohort. When looking at the 22 significantly mutated genes and the 60 genes located on the 13 significantly altered CNV regions, eight of them can be targeted by known drugs (Supplementary Table S8). Among these eight genes, four can be targeted by drugs with known tumor-inhibiting activities. The oncogenes *EGFR*, *ERBB2*, *PIK3CA*, and *TERT* can be targeted by tailored drugs such as erlotinib (15), trastuzumab (31), NVP-BEZ235 (32), and GRN163L (33), respectively. These target drugs and their respective target genes are mostly present in different pathways (Supplementary Fig. S2). The expanded list of novel driver candidates discovered in this study suggests that additional investigations into these candidates might lead to the discovery of novel drug targets for the treatment of lung adenocarcinoma.

Association between RB pathway alterations and poor prognosis in early-stage lung adenocarcinoma patients

We performed pathway enrichment tests on the 22 significantly mutated genes and 60 genes within significantly altered CNVs. We detected 19 significantly enriched pathways (Supplementary Table S9), almost all of which are related to cancer, including non–small cell lung cancer (KEGG pathway, $P = 1.6 \times 10^{-7}$, modified Fisher exact test). This is another indication that these significantly mutated genes and CNVs are involved in lung adenocarcinoma occurrence or progression.

Next, we examined whether genomic alterations accumulated on the genes of known pathways could result in differential clinical outcomes (Fig. 2A–D). We found that genomic alternations in the RB pathway (Supplementary Fig. S3) are strongly associated with significantly shorter disease-free survival in stage I and II lung adenocarcinoma patients ($P = 1.4 \times 10^{-3}$ and $q = 0.031$, log-rank test; Fig. 2A). When we excluded early-stage lung adenocarcinoma patients who were treated with adjuvant chemotherapy from the analysis, we observed an even stronger association ($P = 6.7 \times 10^{-4}$ and $q = 0.026$, log-rank test; Supplementary Fig. S4A). On the other hand, this association was not observed in stage III or IV lung adenocarcinoma patients, regardless of including patients treated with adjuvant chemotherapy ($P = 0.11$ and $q = 0.21$, log-rank test; Fig. 2C) or not ($P = 0.28$ and $q = 0.25$, log-rank test; Supplementary Fig. S4C).

To rule out the possibility that this discrepancy originated from the small number of late-stage patients included in our discovery cohort, we repeated our analysis using the same number of randomly sampled early-stage patients and found that the observed association between RB pathway mutations and poor disease-free survival is specific to early-stage lung adenocarcinoma patients, regardless of including patients treated with adjuvant chemotherapy ($P = 0.025$, log-rank test; Supplementary Fig. S5) or not ($P = 0.020$, log-rank test; Supplementary Fig. S4E). Furthermore, our independent validation cohort of 77 patients with lung adenocarcinoma confirmed this association ($P = 0.017$ and 4.0×10^{-3} respectively, log-rank test; Fig. 2B and Supplementary Fig. S4B). When the discovery and validation cohorts were combined, the association was statistically significant ($P = 5.2 \times 10^{-5}$, log-rank test).

To verify that the association between RB pathway alterations and poor prognosis was not due to confounding effects (e.g., sex, age at diagnosis, node metastasis), we performed univariable and multivariable Cox regression analyses. Accordingly, we found that RB pathway alterations were significantly associated with poor prognosis in the discovery cohort, regardless of correcting for the confounding factors in the survival prediction model ($P = 4.6 \times 10^{-3}$, multivariable Cox regression analysis; Supplementary Table S10) or not ($P = 1.9 \times 10^{-3}$, univariable Cox regression analysis; Supplementary Table S11). These results were also reproduced in the validation cohort ($P = 9.4 \times 10^{-3}$ and 0.024 according to multivariable and univariable Cox regression analyses, respectively). In contrast with the current treatment guidelines for early-stage lung adenocarcinoma patients, which recommends follow-up without adjuvant therapy after complete resection except for high-risk patients (10), these results raise the interesting possibility that early-stage lung adenocarcinoma patients with genomic alterations in the RB pathway might benefit from additional medical interventions.

Table 2. Twenty-two significantly mutated genes ($q < 0.05$) expressed in lung adenocarcinoma

Gene symbol	RefSeq ID	q[a]	Prevalence[b]	Number of affected patients[c]	Remark[d]
EGFR	NM_005228	$<2.2 \times 10^{-16}$	42.4%	72	D, I, S
TP53	NM_000546	$<2.2 \times 10^{-16}$	25.3%	43	D, I, S
KRAS	NM_033360	3.9×10^{-9}	11.8%	20	D, I, S
PIK3CA	NM_006218	5.4×10^{-8}	8.2%	14	I, S
ACACB	NM_001093	9.7×10^{-6}	5.3%	9	
COL6A3	NM_004369	2.3×10^{-5}	4.1%	7	
FRAS1	NM_025074	8.6×10^{-5}	5.9%	10	
PRRC2C	NM_015172	1.7×10^{-4}	2.9%	5	
LAMA2	NM_000426	2.7×10^{-4}	3.5%	6	
SETD2	NM_014159	1.1×10^{-3}	5.3%	9	
PDE4DIP	NM_001198834	1.3×10^{-3}	4.7%	8	I, S
SYNE1	NM_182961	5.4×10^{-3}	6.5%	11	
CENPF	NM_016343	7.0×10^{-3}	4.7%	8	
HMCN1	NM_031935	9.8×10^{-3}	5.9%	10	
COL11A1	NM_080629	1.5×10^{-2}	5.3%	9	
RB1	NM_000321	3.6×10^{-2}	5.9%	10	D, I
COPA	NM_001098398	4.4×10^{-2}	4.1%	7	
LRBA	NM_006726	4.4×10^{-2}	3.5%	6	
F8	NM_000132	4.4×10^{-2}	2.9%	5	
SLIT2	NM_004787	4.4×10^{-2}	2.9%	5	
GCN1L1	NM_006836	4.5×10^{-2}	2.9%	5	
MAP1B	NM_005909	4.5×10^{-2}	2.9%	5	

[a]Calculated after taking into account the BMR, which varies with respect to genomic region and base type.
[b]The number of affected patients divided by the number of patients in the discovery cohort (170 patients).
[c]Indicates how many patients demonstrated substitutions and/or small indels on the gene.
[d]D, I, and S indicate previous studies that report candidate driver genes: D, Ding et al. (6); I, Imielinski et al. (8); S, Seo et al. (9).

RB1 inactivation alters expression of cell-cycle–related proteins in lung adenocarcinoma tumors

We evaluated whether *RB1* inactivation, resulting from mutations or CNV deletion, is associated with protein level changes of other genes related to cell-cycle regulation. Protein expression levels of pRB1, E2F1, cyclin D1, and cyclin E1 within 247 tumor samples were evaluated by IHC and compared for statistical differences (Fig. 3A). Tumors with inactivated *RB1* showed significantly lower pRB1 expression but higher E2F1 expression ($P = 3.2 \times 10^{-5}$ and 1.0×10^{-4}, respectively, rank-sum test; Fig. 3B), consistent with previous reports (34, 35). Higher E2F1 expression correlated with higher expression of cyclin E1 ($P = 5.5 \times 10^{-4}$; Fig. 3C). In addition, tumors with inactivated *RB1* exhibited

Published OnlineFirst October 7, 2014; DOI: 10.1158/1078-0432.CCR-14-0519

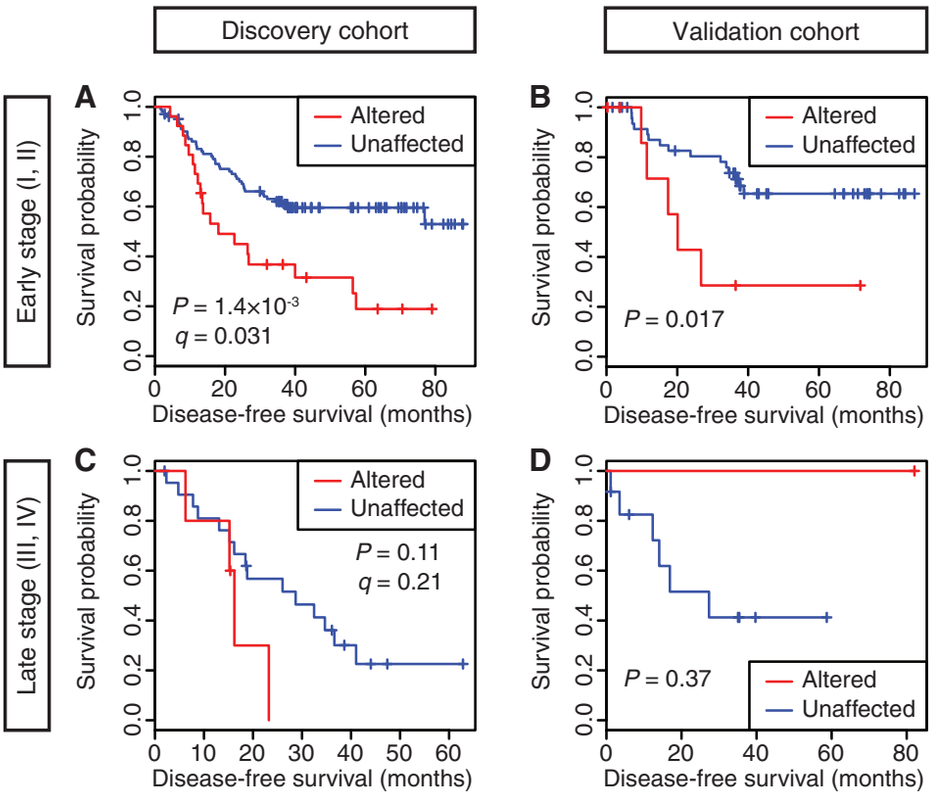

Figure 2.
Association between RB pathway alterations and poor prognosis in early-stage lung adenocarcinoma patients. A, Kaplan–Meier curves of disease-free survival (DFS) for early-stage lung adenocarcinoma patients in the discovery cohort ($n = 129$). The survival curves of the patients with RB pathway mutations and the remaining patients are shown in red and blue, respectively. The log-rank P value and q value are also shown. B, DFS curve for early-stage lung adenocarcinoma patients in the validation cohort ($n = 59$). Otherwise, as in A. C, DFS curve for late-stage lung adenocarcinoma patients in the discovery cohort ($n = 27$). Otherwise, as in A. D, DFS curve for late-stage lung adenocarcinoma patients in the validation cohort ($n = 13$). Otherwise, as in A.

significantly lower expression of cyclin D1 ($P = 1.6 \times 10^{-5}$; Fig. 3B), consistent with a previous report that demonstrated that *RB1* mutations promote cyclin degradation, leading to lower expression of cyclin D1 (36). Taken together, these results indicate that tumors with inactivated *RB1* have altered expression of other proteins essential to cell-cycle regulation, providing a possible link between RB pathway alterations and poor prognosis in early-stage lung adenocarcinoma patients.

Molecular definitions of lung adenocarcinoma subtypes

To gain more comprehensive insight into the molecular definitions of lung adenocarcinoma subtypes, we investigated the associations between significantly mutated genes and tumor subtypes, which were classified on the basis of the most up-to-date tumor classification criteria (13).

The overall molecular definition demonstrated the similarities between the acinar and the papillary subtypes, and between the micropapillary and the solid subtypes (Fig. 4A). Consistent with this observation, when subjected to hierarchical clustering according to the mutation profile, the acinar and the papillary subtypes were included in the same cluster, whereas the micropapillary and the solid subtypes formed another cluster (Fig. 4B). These clustering patterns agree with the known histologic and morphologic characteristics of these tumor subtypes (13).

The molecular definitions of the acinar and the papillary subtypes seem to be dramatically different from those of the micropapillary and the solid subtypes. One of main reasons for this discrepancy stems from *EGFR* mutations, which are strongly associated with the papillary subtype ($P = 4.7 \times 10^{-6}$, Fisher exact test) and inversely associated with the solid ($P = 1.5 \times 10^{-3}$, Fisher exact test), the mucinous ($P = 2.7 \times 10^{-3}$, Fisher exact test), and the micropapillary subtypes ($P = 8.0 \times 10^{-3}$, Fisher exact test). Some of these associations have been previously reported (16), but our large-scale dataset enabled us to look into these associations more completely. Interestingly, the mucinous subtype demonstrated a unique mutation profile, in that the most prevalent mutations accumulated in *KRAS* (Fig. 4A). It is also noteworthy that the solid subtype demonstrated a very heterogeneous mutational profile (Fig. 4A), which is consistent with the observation that this subtype often demonstrates unusual histologic and morphologic characteristics that are fairly different from other subtypes (13). Taken together, our detailed analyses call for a more careful investigation of the molecular definitions of the lung adenocarcinoma subtypes and indicate that a better tailored therapeutic approach should be used when treating patients with lung adenocarcinoma to improve overall therapeutic outcomes.

Discussion

For this study, we generated the whole exome and CNV dataset together with fully annotated clinical data on 247 lung adenocarcinomas, which provide an unprecedented opportunity for investigating the genomic architecture of lung adenocarcinoma and genomic alternations associated with clinical phenotypes and outcomes.

By processing the exome data according to our stringent statistical approaches, we detected 22 significantly mutated genes

Published OnlineFirst October 7, 2014; DOI: 10.1158/1078-0432.CCR-14-0519

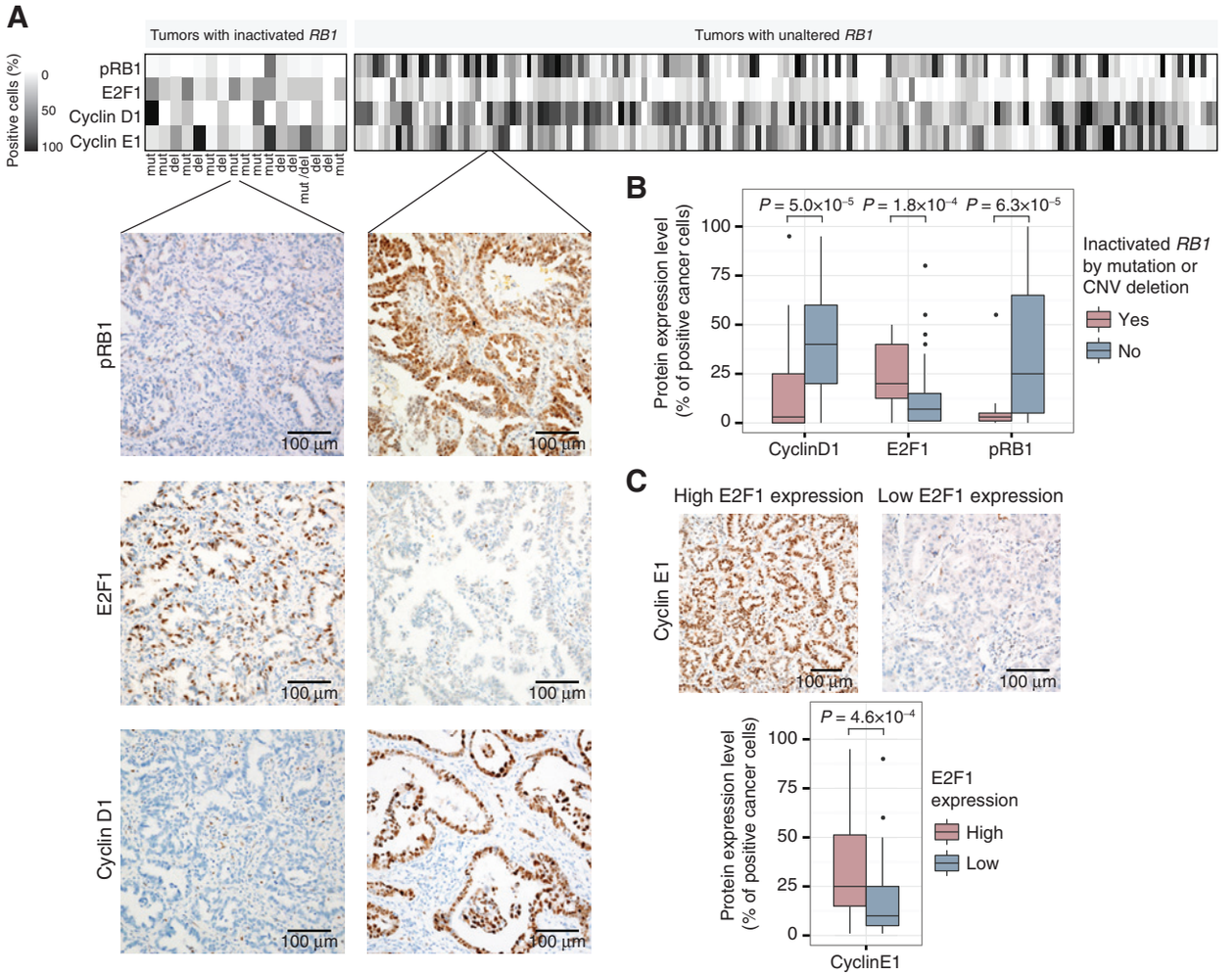

Figure 3.
Rb1 gene inactivation associated with expression changes of cell-cycle–related proteins in lung adenocarcinoma tumors. A, protein expression of cell-cycle–related proteins in 170 tumor samples as measured by IHC (see Materials and Methods). B, altered protein expression of cyclinD1, E2F1, and pRB1 in tumors with inactivated *RB1* compared with tumors with unaltered *RB1*. C, elevated protein expression of cyclinE1 in tumors with high expression of E2F1.

with either base substitutions or small indels, which account for 78% of our discovery cohort. Sixty-nine percent of the discovery cohort had such mutations in six previously reported genes (*EGFR*, *TP53*, *KRAS*, *PIK3CA*, *RB1*, and *SETD2*). The remaining 9% of cases had base substitutions or small indels that accumulated in the 16 novel driver candidates, including *COL11A1*, *CENPF*, *COL6A3*, *LRBA*, and *SLIT2*. Another 6% of the discovery cohort demonstrated $\geq$1 significant CNV(s), leaving 15% of lung adenocarcinomas unexplained by our exome or CNV data (Supplementary Fig. S6). Potential reasons for the 15% of unexplained lung adenocarcinomas include tumor impurity, relatively low overall depth of exome sequencing, translocation or fusion genes such as *EML4-ALK* (37, 38), genomic alterations accumulating in noncoding regions, epigenetic changes (39, 40), and technical underrepresentation by exome capture. This large-scale effort extensively utilized whole-exome sequencing, Sequenom technology, and CNV arrays to detect dozens of novel driver candidates and CNVs, and yet only a small fraction (9% + 6% = 15%) of lung adenocarcinomas can be additionally explained by these genomic alternations compared with 69% of lung adenocarcinomas that can be explained by the genomic alternations that accumulated on six previously reported genes (Supplementary Fig. S6). To comprehensively verify the molecular causes responsible for the 15% of unexplained lung adenocarcinomas, we speculate that future efforts should be directed toward utilizing multiplatform technologies that concurrently monitor the transcriptome, methylome, and noncoding genome, in addition to coding genomes and CNVs.

Although we detected a large number of significantly mutated genes in most of the 170 patients in our discovery cohort, the overall prevalence of mutations in a few genes was lower than previously reported. For instance, *KRAS* has been reported as mutated in about 12% of Asian patients with lung adenocarcinoma (9), whereas in the present study, *KRAS* was only affected in 6% of patients in both the discovery and validation cohorts. To validate the prevalence of *KRAS* and a few other genes in our dataset, we used the more sensitive Sequenom approach, which utilizes mass spectroscopy-based genotyping. Accordingly, 81%

Published OnlineFirst October 7, 2014; DOI: 10.1158/1078-0432.CCR-14-0519

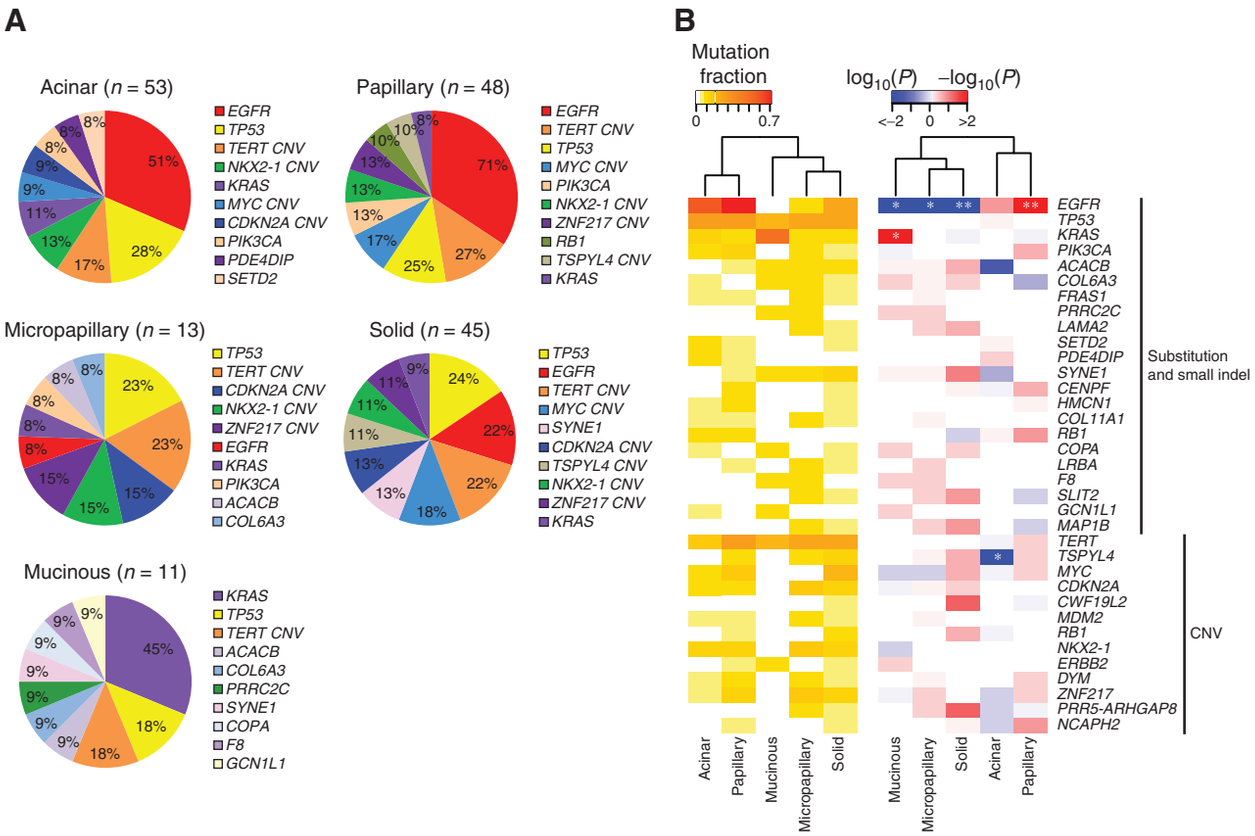

Figure 4.

Molecular definitions of lung adenocarcinoma subtypes. A, for each of the five tumor subtypes (acinar, papillary, solid, micropapillary, and mucinous), the frequencies of the most prevalent 10 significantly mutated genes and significantly altered CNV regions are illustrated. Each CNV region was shown as its representative gene (see Supplementary Table S7 and Supplementary Methods). B, lung adenocarcinoma subtypes are clustered according to their overall mutation profile. In the left panel, tumor subtypes are clustered by the frequencies of mutated genes. For each gene, its mutation fraction was calculated by dividing the number of patients with the mutated gene by the total number of patients included in the discovery cohort. In the right panel, tumor subtypes are clustered by the degree of the association between each subtype and the mutated genes in the subtype, as measured by Fisher exact test. Positive associations and negative associations are represented as $-\log_{10}(P$ value; red) or $\log_{10}(P$ value; blue), respectively. Asterisks indicate the level of significance for each association (*, $P < 0.05$; **, Bonferroni-adjusted $P < 0.05$).

of *EGFR* substitutions (29 of 36) and 100% of *BRAF* substitutions (3 of 3) detected by the Sequenom technology were also detected by exome sequencing (Supplementary Table S12). However, only 21% of *KRAS* substitutions (3 of 18) detected using this technology were detected in our exome dataset.

To understand the underlying reasons for this discrepancy, we looked at the depths of the sequence reads and found that the median read depth of the first *KRAS* exon (where the recurrent mutations were located) was significantly lower than that in other *KRAS* regions ($P < 2.2 \times 10^{-16}$, rank-sum test). However, this phenomenon seems to be specific to our exome data because another exome sequencing study on AML-M5 did not show a lower read depth for the first *KRAS* exon (Supplementary Fig. S7; ref. 41). Therefore, we speculate that the low detection sensitivity of our exome dataset may be due to low capturing efficiency in specific regions of the exome, including the first *KRAS* exon. In support of this hypothesis, the read depth of the first *KRAS* exon in patients in whom we were unable to detect *KRAS* mutations was significantly lower than in patients who have detectable *KRAS* mutations ($P < 2.2 \times 10^{-16}$, rank-sum test; Supplementary Fig. S7C). These findings illustrate that, depending on the exome-capturing platform, genomic studies based on exome sequencing might include significantly underrepresented regions. Thus, alternative technologies should be used to compensate for low detection sensitivity in such regions. In this study, we complemented *EGFR*, *KRAS*, and *PI3KCA* mutations using Sequenom-detected mutations. This increased the overall prevalence of *EGFR*, *KRAS*, and *PIK3CA* mutations from 28%, 6%, and 5% to 44%, 14%, and 6%, respectively.

There are several genes that have been previously reported to be driver candidates in lung adenocarcinoma but that were not detected in our analysis, such as *KEAP1* and *STK11* (6, 8, 20). Of the 247 patients, only 4 patients had *KEAP1* mutations, and only 2 patients had *STK11* mutations. The lower prevalence of these mutations, compared with a previous study done in the United States where 22 and 27 out of 183 patients bore mutations in *KEAP1* and *STK11*, respectively (8), may be the major reason these genes were not detected as driver candidates in our analysis. The low prevalence of mutated *KEAP1* and *STK11* may be due to the low capture efficiency discussed above or to the different ethnic backgrounds of the study populations. The median sequence depths of *KEAP1* and *STK11* (22 and 7, respectively;

Published OnlineFirst October 7, 2014; DOI: 10.1158/1078-0432.CCR-14-0519

Supplementary Fig. S8) are much lower than the median depth of 40 for the overall tumor samples and are lower than the depths of known driver genes such as *EGFR*, *KRAS*, and *BRAF* (45, 48, and 43, respectively), indicating that the lower capture efficiency may be at least partially responsible for the low prevalence. It is also possible that the prevalence of these genes in the Asian population may differ from that of Western populations. Supporting this possibility, another independent exome study on patients with lung adenocarcinoma in South Korea reported low prevalence of mutated *STK11* (2 out of 87) and *KEAP1* (3 out of 87; ref. 9).

The Cancer Genome Atlas (TCGA) recently reported a comprehensive molecular profile of 230 patients with lung adenocarcinoma with diverse ethnic backgrounds (20). Although the TCGA study agrees well with ours on many key findings such as a higher overall mutation rate in smokers than in never-smokers and the negative association between *EGFR* and *KRAS* mutations, there are noticeable differences between the two studies. For example, among the 18 and 22 significantly mutated genes detected in the studies, TCGA and ours, respectively, only six genes (*EGFR*, *TP53*, *KRAS*, *PIK3CA*, *RB1*, and *SETD2*) were found to be present in both lists. Also, the prevalence of *EGFR* mutations was much higher in our study (42.4%) than in the TCGA study (14.3%). These differences may have originated from the differences in ethnic backgrounds and the relative fraction of smokers included in the analysis. Indeed, the TCGA study had a smaller fraction of never-smokers (14.3%, $n = 33$) than in ours (49%, $n = 122$) and the different mutation profiles related to ethnic backgrounds have been previously reported (9, 42). These two studies should be complementary to each other, providing a more comprehensive molecular profile of lung adenocarcinoma and a useful resource for future comparative studies that aim to analyze the complexity and the heterogeneity of lung adenocarcinoma.

To reduce potential false positives in this study, we used a more stringent MuTect parameter, and by doing so, we might have sacrificed the overall detection sensitivity. To estimate the number of mutations we might be missing due to our stringent parameters, we compared the overall detection sensitivity between our current stringent parameter (MuTect $\theta_T = 8.0$) and the more lenient default detection parameter (MuTect $\theta_T = 6.3$). The more lenient cutoff increased the total number of detected mutations by 50%, from 25,065 to 37,548. However, the prevalence of known driver genes was not significantly affected. For instance, the prevalence of *EGFR* was 42.9% using the lenient cutoff compared with 42.4% using our current stringent cutoff. Similarly, the prevalence of *TP53* and *PIK3A* only increased marginally, from 25.3% to 28.8% for *TP53* and from 14.1% to 14.7% for *PIK3CA*. The prevalence of other driver genes was similarly unchanged (Supplementary Table S13). For 45 known driver genes, using the more stringent cutoff reduced the average prevalence only by 0.5% in comparison with the default parameter, indicating that the additional mutations detected using the lenient cutoff might include a high fraction of false positives.

It has been reported that lung adenocarcinoma has a high overall mutation rate (MR), and that the MR is higher in smokers than in never-smokers (8, 43). In comparison with previous reports, our overall MR was a bit lower, probably due to our high detection stringency. Using the current stringent cutoff, we obtained MRs of 5.5/Mbp and 1.9/Mbp for smokers and never-smokers, respectively (Supplementary Fig. S9A). When the lenient cutoff was applied, the MR of smokers increased to 7.7/Mbp. However, the MR of never-smokers also increased, to 4.4/Mbp, resulting in too little difference in MR between smokers and never-smokers (Supplementary Fig. S9B). Repeated analysis with an even more lenient parameter ($\theta_T = 5.5$) resulted in nearly equal MRs (11.1/Mbp for smokers and 8.5/Mbp for never-smokers, Supplementary Fig. S9C). These analyses also suggest that the additional substitutions detected using the lenient cutoff might include many false positives. We concluded that the stringent cutoff results in a more robust set of substitutions at the expense of lowered detection sensitivity. Therefore, we decided to continue to use the more stringent parameter instead of the lenient parameter.

It is noteworthy that the current treatment guidelines for early-stage lung adenocarcinoma provided by the National Comprehensive Cancer Network recommend follow-up without adjuvant therapy after complete resection, except for patients who have residual tumors or those considered at high risk (10). However, because the long-term treatment failure rate is close to 50% (10), it is important to identify subgroups of patients who may benefit from adjuvant treatments such as chemotherapy, even in early-stage lung adenocarcinoma. Our discovery that RB pathway mutations are associated with poor prognosis calls for additional clinical interventions that could improve the clinical outcomes of early-stage lung adenocarcinoma patients. In particular, a set of validated diagnostic markers that could detect genomic alterations in the RB pathway would be useful for patient stratification.

Rb protein acts as a tumor suppressor that inhibits cell-cycle progression by binding and inhibiting transcription factors of the E2F family. When Rb is phosphorylated to pRb by cyclin-dependent kinases, it is unable to form a complex with E2F, which allows E2F to promote cell-cycle progression (44). In this study, lower levels of pRb and cyclin D1 and higher levels of E2F1, a member of the E2F family, were detected in tumors with *RB1* alterations compared with tumors without *RB1* alterations. These findings agree with a previous report showing a lack of cyclin D1 overexpression and high E2F1 levels in tumors with Rb loss (36). Lack of cyclin D1 overexpression in Rb loss can be explained by the increased disassembly of cyclin D-cdk4 complexes and increased cyclin D turnover (36, 45). We also observed that tumors with *RB1* alterations are characterized by low pRb levels and high E2F1 levels, which may explain the poor prognosis of these patients. In addition, our study reveals that high E2F1 protein expression is associated with high cyclin E1 levels, which is consistent with the fact that E2F binds the promoter of the cyclin E gene and increases its expression. High levels of E2F1 and cyclin E1 in tumors with *RB1* alterations suggest that the prognostic effect of *RB1* alterations may be related to the actions of these downstream molecules of Rb1. Our analysis suggests that either pan-CDK inhibitors or selective CDK inhibitors, being used more frequently in clinical trials (46), might be effective to target these downstream molecules of Rb1.

Our analysis of the whole exome, CNV dataset, and fully annotated clinical data provide comprehensive insights into the genomic portrait of lung adenocarcinoma and help define the molecular signatures of lung adenocarcinoma subtypes. These findings could lead to the discovery of useful prognostic markers and help devise personalized treatment strategies for patients with early-stage lung adenocarcinoma.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Published OnlineFirst October 7, 2014; DOI: 10.1158/1078-0432.CCR-14-0519

Authors' Contributions

Conception and design: H.R. Kim, C.-M. Choi, E.K. Choi, S.-W. Kim, Y.-H. Kim, C. Lee, E. Yu, G. Kong, D. Baek, S.J. Jang

Development of methodology: S. Choi, J. Kim, S. Kim, E.K. Choi, Y.-H. Kim, J.-E. Lee, C. Lee, D. Baek, S.J. Jang

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H.R. Kim, C.O. Sung, C.-M. Choi, Y.-H. Kim, J.-Y. Lee, J.-E. Lee, W.-R. Jung, H.Y. Jang, E. Yang, C. Lee, G. Kong, S.J. Jang

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): S. Choi, C.O. Sung, J. Kim, S. Kim, S.-M. Ahn, C.-M. Choi, S.-M. Chun, Y.-H. Kim, J.-Y. Lee, J.S. Song, F. Haq, W.-R. Jung, E. Yang, C. Lee, G. Kong, D. Baek

Writing, review, and/or revision of the manuscript: S. Choi, H.R. Kim, C.O. Sung, S.-M. Ahn, E.K. Choi, S.-W. Kim, Y.-H. Kim, C. Lee, G. Kong, D. Baek, S.J. Jang

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H.R. Kim, Y.-H. Kim, D. Kim S.Y. Lee, C. Lee, E. Yu, G. Kong, D. Baek, S.J. Jang

Study supervision: H.R. Kim, E.K. Choi, S.-W. Kim, C. Lee, G. Kong, D. Baek, S.J. Jang

Grant Support

This work was supported by the National Project for Personalized Genomic Medicine, Ministry of Health and Welfare of Korea (A111218-SC01), by the Leading Foreign Research Institute Recruitment Program (2011-0030105) and the Basic Science Research Program (2011-0014523), the Ministry of Education, Science, and Technology of Korea, and by the Institute for Basic Science (EM1302).

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 March 11, 2014; revised September 11, 2014; accepted September 19, 2014; published OnlineFirst October 7, 2014.

References

1. WHO. World Health Statistics. 2012 ed; 2012.
2. In KH, Kwon YS, Oh IJ, Kim KS, Jung MH, Lee KH, et al. Lung cancer patients who are asymptomatic at diagnosis show favorable prognosis: a Korean Lung Cancer Registry Study. Lung Cancer 2009;64:232–7.
3. Asamura H, Goya T, Koshiishi Y, Sohara Y, Eguchi K, Mori M, et al. A Japanese Lung Cancer Registry study: prognosis of 13,010 resected lung cancers. J Thorac Oncol 2008;3:46–52.
4. Straus GM, Jemal A, McKenna MB, Strauss JA, Cummings KM. The epidemic of smoking-related adenocarcinoma of the lung: The role of the tobacco industry and filtered and low-tar cigarettes. J Thorac Oncol 2007;2: S305.
5. Weir BA, Woo MS, Getz G, Perner S, Ding L, Beroukhim R, et al. Characterizing the cancer genome in lung adenocarcinoma. Nature 2007;450: 893–8.
6. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature 2008;455:1069–75.
7. Lee W, Jiang Z, Liu J, Haverty PM, Guan Y, Stinson J, et al. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. Nature 2010;465:473–7.
8. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell 2012;150:1107–20.
9. Seo JS, Ju YS, Lee WC, Shin JY, Lee JK, Bleazard T, et al. The transcriptional landscape and mutational profile of lung adenocarcinoma. Genome Res 2012;22:2109–19.
10. Ettinger DS, Akerley W, Borghaei H, Chang AC, Cheney RT, Chirieac LR, et al. Non-small cell lung cancer, version 2.2013. J Natl Compr Canc Netw 2013;11:645–53.
11. Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. J Thorac Oncol 2007;2:706–14.
12. Kopp R, Weidenhagen R, Reinmiedl J, Muller C, Furst H, Bittmann I, et al. Outcome following lung resections for pT1 non-small cell lung cancer. Eur J Surg Oncol 2006;32:329–34.
13. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger K, Yatabe Y, et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society: international multidisciplinary classification of lung adenocarcinoma: executive summary. Proc Am Thorac Soc 2011;8:381–5.
14. Kim YH, Ishii G, Goto K, Nagai K, Tsuta K, Shiono S, et al. Dominant papillary subtype is a significant predictor of the response to gefitinib in adenocarcinoma of the lung. Clin Cancer Res 2004;10:7311–7.
15. Byrne BJ, Garst J. Epidermal growth factor receptor inhibitors and their role in non-small-cell lung cancer. Curr Oncol Rep 2005;7:241–7.
16. Shim HS, Lee da H, Park EJ, Kim SH. Histopathologic characteristics of lung adenocarcinomas with epidermal growth factor receptor mutations in the International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society lung adenocarcinoma classification. Arch Pathol Lab Med 2011;135:1329–34.
17. Finberg KE, Sequist LV, Joshi VA, Muzikansky A, Miller JM, Han M, et al. Mucinous differentiation correlates with absence of EGFR mutation and presence of KRAS mutation in lung adenocarcinomas with bronchioloalveolar features. J Mol Diagn 2007;9:320–6.
18. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013;31:213–9.
19. Kan Z, Jaiswal BS, Stinson J, Janakiraman V, Bhatt D, Stern HM, et al. Diverse somatic mutation patterns and pathway alterations in human cancers. Nature 2010;466:869–73.
20. Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. Nature 2014;511:543–50.
21. Chong IW, Chang MY, Chang HC, Yu YP, Sheu CC, Tsai JR, et al. Great potential of a panel of multiple hMTH1, SPD, ITGA11 and COL11A1 markers for diagnosis of patients with non-small cell lung cancer. Oncol Rep 2006;16:981–8.
22. O'Brien SL, Fagan A, Fox EJ, Millikan RC, Culhane AC, Brennan DJ, et al. CENP-F expression is associated with poor prognosis and chromosomal instability in patients with primary breast cancer. Int J Cancer 2007; 120:1434–43.
23. Sherman-Baust CA, Weeraratna AT, Rangel LB, Pizer ES, Cho KR, Schwartz DR, et al. Remodeling of the extracellular matrix through overexpression of collagen VI contributes to cisplatin resistance in ovarian cancer cells. Cancer Cell 2003;3:377–86.
24. Tseng RC, Lee SH, Hsu HS, Chen BH, Tsai WC, Tzao C, et al. SLIT2 attenuation during lung cancer progression deregulates beta-catenin and E-cadherin and associates with poor prognosis. Cancer Res 2010;70: 543–51.
25. Peifer M, Fernandez-Cuesta L, Sos ML, George J, Seidel D, Kasper LH, et al. Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nat Genet 2012;44:1104–10.
26. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 2011;12:R41.
27. Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, et al. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. Nat Genet 2012;44:685–9.
28. Nakshatri H, Badve S. FOXA1 as a therapeutic target for breast cancer. Expert Opin Ther Targets 2007;11:507–14.
29. Pikor LA, Lockwood WW, Thu KL, Vucic EA, Chari R, Gazdar AF, et al. YEATS4 is a novel oncogene amplified in non-small cell lung cancer that regulates the p53 pathway. Cancer Res 2013;73:7301–12.
30. Krig SR, Miller JK, Frietze S, Beckett lung adenocarcinoma, Neve RM, Farnham PJ, et al. ZNF217, a candidate breast cancer oncogene amplified at 20q13, regulates expression of the ErbB3 receptor tyrosine kinase in breast cancer cells. Oncogene 2010;29:5500–10.

Published OnlineFirst October 7, 2014; DOI: 10.1158/1078-0432.CCR-14-0519

31. Stephens P, Hunter C, Bignell G, Edkins S, Davies H, Teague J, et al. Lung cancer: intragenic ERBB2 kinase mutations in tumours. Nature 2004; 431:525–6.
32. Engelman JA, Chen L, Tan X, Crosby K, Guimaraes AR, Upadhyay R, et al. Effective use of PI3K and MEK inhibitors to treat mutant Kras G12D and PIK3CA H1047R murine lung cancers. Nat Med 2008;14:1351–6.
33. Asai A, Oshima Y, Yamamoto Y, Uochi TA, Kusaka H, Akinaga S, et al. A novel telomerase template antagonist (GRN163) as a potential anticancer agent. Cancer Res 2003;63:3931–9.
34. Chen HZ, Tsai SY, Leone G. Emerging roles of E2Fs in cancer: an exit from cell cycle control. Nat Rev Cancer 2009;9:785–97.
35. Xing EP, Yang GY, Wang LD, Shi ST, Yang CS. Loss of heterozygosity of the Rb gene correlates with pRb protein expression and associates with p53 alteration in human esophageal cancer. Clin Cancer Res 1999;5:1231–40.
36. Wikenheiser-Brokamp KA. Retinoblastoma regulatory pathway in lung cancer. Curr Mol Med 2006;6:783–93.
37. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 2007;448:561–6.
38. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. N Engl J Med 2010;363:1693–703.
39. Tsou JA, Shen LY, Siegmund KD, Long TI, Laird PW, Seneviratne CK, et al. Distinct DNA methylation profiles in malignant mesothelioma, lung adenocarcinoma, and non-tumor lung. Lung Cancer 2005;47: 193–204.
40. Tsou JA, Galler JS, Siegmund KD, Laird PW, Turla S, Cozen W, et al. Identification of a panel of sensitive and specific DNA methylation markers for lung adenocarcinoma. Mol Cancer 2007;6:70.
41. Yan XJ, Xu J, Gu ZH, Pan CM, Lu G, Shen Y, et al. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. Nat Genet 2011;43:309–15.
42. Shigematsu H, Lin L, Takahashi T, Nomura M, Suzuki M, Wistuba II, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 2005;97: 339–46.
43. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 2013;499:214–8.
44. Gorgoulis VG, Zacharatos P, Kotsinas A, Mariatos G, Liloglou T, Vogiatzi T, et al. Altered expression of the cell cycle regulatory molecules pRb, p53 and MDM2 exert a synergetic effect on tumor growth and chromosomal instability in non-small cell lung carcinomas (NSCLCs). Mol Med 2000;6:208–37.
45. Pikor LA, Lockwood WW, Thu KL, Vucic EA, Chari R, Gazdar AF, et al. YEATS4 is a novel oncogene amplified in non-small cell lung cancer that regulates the p53 pathway. Cancer Res 2013;73:7301–12.
46. Canavese M, Santo L, Raje N. Cyclin dependent kinases in cancer: potential for therapeutic intervention. Cancer Biol Ther 2012;13:451–7.