Relationship of *EGFR* Mutations, Expression, Amplification, and Polymorphisms to Epidermal Growth Factor Receptor Inhibitors in the NCI60 Cell Lines

Wanqing Liu,[1] Xiaolin Wu,[1] Wei Zhang,[1] Raquel C. Montenegro,[1] Donna Lee Fackenthal,[2] Jared A. Spitz,[1] Lyn Mickley Huff,[5] Federico Innocenti,[1,3,4] Soma Das,[2,3] Edwin H. Cook, Jr.,[5] Nancy J. Cox,[1,2,3] Susan E. Bates,[6] and Mark J. Ratain[1,3,4]

Abstract

Purpose: The mechanism of sensitivity and resistance to epidermal growth factor receptor (EGFR) inhibitors is incompletely understood, particularly in cancers other than non–small-cell lung cancer (NSCLC). To understand the variable response to this class of drugs, we used the NCI60 cancer cell lines. We aimed to determine if there are interactions between *EGFR* expression, mutations, polymorphisms, and gene amplification, and whether these factors are associated with variability in response to EGFR inhibitors.

Experimental Design: The EGFRVIII and tyrosine kinase (TK) domain mutations were examined in the NCI60 cancer cell lines. Five polymorphisms, -216G/T, -191C/A, intron 1 (CA)n, R497K, and 2607A/G, were genotyped. *EGFR* amplification was also assessed with high-density single-nucleotide polymorphism chip and real-time PCR, respectively. The results were correlated with cytotoxicity data for erlotinib and other 11 EGFR inhibitors, as well as other publicly available data for these lines.

Results: All 12 inhibitors behaved similarly. No EGFRVIII but putative TK mutations in two cell lines were found. Both mutant cell lines were insensitive to all inhibitors. Meanwhile, response did not correlate with *EGFR* amplification but with *EGFR* gene expression, especially in the cell lines with relatively normal gene status. In addition, *EGFR* expression was associated with the -216G/T polymorphism but not with the intron 1 (CA)n polymorphism. A combination of -216G/T and R497K polymorphisms was weakly associated with drug response.

Conclusions: These observations suggest that in addition to TK mutations, germ-line variability may also contribute to the pharmacodynamics of EGFR inhibitors, particularly when *EGFR* is genetically normal.

Epidermal growth factor receptor (EGFR) plays an important role in carcinogenesis and is therefore an intriguing target for cancer therapy. A group of quinazoline derivatives have been shown to be potent inhibitors of EGFR tyrosine kinase (TK) activity in human cancer cells in both *in vitro* and *in vivo* models (1–4). One of them, erlotinib, was recently approved for the treatment of non–small-cell lung cancer (NSCLC). To date, a number of clinical trials have been done in which patients have been treated with these drugs either as single agents or in combination with other anticancer drugs. The response rate to these inhibitors varies according to the type of cancer, ethnicity, gender, and other clinical features, but in general has been low (3–7). Some significant factors affecting sensitivity or resistance to EGFR inhibitors have been identified (8). Somatic changes in the *EGFR* gene including gene amplification, an in-frame deletion mutation of exon 2 to 7 (*EGFRVIII*), and TK domain mutations have been recently associated with better response (5–12). Resistant mutations in the EGFR domain were also identified (8, 13). Although significant, these changes do not completely explain the variable clinical outcomes (8). In both *in vitro* and *in vivo* studies, a subset of human cancer cell lines and patients with normal *EGFR* status also responded to these agents (3, 14), suggesting that other factors may also play an important role. Recent studies showed that germ-line DNA variations such as a promoter single nucleotide polymorphism (SNP; -216G/T) and an intron 1 (CA)n polymorphism of *EGFR* might predict drug response (15–19).

Authors' Affiliations: Departments of [1]Medicine and [2]Human Genetics, [3]Committee on Clinical Pharmacology and Pharmacogenomics, and [4]Cancer Research Center, The University of Chicago; [5]Institute for Juvenile Research, Department of Psychiatry, University of Illinois at Chicago, Chicago, Illinois; and [6]Cancer Therapeutics Branch, National Cancer Institute, NIH, Bethesda, Maryland

Received 3/6/07; revised 8/3/07; accepted 9/5/07.

Grant support: Pharmacogenetics of Anticancer Agents Research Group, NIH/National Institute of General Medical Sciences grant U01GM61393 and the William F. O'Connor Foundation. Data have been deposited into PharmGKB supported by NIH/National Institute of General Medical Sciences grant U01GM61374.

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.

Requests for reprints: Mark J. Ratain, The University of Chicago, 5841 South Maryland Avenue, MC 2115, Chicago, IL 60637. Phone: 773-702-4400; Fax: 773-702-3969; E-mail: mratain@medicine.bsd.uchicago.edu.

doi:10.1158/1078-0432.CCR-07-0547

The variants -216G/T (rs712829) and -191C/A (rs712830) have been discovered in the essential promoter region of *EGFR* and have been associated with increased *EGFR* promoter activity and gene expression mediated by an altered interaction between Sp1 and the context sequence of the polymorphism (20). In addition, a SNP at codon 497 of *EGFR* (exon 13, rs11543848), an A to G alteration resulting in a substitution of the amino acid Arg (R) by Lys (K; ref. 21), is the only common missense polymorphism of *EGFR* reported to date, and the *K* allele seems to decrease the activity of *EGFR* (22). There is also a common synonymous A/G SNP in codon 787 in the TK domain (mRNA position 2,607; exon 20; rs17337198). Whether there are interactions between mutations, expression, activity, gene amplification, and these polymorphisms in *EGFR* and whether these factors confer responsiveness to EGFR tyrosine kinase inhibitors (TKI) in cancer cells are largely unknown.

As an enormous amount of cytotoxicity and molecular data have been collected using the NCI60 cancer cell lines, they provide an ideal platform to evaluate interrelationships between molecules and drugs (23). Previous studies of the NCI60 suggested that *EGFR* expression could be correlated with response to EGFR inhibitors (1, 2). In particular, a low level of EGFR predicted for absence of response, although a high level did not guarantee response to the TKIs. Meanwhile, the expression and/or mutation status of Ras, Raf-1, p53, Mdm2, p21, and cyclin D1 could not explain the difference between sensitive and insensitive cell lines (1). However, genetic alterations in the *EGFR* gene have not yet been fully investigated in this panel thus far. In this study, we screened *EGFR* mutations and genotyped five DNA polymorphisms, -216G/T, -191C/A, intron 1 (CA)n, R497K, and 2607A/G, in the NCI60. We also assessed *EGFR* gene copy number in these cells. The interrelationships of these polymorphisms with *EGFR* expression and the activity of erlotinib and 11 other quinazoline derivatives are discussed.

Materials and Methods

RNA, DNA, and cDNA preparation from NCI60 cell lines. Fifty-eight cell pellets of the NCI60 panel were obtained from the National Cancer Institute. Two cell lines, MDA-N and RXF-393, had been discontinued when requested. Total RNA and DNA for all 58 samples (except for the RNA from one leukemia cell line, SR) were extracted using the RNeasy Mini Kit (Qiagen) and the PUREGENE DNA Purification Kit (Gentra Systems), respectively. To prepare cDNA, 1 μg of total RNA was reverse transcribed using SuperScript II reverse transcriptase with both oligo$(dT)_n$ and random primers according to the manufacturer's protocol (Invitrogen). The cDNA was then diluted five times for PCR reactions.

Screening of EGFRVIII and EGFR TK domain mutations. The EGFRVIII mutation was screened according to a protocol described recently (9). Briefly, cDNA was amplified by PCR with primers spanning exons 1 to 8, after which PCR products were separated in a 5% polyacrylamide gel and visualized by ethidium bromide staining. *EGFR* exons 18 to 21 (the *EGFR* TK domain) were examined for mutations in all 58 cell lines using multiplex PCR and direct sequencing (24). Briefly, PCR was done on DNA isolated (as described above) from each of the 58 cell lines. PCR product was treated with shrimp alkaline phosphatase (Roche Diagnostics Corporation) and exonuclease I (U.S. Biochemical Corp.) to remove excess deoxynucleotide triphosphates and primers followed by sequencing on the ABI 3830XL genetic analyzer using ABI Big Dye Terminator v1.1 sequencing kit. Putative alterations of identified misssense DNA variations in EGFR function were assessed by Sort Intolerant from Tolerant (SIFT).[7]

Genotyping of polymorphisms. PCR was done to amplify the DNA sequences containing the polymorphisms of interest. For -191C/A and -216G/T, DNA was amplified using primers 5′-TCTGCTC-CTCCCGATCCCTCCT-3′ and 5′-CAGGTGGCCTGTCGTCCGGTCT-3′. PCRs were set up in a 40-μL volume containing 3 mmol/L $MgCl_2$, 1× Q-solution (Qiagen), 100 μmol/L of each deoxynucleotide triphosphate, 125 nmol/L of forward and reverse primers, 1 unit of Hotstar Taq DNA polymerase (Qiagen), and 25 ng of DNA. Reactions were denatured initially at 98°C for 10 min, and then cycled 35 times at 98°C for 15 s and annealed at 62°C for 15 s and 72°C for 20 s. PCR products were then purified and directly sequenced. For R497K, PCR was done according to the protocol reported previously (25). Genotyping was done by using single base extension and denaturing high-performance liquid chromatography (26) with a probe (5′-TCCGGCAAGAGACGCAGTCC-3′). Genotyping of the intron 1 (CA)n polymorphism was done as previously described (27). Genotypes of 2607A/G were obtained from the sequencing data when screening the TK mutations.

To confirm the accuracy of genotyping, the HapMap Caucasian (HMC) samples (Coriell panel number: HAPMAPPT01; unrelated grandparents, $n = 60$) were used as control. Lymphoblastoid cell lines of the samples were purchased from the Coriell Cell Repository. DNA was extracted as above by the Lymphoblastoid Cell Core of the Pharmacogenetics of Anticancer Agents Research Group at the University of Chicago. The genotypes for the 2607A/G polymorphism were downloaded from the HapMap database,[8] whereas -216G/T, -191C/A, intron 1 (CA)n, and R497K were genotyped as described above.

Copy number estimation using the Affymetrix 100K SNP chip. The *EGFR* copy number was extracted from the whole-genome duplication/deletion assessment after genotyping with the Affymetrix 100K SNP chip. Briefly, DNA isolated from 58 NCI60 lines was divided into three batches of hybridizations, with each batch containing one control (Affymetrix 100K SNP chip reference sample). After scanning with the Affymetrix Scanner 3000 (Affymetrix), the intensities of each probe on a specific cell line chip were globally scaled by multiplying by the ratio of the sum of all probe intensities of the control sample in the batch divided by the sum of all probe intensities of each cell line chip. The estimated copy number (estCN) for each SNP was estimated based on the ratio of the previously chip-wide scaled intensity of the sum of 20 perfect match—corresponding mismatch probe intensities for each SNP in each sample and the batched control and multiplying by 2 for autosomes. A coefficient of variation (CV) among 40 probes of each SNP in the control was determined. To minimize the effect of variability in the intensity estimate at a single SNP, the following processes were used to detect duplication/deletion. (*a*) Seeding: Starting with the 5′ end of each chromosome, the first SNP whose estCN was out of the range of $2 \pm 2 \times$ CV was considered a seed. (*b*) Extension: If a seed was found, we then checked the next 3′ SNP to estimate if the average estCN of the next two consecutive SNPs was still beyond the $2 \pm 2 \times$ CV threshold. (*c*) Termination: Extension was repeated for each downstream SNP until the difference between 2× average estCN of the seed SNP and consecutively extended SNPs and the sum of the estCN of the next two SNPs was larger or smaller than 1. (*d*) Estimation: For regions in which estCN for at least three contiguous SNPs was found to differ from $2 \pm 2 \times$ CV, the estCN was indicated by the intensity ratio relative to the control × 2.

Quantification of EGFR copy number and EGFR mRNA expression by real-time PCR. Real-time PCR with the Mx3000P system (Stratagene) and iQTMSYBR Green Supermix (Bio-Rad Laboratories) was used to

[7] http://blocks.fhcrc.org/sift/SIFT.html
[8] http://www.hapmap.org

Table 1. Comparison of allele frequencies of the five polymorphisms between NCI60 and the HapMap CEU samples

Polymorphism	Allele	NCI60		HMC		*P*
		n	Frequency (%)	*n*	Frequency (%)	
-216G/T	G	69	59.5	82	68.3	0.18
	T	47	40.5	38	31.7	
-191C/A	C	107	92.2	95	79.2	<0.01
	A	9	7.8	25	20.8	
R497K	R	83	71.6	85	70.8	1.00
	K	33	28.4	35	29.2	
2607A/G	A	70	60.3	70	58.3	0.79
	G	46	39.7	50	41.7	
Intron 1 (CA)n	14	1	0.9	2	1.7	12.76, 9, 0.17
	15	5	4.4	1	0.8	
	16	53	46.5	44	36.7	
	17	7	6.1	5	4.2	
	18	13	11.4	28	23.3	
	19	4	3.5	4	3.3	
	20	24	21.1	32	26.7	
	21	5	4.4	4	3.3	
	22	1	0.9	—	—	
	23	1	0.9	—	—	

NOTE: Fisher's exact test was used to compare the four SNPs and the *P* values are shown. χ^2 test was used to compare the intron 1 (CA)n polymorphism and the χ^2, *df*, and *P* values are shown.

quantify the number of copies of the EGFR TK domain according to the protocol described by Moroni et al. (10). *EGFR* mRNA level was also quantified with real-time PCR. Briefly, a 120-bp amplicon was amplified by PCR with primers spanning *EGFR* exon 19 (5′-GGACTCTGGATCCCAGAAGGTG-3′) and exon 20 (5′-GCTGGCCATCACGTAGGCTT-3′). *β-Actin* gene was used as internal control. Primer sequences for *β-actin* were 5′-ACGTGGACATCCGCAAAGAC-3′ and 5′-CAAGAAAGGGTGTAACGCAACTA-3′. Reactions were done with standard curves for *EGFR* or *β-actin* gene and repeated thrice. Real-time PCR was cycled 45 times at 95°C for 30 s, 55°C for 1 min, and 72°C for 30 s after preheating at 95°C for 10 min. Expression of *EGFR* was then normalized to that of the *β-actin* gene.

Linkage disequilibrium. Pairwise linkage disequilibrium between SNPs was estimated using LDPlotter[9] and Arlequin 2.000.[10]

Extraction of data from the NCI60 database. GI_{50}s [molar concentrations expressed as $\log_{10}(IC_{50})$] of 41 cell lines treated with erlotinib (NSC718781) were obtained from the NCI60 database (gefitinib was not included in the database). Data for the other 11 quinazoline derivatives [with number of cell lines tested, NSC691853 (58), NSC691854 (58)], NSC691856 (57), NSC691857 (57), NSC691858 (57), NSC691859 (55), NSC691860 (57), NSC709239 (58), NSC669364 (58), NSC669365 (53), and NSC693255 (58)] were reported previously (1, 2). Data for the genes of interest were extracted from the website of the Developmental Therapeutics Program of the National Cancer Institute/NIH.[11] Two main data sets of expression data (either EGFR protein or mRNA) from a total of seven experiments were extracted from the database. These included EGFR protein level by Western blot (MT1147) and *EGFR* mRNA level by PCR slot blot (MT976) and RNase protection (MT478), as well as microarrays (GC16212, GC18283, GC27729, and GC33544). Correlations below using growth inhibition assays, protein expression, and RNA expression relied on the data from these data sets. The microarray data sets from either Stanford (GC16212 and GC18283) or Novartis Pharmaceuticals (GC27729 and GC33544) were combined and their respective means were used to perform correlations.

Karyotypic aberrations of NCI60 have previously been assessed with the spectral karyotyping (SKY; ref. 28). The copy number data of chromosome 7p12 containing the *EGFR* gene (MT2066) were extracted.

Statistics. χ^2 and Fisher's exact tests were used to compare allele frequencies between populations and test Hardy-Weinberg equilibrium and allelic association. For correlations between polymorphisms and *EGFR* expression, the Spearman correlation test was used to test the allelic effect on gene expression and drug action. Three univariate genetic models (additive, dominant, and recessive) were tested, as the modes of inheritance were either unknown or not clearly identifiable from the plotted data. For example, for the hypothetical variant A > B, we have genotypes A/A, A/B, and B/B. The additive model would be A/A = 0, A/B = 1, and B/B = 2. The dominant model would be A/A = 0, A/B = 1, and B/B = 1. The recessive model would be A/A = 0, A/B = 0, and B/B = 1. For the intron 1 (CA)n polymorphism, the sum of the CA repeats in both alleles was used as the variable. When analyzing the diplotypes of -216G/T and R497K polymorphisms, a one-way ANOVA or the nonparametric Kruskal-Wallis test was used to compare three or more groups. The Spearman correlation test was also used to investigate the correlation between gene expression, copy number, and drug activity. Correlation tests with GI_{50}s were carried out separately for each of 12 TKIs and the median, maximum, and minimum Spearman *r* values for 12 TKIs are presented. A Spearman correlation coefficient of 0.30 was considered to be indicative of a weak but real association unless a specific *P* value ($P = 0.05$ was considered as statistically significant) was shown. *EGFR* mRNA level estimated by real-time PCR was log transformed before analysis. All tests were done using GraphPad Prism version 4.0 for Windows (GraphPad Software).[12]

Results

Mutation detection in the EGFR gene. No deletion mutation (*EGFRVIII*) in the NCI60 was detected. In the TK domain, two

[9] https://innateimmunity.net/IIPGA2/Bioinformatics/
[10] http://anthro.unige.ch/arlequin
[11] http://dtp.nci.nih.gov/docs/dtp_search.html
[12] http://www.graphpad.com

putative mutations in exon 19 were found. In the leukemia cell line RPMI-8226, a heterozygous C > T change was detected at the second base of codon 751, leading to a misssense substitution of threonine (T) to isoleucine (I). A homozygous C > T change was found in the melanoma cell line SK-MEL-28 at the first base of codon 753, leading to a misssense change of proline (P) to serine (S). Because of the absence of matched normal DNA, we are not able to determine whether these two misssense amino acid changes are oncogenic mutations. We therefore predicted the putative alteration in the function of EGFR by these two mutations with the SIFT program. SIFT scores of 0.00 and 0.01 were predicted for T751I and P753S, respectively (the threshold for a deleterious substitution is 0.05). This strongly suggests that both variations are oncogenic mutations.

The synonymous polymorphism 2607A/G was observed in exon 20 when sequencing the EGFR TK domain, and the genotypes were obtained.

Frequency of polymorphisms in the NCI60 cell lines. Five polymorphisms were genotyped in all 58 samples. The intron 1 (CA)n polymorphism could not be genotyped in one cell line (MOLT-4). We also genotyped the -216G/T, -191C/A, intron 1 (CA)n, and R497K polymorphisms in the HMC samples whereas genotypes for 2607A/G were downloaded from the HapMap database. Given the Caucasian origin of the NCI60 cell lines (29), the allelic distributions of the -216G/T, intron 1 (CA)n, R497K, and 2607A/G polymorphisms are similar to those of the HMC samples (Fisher's exact test, $P > 0.05$ for all tests; Table 1). The allele frequency of -191A is higher in the HMC samples compared with the NCI60 samples (20.8% versus 7.8%; Fisher's exact test, $P = 0.005$). However, the latter is very close to what we have previously reported (20). No deviation from Hardy-Weinberg equilibrium was observed ($P > 0.05$) for all polymorphisms in both populations. This suggests that the allele distribution of these polymorphisms in the cancer genomes (NCI60) is similar to that in the normal genomes (HMC).

Linkage disequilibrium and association between polymorphisms. A low level of linkage disequilibrium among the four SNPs was observed in both populations ($r^2 < 0.06$ for all comparisons in NCI60 and $r^2 < 0.13$ in HMC samples). However, a significant association was found between the (CA)n polymorphism and both -191C/A (test of pairwise linkage disequilibrium by Arlequin software, $P < 0.0001$ and $P = 0.002$ in HMC and NCI60, respectively) and -216G/T ($P = 0.001$ and $P = 0.06$ in HMC and NCI60, respectively). An association was not observed between the (CA)n and R497K variants ($P = 0.63$ in HMC and $P = 0.78$ in NCI60). Interestingly, there is also significant linkage disequilibrium between the (CA)n and 2607A/G in NCI60 ($P = 0.0002$) but not in HMC ($P = 0.44$). An obvious allelic cosegregation of *-191A* and the allele *18* of (CA)n was observed in both populations (χ^2 test, $P < 0.01$). The repeat number of the (CA)n polymorphism is also nonrandomly associated with the genotypes of the -216G/T variant with an additive decrease of the CA repeat numbers from the G/G to T/T genotypes (Spearman $r = -0.26$, $P = 0.047$ in NCI60 and $r = -0.24$, $P = 0.06$ in HMC, respectively). This suggests a clustering of shorter alleles of the (CA)n polymorphism among T carriers of -216G/T.

Correlation between EGFR mutations, copy number, expression, and polymorphisms. *EGFR* copy numbers were evaluated by high-density SNP Chip (estCN). To further confirm this, real-time PCR was done to assess the amplification of EGFR TK domain. SKY copy number (28) of chromosome 7p12 (MT2066) was also extracted from the NCI60 database. There were significant correlations between the three data sets (estCN versus real-time PCR, $r = 0.56$; estCN versus SKY, $r = 0.61$; $P < 0.0001$ for both). We therefore adopted the mean values (meanCN) of the three data sets for further analysis. The median, min, and max of meanCNs in all cell lines were 2.85, 1.3, and 5.5, indicating a very moderate amplification of the *EGFR* gene in this panel. One NSCLC cell line (NCI-H460) showed obvious loss of heterozygosity with meanCN <1.5 copies, which was confirmed by homozygosity of all genotypes tested. Six cell lines [SK-MEL-5 (melanoma), A498 (renal), HT-29 (colon), SF-539 (central nervous system), SF-295 (central nervous system), and HOP-92 (NSCLC)] amplified to ≥4 copies.

EGFR mRNA expression levels were successfully quantified in 55 of 58 cell lines (in all except in SR, NCI-H322, and SK-MEL-2 due to bad RNA quality). These measurements were highly correlated with previously reported EGFR protein or mRNA levels deposited into the NCI60 database (Spearman correlations with the five data sets in the NCI60 database: $r = 0.55$-0.84, $P < 0.0001$ for all tests).

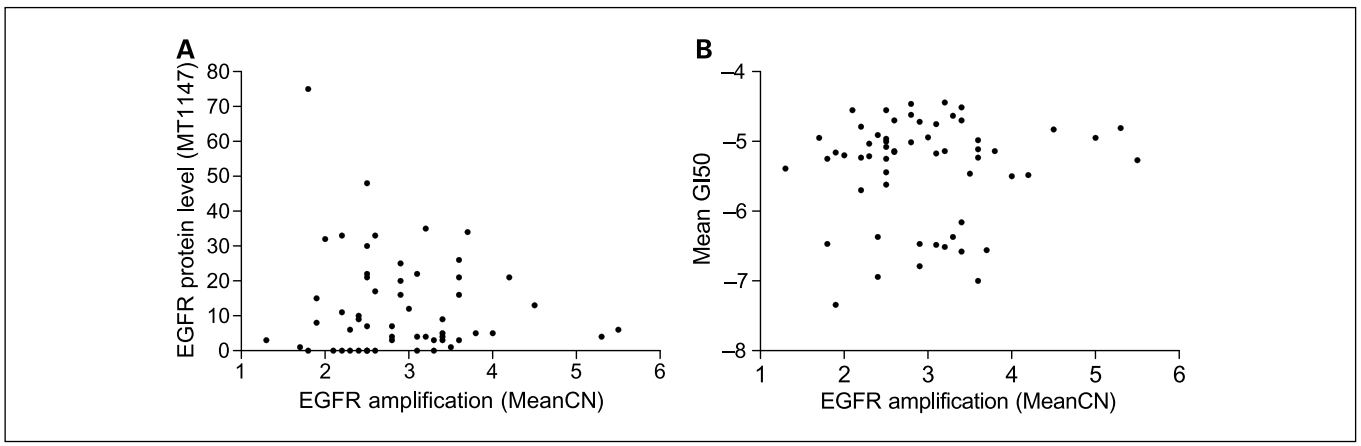

Fig. 1. Correlations between estimated copy number (meanCN) of *EGFR* and gene expression or cytotoxicity of TKIs. *A,* correlation between meanCN and EGFR protein level (MT1147) as an example (Spearman $r = -0.004$, $P = 0.98$). *B,* correlation between meanCN and the mean of GI_{50}s of 12 TKIs (Spearman $r = -0.11$, $P = 0.43$).

There was no significant correlation between *EGFR* copy number (meanCN) and any of the six measurements for *EGFR* expression mentioned above (median Spearman r = -0.02-0.15, $P > 0.05$ for all tests). A considerable number of the samples with obvious genomic amplification (meanCN, ≥3) maintained a low level of *EGFR* expression (Fig. 1A).

To evaluate the functional properties of the five polymorphisms, their genotypes were correlated with each *EGFR* expression data set. The -216G/T is the only polymorphism consistently associated with both *EGFR* protein (MT1147) and mRNA levels (real-time PCR, MT976, and microarray data sets; Table 2). The associations clearly showed an additive effect of the T allele with a positive linear trend (Fig. 2A). There was no evidence for a relationship between the intron 1 (CA)n polymorphism and *EGFR* expression (Table 2).

EGFR copy number is not associated with any of the polymorphisms investigated (data not shown). However, when the gene expression levels were stratified by the level of amplification, the correlation with -216G/T was greater in the cells with low or no amplification (meanCN, <3; Fig. 2B).

With regard to the two cell lines carrying EGFR TK mutations, a slight increase of *EGFR* copy number (meanCN, 2.8 for SK-MEL-28 and 3.3 for RPMI-8226) was observed. Furthermore, both cell lines are EGFR negative.

Correlation between EGFR and cytotoxicity of EGFR inhibitors. As previously reported, the 11 EGFR inhibitors showed similar GI_{50} fingerprints across the NCI60 cell line series (1). The data set of erlotinib showed a weaker but statistically significant correlation with the other 11 drugs (erlotinib versus the mean of the other 11 TKIs, Spearman r = 0.52, P = 0.0004).

Table 2. Correlation between *EGFR* polymorphisms and gene expression

NCI60 ID	Method	Level	Polymorphisms	Spearman correlation						*P**
				Additive		Dominant		Recessive		
				r	*P*	*r*	*P*	*r*	*P*	
MT1147	Western blot	Protein	-216G/T	0.37	<0.01	0.29	0.03	0.36	0.01	
			-191C/A	-0.26	0.06	-0.25	0.06	—	—	
			(CA)n—genotype	—	—	—	—	—	—	0.27
			(CA)n—sum	-0.21	0.12	—	—	—	—	
			R497K	0.01	0.95	-0.03	0.85	0.13	0.34	
			2607A/G	-0.03	0.80	-0.05	0.73	0.02	0.90	
MT976	PCR slot blot	RNA	-216G/T	0.27	0.05	0.17	0.21	0.31	0.02	
			-191C/A	-0.17	0.21	-0.18	0.19	—	—	
			(CA)n—genotype	—	—	—	—	—	—	0.35
			(CA)n—sum	-0.05	0.72	—	—	—	—	
			R497K	0.13	0.33	0.15	0.28	0.00	0.97	
			2607A/G	-0.11	0.41	-0.15	0.29	0.04	0.79	
MT478	RNase protection	RNA	-216G/T	0.23	0.09	0.16	0.23	0.24	0.07	
			-191C/A	-0.12	0.37	-0.12	0.38	—	—	
			(CA)n—genotype	—	—	—	—	—	—	0.70
			(CA)n—sum	-0.05	0.73	—	—	—	—	
			R497K	0.19	0.17	0.19	0.17	0.07	0.59	
			2607A/G	-0.18	0.19	-0.19	0.17	-0.05	0.70	
GC18283, GC16212 (Stanford)	Microarray	RNA	-216G/T	0.29	0.03	0.23	0.09	0.28	0.03	
			-191C/A	-0.20	0.14	-0.20	0.14	—	—	
			(CA)n—genotype	—	—	—	—	—	—	0.56
			(CA)n—sum	-0.12	0.37	—	—	—	—	
			R497K	0.16	0.24	0.16	0.24	0.06	0.66	
			2607A/G	-0.09	0.49	-0.12	0.38	0.03	0.85	
GC27729, GC33544 (Novartis)	Microarray	RNA	-216G/T	0.31	0.02	0.24	0.07	0.30	0.02	
			-191C/A	-0.19	0.15	-0.19	0.14	—	—	
			(CA)n—genotype	—	—	—	—	—	—	0.81
			(CA)n—sum	-0.16	0.23	—	—	—	—	
			R497K	0.21	0.11	0.23	0.08	0.02	0.88	
			2607A/G	-0.14	0.31	-0.15	0.27	-0.03	0.84	
—	Real-time PCR	RNA	-216G/T	0.25	0.06	0.20	0.13	0.24	0.08	
			-191C/A	0.00	0.95	-0.02	0.88	—	—	
			(CA)n—genotype	—	—	—	—	—	—	0.90
			(CA)n—sum	-0.11	0.45	—	—	—	—	
			R497K	0.17	0.21	0.19	0.17	0.01	0.92	
			2607A/G	-0.18	0.19	-0.22	0.10	0.04	0.75	

NOTE: A total of six data sets were included. For -191C/A, because there was only one A/A homozygote, only additive and dominant models were tested. The intron 1 (CA)n polymorphism was tested with the additive model only when the sum of repeat numbers was used as a variable. Kruskal-Wallis test was also used to test the genotype-phenotype correlation of the (CA)n polymorphism. For this, only the genotype groups with more than two samples each were included (16/16, 16/17, 16/18, 16/20, and 20/20).
*Kruskal-Wallis test.

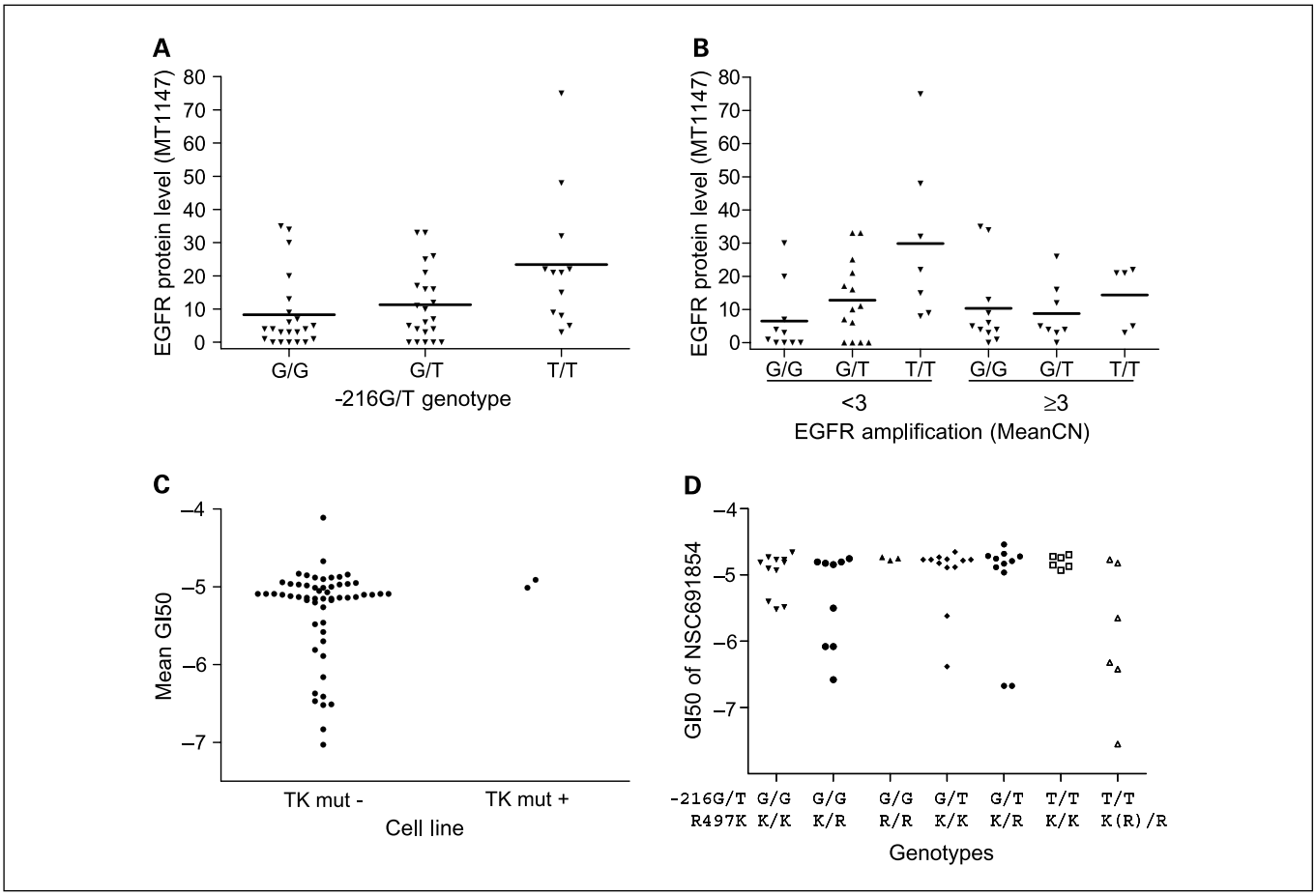

Fig. 2. Correlation between polymorphisms and *EGFR* expression or drug cytotoxicity. *A,* -216G/T and *EGFR* level by Western blot (MT1147) is presented as an example: Spearman $r = 0.37$, $P = 0.005$. *B,* data categorized by *EGFR* amplification status. Cells were divided into groups with low or no amplification (meanCN, <3) and higher amplification (meanCN, ≥3). One cell line (NCI-H460) with *EGFR* gene deletion was not included. Horizontal bars, mean of each group. Spearman $r = 0.50$, $P = 0.005$ for the low amplification groups; $r = 0.14$, $P = 0.51$ for the higher amplification ones. *C,* cytotoxicity in the cell lines with or without EGFR TK mutations. *D,* association between the drug cytotoxicity of NSC691854 and the combinations of -216G/T and R497K genotypes (ANOVA, $P = 0.03$). The only sample with -216T/T and 497R/R (SK-OV-3) genotypes was combined into the last group. Similar patterns were also observed for other four drugs: NSC691856 ($P = 0.09$), NSC691859 ($P = 0.04$), NSC691860 ($P = 0.03$), and NSC709239 ($P = 0.06$).

An earlier report also suggested that *EGFR* expression levels were significantly associated with drug sensitivity of the 11 TKIs (1). In this study, the cytotoxicity of 11 TKIs showed a significant correlation with the new measurement of *EGFR* mRNA expression by real-time PCR (level of mRNA versus the mean of 11 TKIs, Spearman $r = -0.50$, $P < 0.0001$). Compared with the protein level measured by Western blot, better correlations between mRNA measurements (RNase protection, microarray, and real-time PCR) and the 11 TKIs response were observed (Table 3). Similarly, the cytotoxicity of erlotinib correlated with all *EGFR* mRNA expression data sets ($r = -0.43$ to -0.35, $P < 0.03$ for all tests) but not with protein level (erlotinib versus MT1147, $r = -0.20$, $P = 0.23$). For all 12 TKIs, despite the positive correlation between sensitivity and increased gene expression, there is still a considerable subgroup of resistant cells that have high *EGFR* expression based on both mRNA and protein levels. Nevertheless, the data confirmed that cells with no detectable *EGFR* protein are indeed resistant to these agents (data not shown).

Next, we searched for possible links between sensitivity to the 12 TKIs and *EGFR* genomic alterations. Both cell lines carrying TK mutations are insensitive to all 12 TKIs (Fig. 2C). No significant association was found between gene copy number and drug cytotoxicity (meanCN versus 12 TKIs, $r = -0.16$ to 0.17, $P > 0.19$ for all tests; Fig. 1B) as well. However, the correlation coefficients between *EGFR* expression and GI_{50}s were consistently higher in the cells with low or no amplification (meanCN, <3) when compared with the cells with increased *EGFR* copy number (meanCN, ≥3; Table 3).

Although -216G/T is strongly associated with *EGFR* expression, there is no obvious correlation between this polymorphism (and any of the other four polymorphisms) and drug cytotoxicity. However, when combining -216G/T and R497K, there are weak associations between diplotypes and the response of 5 of 12 drugs (Fig. 2D), and the G-K and T-K haplotypes are associated with relative resistance (compared with the T-R haplotype).

Discussion

In this study, we tested the relationships between *EGFR* germ-line polymorphisms, gene amplification, TK domain mutations, gene expression, and cytotoxicity of EGFR inhibitors in NCI60. We found that whereas TK mutations are rare and

associated with resistance, higher *EGFR* expression, particularly in the cells with relatively normal *EGFR* genomic status, is correlated with sensitivity of EGFR inhibitors. However, *EGFR* gene amplification and DNA polymorphisms are poor predictors for response, although the latter are strongly associated with *EGFR* expression.

Mutations in the *EGFR* TK domain have been discovered in NSCLC and have been associated with response to gefitinib and erlotinib (5–8, 13). By sequencing the exons encoding the *EGFR* TK domain, we found that *EGFR* TK mutations are absent in the NCI60 NSCLC lines but occur in a leukemia (RPMI-8226) and a melanoma cell line (SK-MEL-28). This is in agreement with previous studies (6, 30–35) wherein the EGFR TK domain of either single NSCLC cell lines in this panel or the whole NCI60 has been sequenced. Interestingly, both mutant cell lines produce low or no detectable EGFR and are insensitive to EGFR inhibitors. Whether these mutations led to low *EGFR* expression and therefore resulted in resistance requires further investigation.

Gene amplification is one of the important mechanisms leading to overexpression of *EGFR* in cancer. In colon cancer, *EGFR* amplification has been shown to be a predictive factor for the response to cetuximab or panitumumab, two EGFR antibodies (10). A similar mechanism has been shown for glioma (12) and NSCLC (3, 11, 36) patients treated with either erlotinib or gefitinib. Although amplification of *EGFR* is common in the NCI60 panel, it is correlated with neither gene expression nor cytotoxicity. Unlike the higher copy number observed in cancer patients (10–12, 36), the *EGFR* amplification in these cells is very moderate. A majority of the cell lines have an increased gene copy number but a relatively low expression level, suggesting that the 2- to 3-fold *EGFR* amplification observed may not be sufficient to generate high levels of EGFR. Because there was no correlation between gene amplification and protein expression, and most of the samples with amplification did not overexpress EGFR, it is perhaps not surprising that drug cytotoxicity did not correlate with amplification.

Table 3. Spearman correlation between *EGFR* expression and cytotoxicity of TKIs stratified by the gene amplification

EGFR pattern		Spearman *r*		
Amplification (meanCN)	**Expression**	**Median**	**Min**	**Max**
<3	MT1147	-0.25	-0.68	0.05
	MT976	-0.36	-0.59	0.01
	MT478	-0.36	-0.70	-0.02
	Microarray (STF)	-0.44	-0.76	-0.11
	Microarray (MOV)	-0.38	-0.71	-0.08
	Real-time PCR	-0.41	-0.70	-0.17
≥3	MT1147	-0.24	-0.49	-0.03
	MT976	-0.01	-0.33	0.21
	MT478	-0.29	-0.59	0.07
	Microarray (STF)	-0.28	-0.50	-0.06
	Microarray (MOV)	-0.29	-0.60	-0.09
	Real-time PCR	-0.31	-0.56	-0.03
All	MT1147	-0.24	-0.54	-0.03
	MT976	-0.18	-0.37	0.05
	MT478	-0.33	-0.59	-0.11
	Microarray (STF)	-0.37	-0.57	-0.07
	Microarray (MOV)	-0.32	-0.62	-0.06
	Real-time PCR	-0.36	-0.63	-0.14

NOTE: Median, min, and max Spearman *r* values for 12 drugs are shown.

EGFR expression, in particular the mRNA level, was a determinant of sensitivity in this panel because high levels of *EGFR* seem to be necessary (but not sufficient) for sensitivity. The lesser predictive value of EGFR protein level might be due to the limited sensitivity of Western blot compared with other techniques measuring mRNA. The correlation, however, may only be relevant to a small subset of cells with genetically normal EGFR. Correlations between *EGFR* expression and drug sensitivity were increased when the cell lines with gene amplification were removed from the analysis, suggesting that in cancer cells, when somatic changes of EGFR are absent, germ-line variability in regulation of *EGFR* expression may confer some EGFR-dependent environment. In clinical studies, it seems that a subgroup of patients with normal *EGFR* status respond to the TKIs (3, 14). Recent studies in NSCLC also suggested a significant relationship between *EGFR* expression and better response to gefitinib or erlotinib administration (37). Thus, our findings may be relevant to the patients with genetically normal EGFR. Further investigations on this in clinical studies would be obviously helpful to understand the pharmacodynamics of EGFR inhibitors completely.

Variability in *EGFR* expression could be caused by germ-line polymorphisms. An intronic microsatellite polymorphism has been associated with *EGFR* expression, with the repeat length of CA nucleotides being inversely correlated with *EGFR* mRNA and protein levels (38, 39). A recent study showed that in the absence of *EGFR* amplification, there was an association between the shorter alleles of intron 1 (CA)n and better erlotinib response in head and neck cancer cell lines (15). This polymorphism has also been suggested to be correlated with skin rash, a major adverse effect of EGFR inhibitors but a potential marker for improved response in cancer patients (15). Among lung cancer patients treated with gefitinib or erlotinib, shorter alleles were also found to be associated with better clinical outcomes (16–18). We recently showed that the -216G/T polymorphism might independently modulate *EGFR* promoter activity and gene transcription (20). In this study, we confirmed that the -216G/T was significantly associated with both protein and mRNA expression of *EGFR*, with the T allele carriers showing greater expression compared with the G allele. In contrast, we did not observe any correlation between either genotype or sum of allele numbers of the intron 1 (CA)n and *EGFR* expression. Interestingly, a significant correlation between the repeat length of the (CA)n and -216G/T genotypes was observed with the shorter alleles clustering in the T allele carriers in both NCI60 and the HapMap samples. Thus, it is unlikely that the intron 1 (CA)n polymorphism alone regulates *EGFR* transcription, and the previously reported association of the intron 1 (CA)n polymorphism with *EGFR* expression or pharmacodynamics of EGFR-targeted agents may be due to the association of this microsatellite polymorphism with the functional variant in the promoter. This has been further confirmed in a recent study with advanced NSCLC patients treated with gefitinib. Improved progression-free survival was found in patients with either shorter length alleles of the (CA)n polymorphism or at least one *-216T* allele, but the latter had a stronger association. The

-216G/T was also significantly associated with better response and treatment-related toxicities (19).

However, the relationship of these polymorphisms with the GI_{50}s of 12 EGFR inhibitors in the NCI60 is less clear. The best associations were observed when combining -216G/T and R497K polymorphisms. The G-K and T-K haplotypes were associated with resistance and the T-R haplotype seemed to be associated with sensitivity. Because the 497R allele seems to increase EGFR activity (22), the T-R haplotype may confer both higher expression and increased EGFR activity. Although this might result in a better response, the limited number of samples in this study precludes any definitive conclusions.

We recognize many limitations of this *in vitro* study, including the heterogeneity of the NCI60 panel. The data, however, warrant further explorations in both preclinical and clinical areas, particularly with regard to studies of the relationship of germ-line polymorphisms (in *EGFR* and related genes) to the pharmacodynamics of this important group of agents.

Acknowledgments

We thank the National Cancer Institute for their contribution in the development of the 60 cell line screening panel and for their willingness to provide cells for analysis; Pfizer, Inc., for providing permission to use the data for 11 EGFR inhibitors; and Jacqueline Ramirez for her assistance in the preparation of the manuscript.

References

1. Bishop PC, Myers T, Robey R, et al. Differential sensitivity of cancer cells to inhibitors of the epidermal growth factor receptor family. Oncogene 2002;21:119–27.
2. Wosikowski K, Schuurhuis D, Johnson K, et al. Identification of epidermal growth factor receptor and c-erbB2 pathway inhibitors by correlation with gene expression patterns. J Natl Cancer Inst 1997;89:1505–15.
3. Tsao MS, Sakurada A, Cutz JC, et al. Erlotinib in lung cancer—molecular and clinical predictors of outcome. N Engl J Med 2005;353:133–44.
4. Tang PA, Tsao MS, Moore MJ. A review of erlotinib and its clinical use. Expert Opin Pharmacother 2006;7:177–93.
5. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004;304:1497–500.
6. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129–39.
7. Pao W, Miller V, Zakowski M, et al. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A 2004;101:13306–11.
8. Sequist LV, Bell DW, Lynch TJ, et al. Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. J Clin Oncol 2007;25:587–95.
9. Ji H, Zhao X, Yuza Y, et al. Epidermal growth factor receptor variant III mutations in lung tumorigenesis and sensitivity to tyrosine kinase inhibitors. Proc Natl Acad Sci U S A 2006;103:7817–22.
10. Moroni M, Veronese S, Benvenuti S, et al. Gene copy number for epidermal growth factor receptor (EGFR) and clinical response to anti-EGFR treatment in colorectal cancer: a cohort study. Lancet Oncol 2005;6:279–86.
11. Cappuzzo F, Hirsch FR, Rossi E, et al. Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. J Natl Cancer Inst 2005;97:643–55.
12. Haas-Kogan DA, Prados MD, Tihan T, et al. Epidermal growth factor receptor, protein kinase B/Akt, and glioma response to erlotinib. J Natl Cancer Inst 2005;97:880–7.
13. Pao W, Miller VA, Politi KA, et al. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. PLoS Med 2005;2:e73.
14. Bell DW, Lynch TJ, Haserlat SM, et al. Epidermal growth factor receptor mutations and gene amplification in non-small-cell lung cancer: molecular analysis of the IDEAL/INTACT gefitinib trials. J Clin Oncol 2005;23:8081–92.
15. Amador ML, Oppenheimer D, Perea S, et al. An epidermal growth factor receptor intron 1 polymorphism mediates response to epidermal growth factor receptor inhibitors. Cancer Res 2004;64:9139–43.
16. Nie Q, Wang Z, Zhang GC, et al. The epidermal growth factor receptor intron1 (CA) n microsatellite polymorphism is a potential predictor of treatment outcome in patients with advanced lung cancer treated with gefitinib. Eur J Pharmacol 2007;7:177–93.
17. Tzeng CW, Frolov A, Frolova N, et al. Pancreatic cancer epidermal growth factor receptor (EGFR) intron 1 polymorphism influences postoperative patient survival and *in vitro* erlotinib response. Ann Surg Oncol 2007;14:2150–8.
18. Han SW, Jeon YK, Lee KH, et al. Intron 1 CA dinucleotide repeat polymorphism and mutations of epidermal growth factor receptor and gefitinib responsiveness in non-small-cell lung cancer. Pharmacogenet Genomics 2007;17:313–9.
19. Liu G, Gurubhagavatula S, Zhou W, et al. Epidermal growth factor receptor polymorphisms and clinical outcomes in non-small-cell lung cancer patients treated with gefitinib. Pharmacogenomics J. Epub 2007.
20. Liu W, Innocenti F, Wu MH, et al. A functional common polymorphism in a Sp1 recognition site of the epidermal growth factor receptor gene promoter. Cancer Res 2005;65:46–53.
21. Moriai T, Kobrin MS, Korc M. Cloning of a variant epidermal growth factor receptor. Biochem Biophys Res Commun 1993;191:1034–9.
22. Moriai T, Kobrin MS, Hope C, Speck L, Korc M. A variant epidermal growth factor receptor exhibits altered type α transforming growth factor binding and transmembrane signaling. Proc Natl Acad Sci U S A 1994;91:10217–21.
23. Shoemaker RH. The NCI60 human tumour cell line anticancer drug screen. Nat Rev Cancer 2006;6:813–23.
24. Bode B, Frigerio S, Behnke S, et al. Mutations in the tyrosine kinase domain of the EGFR gene are rare in synovial sarcoma. Mod Pathol 2006;19:541–7.
25. Martin K, Radlmayr M, Borchers R, Heinzlmann M, Folwaczny C. Candidate genes colocalized to linkage regions in inflammatory bowel disease. Digestion 2002;66:121–6.
26. Innocenti F, Liu W, Chen P, Desai AA, Das S, Ratain MJ. Haplotypes of variants in the UDP-glucuronosyltransferase1A9 and 1A1 genes. Pharmacogenet Genomics 2005;15:295–301.
27. Liu W, Innocenti F, Chen P, Das S, Cook EH, Jr., Ratain MJ. Interethnic difference in the allelic distribution of human epidermal growth factor receptor intron 1 polymorphism. Clin Cancer Res 2003;9:1009–12.
28. Roschke AV, Tonon G, Gehlhaus KS, et al. Karyotypic complexity of the NCI-60 drug-screening panel. Cancer Res 2003;63:8634–47.
29. Adams S, Robbins FM, Chen D, et al. HLA class I and II genotype of the NCI-60 cell lines. J Transl Med 2005;3:11.
30. Ikediobi ON, Davies H, Bignell G, et al. Mutation analysis of 24 known cancer genes in the NCI-60 cell line set. Mol Cancer Ther 2006;5:2606–12.
31. Mukohara T, Engelman JA, Hanna NH, et al. Differential effects of gefitinib and cetuximab on non-small-cell lung cancers bearing epidermal growth factor receptor mutations. J Natl Cancer Inst 2005;97:1185–94.
32. Mantha AJ, Hanson JE, Goss G, Lagarde AE, Lorimer IA, Dimitroulakos J. Targeting the mevalonate pathway inhibits the function of the epidermal growth factor receptor. Clin Cancer Res 2005;11:2398–407.
33. Conde E, Angulo B, Tang M, et al. Molecular context of the EGFR mutations: evidence for the activation of mTOR/S6K signaling. Clin Cancer Res 2006;12:710–7.
34. Petty WJ, Dragnev KH, Memoli VA, et al. Epidermal growth factor receptor tyrosine kinase inhibition represses cyclin D1 in aerodigestive tract cancers. Clin Cancer Res 2004;10:7547–54.
35. Zhao X, Weir BA, LaFramboise T, et al. Homozygous deletions and chromosome amplifications in human lung carcinomas revealed by single nucleotide polymorphism array analysis. Cancer Res 2005;65:5561–70.
36. Hirsch FR, Varella-Garcia M, Bunn PA, Jr., et al. Molecular predictors of outcome with gefitinib in a phase III placebo-controlled study in advanced non-small-cell lung cancer. J Clin Oncol 2006;24:5034–42.
37. Dziadziuszko R, Witta SE, Cappuzzo F, et al. Epidermal growth factor receptor messenger RNA expression, gene dosage, and gefitinib sensitivity in non-small cell lung cancer. Clin Cancer Res 2006;12:3078–84.
38. Gebhardt F, Zanker KS, Brandt B. Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1. J Biol Chem 1999;274:13176–80.
39. Buerger H, Gebhardt F, Schmidt H, et al. Length and loss of heterozygosity of an intron 1 polymorphic sequence of egfr is related to cytogenetic alterations and epithelial growth factor receptor expression. Cancer Res 2000;60:854–7.

Clinical Cancer Research

Relationship of *EGFR* Mutations, Expression, Amplification, and Polymorphisms to Epidermal Growth Factor Receptor Inhibitors in the NCI60 Cell Lines

Wanqing Liu, Xiaolin Wu, Wei Zhang, et al.

Clin Cancer Res 2007;13:6788-6795.