RESEARCH ARTICLE

Next-Generation Sequencing of Stage IV Squamous Cell Lung Cancers Reveals an Association of PI3K Aberrations and Evidence of Clonal Heterogeneity in Patients with Brain Metastases

Paul K. Paik1,2, Ronglai Shen3, Helen Won4, Natasha Rekhtman4, Lu Wang4, Camelia S. Sima3, Arshi Arora3, Venkatraman Seshan3, Marc Ladanyi4,5, Michael F. Berger2,4,5, and Mark G. Kris1,2
ABSTRACT

Large-scale genomic characterization of squamous cell lung cancers (SQCLC) has revealed several putative oncogenic drivers. There are, however, little data to suggest that these alterations have clinical relevance. We performed comprehensive genomic profiling (including next-generation sequencing) of 79 stage IV SQCLCs and analyzed differences in the clinical characteristics of two major SQCLC subtypes: FGFR1 amplified and PI3K aberrant. Patients with PI3K-aberrant tumors had aggressive disease marked by worse survival (median overall survival, 8.6 vs. 19.1 months, \( P < 0.001 \)), higher metastatic burden (>3 organs, 18% vs. 3%, \( P = 0.025 \)), and greater incidence of brain metastases (27% vs. 0% in others, \( P < 0.001 \)). We performed whole-exome and RNA sequencing on paired brain metastases and primary lung cancers to elucidate the metastatic process to brain. SQCLC primaries that gave rise to brain metastases exhibited truncal PTEN loss. SQCLC brain metastases exhibited a high degree of genetic heterogeneity and evidence of clonal differences between their primary sites.

SIGNIFICANCE: We performed next-generation sequencing of metastatic SQCLCs and primary lung–brain metastasis pairs, identifying PI3K-aberrant tumors as an aggressive subset associated with brain metastases. We identified genetic heterogeneity between lung primaries–brain metastases as well as clonal populations that may highlight alterations important in the metastatic process. Cancer Discov; 5(6); 610–21. © 2015 AACR.

INTRODUCTION

Squamous cell lung cancers (SQCLC) account for 20% of all lung cancers. This equates to 350,000 patients diagnosed with this disease worldwide every year. Half of these patients present with metastatic disease, and most will not be cured. These outcomes reflect the underlying aggressiveness of the disease and the paucity of new advances in the field, where the last substantial improvement in survival came with the introduction of cisplatin in the 1980s (1).

The year 2004 marked a divergence in the management and conceptualization of the two most common histologic subtypes of lung cancer: SQCLCs and lung adenocarcinomas. The identification of EGFR tyrosine kinase inhibitor (TKI)–sensitizing mutations by three groups that year presaged the discovery of other druggable oncogenic events, primarily in adenocarcinomas, and including mutations in HER2 and BRAF and oncogenic rearrangements of ALK, RET, and ROS1. The breadth of these discoveries led to a rethinking of adenocarcinomas as a set of related but distinct oncogene-addicted cancers, each of which could be treated with a targeted therapy (2).

The majority of these events were found not to occur with any appreciable frequency in SQCLCs. Comprehensive molecular analyses of resected SQCLC tumors by The Cancer Genome Atlas (TCGA) Consortium did, however, identify somatic variants, copy-number alterations, and gene expression changes that circumscribe genetic subsets of SQCLCs (3). Prior work showed that some of these key targets were both oncogenic and responsive to pharmacologic inhibitors, including FGFR1 amplification (4), DDR2 mutations (5), and hyperactivating PI3K pathway alterations (6, 7). In aggregate, these potential therapeutic targets are thought to occur in at least 50% of SQCLC cases (3, 8).

The TCGA analysis involved resection specimens from early-stage (I–III) patients to ensure a precise histologic diagnosis and sufficient tumor content. The applicability of the findings to stage IV disease is, as a result, unknown. In addition, there were limited clinical data available for analysis in TCGA. As such, the clinical sequelae of biologic differences could not be assessed. These differences can provide necessary prognostic information to patients and the hypothetical groundwork to unravel the complex manifestations of these cancers.

In light of these gaps in our knowledge, we sought to identify differences in the clinical characteristics of patients with stage IV SQCLCs whose tumors were comprehensively tested for oncogenic drivers and to shed light on the biologic underpinnings of these differences. We focused, as the starting point for our study, on overall survival and patterns of metastasis in two distinct molecularly defined groups with alterations affecting either FGFR1 or PI3K, both high-frequency pathways scored as significantly altered by TCGA and for which preclinical data support potential roles as actionable drivers (4, 6, 7).

1 Thoracic Oncology Service, Division of Solid Tumor Oncology, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York. 2 Weill Cornell Medical College, New York, New York. 3 Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York. 4 Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York. 5 Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, New York.

Note: Supplementary data for this article are available at Cancer Discovery Online (http://cancerdiscovery.aacrjournals.org/).

PK. Paik and R. Shen contributed equally to this article.

Corresponding Author: Paul K. Paik, Memorial Sloan Kettering Cancer Center, 300 East 66th Street, New York, NY 10065. Phone: 646-888-4202; Fax: 646-888-4263; E-mail: paikp@mskcc.org

doi: 10.1158/2159-8290.CD-14-1129

©2015 American Association for Cancer Research.
RESULTS

Squamous Cell Lung Cancer Mutation Analysis Program Results

Seventy-nine patients with stage IV disease had sufficient archived tumor material for FISH for FGFR1 amplification, immunohistochemistry (IHC) for PTEN loss, Sequenom genotyping for hotspot mutations in eight oncogenes, including PIK3CA, and exon capture/next-generation sequencing of 279 oncogenes and tumor suppressors by IMPACT (Integrated Mutation Profiling of Actionable Cancer Targets; see Methods). For the purposes of this analysis, upstream aberrations in PI3K were defined as PIK3CA and PTEN mutations and PTEN loss, all relatively high-frequency events in the TCGA analysis that have been shown to predict for response to pharmacologic inhibition in vitro or in vivo (7) and which are the exploratory predictive biomarkers used in a number of clinical trials (NCT01297491 and NCT01655225).

As shown in Fig. 1, 20% of patients had FGFR1 amplification [95% confidence interval (CI), 12%–31%], 29% had complete loss of PTEN (95% CI, 28%–30%), 10% had PTEN mutations (95% CI, 7%–13%), and 11% had PIK3CA mutations (95% CI, 6%–20%). Thirteen percent (2/16) of FGFR1-amplified tumors had overlapping PTEN loss. Seventeen percent of tumors that had PTEN loss by IHC also had concurrent PTEN mutations, PTEN loss, or PIK3CA mutations. In total, 61% (95% CI, 49%–71%) of patients had tumors that bore an aberration in either FGFR1 or the upstream PI3K pathway.

IMPACT results of selected genes, including those found significantly mutated by TCGA, are shown in Fig. 2. Genes mutated in multiple patients included TP53 (89%), CDKN2A (11%), NFE2L2 (17%), and KEAP1 (20%). NFE2L2 and KEAP1 mutations appeared in nonoverlapping cases, as expected (3). DDR2 mutations were found in 6% of patients (95% CI, 3%–14%). Gene-specific copy-number alterations were found at relatively high frequencies in chromosome 3 (SOX2, PIK3CA). A complete list of somatic mutations and copy-number alterations detected by IMPACT is provided in Supplementary Table S1. A comparison of the frequency of selected genes found by the Squamous Cell Lung Cancer Mutation Analysis Program (SQ-MAP) versus TCGA is provided in Supplementary Table S2. A number of gene frequencies were significantly different by the Fisher exact test but lost significance when adjusted for multiple testing effects.

Patient Characteristics by Genotype

Patient characteristics are shown in Table 1. Characteristics were by and large the same. There were no significant differences in age, sex, smoking history, Karnofsky performance status, or treatment with a platinum doublet by genotype. There were differences in the proportion of patients treated with genotype-directed targeted therapies. Thirty-one percent of patients with FGFR1 amplifications were treated with AZD4547, a specific and potent FGFR1–3 TKI. One patient with a PI3K alteration (PTEN loss) was treated with a PI3Kα inhibitor (buparlisib). No patient had a response (either partial or complete) to targeted therapy.

Overall Survival by Genotype

Median overall survival (OS) was, on the other hand, significantly different among the genotypes: FGFR1 amplified, 18.8 months (95% CI, 10.8–not reached); PI3K aberrant, 8.6 months (95% CI, 6.9–10.7); others, 21.3 months (95% CI, 15.9–27). Patients with PI3K aberrations had significantly worse OS than patients without these aberrations (median OS, 8.6 vs. 19.1 months, P < 0.001; Fig. 3A). OS was also worse compared with patients with FGFR1 amplification (median OS, 8.6 vs. 18.8 months, P < 0.001). Patients with FGFR1 amplification had a trend toward improved OS versus those without FGFR1 amplification (median OS, 18.8 vs. 10.8 months, P = 0.064; Fig. 3B).

Multivariate analysis using genotype, age, sex, performance status, and number of somatic changes as an additional genomic covariate (mutations + copy number) similarly demonstrated significantly worse OS for patients harboring a PI3K aberration (HR for death, 5.6; 95% CI, 2.8–11.2; P < 0.0001). There was no significant association between OS and number of somatic changes as a continuous variable. Exclusion of the 2 patients who had overlapping FGFR1 amplification and PTEN loss or reassignment of them to the different genotype categories did not affect our results.
**Figure 2.** Summary of somatic mutations and copy-number alterations in 79 stage IV SQCLC tumors tested by targeted exon sequencing through IMPACT.

**Table 1. Clinical characteristics of 79 stage IV SQCLC patients**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>PI3K (n = 33)</th>
<th>FGFR1 (n = 16)</th>
<th>Other (n = 30)</th>
<th>P (PI3K vs. FGFR1)</th>
<th>P (PI3K vs. other)</th>
<th>P (FGFR1 vs. other)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>68 (47–85)</td>
<td>73 (48–82)</td>
<td>69 (45–87)</td>
<td>0.10</td>
<td>0.36</td>
<td>0.42</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>48%</td>
<td>44%</td>
<td>30%</td>
<td>&gt;0.95</td>
<td>0.20</td>
<td>0.52</td>
</tr>
<tr>
<td>Smoking history (pack years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>40 (0–90)</td>
<td>43 (1–100)</td>
<td>27 (0–120)</td>
<td>0.63</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>Current or former</td>
<td>94%</td>
<td>100%</td>
<td>80%</td>
<td>&gt;0.95</td>
<td>0.14</td>
<td>0.08</td>
</tr>
<tr>
<td>Never</td>
<td>6%</td>
<td>0%</td>
<td>20%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KPS ≤70%</td>
<td>36%</td>
<td>38%</td>
<td>43%</td>
<td>&gt;0.95</td>
<td>0.79</td>
<td>0.052</td>
</tr>
<tr>
<td>Prior lines of therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>1 (0–4)</td>
<td>2 (0–4)</td>
<td>2 (0–4)</td>
<td>0.19</td>
<td>0.002</td>
<td>0.15</td>
</tr>
<tr>
<td>Treated with platinum doublet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion (%)</td>
<td>21/27 (78%)</td>
<td>9/13 (69%)</td>
<td>20/25 (80%)</td>
<td>0.70</td>
<td>&gt;0.95</td>
<td>0.069</td>
</tr>
<tr>
<td>Treated with targeted therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion (%)</td>
<td>1/27 (4%)</td>
<td>4/13 (31%)</td>
<td>0</td>
<td>0.03</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Number of metastatic sites &gt;3</td>
<td>6 (18%)</td>
<td>1 (6%)</td>
<td>0</td>
<td>0.40</td>
<td>0.025</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Abbreviation: KPS, Karnofsky performance status.
Because PIK3CA amplification was found, post hoc, in a substantial number of patients (23%), we also performed an OS analysis of this group of patients versus those without PI3K pathway alterations. Median OS for those with PIK3CA-amplified tumors versus those without a PI3K pathway aberration was 10.8 months (95% CI, 8.6–19.6) versus 20.2 months (95% CI, 14.1–24.2); P = 0.15. There was, however, overlap with other PI3K aberrations: 40% of patients whose tumors bore amplified PIK3CA also had tumors that exhibited complete loss of PTEN.

Metastatic Sites by Genotype

Sites of metastatic disease are listed in Table 2. Brain metastases were the only site that varied significantly by genotype, occurring in 27% (N = 9/33) of those with PI3K aberrations and 0% of all others (N = 0/46; P < 0.001). Overall, brain metastases were relatively uncommon, occurring in 11% of the entire cohort. In addition, patients with PI3K aberrations exhibited a significantly higher burden of metastatic disease than other patients (>3 organ sites involved: 18% vs. 3%, P = 0.025). Exclusion of the 2 patients who had overlapping FGFR1 amplification and PTEN loss or reassignment of them to the different genotype categories did not affect our results.

SQCLC Brain Metastases Also Exhibit PTEN Loss

In an effort to better understand the biologic underpinning of the higher incidence of brain metastases in SQCLC patients harboring PI3K aberrations, we analyzed 6 surgically resected SQCLC brain metastases that had frozen tissue available for testing by whole-exome sequencing, RNA sequencing, and IHC. Four of these patients had matched archived formalin-fixed paraffin-embedded (FFPE) samples of their primary lung cancers. The clinical characteristics and treatment course of these patients are provided in Supplementary Table S3.

Because PIK3CA amplification was found, post hoc, in a substantial number of patients (23%), we also performed an OS analysis of this group of patients versus those without PI3K pathway alterations. Median OS for those with PIK3CA-amplified tumors versus those without a PI3K pathway aberration was 10.8 months (95% CI, 8.6–19.6) versus 20.2 months (95% CI, 14.1–24.2); P = 0.15. There was, however, overlap with other PI3K aberrations: 40% of patients whose tumors bore amplified PIK3CA also had tumors that exhibited complete loss of PTEN.

Metastatic Sites by Genotype

Sites of metastatic disease are listed in Table 2. Brain metastases were the only site that varied significantly by genotype, occurring in 27% (N = 9/33) of those with PI3K aberrations and 0% of all others (N = 0/46; P < 0.001). Overall, brain metastases were relatively uncommon, occurring in 11% of the entire cohort. In addition, patients with PI3K aberrations exhibited a significantly higher burden of metastatic disease than other patients (>3 organ sites involved: 18% vs. 3%, P = 0.025). Exclusion of the 2 patients who had overlapping FGFR1 amplification and PTEN loss or reassignment of them to the different genotype categories did not affect our results.

SQCLC Brain Metastases Also Exhibit PTEN Loss

In an effort to better understand the biologic underpinning of the higher incidence of brain metastases in SQCLC patients harboring PI3K aberrations, we analyzed 6 surgically resected SQCLC brain metastases that had frozen tissue available for testing by whole-exome sequencing, RNA sequencing, and IHC. Four of these patients had matched archived formalin-fixed paraffin-embedded (FFPE) samples of their primary lung cancers. The clinical characteristics and treatment course of these patients are provided in Supplementary Table S3.

PTEN IHC and whole-exome sequencing of both the primary lung tumors and brain metastases demonstrated complete loss of PTEN in 67% of cases (N = 4/6). The other 2 cases exhibited weak PTEN staining (1+). There were no differences in PTEN loss between primary and metastasis pairs. A PIK3CA E542K mutation was present in 1 of the 6 patient brain metastasis samples (PP5), coincident with PTEN loss. A PIK3CG D521Y mutation was present in both the primary lung tumor and brain metastasis in 1 of 6 samples (PP4). Whole-exome sequencing detected heterozygous loss of PTEN in all of the brain metastases. No other functionally validated mutations within the canonical PI3K pathway in either the primary lung tumors or brain metastases were found. No other common oncopgenic drivers (e.g., EGFR, BRAF, KRAS, NRAS, HER2, MEK, DDR2, KEAP1, FGFR1–4) were found. FISH demonstrated 8p11 amplification in PP4, but sequencing did not demonstrate focal amplification of FGFR1 in either the primary lung tumor or brain metastasis. The mutational burden and spectrum (missense, silent, noncoding, truncating) were similar in all cases, save for the primary lung tumor from PP3, which showed a mutation burden roughly 10-fold that of the other samples. This was the only tumor biopsied after treatment with radiotherapy (concurrent cisplatin + etoposide and radiotherapy).

To further validate the IHC-based PTEN loss and heterozygous loss of PTEN signals, we performed RNA sequencing on each of the brain metastases and compared the gene expression results with an existing PTEN loss RNA signature derived by Saal and colleagues (9). This analysis yielded PTEN signature scores in each brain metastasis commensurate with PTEN loss. These scores were significantly different from those derived from TCGA expression data for resected early-stage primary lung cancers (P = 0.003; Fig. 4A).
Table 2. Sites of metastasis by genotype

<table>
<thead>
<tr>
<th>Site</th>
<th>PI3K (n = 33)</th>
<th>%</th>
<th>P (vs. all others)</th>
<th>FGFR1 (n = 16)</th>
<th>%</th>
<th>P (vs. all others)</th>
<th>Other (n = 30)</th>
<th>%</th>
<th>Total (n = 79)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung</td>
<td>11</td>
<td>33</td>
<td>0.81</td>
<td>7</td>
<td>44</td>
<td>0.56</td>
<td>10</td>
<td>33</td>
<td>28</td>
<td>35</td>
</tr>
<tr>
<td>Brain</td>
<td>9</td>
<td>27</td>
<td>&lt;0.001</td>
<td>0</td>
<td>0</td>
<td>0.19</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Bone</td>
<td>12</td>
<td>36</td>
<td>0.63</td>
<td>3</td>
<td>19</td>
<td>0.24</td>
<td>11</td>
<td>37</td>
<td>26</td>
<td>33</td>
</tr>
<tr>
<td>Extrathoracic lymph node</td>
<td>12</td>
<td>36</td>
<td>0.31</td>
<td>6</td>
<td>18</td>
<td>0.54</td>
<td>5</td>
<td>17</td>
<td>23</td>
<td>29</td>
</tr>
<tr>
<td>Pleura</td>
<td>6</td>
<td>18</td>
<td>0.13</td>
<td>5</td>
<td>31</td>
<td>0.76</td>
<td>11</td>
<td>37</td>
<td>22</td>
<td>28</td>
</tr>
<tr>
<td>Liver</td>
<td>4</td>
<td>12</td>
<td>&gt;0.95</td>
<td>1</td>
<td>6</td>
<td>0.68</td>
<td>5</td>
<td>17</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Adrenal</td>
<td>7</td>
<td>21</td>
<td>0.37</td>
<td>3</td>
<td>19</td>
<td>0.72</td>
<td>3</td>
<td>10</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Soft tissue</td>
<td>4</td>
<td>12</td>
<td>&gt;0.95</td>
<td>1</td>
<td>6</td>
<td>0.68</td>
<td>4</td>
<td>13</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Pericardium</td>
<td>1</td>
<td>3</td>
<td>&gt;0.95</td>
<td>1</td>
<td>6</td>
<td>0.50</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Pancreas</td>
<td>0</td>
<td>0</td>
<td>&gt;0.95</td>
<td>1</td>
<td>6</td>
<td>0.20</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 4. A, PTEN loss signature score in resected SQCLC brain metastases. Values plotted to the right are more strongly correlated with PTEN loss. Expression of PIK3CA and PTEN from individual samples in the TCGA SQCLC analysis is plotted for reference, confirming a trend toward decreased PTEN expression and increased PIK3CA expression and signature scores consistent with PTEN loss. PTEN signature activities in the brain metastases were compared with the signatures in the TCGA primary tumors using a one-sided Wilcoxon rank-sum test. B, prometastatic WNT gene signature score in resected SQCLC brain metastases. Values plotted to the right are more strongly correlated with WNT activity. Expression of HOXB9 and LEF1, key promoters of the brain metastatic process in lung adenocarcinomas, from individual samples in the TCGA SQCLC analysis is plotted as an SQCLC-specific reference. WNT signature activities in the brain metastases were compared with the signatures in the TCGA primary tumors using a one-sided Wilcoxon rank-sum test. C, plasminogen activator (PA) inhibitor serpin gene expression levels from the TCGA SQCLC analysis and SQCLC brain metastases. There was no substantial increase in PA serpin expression in the brain metastases versus early stage tumors from TCGA. LUSC, lung squamous cell carcinoma.
SQCLC Brain Metastases Are Not Marked by Increased WNT Signaling or Serpin Overexpression

Two mechanisms underlying the development of brain metastases in lung adenocarcinomas have been previously published, including activation of WNT signaling (with HOXB9 and LEF1 expression strongly correlated with metastatic potential) and upregulation of plasminogen activator (PA) inhibitory serpins (10, 11). Gene expression analysis of the SQCLC brain metastases did not show an association with increased WNT signaling (Fig. 4B) or PA serpin overexpression (Fig. 4C). We did confirm, as an internal control, that there was an association between HOXB9 and LEF1 expression and WNT signature activity in the TCGA cases.

SQCLC Brain Metastases and Lung Primaries Are Highly Genetically Divergent

In order to better understand the degree and manner of divergence that might exist between brain metastases and their lung primaries, we performed whole-exome sequencing of paired lung primaries and brain metastases from 4 SQCLC patients (PP1-4) and whole-exome sequencing of an unpaired brain metastasis from one other (PP5). Tumor and normal DNA were sequenced at 100× coverage depth. In the PP1 paired set, 3,652 somatic mutations were detected, of which 610 were nonsilent. Among these (Fig. 5A). In the PP4 paired set, a total of 2,833 mutations were shared between the lung primary and brain metastasis (Fig. 5A). In the PP4 paired set, a total of 2,833 mutations were detected, of which 610 were nonsilent. Among these nonsilent mutations, 26% were shared between the lung primary and brain metastasis. PP3 was excluded from paired analysis as a result of the hypermutated phenotype of the primary lung tumor; PP2 did not have a matched normal counterpart for the paired analysis.

To explore evidence of clonal subpopulations in the lung primaries and brain metastases, we estimated the cancer cell fraction for each mutation, as described in the Methods section. The dominant clusters of mutations private to the lung primaries and brain metastases, respectively, showed clear evidence of subclonality, indicating a high degree of intratumor heterogeneity (Fig. 5B, dense clouds along the x-axes and y-axes). Two independent methods of calculating cancer cell fractions (pyClone and FACETS) were used, and the results were highly concordant. These patterns in PP1 and PP4 were also seen when the analysis was restricted to mutations with ≥80× depth of coverage. Finally, construction of phylogenetic trees for all somatic mutations detected in the PP1 and PP4 paired sets yielded evidence of a branched evolutionary pattern (Fig. 5C). Genes that were clonal in the brain metastases but subclonal in the lung primaries are shown in Supplementary Table S4.

Analysis of the copy-number data from the paired sets also yielded evidence of great divergence between the lung primaries and brain metastases. In the PP1 paired set, only 24% of the copy-number alterations were shared (Supplementary Fig. S1). Major alterations observed in the primary tumor from PP1 included a gain in 7p and loss in 6q that were absent in the brain metastasis (Supplementary Fig. S2A). A prominent gain of 5p and loss of 5q were observed in all of the brain metastases analyzed (Supplementary Fig. S2B–S2D). In contrast, not all of the primary tumors carried these aberrations; in PP1, the 5p gain was not present in the lung primary. FISH confirmed the presence of multiple copies of 5p in nearly all tumor cells analyzed from the brain metastases, with most harboring >6 copies (Supplementary Fig. S3). This degree of aneuploidy was not seen in the PP1 primary lung tumor, which was the only primary tumor with sufficient material left over for 5p assessment by FISH (disomy in 80% of cells).

Figure 5. Patterns of shared and unique somatic mutations in pairs of lung cancer primary and brain metastasis tumors. A, nonsilent mutations (present, blue; absent, gray) in the paired primary tumor and brain metastasis from PP1 (1P:1M) and PP4 (4P:4M). Known cancer genes are annotated on the right side of each plot. The total number of nonsilent mutations (percent shared mutations) is indicated at the bottom of each plot. B, two-dimensional plots of cancer cell fraction (CCF) estimated from the variant allele frequencies of all mutations, adjusted for tumor purity and local copy-number states. Increasing red intensity indicates higher mutation density. C, phylogenetic trees derived from all mutations by neighbor-joining method.
DISCUSSION

Our data demonstrate that somatic changes linked to activation of PI3K signaling in patients with stage IV SQCLCs are associated with a unique natural history of disease, reinforcing the preclinical data that show that these changes can engender a distinct biologic program that can be targeted for therapeutic gain.

The frequency of PI3K and FGFR1 pathway changes in our stage IV dataset is similar to those reported by TCGA. TCGA found upstream changes in PI3K signaling (PIK3CA mutations and PTEN mutations) in 21% of the resected early-stage cancers and FGFR1 amplification in 17% (3). We found PIK3CA and PTEN mutations in 21% of patients and FGFR1 amplification in 20%. In addition, we found PTEN loss by IHC in 29% of stage IV cases. A comparable frequency is unavailable from TCGA, as IHC was not performed, although Soria and colleagues (12) found PTEN loss by IHC in 24% of early-stage SQCLCs. The discrepancy between the frequencies of PTEN protein loss and genetic alterations such as copy-number loss or PTEN mutations suggests that transcriptional or posttranscriptional mechanisms are likely at play, such as promoter hypermethylation (12).

Although the mutation frequencies of some genes detected by IMPACT were significantly different from TCGA by the Fisher exact test, none reached significance when adjusted for multiple testing. An example is AKTI, which was mutated in 6% of patients in our series and 0.6% in TCGA (Fisher exact test \( P = 0.009\), adjusted \( P = 1 \)). These mutations occurred in either the pleckstrin homology domain or the kinase domain. Redistributing these patients into the PI3K-aberrant group (only two of whom did not have any other upstream PI3K pathway alteration and so were not already in this category) did not alter our results. A larger sample size will be required to determine whether these differences in frequency are indeed significant, as some variation is expected from the greater depth of coverage associated with targeted exon sequencing. Interestingly, many of the genes with the greatest differences in frequencies are involved in transcriptional regulation rather than signal transduction. These are, to our knowledge, the first published next-generation sequencing results from patients with metastatic SQCLCs.

Several groups have reported survival differences associated with FGFR1 amplification in early-stage SQCLCs with divergent results (13–16). Although preclinical modeling has clearly shown that FGFR1-amplified cell lines and patient-derived xenografts do respond to FGFR1 inhibition (4, 17) and preliminary data from clinical trials have shown modest activity of selective FGFR1 inhibitors in patients with FGFR1-amplified SQCLCs (18, 19), responses have not been uniform and amplification does not translate to increased protein expression in many cases (20). We therefore note that within this genomic subgroup, heterogeneity in the activation of FGFR1 signaling could exist that FGFR1 amplification by itself might not adequately capture. This may be a confounder in our analysis of this genotype. As this is the first study to address prognosis of FGFR1-amplified tumors in metastatic disease, a direct comparison with prior clinical data, which are exclusively in early-stage disease, is not possible.

To date, no group has yet reported clinical outcomes for patients with PI3K-aberrant SQCLCs. We showed, for the first time, that SQCLC patients with upstream PI3K aberrations have a significantly poorer survival compared with others. This holds true when taking into account other clinical and genomic variables on multivariate analysis. There were no differences in the number of lines of therapy between molecular subgroups. More patients received a targeted therapy in the FGFR1-amplified group, although the preliminary clinical activity of these agents is poor and this difference is unlikely to have favorably affected survival as a result (18, 19). The magnitude of the survival difference for patients with PI3K aberrations is substantial (HR for death, 5.6) and exceeds any therapeutically driven improvement in survival to date, arguing that stratification by this genotype category should be considered and PI3K status reported in future studies.

Consistent with this more aggressive phenotype, SQCLC patients with PI3K-aberrant cancers had a surprisingly higher burden of metastatic disease in general and a higher incidence of brain metastases specifically. At 27%, this incidence approaches that seen in lung adenocarcinomas (21). This observation was reproduced in the resected SQCLC brain metastases we analyzed, where two thirds exhibited complete loss of PTEN protein, all exhibited heterozygous loss of PTEN, and all had a pattern of gene expression consistent with PTEN loss.

With regard to functional implications, the retention of PTEN loss and PIK3CA mutations between the primary and metastatic sites suggests that PI3K signaling activation is necessary but not sufficient for the metastatic process to the brain. As for evidence implicating PI3K signaling in the development of brain metastases, this is by and large correlative. Saal and colleagues (9) demonstrated a correlation between PTEN loss and the development of distant metastases in breast cancer. Data from melanoma suggest that hyperactivation of AKT and PTEN loss is more common in brain metastases than extracranial metastases (22). Similar results correlating AKT and PTEN loss were found in an analysis of 52 breast cancer brain metastases (25% exhibited PTEN loss), though the lack of comparison to primary tumors that did not harbor brain metastases limits the interpretation of these data (23). SNP analysis of AKTI, AKT2, PIK3CA, PTEN, and FRAPI from predominantly lung adenocarcinoma (68%) primary tumors yielded an association between certain SNPs in AKTI, AKT2, and PIK3CA and risk of brain metastasis (24). The functional relevance of these SNPs is unknown, however. The probrain metastatic program triggered by PI3K pathway activation remains, as a result, undefined, although PTEN loss in particular is known to facilitate integrin-mediated cell spreading through impairment of cytoskeletal turnover (25). Other mechanisms that are important in the development of metastases in lung adenocarcinomas, including upregulated WNT signaling and overexpression of serpins, were not present in our series, suggesting that the mechanism that drives the metastatic process to the brain differs by histology (10, 11). Potential prometastatic targets of interest from our study include \( \gamma \)- and \( \beta \)-clustered protocadherins on 5q31 (PCDHGAA4, PCDHGA10, PCDHGB2, and PCDHIB6) that are known to play a role in cell adhesion (26), which were clonal in the brain metastases and which were expressed at low levels relative to the TCGA primary lung dataset (Supplementary Fig. S4).
Intriguingly, although there were no clonal differences in PTEN loss or PIK3CA mutations between the primary lung cancer and brain metastasis pairs, deeper genomic analysis did demonstrate evidence of clonal heterogeneity. The low degree of shared events (mutation and copy number) between the paired samples and the presence of multiple subclonal populations within each tumor sample indicates a high degree of genetic heterogeneity both within and between the lung primaries and the brain metastases analyzed. Thus, the originating brain metastatic clone may very well be different from the predominant primary lung cancer clone. Indeed, two recent studies of early-stage NSCLCs reported a significant amount of mutational and copy-number heterogeneity by multiregion sequencing in surgical samples of early-stage patients (27, 28). Our data show evidence of even greater heterogeneity in the metastatic setting, with as little as 15% of mutations shared between sites of disease, substantially less than the 76% of mutations shared across regions of primary lung tumors in early-stage disease (27). Notwithstanding blood–brain barrier penetration by drugs, this offers an additional explanation—one which has the potential to be better elucidated—as to why metastatic brain disease is often difficult to treat.

In conclusion, our data provide evidence of a unique biologic program associated with aberrations known to activate PI3K signaling in SQCLCs. This is the first such evidence of its kind in this disease, and it simultaneously validates the analysis performed by TCGA in resection specimens and supports the current efforts to target this pathway pharmacologically. Although it follows that targeting PI3K signaling may help to prevent or treat brain metastases in patients with SQCLCs, functional studies will be needed to determine the exact role of this pathway in the metastatic process. Our analysis of paired SQCLC primary lung and brain metastasis samples revealed the presence of copy-number differences and subclonal somatic mutations between the lung primary and brain metastasis tumors as well as a low proportion of shared events between the two, indicating a clonally divergent evolutionary process and the requirement for multiregion sampling to track the origin of the metastatic subclone and to identify other alterations that contribute to the metastatic process to the brain.

**METHODS**

**Study Design and Patients**

Patients with pathologically verified SQCLC with available tissue were approached for informed consent to an Institutional Review Board (IRB)/Privacy Board–approved protocol for testing of archived FFPE tumor specimens through a prospective testing platform termed SQ-MAP. This study was conducted in accordance with the Declaration of Helsinki. There were no other eligibility criteria for this. Although no stage restrictions were present for testing, the current analysis is limited to patients with stage IV disease. Seventy-nine patients were identified for this study who had tumors sufficient for FISH for FGFR1 amplification, mass spectrometry–based genotyping (Sequenom) for PIK3CA mutations (exons 9 and 20), IHC for PTEN loss, and targeted-exon sequencing for a large panel of mutations and copy-number changes as described below.

Resected SQCLC brain lesions with FFPE tumor blocks and frozen specimens as well as their paired primary lung tumor FFPE biopsies/resections were identified through an IRB/Privacy Board–approved Biospecimen Utilization and Waiver of Authorization.

**Pathologic Verification**

All pathologic samples were reviewed by thoracic pathologists at our institution using light microscopy and IHC for TTF1 and p63 (AN isoform) and selectively for Napsin A and muscularin as described previously (29).

**FISH for FGFR1 Amplification and Chromosome 5p Gain**

FGFR1 amplification was determined through the use of an FGFR1/CEN8 dual-color FISH probe from Zytovision (#Z-2072-200) in a CLIA laboratory. Sections (4 μm) were generated from FFPE blocks for testing. Tumor specimens were deparaffinized in xylene and dehydrated in ethanol. FISH was performed according to the manufacturer’s instructions with minor modifications. FISH analysis and signal capture were performed on fluorescence microscopes (AXIO, Zeiss) coupled with the ISIS FISH Imaging System (Metasystems). A minimum of 50 interphase nuclei from each tumor specimen were scored. An FGFR1/CEN8 ratio ≥2 in ≥20% tumor cells was used as the criterion to define FGFR1 amplification.

Gain of chromosome 5p was verified through the use of a 5p13 FISH probe, RP11-767K20, from Empire Genomics. The BAC clone RP11-767K20 was labeled by TAMRA (red signal). Sections (4 μm) were generated from FFPE blocks for testing. FISH analysis and signal capture were performed on fluorescence microscopes (AXIO, Zeiss) coupled with the ISIS FISH Imaging System (Metasystems). A minimum of 100 interphase nuclei from each tumor specimen were scored. There was sufficient archival material to assess chromosome 5p status in the PP1, PP2, PP3, and PP4 brain metastases as well as the PP1 primary lung tumor.

**PTEN Loss**

IHC for PTEN expression was performed using a mouse monoclonal antibody (Cell Signaling; clone 138G6) on archival FFPE tissue sections in a CLIA laboratory. Heat-induced epitope retrieval at 97 to 99°C for 40 minutes was accomplished using Target Retrieval Solution pH 9.0 (DAKO) followed by incubation with primary PTEN antibody (dilution 1:50). Diaminobenzidin was used as the chromogen. Positive cytoplasmic and/or nuclear staining of blood vessel endothelial cells was used as the internal positive control. H-score was calculated according to the following formula: H-Score = ∑(I × N) / (N), where I is the intensity (range, 1–3), where an H-score ≥600 was considered PTEN negative.

**Mutation Analysis by Sequenom**

Tumors were genotyped by the Sequenom Mass ARRAY system (Sequenom, Inc.). Briefly, samples were tested in duplicate using a series of multiplexed assays designed to interrogate hot-spot mutations in 8 oncoproteins: EGFR, KRAS, BRAF, PIK3CA, NRAS, ERBB2, HER2, MAP2K1, MEK1, and AKT1. A total of 92 nonsynonymous mutations were tested in 6 multiplex reactions. Genomic DNA amplification and single base pair extension steps were performed using specific primers designed with the Sequenom Assay Designer v3.1 software. The allele-specific single base extension products were then quantitatively analyzed using matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS) on the Sequenom MassArray Spectrometer. All automated mutation calls were confirmed by manual (visual) review of the spectra.

**Next-Generation Sequencing through IMPACT**

Genomic alterations in key cancer-associated genes were also profiled using exon capture by hybridization followed by next-generation sequencing (30). This assay, termed IMPACT, encompasses all protein-coding exons and select introns of 279 cancer genes,
PI3K and Brain Metastases in Squamous Cell Lung Cancers

...and was the basis for a clinical version of this assay, termed MSK-IMPACT, for which methods and validation data are presented in detail (31). Genes were selected to include commonly implicated oncogenes, tumor suppressor genes, and components of pathways deemed actionable by current targeted therapies (Supplementary Table S3). We sheared DNA isolated from FFPE tissue for 6 minutes on a Covaris E220 instrument, prepared barcoded sequence libraries (New England Biolabs; Kapa Biosystems), and performed exon capture on barcoded pools by hybridization (Nimblegen SeqCap) using custom oligonucleotides. The amount of input DNA ranged from 112 to 500 ng (mean, 250 ng). Barcoded libraries were pooled at 100 ng per sample and used in a single exon capture reaction as previously described (32). To prevent off-target hybridization, we spiked in a pool of blocker oligonucleotides complementary to the full sequences of all barcoded adaptors to a final concentration of 10 μmol/L. Hybridized DNA was subsequentially sequenced on a single lane of an Illumina HiSeq to generate paired-end 75-bp reads or 100-bp reads. Data were demultiplexed using CASAVA, and reads were aligned to the reference human genome (hg19) using the Burrows-Wheeler Alignment tool (33). Local realignment and quality score recalibration were performed using the Genome Analysis Toolkit (GATK) according to GATK best practices (34). We obtained a mean unique on-target sequence depth of 355-fold (range, 60–1,249-fold).

Sequence data were analyzed to identify three classes of DNA alterations: single-nucleotide variants, small insertions/deletions (indels), and copy-number alterations. Single-nucleotide variants were called using muTect (35). For 31 of 79 tumors with patient-matched normal DNA, the somatic status could be inferred directly. For 48 of 79 tumors without matched normal DNA, we filtered out all silent variants and all additional variants in dbSNP but not in COSMIC (Catalogue of Somatic Mutations in Cancer; ref. 36). Indels were called using the SomaticIndelDetector tool in GATK. All candidate mutations and indels were reviewed manually using the Integrative Genomics Viewer. The mean sequence coverage was calculated using the DepthOfCoverage tool in GATK, which was used to compute copy number as described previously (32). For this analysis, all copy-number alteration calls were normalized against the average ploidy of the tumor sample. Binomial exact CIs for select gene alterations were calculated with the Clopper-Pearson method.

Whole-Exome Sequencing Somatic Mutation Analysis

Paired-end whole-exome sequencing was performed on tumor trios (primary tumor, brain metastasis, and matched normal; database of Genotypes and Phenotypes accession number phs000907.v1.p1). Somatic mutation detection was performed using standard algorithms by the Memorial Sloan Kettering Cancer Center (MSKCC) Bioinformatics Core. Briefly, the raw reads were aligned to the hg19 b37 version of the human reference genome using the Burrows-Wheeler Alignment algorithm (33). Preprocessing steps include removal of PCR duplicates using the MarkDuplicates tool in Picard (37), realignment around indels, and quality scores recalibration using the GATK analysis toolkit (34). Somatic mutation calls were generated by the MuTect algorithm (35), and variant annotation was done using Oncotator (38). Each gene was scored and ranked by functional categories for the somatic mutation observed in the gene (truncation, deleterious missense, nondeleterious missense, silent) and the number of mutations observed across samples. Low expression genes as inferred from the RNA-sequencing data were removed. Gene set enrichment analysis was then performed to look for pathways and biologic sets significantly enriched in the top-ranking mutated genes.

Whole-Exome Sequencing Copy-Number Analysis

To identify copy-number alterations from whole-exome sequencing data, we applied seqDNAcopy (39), which extends the circular binary segmentation algorithm to next-generation sequencing data. Similar to the log-ratio (logR) from copy-number array, logR from whole-exome sequencing is computed based on binned read counts between the tumor and the matched normal. Copy-number abnormalities in the tumor sample cause changes in the logR values reflecting the true underlying copy-number ratio. The logR measure was scaled by the ratio of total read counts in the tumor and normal. Guanine-cytosine bias observed in logR was corrected using loess regression. Off-target reads were removed to reduce noise. The variability of the logR from whole-exome sequencing is expected to be inversely proportional to the average read count. Therefore, an inverse variance weighted version of circular binary segmentation was used for copy-number segmentation. seqDNAcopy was implemented in R. The Integrative Genomics Viewer was then used to visualize genome-wide copy-number profiles across samples.

Analysis of Clonality

We developed a statistical and computational pipeline called FACETS that leverages rich information from germline polymorphic sites to estimate tumor purity, ploidy, and allele-specific copy number in DNA sequencing data of tumor-normal sample pairs (40). The output of FACETS also includes cellular fraction estimates to identify clonal versus subclonal copy-number aberrations. Purity estimation by FACETS relative to pathologic determination of tumor purity is shown in Supplementary Table S6. The purity, ploidy, and allele-specific copy-number estimates were then used as PyClone (41) input to obtain estimates of cancer cell fraction for somatic mutations by correcting for tumor purity and local copy-number states. Phylogenetic tree analysis was performed using Hamming distance computed using all somatic mutations and neighbor-joining method.

PTEN, WNT, and Serpin Gene Expression Analyses

The PTEN loss signature genes and WNT lung cancer metastasis signature genes were obtained from Saal and colleagues (9) and Nguyen and colleagues (10), respectively. Log-transformed normalized counts were obtained from RNA-sequencing data and z-score transformed. RNA-seq gene expression data from the TCGA study (primary lung tumors) were similarly normalized. The PTEN and WNT signature scores were then computed by aggregating signed expression values of the signature genes (signs obtained from the respective studies). Positive signs indicate positive correlation of signature gene expression with PTEN loss/WNT signaling, and negative signs indicate negative correlation of signature gene expression with PTEN loss/WNT signaling. Samples were then sorted by the PTEN and WNT signature scores. PTEN and WNT signature activities in the brain metastases were compared with the signatures in the TCGA primary tumors using a one-sided Wilcoxon rank-sum test. Gene expression for key prometastatic serpins described in Valiente and colleagues (11) was assessed in the brain metastases and compared with expression in the TCGA primary tumors using a one-sided Wilcoxon rank-sum test.

Disclosure of Potential Conflicts of Interest

P.K. Paik reports receiving commercial research support from AstraZeneca and Novartis. N. Rekhtman is a member of the NCCN/BMS pathology advisory panel. L. Wang is a consultant/advisory board member for Exelixis. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions

Conception and design: P.K. Paik, R. Shen, M.F. Berger
Development of methodology: P.K. Paik, R. Shen, H. Won, N. Rekhtman, L. Wang, M.G. Kris
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): P.K. Paik, N. Rekhtman, M. Ladanyi, M.F. Berger, M.G. Kris

...
RESEARCH ARTICLE

Analysis and interpretation of data (e.g., statistical analysis, bioinformatics, computational analysis): P.K. Paik, R. Shen, H. Won, N. Rekhtman, C.S. Sima, A. Arora, V. Seshan, M. Ladanyi, M.F. Berger, M.G. Kris

Writing, review, and/or revision of the manuscript: P.K. Paik, R. Shen, H. Won, N. Rekhtman, L. Wang, C.S. Sima, V. Seshan, M. Ladanyi, M.G. Kris

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): P.K. Paik, H. Won, M. Ladanyi, M.G. Kris

Study supervision: P.K. Paik, M.F. Berger, M.G. Kris

Acknowledgments

The authors thank Dr. Laetitia Boursu for supervision of Sequenom assays, Edyta Brzostowska for her administrative supervision of SQ-MAP testing, the members of the MSKCC Bioinformatics Core, and the MSKCC Brain Tumor Center for provision of resected metastatic SQCLC specimens.

Grant Support

This study was supported by the Geoffrey Beene Cancer Research Fund at Memorial Sloan Kettering and a Conquer Cancer Foundation of ASCO Career Development Award. The MSKCC Sequenom facility was supported by the Anbinder Fund.

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 September 24, 2014; revised March 12, 2015; accepted March 27, 2015; published OnlineFirst April 30, 2015.

REFERENCES


620 | CANCER DISCOVERY | JUNE 2015 www.aacrjournals.org

Published OnlineFirst April 30, 2015; DOI: 10.1158/2159-8290.CD-14-1129

Downloaded from cancerdiscovery.aacrjournals.org on June 30, 2017. © 2015 American Association for Cancer Research.
Next-Generation Sequencing of Stage IV Squamous Cell Lung Cancers Reveals an Association of PI3K Aberrations and Evidence of Clonal Heterogeneity in Patients with Brain Metastases

Paul K. Paik, Ronglai Shen, Helen Won, et al.

Cancer Discovery 2015;5:610-621. Published OnlineFirst April 30, 2015.

Updated version
Access the most recent version of this article at:
doi:10.1158/2159-8290.CD-14-1129

Supplementary Material
Access the most recent supplemental material at:
http://cancerdiscovery.aacrjournals.org/content/suppl/2015/04/30/2159-8290.CD-14-1129.DC1

Cited articles
This article cites 36 articles, 16 of which you can access for free at:
http://cancerdiscovery.aacrjournals.org/content/5/6/610.full#ref-list-1

Citing articles
This article has been cited by 11 HighWire-hosted articles. Access the articles at:
http://cancerdiscovery.aacrjournals.org/content/5/6/610.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.