Oncogenic Activating Mutations Are Associated with Local Copy Gain


Departments of 1Bioinformatics, 2Research Oncology, 3Molecular Biology, and 4Pathology, Genentech, Inc., South San Francisco, California; and 5Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts

Abstract

Although activating mutations and gains in copy number are key mechanisms for oncogene activation, the relationship between the two is not well understood. In this study, we focused on KRAS copy gains and mutations in non–small cell lung cancer. We found that KRAS copy gains occur more frequently in tumors with KRAS activating mutations and are associated with large increases in KRAS expression. These copy gains tend to be more focal in tumors with activating mutations than in those with wild-type KRAS.

Fluorescence in situ hybridization analysis revealed that some tumors have homogeneous low-level gains of the KRAS locus, whereas others have high-level amplification of KRAS, often in only a fraction of tumor cells. Associations between activating mutation and copy gains were also observed for other oncogenes (EGFR in non–small cell lung cancer, BRAF and NRAS in melanoma). Activating mutations were associated with copy gains only at the mutated oncogene locus but not other oncogene loci. However, KRAS activating mutations in colorectal cancer were not associated with copy gains. Future work is warranted to clarify the relationship among the different mechanisms of oncogene activation. (Mol Cancer Res 2009;7(8):1244–52)

Introduction

Oncogene activation is one of the key processes underlying the development of cancer. Oncogenes can be activated by a variety of mechanisms, including overexpression (due, for example, to translocation or copy gain) and protein mutation leading to elevated or altered activity (1, 2). New technologies allow these phenomena to be studied on an unprecedented scale (3), but the ways in which different mechanisms of oncogene activation may interact within the same tumor are less well-understood.

RAS genes (including KRAS and NRAS) encode founding members of an extensive family of small GTPases, which regulate a variety of key cellular processes through their function as molecular switches (4, 5). Wild-type Ras proteins cycle between GDP- and GTP-bound forms, regulated by GTPase-activating proteins, which activate their intrinsic catalytic activity and by guanine nucleotide exchange factors, which stimulate the exchange of GDP for GTP. In their GTP-bound forms, these proteins engage a variety of effectors, such as RAF and phosphoinositide 3-kinase, which mediate both normal and oncogenic signaling effects. Common RAS-activating mutations are located at codons 12, 13, and 61, rendering the proteins insensitive to the action of GTPase-activating proteins and leading to constitutively GTP-bound, actively signaling forms. RAS genes have also been found to undergo copy gains at low frequencies in several cancer types (6-11). A recent high-resolution study of copy number alterations in non–small cell lung cancer (NSCLC) revealed recurrent copy gains centered on the KRAS locus, providing evidence that these gains are functionally significant and under positive selection (12). Although Ras proteins mediate a variety of oncogenic stimuli (including proliferative, survival, and antiapoptotic signals), they also have context-dependent activities that can oppose the requirements for tumor growth. In particular, the activity of oncogenic mutant Ras can induce growth arrest or apoptosis, for example, in primary cells (5, 13), and wild-type Ras has been shown to act like a tumor suppressor in some models (5). These diverse, context-dependent activities make Ras proteins a particularly interesting subject for the study of oncogenic alterations.

It has been established for >20 years that loss of one tumor suppressor allele by deletion frequently accompanies inactivation of the other by mutation, consistent with Knudson’s “two-hit” model of hereditary cancers (14). However, the relationship between copy number alteration and activating mutation for oncogenes has not been extensively studied. In perhaps the best characterized example, activating mutations in EGFR (e.g., L858R and exon 19 deletions) flank the ATP binding site inside the tyrosine kinase domain and may be associated with NSCLC response to tyrosine kinase inhibitors (15, 16). EGFR can also be activated in NSCLC by gene amplification, and careful examination has revealed that mutation and amplification tend to occur together (17, 18). Furthermore, EGFR copy gains in tumors with mutations are associated with an increasing ratio of mutant to wild-type allele, implying that the mutant allele is selectively gained (18). In glioblastoma, EGFR amplification...
occurs in 40% of tumors and a majority of amplified tumors have mutations in EGFR. Furthermore, in these tumors, EGFR mutations are rarely observed in the absence of EGFR amplification (19). HRAS activating mutations have also been associated with HRAS copy gains in Spitz nevi (20).

In this study, we further examine the relationship between activating mutations and copy gains for well-characterized oncoproteins with high mutation prevalence. For oncoproteins that also undergo copy gains, it is possible to characterize a series of cancers and determine how frequently mutation and copy gain occur in the same cancer. We might expect one of three possible relationships. First, the mutations and copy gains might occur independently of each other, for example if the copy gains are the consequence of random genomic instability and functionally neutral. Second, if copy gains confer additional oncogene activity, then they might be functionally redundant with activating mutations and the two types of alterations would tend to be mutually exclusive. Finally, copy gains might tend to occur in the same cancers that have activating mutations, as has been reported for EGFR in NSCLC. We focus primarily on KRAS mutations in NSCLC, both because the observed tumor suppressor and proapoptotic activities of Ras proteins present some biologically interesting possibilities, and because copy number alterations at the KRAS locus seem to be genuine but poorly understood. Both for KRAS in NSCLC and other oncogenes, we observe a recurrent positive association between activating mutation and copy gain, suggesting that selection for increased copy number of mutated oncogenic loci may be a common phenomenon.

**Results**

**KRAS Mutant NSCLC Has More Frequent KRAS Copy Gain**

We set out to characterize copy gains at the KRAS locus in NSCLC, and to shed light on the relationship between these gains and other genetic alterations, particularly KRAS activating mutations. We sequenced KRAS for mutations and assessed KRAS copy number by qPCR in an initial series of 166 primary NSCLC tumor specimens (Supplementary Table S1). KRAS activating mutations were observed in 27 of 166 (16.3%) of the samples. Some tumors showed evidence of copy gains at the KRAS locus (28 of 166 had a copy number of >2.5), but most gains were modest. Only four tumors had a copy number of >4, although one had clear evidence of high-level amplification (24 copies). We did not find any association between histopathologic subtypes of NSCLC and KRAS copy number (Supplementary Table S1) or between different KRAS mutations and KRAS copy number. However, when we examined whether copy numbers differed between tumors with and without KRAS activating mutations, we found that KRAS mutant tumors had significantly higher copy number than KRAS wild-type tumors ($P = 1.7 \times 10^{-3}$; relative risk of gain, 3.33; Fig. 1). This difference was not accounted for by differences in DNA content between mutants and wild-type tumors (due, for example, to differences in cell cycle state), as we calculated KRAS copy number relative to total DNA content.

![Figure 1](image)

**KRAS copy gains are associated with activating mutations in NSCLC tumors.** Mutant tumors ($n = 27$) have higher copy number than wild-type tumors ($n = 139$; $P = 1.7 \times 10^{-3}$). One wild-type sample with 24 copies of KRAS was excluded from the figure for clarity but was included in statistical analysis. Bars, median of each group.

Although copy gains were found predominantly in tumors with KRAS activating mutations, some KRAS wild-type tumors also exhibited gains (including the sample with apparent high-level amplification). We compared the breadth of regions of gain between wild-type and mutant tumors, using 100k single nucleotide polymorphism (SNP) array data generated for a subset of the tumors. This subset was selected without regard to KRAS mutation or copy number status. The SNP arrays yielded copy number values very consistent with those obtained by qPCR (Pearson correlation coefficient $r = 0.71$; $P = 7.8 \times 10^{-12}$; Fig. 2A). We considered adjacent nonoverlapping 1-Mbp-wide windows across the chromosome and calculated the median copy number within each window for either wild-type or mutant tumors that exhibited KRAS gains (SNP array copy number, >2.5; Fig. 2B). KRAS mutant tumors had more focal gains spanning the KRAS locus, whereas wild-type tumors had broader gains, with similar magnitude of gains across chromosome 12p. These distinct patterns suggest that the copy gains found in the presence of activating mutations are different in character from those found in the absence of such mutations.

**KRAS Copy Gain Is Associated with Increased KRAS mRNA**

The association of copy number with mutation status is suggestive of increased selective pressure in favor of copy gains in KRAS mutant tumors. However, the modest amplitude of the apparent gains led us to question whether they could have significant functional consequences. We used gene expression microarray data for KRAS to examine whether these gains were associated with increases in mRNA levels. As illustrated in Fig. 3, KRAS mRNA levels were strongly correlated with copy number (Pearson correlation coefficient $r = 0.64$; $P = 1.1 \times 10^{-7}$), and even modest gains in copy number were associated with increases in mRNA levels.

**Chromosomal Aberrations Affecting KRAS Are Heterogeneous Both within and between Tumors**

To further characterize the KRAS copy number alterations in NSCLC in more detail, we developed a fluorescence *in situ*
hybridization (FISH) assay for the KRAS locus. A major advantage of the FISH assay is its ability to quantify the copy number of KRAS in individual cells, allowing tumor heterogeneity and stromal admixture to be characterized. The KRAS FISH probes we developed localized to chromosom al region 12p12.1 with no cross-hybridization to any other chromosome. The hybridization efficiency of the FISH probe was >95%.

We carried out FISH analysis for five wild-type and seven mutant tumors using KRAS and centromere 12 (CEP12) probes. These tumors had a wide range of KRAS copy numbers as determined by qPCR. The FISH analysis confirmed the copy number alterations observed by qPCR and provided more detailed information (Supplementary Table S2). Three wild-type and three mutant tumors exhibited tight clusters of KRAS loci, indicating high-level amplification. Two of these had homogeneous amplification of KRAS throughout the sections examined: tumor 126, the tumor for which qPCR reported 24 copies of KRAS (Fig. 4A), and tumor 92. Tumor 92 also had tight clusters of the CEP12 FISH probe, suggesting that sequence close to the centromere was coamplified with the KRAS locus (Fig. 4B). The remaining four amplified tumors were heterogeneous, with amplification observed in only 15% to 30% of tumor cells. For example, tumor 33 had widely variable KRAS copy numbers within a single region (Fig. 4C), including cells with amplification presenting as tight clusters (top right) and cells with no apparent copy gain (bottom left). Such heterogeneity, along with stromal admixture, may explain the limited magnitude of copy gains indicated by qPCR. The six tumors without amplified clusters of KRAS loci had between two and seven copies of KRAS per cell (Supplementary Table S2). These absolute copy numbers were not very similar to our qPCR copy number estimates, which were calculated relative to total DNA content and represent the number of KRAS copies per diploid genome equivalent. We used CEP12 counts as an estimate of ploidy to calculate FISH-derived numbers of KRAS copies per diploid genome equivalent, and got results that were quite similar to those obtained by qPCR (Supplementary Table S2). These data suggest that the modest copy number gains detected by qPCR or microarray analysis resulted in some cases from high-level gains in a subset of tumor cells and in some cases from homogeneous low-level gains.

Either the Mutant or Wild-Type Allele of KRAS Can Predominate in Cancers with KRAS Copy Gains

We attempted to determine, for tumors with increased KRAS copy number, whether the wild-type or mutant allele is preferentially gained. We had sufficient material for seven tumors to carry out TaqMAMA allele–specific qPCR, which uses distinct primers to determine copy number for wild-type and mutant DNA. We validated TaqMAMA by comparing the sum of wild-type and mutant allele–specific copy numbers with the total KRAS copy number determined with allele-nonspecific primers. The results were highly consistent (Supplementary Fig. S1A), indicating that TaqMAMA could accurately determine allele-specific copy number. However, they did not indicate a clear preference for gain of either the wild-type or mutant alleles (Supplementary Fig. S1B). To obtain data on tumors that yielded insufficient DNA for TaqMAMA, we estimated the wild-type to mutant ratio by quantitative analysis of sequencing trace peak heights. The results obtained from this analysis were very similar to those from TaqMAMA (Supplementary Fig. S1C) and the forward and reverse sequencing traces were highly consistent (Supplementary Fig. S1D), showing that peak height analysis is also an effective method for determining allele ratios. This approach enabled us to examine the numbers of wild-type and mutant copies in a total of 27 tumors (Fig. 5A). However, we still did not see a clear preference for the gain of either wild-type or mutant KRAS copies, even taking into account the presence of additional wild-type DNA from nontumor cells.

Wild-type DNA from stromal cells would reduce the apparent ratio of mutant to wild-type copies observed in tumor

![Image](mcr.aacrjournals.org)
samples. To gain additional information, we determined allele-specific copy numbers for four NSCLC cell lines in which both wild-type and mutant KRAS are present (Calu-1, H23, H358, and HOP62). All four cell lines had more mutant copies of KRAS than wild-type copies (Fig. 5B).

H-ras has been shown to induce cell cycle arrest and senescence in mouse embryonic fibroblasts in the presence of p53 and p16, but not when either tumor suppressor is removed (13). We reasoned that KRAS copy gains might be more common in tumors with p53 or p16 mutations, but found no association between KRAS copy number and the mutation status of either tumor suppressor (P > 0.3; Supplementary Table S1).

Associations between Mutation and Copy Gain Are Also Observed for EGFR, BRAF, and HRAS

We then used Affymetrix SNP array copy number data to investigate the possibility that the association between copy gain and activating mutation extends to other oncogenes. First, the association between copy number gains and activating mutation was confirmed for EGFR (P = 3.1 x 10^{-4}; Fig. 6A). We did not find a significant difference in copy number between different types of EGFR mutations. Then we examined BRAF and NRAS copy number in a series of 33 short-term melanoma cultures (21). Higher copy numbers were observed in samples with activating mutations than in samples in which no mutations were found, both for BRAF (P = 9.1 x 10^{-3}; Fig. 6B) and NRAS (P = 5.5 x 10^{-3}; Fig. 6C). These results suggest that it is common for activating sequence mutations to be associated with copy gains. However, we did not observe evidence for more focal gains in mutant than in wild-type samples for EGFR, BRAF, or NRAS (data not shown). We also examined KRAS copy number in 57 primary colorectal tumor specimens and found no significant difference in copy number between samples with and without activating mutations (P = 0.20; Fig. 6D). There was little evidence of copy gains at the KRAS locus in these tumors, with the exception of one wild-type sample with evidence of amplification. These results indicate that the co-occurrence of mutation and copy gains, although not universal, might be widespread.

One possible explanation for the associations we observed between oncogene activating mutation and copy number is that the mutations we examined are associated with copy gains throughout the genome and not specifically at the oncogene locus. To test this possibility, we examined whether activating mutations in a given oncogene were also associated with copy gains at the loci of other oncogenes. For example, were activating mutations in KRAS in NSCLC also associated with copy gain of BRAF? As shown in Table 1, activating mutations were not associated with copy gains at unrelated oncogene loci. One exception was that BRAF-activating mutations in melanoma were associated with copy gains at the EGFR locus. This may be explained by EGFR and BRAF both being located on chromosome 7. Melanomas with elevated BRAF copy number tend to have gains of broad regions of chromosome 7, in some cases the entire chromosome (21, 22). Overall, mutation of a given oncogene was selectively associated with copy gains at the genomic locus of that particular oncogene.

Discussion

The activation of oncogenes by sequence mutations or copy gains are often considered independently. Our results suggest that activating mutations and copy gains may often co-occur within the same tumor. We found that cancers with activating mutations in an oncogene are more likely to have copy gains at the locus of that oncogene for KRAS and EGFR in NSCLC, and for BRAF and NRAS in melanoma. These gains were locus specific and not observed at the loci of unrelated oncogenes. The magnitude of the copy gains measured by qPCR or SNP arrays was quite modest in most samples, but when we examined KRAS expression in NSCLC, we found that even small increases in apparent copy number were accompanied by increases in mRNA level. This lends support to the notion that even apparently modest gains can be functionally significant. Also, when we examined KRAS copy number by FISH, we found that several tumors had high-level amplification apparent in a small fraction of cells. Both this tumor heterogeneity and the presence of stromal DNA can cause misleadingly low copy number estimates by methods that interrogate bulk tumor DNA.

Ding et al. (23) recently reported that copy gains were associated with mutation and increased expression for EGFR, KRAS, and a third gene, EPHB1, in a large study of lung adenocarcinoma. This suggests that our observations extend beyond our own data set and to additional genes. Some other oncogenes did not have statistically significant associations between copy gains, mutation, and expression in the data set of Ding et al. (23). These genes were mutated at lower frequencies than EGFR and KRAS, resulting in lower statistical power to detect associations. More data are required to determine whether mutations in genes with low mutation frequencies are associated with increases in copy number.

The relationship between gene mutation and copy gain requires more complete mechanistic explanation. The propagation
of cancer cells in the stromal microenvironment can be viewed as an evolutionary process (24). Genetic alterations occur at random in individual cells, and those that confer a selective advantage may drive clonal expansion and become fixed in the tumor cell population. Selectively neutral alterations may also be fixed if they happen to be present in an expanding clone (becoming so-called “passenger” mutations). This evolutionary model provides a framework within which our findings may be interpreted. The observation that copy gains are more common in the presence of activating mutations implies that these gains are under positive selection and that this selective pressure is stronger in mutant tumors. This is consistent with the more focal KRAS gains observed in mutant tumors, and with the increase in KRAS mRNA that accompanies copy gains. It may be that a single mutation gene copy is sometimes insufficient to drive complete activation of an oncogenic signaling pathway without increased protein expression. However, our data cannot preclude alternative models, and suggest that different selective pressures are at work in different tumors based on their individual genetic, epigenetic and other features.

Ras proteins are best known for their oncogenic activities, but can also act in ways that oppose the requirements for tumor growth. As a result, the relationship between the level of KRAS copy gains and proliferative capacity is hard to predict. The appropriate level of amplification to achieve the optimal level of pathway activity will be dependent on the abundance and stoichiometry of signaling partners, which will differ among tumors. In fibroblasts, overexpression of a RAS mutant allele leads to senescence, whereas physiologic levels promote proliferation (13, 25). In contrast, single copy expression of mutant RAS in breast epithelial cells does not lead to proliferation, whereas overexpression does (26). Variation in the mutational context found in individual tumors will also lead to differing intensity of selection pressure for oncogene activity. For example, H-ras has been shown to induce cell cycle arrest and senescence in mouse embryonic fibroblasts in the presence of p53 and p16, but not when either tumor suppressor is removed (13). Such variation in selection pressure could explain why EGFR-activating mutations in NSCLC are associated with robust EGFR amplification, whereas KRAS activating mutations are associated with more moderate copy gains.

The consequence of copy gains can also depend on the allele that is gained. Due to the clear oncogenic potential of mutant Ras, as well as the fact that wild-type KRAS can act as a tumor suppressor in some contexts (27, 28), we expected to see a preferential gain of the mutant allele. Indeed, we found more copies of the mutant KRAS allele than the wild-type allele in all four of the NSCLC cell lines we examined that had both alleles present. However, we did not observe any clear preference for either allele to be gained in tumors. Furthermore, we found some copy gains in wild-type tumors, including an instance of high-level KRAS amplification in NSCLC. Although wild-type Ras alleles can sometimes act as tumor suppressors, they can also act as tumor promoters, depending on the tissue and oncogenic stimulus (reviewed in ref. 5). Selection pressure in individual tumors may modulate the copy number of each allele to fine tune the level of KRAS activity.

We did not find any evidence that copy gains were associated with KRAS mutation in colorectal cancer. This could be due to a lack of selective pressure in favor of gains (for one of the reasons discussed above) or to a lack of copy number alterations on which such pressure could act. We found fewer copy number alterations genome wide in our colorectal cancer than in our NSCLC samples (data not shown), and there were very few gains at the KRAS locus. However, it is interesting that one wild-type tumor had over 10 copies of KRAS by SNP array.

This study suggests that the co-occurrence of activating mutations and copy gains may be quite common. Accurately assessing the activation status of an oncogene may require consideration of both sequence mutations and copy number, and ultimately other factors such as epigenetic regulation. The genetic state of each individual tumor (and of individual cells within that tumor) is dynamic and determined by complex interactions between the stromal environment and both aberrant and normal signaling pathways. Future research will clarify the detailed mechanisms by which diverse patterns of oncogene activation occur.

Materials and Methods

Samples

NSCLC and colon cancer fresh frozen samples consented for research use were obtained from several anonymized sample banks (Supplementary Table S1). Cell lines were obtained from the American Type Culture Collection. Tumor sections for research use were obtained from several anonymized sample banks (Supplementary Table S1). Cell lines were obtained from the American Type Culture Collection. Tumor sections were stained with H&E to confirm diagnosis and assess tumor content. DNA was isolated using the Qiagen DNeasy kit from 166 NSCLC and 57 colorectal tumor specimens, which had >70% tumor content, and from the NSCLC cell lines Calu-1, 

![FIGURE 4. Examples of KRAS FISH results for NSCLC tumors. Blue, 4',6-diamidino-2-phenylindole; red, KRAS (spectrumOrange); green, CEP12 (spectrumGreen). A. Tumor 126. B. Tumor 92. C. Tumor 33.](Image 305x49 to 534x346)
DNA from 33 short-term melanoma cultures was also obtained, as previously described (29).

Mutational Status Analysis

The following regions were Sanger sequenced (by Polymorphic DNA Technologies, Inc.) as follows: KRAS coding exons in NSCLC (166 samples), EGFR exons 1-28 in NSCLC (68 samples), TP53 coding exons in NSCLC (72 samples), CDKN2A coding exons in NSCLC (153 samples), and KRAS exon 2 in colorectal tumors (57 samples). Mutations were identified by comparing the results to the R35 public human genome reference assembly (30) with Mutation Surveyor software (SoftGenetics) and confirmed by manual review. For melanoma short-term cultures, mass spectrometric genotyping was done in all samples for common activating mutations in BRAF and NRAS as previously described (29).

Real-time PCR Analysis

KRAS allele non-specific copy number was determined by quantitative PCR using the 7900HT Fast Real-Time PCR system (Applied Biosystems). The KRAS primers were 5′-AAGGTGCACTGTAATAATCCAG-3′ and 5′-AGACAGGTTTCTCCATCAATT-3′ and they amplified exon 5 of KRAS. Total DNA content was estimated by assaying LINE-1 elements for each sample using the primers 5′-AAAGCCGCTCAACTACATGG-3′ and 5′-TGCTTTGAATGCGTCCCAGAG-3′. The KRAS copy number was defined as outlined previously (32).

Allele-specific copy numbers were determined using a modification of the TaqMAMA procedure previously described for PCR-based genotyping (33). A common 3′ primer (5′-CTAGGTTGTTATCTCCATTTAGA-3′) was used, and two 5′ primers were designed for each nucleotide variant of interest as described (33). The 5′ primer sequences were as follows: for 215G>T, 5′-AAACTTGTGGTAGTTGGAGCGG-3′ (wild-type) and 5′-AAACTTGTGGTAGTTGGAGCCT-3′ (mutant); for 216G>A, 5′-AACTTGTGGTAGTTGGAGCTCG-3′ (wild-type) and 5′-AACTTGTGGTAGTTGGAGCTCA-3′ (mutant).

All qPCR reactions were carried out in the same manner. In each 15-μL reaction mixture, 15 ng DNA (except for the standard curve samples) were mixed with the Brilliant SYBR Green QPCR Master Mix (Stratagene), and 3 pmol/L of each primer. After a single step at 95°C 10 min, 40 cycles of amplification were undertaken (95°C for 30 s, 55°C for 1 min, and 72°C for 30 s) followed by an optional step for melting curve analysis (95°C for 1 min, 55°C for 30 s, and 95°C for 30 s).

Three DNA samples were used to establish standard curves for different KRAS alleles. In addition to Human Genomic DNA-Female (Promega), DNA was isolated using the DNeasy kit (Qiagen) from two cell lines: H2030 (American Type Culture Collection) and HEC1A with the wild-type KRAS knocked out (HEC1A-d; kindly provided by Dr. Todd Waldman, Georgetown University School of Medicine, Washington, D.C.; ref. 34). Human Genomic DNA-Female was assumed to have two wild-type copies of KRAS. Sequencing and qPCR data indicated that H2030 has two copies of KRAS and is homozygous.
for the 215G>T (G12C) allele. The HEC1A-d cell line is derived from HEC1A, a previously described near diploid cell line heterozygous for KRAS 216G>A (G12D), by knocking out the wild-type KRAS exon 2 (34). Thus, HEC1A-d has one copy of KRAS exon 2 with the 216G>A mutation.

For each KRAS TaqMan primer set, a standard curve was made using nominal 64, 32, 16, 8, 4, and 2 ng quantities of the appropriate DNA (215G wild-type, Human Genomic DNA-female; 215T mutant, H2030; 216G wild-type, human genomic DNA-female; 216A mutant, HEC1A-d). qPCR of LINE-1 elements was used to estimate the total DNA quantity in each sample. All tumor samples and standard curves were run in triplicate. The copy numbers for each allele were calculated as follows:

\[
\begin{align*}
215G \text{ wild-type copy number} &= \left[ \frac{\text{Qty}(215 \text{ G DNA})}{\text{Qty}(\text{LINE-1 DNA})} \right] \times 2 \\
215T \text{ mutant copy number} &= \left[ \frac{\text{Qty}(215 \text{ T DNA})}{\text{Qty}(\text{LINE-1 DNA})} \right] \times 2 \\
216G \text{ wild-type copy number} &= \left[ \frac{\text{Qty}(216 \text{ G DNA})}{\text{Qty}(\text{LINE-1 DNA})} \right] \times 2 \\
216A \text{ mutant copy number} &= \left[ \frac{\text{Qty}(216 \text{ A DNA})}{\text{Qty}(\text{LINE-1 DNA})} \right] \times 1
\end{align*}
\]

**SNP Array Analysis**

Sixty-nine NSCLC and 57 colon cancer tumor samples were run on GeneChip Human Mapping 100K Set SNP arrays (Affymetrix). Copy number at each SNP was determined using genome-smoothed analysis of copy number (CNAT3.0.0.21; ref. 35). Thirty-three melanoma short-term culture samples were run on GeneChip Human Mapping 500K Set SNP arrays (Affymetrix) and copy number at each SNP was determined by single point analysis (CNAT 1.5.4.2-β; ref. 35). Probes for each oncogene were defined as those within 100 kb of the gene boundaries. Copy number was determined by taking the median copy number for an oncogene’s probes.

To examine the chromosomal extent of KRAS copy gains, NSCLC samples with KRAS gains (>2.5 copies) based on the SNP data were segregated based on their KRAS mutational status (wild-type or mutant). Chromosome 12 was broken up in adjacent nonoverlapping one megabase segments and the median KRAS copy number for each segment was determined.

**Analysis of Sequencing Traces**

For each sample, the raw fluorescence intensity data for the four electrophoresis channels were obtained along with the base calls. At every call position of each channel, the expected height for a homozygous peak was calculated as the Gaussian-weighted average of the heights of neighboring peaks that resulted in a base call from the channel under consideration. The weighting function reached half maximum 10 bases from the call position. Next, each observed peak height was divided by its expected homozygous peak height to obtain a normalized peak height. The relative abundance of the wild-type allele

![FIGURE 6. SNP array–based copy number versus mutation status for (A) EGFR in NSCLC (mutant, n = 6; wild-type, n = 55; P = 0.033), BRAF in melanoma (mutant, n = 20; wild-type, n = 11; P = 4.3 x 10^-3), N-RAS in melanoma (mutant, n = 9; wild-type, n = 24; P = 0.015), and KRAS in colorectal cancer (mutant, n = 12; wild-type, n = 57; P = 0.090). Bars, median of each group.](image-url)
was determined by dividing the normalized peak height of the wild-type allele by the sum of the normalized peak heights of the mutant and wild-type allele. The resulting quantity is an estimate of the fraction of DNA molecules that carried a wild-type residue at the position of interest. The number of wild-type copies was calculated by multiplying this fraction by the total number of \( \text{KRAS} \) copies measured by qPCR.

*Expression Analysis*

Messenger RNA from 55 NSCLC samples was isolated using the RNeasy kit (Qiagen) and run on Human Genome U133 Plus 2.0 Array (Affymetrix) according to the manufacturer's instructions. \( \text{KRAS} \) expression was assessed by the probe set with the highest intensity (214352_s_at).

*FISH Analysis*

Five wild-type and seven mutant tumors exhibiting a range of \( \text{KRAS} \) copy numbers by qPCR were selected and successfully analyzed by FISH.

A bacterial artificial chromosome (BAC) contig comprising of two overlapping clones, RP11-111918 and RP11-29515, covering the entire \( \text{KRAS} \) locus and adjoining areas (based on the UCSC Genome Browser March 2006 assembly) were used as a probe for the FISH experiments. A commercially available probe for CEP12 (Vysis/Abbott Laboratories) was also used to enumerate the copies of chromosome 12 in the FISH experiments.

DNA from the BAC clones was extracted using standard methods. The extracted BAC DNA was labeled with Spectrum Orange (Vysis/Abbott Laboratories) by nick translation using the Vysis Nick Translation kit (Vysis/Abbott Laboratories), according to the manufacturer's instructions. FISH to normal human metaphases (Abbott Laboratories) was used to confirm the genomic location of the BAC clones. Approximately 300 ng of labeled probes were precipitated in excess Human Cot-1 DNA (Invitrogen) and sonicated salmon sperm DNA (Sigma), and labeled probes were precipitated in excess Human Cot-1 DNA (Invitrogen) and resuspended in a 50% formamide, 10% dextran sulfate, and 2×SSC hybridization buffer (Vysis/Abbott Laboratories) for the FISH experiments. Dual color FISH on cytogenetic preparations and formalin-fixed paraffin-embedded tissue were done as described previously (36), with some modifications. After an overnight incubation at 56°C, the slides were deparaffinized in three washes of CitroSolv for 5 min each, followed by two washes in alcohol. After air drying, the slides were incubated in a 1 mol/L solution of sodium sulfocyanate for 30 min at 85°C and then were treated with pepsin before additional washes in water and a series of ethanol. Dried slides were then denatured (76°C for 6 min) with the probe and were hybridized overnight at 37°C (ThermoBrite; Vysis). Posthybridization washes and counter-staining were done in a manner similar to those previously described (36). The slides were visualized using an Olympus BX61 microscope and analyzed using FISHView software (Applied Spectral Imaging) and the ISIS Software (Metasystems, Inc.). For the \( \text{KRAS} \) copy number analysis, the \( \text{KRAS} \) and CEP12 signals were counted in 100 nuclei per tumor.

*Statistical Analysis*

Associations between \( \text{KRAS} \) qPCR copy number and mutational status were assessed using two-sided Wilcoxon rank sum tests. Associations between SNP array copy number and oncogene mutational status (reported in Table 1) were assessed using one-sided Wilcoxon rank sum tests to specifically test the hypothesis that mutation was associated with increased (and not decreased) copy number. Pearson correlation was used to assess the relationship between qPCR and SNP array copy number and between \( \text{KRAS} \) copy number and expression level. Relative risk calculations used a cutoff of >2.5 copies for gains. All analyses were done in R.6

*Disclosure of Potential Conflicts of Interest*

G. Cavet is an employee of Genentech, Inc. and has received a commercial grant from Genentech, Inc.

*Acknowledgments*

We thank Mark Lackner and members of the Lackner laboratory for generously sharing equipment, and the Genentech Histology Lab and Human Tissue Lab for essential assistance with samples.

*References*


Table 1. \( P \) values for Association of Copy Gains and Activating Mutations

<table>
<thead>
<tr>
<th>Gene With Activating Mutation (Tumor Type)</th>
<th>Gene With Copy Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR (chr7)</td>
<td>KRAS (chr12)</td>
</tr>
<tr>
<td>3.1 × 10⁻³</td>
<td>0.99</td>
</tr>
<tr>
<td>KRAS (chr12)</td>
<td>0.51</td>
</tr>
<tr>
<td>BRAF (melanoma)</td>
<td>6.1 × 10⁻³</td>
</tr>
<tr>
<td>NRAS (melanoma)</td>
<td>0.98</td>
</tr>
</tbody>
</table>

6 http://www.r-project.org


35. Affymetrix technical note: Copy Number Analysis Tool.