**EGFR Exon 20 Insertion Mutations in Lung Adenocarcinomas: Prevalence, Molecular Heterogeneity, and Clinicopathologic Characteristics**

Maria E. Arcila1, Khedoudja Nafa1, Jamie E. Chaft2, Natasha Rekhtman1, Christopher Lau1, Boris A. Reva3, Maureen F. Zakowski1, Mark G. Kris2, and Marc Ladanyi1,4

**Abstract**

In contrast to other primary epidermal growth factor receptor (EGFR) mutations in lung adenocarcinomas, insertions in exon 20 of EGFR have been generally associated with resistance to EGFR-tyrosine kinase inhibitors. Their molecular spectrum, clinicopathologic characteristics, and prevalence are not well established. Tumors harboring EGFR exon 20 insertions were identified through an algorithmic screen of 1,500 lung adenocarcinomas. Cases were first tested for common mutations in EGFR (exons 19 and 21) and KRAS (exon 2) and, if negative, further analyzed for EGFR exon 20 insertions. All samples underwent extended genotyping for other driver mutations in EGFR, KRAS, BRAF, ERBB2/HER2, NRAS, PIK3CA, MEK1, and AKT by mass spectrometry; a subset was evaluated for ALK rearrangements. We identified 33 EGFR exon 20 insertion cases [2.2%, 95% confidence interval (CI), 1.6–3.1], all mutually exclusive with mutations in the other genes tested (except PIK3CA). They were more common among never-smokers (P < 0.0001). There was no association with age, sex, race, or stage. Morphologically, tumors were similar to those with common EGFR mutations but with frequent solid histology. Insertions were highly variable in position and size, ranging from 3 to 12 bp, resulting in 13 different insertions, which, by molecular modeling, are predicted to have potentially different effects on erlotinib binding. EGFR exon 20 insertion testing identifies a distinct subset of lung adenocarcinomas, accounting for at least 9% of all EGFR-mutated cases, representing the third most common type of EGFR mutation after exon 19 deletions and L858R. Insertions are structurally heterogeneous with potential implications for response to EGFR inhibitors. *Mol Cancer Ther*; 12(2); 220–9. ©2012 AACR.

**Introduction**

The identification of activating mutations within the tyrosine kinase (TK) domain of the epidermal growth factor receptor (EGFR) has transformed the management of patients with non–small cell lung cancers. Starting with the initial studies, 2 mutation types have been recognized as the most prevalent and clinically significant: in-frame deletions in exon 19 and the point mutation L858R (1–3). Together, these represent approximately 90% of all EGFR mutations, and their association with response to tyrosine kinase inhibitors (TKI) is well characterized. Mutations involving codons G719 and L861 are also associated with sensitivity but their incidence is much lower.

Insertions in exon 20 are included among the rarer activating mutations in the TK domain of EGFR (4–9). They represent a combination of in-frame insertions and/or duplications of 3 to 21 base pairs (bp), predominantly clustered between codons 767 and 774. Importantly, in contrast to the more classic activating EGFR mutations, these insertions have been associated with de novo resistance to approved EGFR-TKIs (erlotinib and gefitinib; refs. 10–14) and to irreversible inhibitors that have recently entered clinical trials (neratinib, afatinib, and dacomitinib; refs. 10–16). In vitro studies show that cells harboring some of the most prevalent insertions require an average of 100-fold higher concentrations of these agents for inhibition, well beyond clinically achievable plasma levels. Clinical studies, although limited, confirm the preclinical findings (6, 8, 9, 12, 15–20), but rare cases with better clinical responses have been reported (8, 18, 20). Importantly, many of the insertions identified in patient samples have not been tested against these inhibitors. Further understanding of...
the biology, prognostic, and predictive implications of these mutations is needed but has remained limited by the small number of patients included in clinical trials and the lack of preclinical models, such as patient-derived cell lines or genetically engineered mouse models.

Despite the importance of EGFR exon 20 insertions as potentially targetable driver mutations, to date, only a few reports have been dedicated to these tumors and most have been confined to East Asian populations. In this setting, with the exception of EGFR-TKI sensitivity, the clinical and pathologic characteristics seem to closely match those of the classic EGFR mutations, including predilection for females, never-smokers, and adenocarcinoma histology. While the true incidence of these mutations is not yet well defined, with reports ranging from 0% to 13% (4, 6–8, 21, 22), reviews have suggested that insertions in exon 20 may represent up to 4% of all EGFR mutations (23). The incidence, clinicopathologic characteristics, and molecular spectrum of these mutant tumors remain to be explored in the U.S. population.

The aim of the current study was (i) to determine the frequency and molecular spectrum of EGFR exon 20 insertions in a large cohort of patients with lung adenocarcinomas, (ii) to assess the clinical and histopathologic characteristics, and (iii) to confirm their mutually exclusive nature with mutations in EGFR, KRAS, BRAF, ERBB2/HER2, NRAS, PIK3CA, MAP2K1/MEK1, and AKT as well as ALK rearrangements.

Materials and Methods

Patients and mutation analysis

Clinical cases of lung adenocarcinomas received for routine EGFR and KRAS testing at Memorial Sloan-Kettering Cancer Center (New York, NY) between January 2009 and January 2011 were selected for the study, under an Institutional Review Board-approved waiver. The study period was chosen to allow a minimum of 1 year of potential follow-up time.

Clinical testing for the detection of major mutations in EGFR (exon 19 deletions and L858R) and KRAS (exon 2) was carried out by fragment analysis and Sanger sequencing, respectively, using previously described methods (24, 25). Extended mutation analysis for other recurrent point mutations in EGFR, KRAS, BRAF, ERBB2/HER2, NRAS, AKT, MAP2K1, and PIK3CA was conducted in all cases by mass spectrometry genotyping (Sequenom) as previously described (26). Briefly, samples were subjected to a series of multiplexed assays designed to interrogate a total of 92 nonsynonymous mutations in 6 multiplex reactions (see Supplementary Table S1 for complete list of tested mutations). Amplification and single base pair extension primers were designed with the Sequenom Assay Designer v3.1 software. Allele-specific single base extension products were quantitatively analyzed using matrix-assisted laser desorption/ionization-time of flight/mass spectrometry (MALDI-TOF/MS) on the Sequenom MassArray Spectrometer. All automated system mutation calls were confirmed by manual review of the spectra. All testing was carried out in duplicate.

When sufficient tissue was available, samples that were EGFR/KRAS wild-type were also tested for ALK rearrangements by fluorescent in situ hybridization (Vysis ALK Break Apart FISH Probe Kit) using standard protocols (Supplementary Fig. S1 outlines the sequential genotyping algorithm).

Testing for EGFR exon 20 insertions

Assessment for insertions in exon 20 of EGFR primarily targeted cases known to be negative for major EGFR (exon 19 del, L858R) and KRAS mutations, given previous reports of their mutual exclusivity and further based on DNA availability. Initial screening was conducted by a sizing assay (24, 27) using primers FW1:5'-TCCTCA-CCTGGAAGGGTCCA-3' and REV1:5'-Fam-TGCCAC-CTCCACTCCGTCTA-3'. Positive cases were characterized by Sanger sequencing using primers FW1:5'-CATGCGTCTTCCGT-3' and REV1:5'-GTATAGGGGTACCGTGGTGAG-3' following previously described protocols.

To confirm the mutually exclusivity of EGFR exon 20 insertions with major EGFR and KRAS mutations, as well as other rarer mutations not well represented in our cohort, we tested a separate set of adenocarcinomas with a known positive mutation profile and sufficient DNA. Also, to confirm that mutations were confined to adenocarcinomas, we tested sets of squamous and small cell carcinomas following similar protocols.

Histopathology

Morphologic analysis was conducted by semiquantitatively recording 6 patterns—lepidic (bronchioalveolar), acinar, papillary, micropapillary, solid, and mucinous (28). The distribution of morphologic patterns in adenocarcinomas with exon 20 insertions was compared with a control group of adenocarcinomas with canonical EGFR mutations. The groups were compared by 2-tailed Fisher exact test.

Statistical analysis

The association between EGFR mutation status and clinical and biologic characteristics was analyzed by Fisher exact test. Age differences were compared using the t test for independent samples. The 2-sided significance level was set at P < 0.05.

Prediction of functional impact of exon 20 insertions

To examine the likelihood that previously unreported mutations identified in our series would have similar impact on the function of the protein, we used a computational biology approach using the publicly available Mutation Assessor software as described in detail elsewhere (29).
### Table 1. EGFR exon 20 insertions identified in the study

<table>
<thead>
<tr>
<th>Size (total)</th>
<th>n</th>
<th>Nucleotide sequence</th>
<th>CDS mutation (inserted sequence)</th>
<th>Amino acid mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td></td>
<td>GAA GCC TAC GTC ATG GCC AGGTG GCC AAC CCC CAT GTC GCC GCT TG</td>
<td>c.2311_2312ins9</td>
<td>p.D770_N771insSVD</td>
</tr>
<tr>
<td>9bp ins (15)</td>
<td></td>
<td>GAA GCC TAC GTC ATG GCC AGGTG GCC AAC CCC CAT GTC GCC GCT TG</td>
<td>[GCGTGGACA duplication]</td>
<td>p.D770_N771insSVD</td>
</tr>
<tr>
<td>6bp ins (9)*</td>
<td></td>
<td>GAA GCC TAC GTC ATG GCC AGGTG GCC AAC CCC CAT GTC GCC GCT TG</td>
<td>c.2319_2320ins9</td>
<td>p.H773_V774insNPH</td>
</tr>
<tr>
<td>3bp ins (6)</td>
<td></td>
<td>GAA GCC TAC GTC ATG GCC AGGTG GCC AAC CCC CAT GTC GCC GCT TG</td>
<td>[CCACGGTGAC duplication]</td>
<td>p.D770_N771insG</td>
</tr>
<tr>
<td>12bp ins (3)</td>
<td></td>
<td>tccag GAAGCCT ATG TAC GTC ATG GCC AGGTG GCC AAC CCC CAT GTC GCC GCT TG</td>
<td>[CCACGGTGAC duplication]</td>
<td>p.A763_Y764insFQA</td>
</tr>
</tbody>
</table>

*One 6-bp insertion could not be characterized by Sanger sequencing due to very low mutant peaks.*
Results

Initial screening
A total of 1,500 adenocarcinomas were reviewed for the study and screened under the clinical and extended mass spectrometry genotyping assays. Of these, 901 were positive for mutations (60%, 901 of 1,500), 36 had ALK rearrangements (437 tested), and 563 had no alterations detected. The latter group of tumors was termed "pan-negative." The detailed distribution of mutations is outlined in Supplementary Table S2.

EGFR exon 20 insertion testing
A total of 600 cases were tested. This included 464 pan-negative tumors (of the 563 above cases negative for all other mutations and ALK rearrangements) and 136 mutation-positive tumors as specified in Supplementary Table S3. We could not test the remaining 99 of the 563 pan-negative cases due to unavailable material. Among the tested group, we detected 33 insertion mutations (6% of tested, 7% of the pan-negative set), all mutually exclusive with other genetic alterations except for 2 with concurrently mutated PIK3CA (both H1047R). Insertions of 9 bps were the most common mutation type (48%, 16 of 32). Sanger sequencing of 32 positive cases showed all mutations were confined to the 5' end of the exon, between codons A763 and C775 comprising duplications and insertions as outlined in Table 1. At the amino acid level, mutations were highly variable with 13 different types identified. One sample harbored a concurrent D770N point mutation. The specific insertion sequence could not be ascertained in one tumor (6-bp insertion) with very low mutant peaks due to very low tumor content. EGFR exon 20 insertions represented 9% (33 of 367) of all EGFR-mutated tumors (Fig. 1).

Testing of the separate set of adenocarcinomas with a known positive mutation profile (n = 311, 70 EGFR ex 19 del, 70 L858R, 120 KRAS G12dG13, 7 NRAS, 3 MAP2K1, 2 AKT, 30 BRAF) identified no EGFR exon 20 insertions, confirming their mutually exclusive relationship with these other driver mutations. No EGFR exon 20 insertions were found among 105 squamous cell carcinomas and 8 small cell carcinomas tested.

Morphologic features
Morphologically, all tumors were highly heterogeneous with a mixture of various patterns. Predominant patterns included acinar/papillary/micropapillary (n = 21; 70%), lepidic (n = 5; 17%), and solid (n = 4; 13%); all tumors were entirely nonmucinous. This distribution of morphologic patterns was similar to the control group of 36 adenocarcinomas with classic sensitizing EGFR mutations in exon 19 and 21, although tumors with exon 20 insertions showed a trend for a greater proportion of solid component, but this was not statistically significant (Supplementary Table S4 and Supplementary Fig. S2).

Predicted functional impact
Computational prediction of the functional impact of exon 20 insertions showed functional scores ranging from 2 to 3.6 corresponding to the medium to high functional impact categories. Insertions in codons 762 and 766 affect residues conserved in the entire family of tyrosine kinases; insertions in codons 768 to 769 affect residues conserved in the large specific subfamily of EGFR homologs; insertions in codons 774 and 775 affect residues that are conserved both across all tyrosine kinase homologs and within the specific EGFR subfamily. The point mutation D770N, concurrently found with one H773_V774insNPH, had a low impact score of 1.2 and not likely to be comparable with the impact of the associated insertion.

Clinical characteristics
The clinical characteristics of patients with tumors harboring EGFR exon 20 insertions are summarized in Table 2. Sixty-seven percent of patients were female, 48% were never-smokers, and 12% were of Asian descent. When compared with patients whose tumors lacked them, EGFR exon 20 insertions were more common among never smokers (P < 0.0001), but there was no significant difference in age, sex, ethnic origin, or stage at diagnosis. No significant differences were noted in comparison to patients with classic sensitizing EGFR mutations (including EGFR ex 19 del, L858R, L861Q, and G719 mutations). The proportion of EGFR exon 20

www.aacnjournals.org  Mol Cancer Ther; 12(2) February 2013 223

Downloaded from mct.aacnjournals.org on June 20, 2017. © 2013 American Association for Cancer Research.
insertions among all *EGFR* mutations was the same (9%) for both the Caucasian and the Asian patient subsets.

**Response to treatment and survival analysis**

Of the 33 patients with exon 20 insertions, 5 received erlotinib for advanced disease: 1 patient with A763_Y764insFQEA was treated with erlotinib in combination with chemotherapy and had partial response; 1 patient with V774_C775insHV was treated with a combination of chemotherapy and gefitinib with partial response followed by 1 year of erlotinib maintenance before disease progression, at which time he was switched to neratinib without benefit; 2 patients with D770_N771insGT and V774_C775insHV were treated with erlotinib as single agent with no response. Finally, one patient with V769_D770insASV was lost to follow-up.

For the 15 patients who presented with advanced disease, the median overall survival was more than 4 years. These patients received combinations of standard chemotherapies including cisplatin or carboplatin with a taxane or pemetrexed. Two patients with remarkable survival had multimodality therapies, 1 with resection of multifocal lung lesions and another with unilateral surgery and contralateral radiation therapy. Both patients had prolonged disease control with these interventions. Other patients received standard chemotherapy agents with typical or less-than-average duration of benefit.

**Discussion**

Insertions in exon 20 are a subset of activating *EGFR* mutations primarily known for their reported association with *de novo* resistance to TKIs. To date, however, few studies have focused on this subset, each confined to a limited number of mutation-positive East Asian patients (4–9). Many of the mutations appear in the literature a single time, some anecdotally associated with response to EGFR inhibitors. It is therefore difficult, even with the combination of these studies, to draw conclusions as to the true prevalence of these mutations, their molecular spectrum, clinicopathologic characteristics or their pattern of resistance. Table 3 summarizes the largest studies and the mutations identified.

To our knowledge, our study represents the largest assessment for *EGFR* exon 20 insertions and the most comprehensive analysis for other mutations in the same cohort. We used an algorithmic approach for our initial screening, focusing on the group negative for major mutations in *EGFR* and KRAS, given the previous reports of their mutually exclusive relationship (7). On the basis of this analysis, we identified 33 patients with insertions, corresponding to 9% of all *EGFR*-mutated samples. We estimate, however, that the true incidence may be even higher, closer to 11%, factoring in the expected positive cases, which would have been detected if our entire driver mutation-negative group had been tested (99 "mutation negative" samples were not tested due to unavailable DNA). This rate has been further validated by our subsequent clinical testing data for the year 2011, following the inclusion of *EGFR* exon 20 insertion analysis as part of our standard reflex testing of clinical samples. During an 8-month period, 19 additional *EGFR* exon 20 insertion cases were detected among 179 *EGFR*-mutant samples (19 of 179 or 11%, 95% confidence interval (CI), 7–16). This rate is consistent with the highest rates previously reported in smaller studies (6, 7), confirming that exon 20 insertions are the third most common *EGFR* mutation after exon 19 deletions and L858R. The overall underestimation of *EGFR* exon 20 insertions in the literature may reflect the fact that many studies have focused on the 2 major mutations and that indels, especially in the setting of low tumor content, may occasionally be mistaken for "high background" on Sanger sequencing traces, a pitfall that is avoided by the simple and more sensitive PCR product sizing assay used in the present study.

We estimate that the overall incidence of *EGFR* exon 20 insertions among all adenocarcinomas is approximately

---

**Table 2. Clinical characteristics of patients with *EGFR* exon 20 insertions and comparison to all other patients and to patients with other *EGFR* mutations**

<table>
<thead>
<tr>
<th></th>
<th><em>EGFR</em> exon 20 insertions (n = 33)</th>
<th><em>No EGFR</em> exon 20 ins (n = 1368)</th>
<th>Other <em>EGFR</em> mutations (n = 291)*a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (female/male)</td>
<td>22/11</td>
<td>867/501</td>
<td>212/79</td>
</tr>
<tr>
<td>Median age (range)</td>
<td>66 (38–85)</td>
<td>66 (20–96)</td>
<td>66 (32–90)</td>
</tr>
<tr>
<td>Smoking status (never/former or current)</td>
<td>16/17b</td>
<td>347/1,021</td>
<td>138/153</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I–II/III–IV</td>
<td>18/15</td>
<td>829/739</td>
<td>121/170</td>
</tr>
<tr>
<td>II/IV</td>
<td>17/16</td>
<td>521/847</td>
<td>106/185</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian/Caucasian</td>
<td>4/28</td>
<td>74/1,272</td>
<td>32/249</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>22</td>
<td>10</td>
</tr>
</tbody>
</table>

*P = 0.0045 (no other comparisons were statistically significant).*
Through concurrent extended mass spectrometry genotyping and additional testing of known positive EGFR and KRAS cases, we also confirmed the mutually exclusive nature of these mutations with all other tested oncogenes, with the exception of PIK3CA. The coexistence of PIK3CA mutations with other oncogene mutations is a frequent event in lung adenocarcinomas (30).

At the molecular level, in agreement with prior studies (7, 8, 11), we found that insertions in exon 20 are all in-frame and confined to the proximal region of the exon, between codons 763 and 775. Compared with other studies, we identified greater heterogeneity with 13 different mutation types within the hotspot region, varying significantly in size and position. Insertions such as V769_D770insASV, previously considered among the most prevalent insertions (7, 8, 23), represented only 12% (4 of 33) of the mutations in our series. In contrast, the rarely reported mutation A763_Y764insFQEA, a duplication spanning the intron–exon junction (Fig. 2), represented 9% of all cases and several other insertions had not been previously reported. On the basis of our experience over a 3-year period, we have identified 20 different insertions, confirming their wide variability. This degree of heterogeneity is unlike other insertions such as those in exon 19 of EGFR (31) or in exon 20 of HER2 (32), where the vast majority of inserted sequences share the same length and amino acid content. Similarly, contrasting features can even be noted in comparison with EGFR deletions, which involve a common defining region within E746 to A750 (Fig. 3). How this structural heterogeneity may impact on biologic behavior and response to targeted therapy is not yet known.

<table>
<thead>
<tr>
<th>Study</th>
<th>Year (EGFR ex20/total)</th>
<th>% total (EGFR ex20/total)</th>
<th>% of EGFR mutant (EGFR ex20/total)</th>
<th>Reported mutations</th>
<th>Cosmic database nomenclature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huang and colleagues</td>
<td>2004 2% (2/101)</td>
<td>5.1% (2/39)</td>
<td>D761_E762insEAFQ(1)</td>
<td>D763_E764insFQEA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S768_D770dup(1)</td>
<td>D770_N771insSVD</td>
<td>p.A761_Y762insEAFQ?</td>
</tr>
<tr>
<td>Kosaka and colleagues</td>
<td>2004 1.4% (4/277)</td>
<td>3.6% (4/111)</td>
<td></td>
<td>D770_N771insSVD</td>
<td>p.A767_S768insTLA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p.V769_D770insASV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p.D770-GY</td>
</tr>
<tr>
<td>Shigematsu and</td>
<td>2005 1.9% (12/617)</td>
<td>8.9% (12/134)</td>
<td>ASV770-772ins (4)</td>
<td>p.V769_D770insASV</td>
<td></td>
</tr>
<tr>
<td>colleagues</td>
<td></td>
<td></td>
<td>H774ins (2)</td>
<td>p.V773_C774insH</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G771ins (1)</td>
<td>p.D770_N771insSVD</td>
<td>p.A767_S768insTLA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NP773-774ins, H775Y(1)</td>
<td>p.H773_V774insPH</td>
<td>p.D770_N771insG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PH774-775ins(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NPH774-776ins(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HV775-776ins(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mitsudomi and</td>
<td>2005 0% (0/59)</td>
<td>0% (0/33)</td>
<td>774_776insNP(2)</td>
<td>p.H773_V774insNPH</td>
<td></td>
</tr>
<tr>
<td>colleagues</td>
<td></td>
<td></td>
<td>770_772insASV(1)</td>
<td>p.V769_D770insASV</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>771_773insSVD(1)</td>
<td>p.D770_N771insSVD</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>772_773insV(1)</td>
<td>p.P772_H773insV</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>772_773insV(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>777ins(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequist and</td>
<td>2007 1.8% (5/278)</td>
<td>7% (5/68)</td>
<td>D770_N771 ins SVD(1)</td>
<td>V774_C775insHV</td>
<td></td>
</tr>
<tr>
<td>colleagues</td>
<td></td>
<td></td>
<td>D770 del ins Gl(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N771 del ins TH(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P772_H773 dup(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>H773_V774 dup(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wu and colleagues</td>
<td>2008 2.5% (13/515)</td>
<td>5% (13/253)</td>
<td>S768_D770dupSVD(3)</td>
<td>p.H773_V774insNPH</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A767_V769dupASV(3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D770_N771insD 1(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P772_H773insYNP, H773Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N771_H773dupNP(2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D770_N771insG 2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Studies specifically dedicated to exon 20 insertions.

*Mutations diagrammed but specific sequence not reported.
analysis of these mutations predicts that all insertions, regardless of length and amino acid composition, would confer a significant functional impact by affecting the evolutionary conserved protein regions.

To further explore the possible functional differences between the different EGFR exon 20 insertions, we examined their effect on the 3-dimensional structure of the EGFR kinase domain. These in silico molecular modeling studies showed that the various insertions are predicted to interact differently with the erlotinib-binding region. Those involving amino acids 764 to 770 showed the least interaction with the drug-binding pocket. In contrast, insertions between A763 and 764 are predicted to cause significant rearrangement of the C helix, which could markedly reduce drug affinity. Insertions in the more distal region of the hotspot, particularly those affecting the 773 to 775 region would affect the drug-binding pocket directly, predicting the most significant obstructive effect on erlotinib binding. These predictions suggest a basis for the observed variability of response to erlotinib in patients with different EGFR exon 20 insertions. A better understanding of the biology of these mutations is therefore needed and will require further research into the structure of a wide variety of insertions and the development of additional preclinical models. On the basis of available literature and our present findings, it appears that insertions between codons 769 and 775, despite variable length or amino acid composition, are associated with resistance...

Figure 2. Modeling of EGFR exon 20 insertions using the 3-dimensional structure of the EGFR kinase domain predicts different interactions with the erlotinib-binding region. The X-ray structure at 2.6 Å resolution (PDB code 1M17) is used to show the drug and the positions of mutations. Yellow, C helix; green, erlotinib; labeled residues identify the region of the insertion.
to currently approved EGFR-TKIs (6, 8, 9, 12) and to irreversible inhibitors entering clinical trials (15, 16). It should be mentioned that mutations such as the A767_Y769dupASV and S768_D770dupAVD, commonly associated with resistance, do not involve codons 767 or 768 but represent a duplication of the indicated wild-type sequence inserted distally between 769 and 770 and 770 and 771, respectively. In contrast, we found no specific literature to support a resistance pattern for mutations in the region between codons 762 and 768, which would encompass mutations A763_Y764insFHEA, A767_S768insTLA, V769_D770insGE, and M776_A777insAI. In fact, 2 patients with tumors harboring the latter 2 mutations have been reported to show prolonged periods of disease control with irreversible inhibitors entering clinical trials (15, 33). Of interest, the point mutation S768I in this region, also associated with resistance, has been reported to show different responses to TKIs depending on the presence of other mutations; specifically, S768I has been associated with resistance when found in conjunction with G719A (8) or V769L (34) but not with L858R (8). In our series, we identified 5 patients with the S768I mutation, one in conjunction with G719A but none were treated with TKIs. Also, we note the incidental finding of a D770N mutation concurrently with an H773_V774insNPH. While this mutation has been previously reported (35), its association with resistance is not established. In our computational assessment, this mutation scored a low score, predicting no functional impact.

In conclusion, we find that EGFR exon 20 insertions are a highly heterogeneous family of activating mutations with an incidence that is notably higher than previously reported, placing them as the third most common EGFR mutation after EGFR exon 19 deletions and the L858R point mutation. The high variability identified in these mutations may confer diversity in biologic behavior and response to targeted therapies, arguing against their blanket designation as “non-responsive” mutations. Given the high incidence of lung adenocarcinomas, we estimate that testing could identify more than 5,000 patients with these mutations every year in the United States alone. Preclinical research and drug development represent unmet needs in this underestimated subgroup of patients with lung cancer.

Disclosure of Potential Conflicts of Interest

M.G. Kris has commercial research grant from Pfizer, Inc. and Boehringer Ingelheim, and is a consultant/advisory board member of Pfizer, Inc., Clovis, Roche/Genentech, and Boehringer Ingelheim. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions
Conception and design: M.E. Arcila, J.E. Chaft, N. Rekhtman, M.G. Kris, M. Ladanyi
Development of methodology: M.E. Arcila, K. Nafa, N. Rekhtman, C. Lau, M. Ladanyi

www.aacnjournals.org Mol Cancer Ther; 12(2) February 2013 227
Acknowledgments
The authors thank Dr. Laetitia Borsu, Angela Marchetti, Talia Mitchell, Justyna Sadowska, and Jacklyn Casanova for technical assistance with molecular testing assays and Edyta Brzostowski for help with the MSKCC Lung Cancer Mutation Analysis Project database.

Grant Support
The study was supported by NIH P01 CA129243 grant (M.G. Kris and M. Ladanyi).

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 June 21, 2012; revised October 18, 2012; accepted November 19, 2012; published OnlineFirst January 31, 2013.

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc): M.E. Arcila, K. Nafa, J.C. Bell, N. Rekhtman, C. Lau, M.G. Kris

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): M.E. Arcila, K. Nafa, J.C. Bell, N. Rekhtman, B.A. Reva, M.G. Kris, M. Ladanyi

Writing, review, and/or revision of the manuscript: M.E. Arcila, K. Nafa, J.C. Bell, N. Rekhtman, M.F. Zakowski, M.G. Kris, M. Ladanyi

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M.E. Arcila, K. Nafa, M.F. Zakowski, M.G. Kris

Study supervision: K. Nafa, M.G. Kris, M. Ladanyi

References


Characteristics, Prevalence, Molecular Heterogeneity, and Clinicopathologic Characteristics

Maria E. Arcila, Khedoudja Nafa, Jamie E. Chaft, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1535-7163.MCT-12-0620

Supplementary Material
Access the most recent supplemental material at:
http://mct.aacrjournals.org/content/suppl/2012/11/27/1535-7163.MCT-12-0620.DC1

Cited articles
This article cites 35 articles, 22 of which you can access for free at:
http://mct.aacrjournals.org/content/12/2/220.full.html#ref-list-1

Citing articles
This article has been cited by 23 HighWire-hosted articles. Access the articles at:
/content/12/2/220.full.html#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.