Homozygous Deletions and Chromosome Amplifications in Human Lung Carcinomas Revealed by Single Nucleotide Polymorphism Array Analysis

Xiaojun Zhao,[1,4] Barbara A. Weir,[1,4] Thomas LaFramboise,[1,4] Ming Lin,[2,6] Rameen Beroukhim,[1,5] Levi Garraway,[1,5] Javad Beheshti,[1,4] Jeffrey C. Lee,[1,4] Katsuhiko Naoki,[11] William G. Richards,[10] David Sugarbaker,[10] Fei Chen,[4,8] Mark A. Rubin,[4,8] Pasi A. Jänne,[1,3,5,9] Luc Girard,[12,13] John Minna,[12,13] David Christiani,[5,7] Cheng Li,[2,6] William R. Sellers,[1,5,14] and Matthew Meyerson[1,4,14]

Departments of [1]Medical Oncology and [2]Biostatistical Sciences, and [3]Lowe Center for Thoracic Oncology, Dana-Farber Cancer Institute; Departments of [4]Pathology and [5]Medicine, Harvard Medical School; Departments of [6]Biostatistics and [7]Environmental Health, Harvard School of Public Health, Boston, Massachusetts; Departments of [8]Pathology and [9]Medicine, and [10]Division of Thoracic Surgery, Brigham and Women's Hospital, Boston, Massachusetts; [11]Department of Respiratory Medicine, Yokohama Municipal Citizen's Hospital, Yokohama, Japan; [12]Hamon Center for Therapeutic Oncology Research; [13]Departments of Internal Medicine and Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas; and [14]Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts

Abstract

Genome-wide copy number changes were analyzed in 70 primary human lung carcinoma specimens and 31 cell lines derived from human lung carcinomas, with high-density arrays representing ~115,000 single nucleotide polymorphism loci. In addition to previously characterized loci, two regions of homozygous deletion were found, one near the *PTPRD* locus on chromosome segment 9p23 in four samples representing both small cell lung carcinoma (SCLC) and non–small cell lung carcinoma (NSCLC) and the second on chromosome segment 3q25 in one sample each of NSCLC and SCLC. High-level amplifications were identified within chromosome segment 8q12-13 in two SCLC specimens, 12p11 in two NSCLC specimens and 22q11 in four NSCLC specimens. Systematic copy number analysis of tyrosine kinase genes identified high-level amplification of *EGFR* in three NSCLC specimens, *FGFR1* in two specimens and *ERBB2* and *MET* in one specimen each. *EGFR* amplification was shown to be independent of kinase domain mutational status. (Cancer Res 2005; 65(13): 5561-70)

Introduction

Mapping copy number alterations in the cancer genome has contributed to the subsequent identification of tumor suppressor genes and oncogenes. The delineation of cancer-specific homozygous deletions enabled the discovery of several different tumor suppressor genes, including *RB1* (1), *CDKN2A* (2, 3), *PTEN* (4, 5), and *SMAD4/DPC4* (6). Similarly, genes that are targets for high-level amplification in cancers are likewise often subject to oncogenic mutations. Examples where identification of cancer-specific amplifications preceded the identification of cancer somatic mutations include *PDGFRA* (7, 8), *EGFR* (9–14), *ERBB2* (15, 16), and *PIK3CA* (17–19).

High-resolution genomic approaches now make it possible to screen for chromosomal copy number alterations in a systematic manner. Array comparative genomic hybridization can provide high-resolution detection of copy number changes (20, 21). Single nucleotide polymorphism (SNP) arrays can measure cancer-specific loss of heterozygosity (LOH) with high accuracy in a genome-wide fashion (22–26). Furthermore, SNP arrays that cover ~10,000 SNP loci can be used to detect DNA copy number changes at the genome level, including high-level amplifications and homozygous deletions (27, 28).

Lung cancer is the leading cause of cancer death in the world and is estimated to result in ~160,000 deaths annually in the United States alone (29). Genomic studies have begun to impact the diagnosis and treatment of human lung carcinoma. For example, mutations in the *EGFR* kinase domain in non–small cell lung carcinoma (NSCLC) specimens were found to correlate with patient responses to gefitinib and erlotinib (11–13).

To discover novel genomic changes in human lung carcinomas, in the hopes of identifying additional pathways active in these diseases, we have used SNP arrays covering ~115,000 SNP loci to investigate copy number changes in a panel of DNA from 77 NSCLC and 24 small cell lung carcinoma (SCLC) lung cell lines and primary tumors. Given the high resolution of the SNP arrays, we have been able to identify several small homozygous deletions and amplifications that have not been detected by previous methods.

Materials and Methods

Primary tumor and cell line specimens. We obtained the following genomic DNA: lung adenocarcinoma (HOP-62, NCI-H23), large cell lung carcinoma (LC; HOP-92), and squamous cell lung carcinoma (NCI-H266) from the National Cancer Institute. We prepared genomic DNA from the following cell lines: adenocarcinoma (H1437, H1819, H1993, H2009, H2087, H2122, H2347, HCC193, HCC461, HCC515, HCC78, HCC827), adenosquamous lung carcinoma (HCC366), LC (H2126, HCC1359, HCC1171), unspecified NSCLC (H2882, H2887), squamous cell lung carcinoma (H157, HCC15, HCC95), bronchioloalveolar carcinoma (BAC; H358), and SCLC (H524, H526, H1184, H1607, H1963). The primary tumors were from anonymous patients and were surgically dissected and frozen at −80°C until use. All primary tumor specimens were examined histologically to ensure at least 70% neoplastic tissue, except SCLC samples that were considered to have high tumor contents. These tumors consisted of 19 SCLC and 51 NSCLC. These include SCLC (S0168T, S0169T, S0170T, S0171T, S0172T, S0173T, S0177T, S0185T, S0187T, S0188T, S0189T, S0190T, S0191T, S0192T, S0193T, S0194T, S0196T, S0198T, S0199T), lung adenocarcinoma (MGH1622T, MGH7T, MGH1028T, S0356T, S0372T, S0377T, S0380T, S0392T,

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

X. Zhao and B. Weir contributed equally to this work.

Requests for reprints: Matthew Meyerson, Department of Medical Oncology, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115. Phone: 617-632-4768; Fax: 617-632-5998; E-mail: matthew_meyerson@dfci.harvard.edu.

S0395T, S0397T, S0405T, S0412T, S0464T, S0471T, S0479T, S0482T, S0488T, S0498T, S0500T, S0502T, S0514T, S0522T, S0524T, S0534T, S0535T, S0539T, AD157T, AD163T, AD309T, AD311T, AD327T, AD330T, AD334T, AD335T, AD336T, AD337T, AD347T), squamous cell lung carcinoma (S0446T, S0449T, S0458T, S0465T, S0480T, S0485T, S0496T, S0508T, S0515T, S0536T), adenocarcinoma/BAC (S0376T), and BAC (S0509T, AD338T, AD362T).

Single nucleotide polymorphism array. For each sample, SNPs were genotyped with two different arrays, CentXba and CentHind, in parallel (Affymetrix, Inc., Santa Clara, CA). Array experiments were done according to the manufacturer's recommendations. In brief, two aliquots of DNA (250 ng each) were first digested with *Xba*I or *Hin*dIII restriction enzyme (New England Biolabs, Boston, MA), respectively. The digested DNA was ligated to an adaptor before subsequent PCR amplification using AmpliTaq Gold (Applied Biosystems, Foster City, CA). Four 100 μL PCR reactions were then set up for each *Xba*I or *Hin*dIII adaptor-ligated DNA sample. The PCR products from four reactions were pooled, concentrated, and fragmented with DNase I to a size range of 250 to 2,000 bp. Fragmented PCR products were then labeled, denatured, and hybridized to the array. After hybridization, the arrays were washed on the Affymetrix fluidics stations, stained, and scanned using the Gene Chip Scanner 3000 and the genotyping software, Affymetrix Genotyping Tools Version 2.0.

Data analysis. Data were normalized to a baseline array with median signal intensity at the probe intensity level with the invariant set normalization method described by Li and Wong (30). After normalization, the signal values for each SNP in each array were obtained with a model-based (PM/MM) method (31). Signal intensities at each probe locus were compared with a set of normal reference samples representing 12 individuals. From raw signal data, the inferred copy numbers at each SNP locus was estimated by applying the hidden Markov model (HMM; ref. 27). We applied the HMM model based on the assumption of diploidy or triploidy; thus, possible normalized copy numbers are (0, 1, 2, 3, 4, . . .; diploid) or (0, 0.67, 1.33, 2, 2.67, 3.33, 4, . . .; triploid), leading to the possible copy number set (0, 0.67, 1, 1.33, 2, 2.67, 3, 3.33, 4, . . .). The analysis methods described above are implemented in the dChip software Version 1.3, which is freely available to academic users (http://biosun1.harvard.edu/~cli/dchip2004.exe). Mapping information of SNP locations and cytogenetic band are based on curation of Affymetrix and University of California Santa Cruz hg 16 (http://genome.ucsc.edu).

We applied the circular binary segmentation algorithm (32) to our raw $\log_2$ ratio data. This algorithm recursively splits chromosomes into subsegments based on a maximum t statistic. The reference distribution for this statistic, estimated by permutation, is used to decide whether or not to split at each stage (see ref. 32 for details). We compared the (rounded) mean raw estimated copy for each segment to our HMM results.

Quantitative real-time PCR. Relative gene copy numbers and gene expression were determined by quantitative real-time PCR using a PRISM 7500 sequence Detection System (Applied Biosystems) and a QuantiTect SYBR Green PCR kit and a QuantiTect SYBR Green RT-PCR kit (Qiagen, Inc., Valencia, CA). The standard curve method was used to calculate target gene copy number in the tumor DNA sample normalized to a repetitive element Line-1 and normal reference DNA. The comparative threshold cycle method was used to calculate gene expression normalized to *β-actin* as a gene reference and normal human lung RNA as an RNA reference. Primers were designed using *Primer 3* (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and synthesized by Invitrogen (Carlsbad, CA). Primer sequences are available upon request.

Interphase fluorescence *in situ* hybridization. Fluorescence *in situ* hybridization (FISH) probes were made from BAC clones RP11-805B16 and RP11-153K21 (Children's Hospital, Oakland Research Institute, Oakland, CA), identified to overlap the *ASPH* locus. BAC DNA was purified and 100 ng of each clone was labeled with digoxigenin-dUTP using random primers. The DNA was then purified with a MicroSpin S-200 HR column, ethanol precipitated, and resuspended in 100 μL hybridization solution. The control FISH probe, CEP 8 SpectrumOrange (Vysis, Downers Grove, IL), detects the centromeric region of chromosome 8. H2171 cells were grown in culture using standard methods and harvested by centrifugation after 3 days of growth. Slides for FISH analysis were prepared according to the control probe manufacturer's directions (Vysis). Briefly, cell pellets were resuspended in a fixative, a 3:1 solution of methanol and acetic acid. The cell suspension was dropped onto slides dipped in fixative and dried for 10 minutes over a 67°C water bath. The slides were pretreated with 2× SSC at 37°C for 1 hour, then digested for 5 minutes with a 1:25 solution of All III and rinsed with PBS. They were then incubated for 1 minute in 10% buffered formalin at room temperature, rinsed with PBS, dehydrated in an ethanol series (70%, 85%, 95%, and 100%), and air dried. Ten microliters of probe solution (6 μL hybridization buffer, 1 μL Cot-1 DNA, 1 μL centromere control probe, and 1 μL each of RP11-805B16 and RP11-153K21 probes) were incubated on the slide under a sealed coverslip for 3.5 minutes at 85°C and then placed in a humidified chamber overnight at 37°C. The slides were then washed in 0.5× SSC at 75°C for 5 minutes. The *ASPH* probes were detected using a 1:500 dilution of FITC antidigoxigenin in 10% normal goat serum, with 4′,6-diamidino-2-phenylindole as a counterstain.

Results

Genome-wide analysis of lung cancer. One hundred one human lung carcinoma DNA samples, including 51 NSCLC primary tumor samples, 26 NSCLC cell line samples, 19 SCLC primary tumor samples, and 5 SCLC cell line samples, were hybridized to SNP arrays containing 115,593 mapped SNP loci. Two independent algorithms, the HMM in dChipSNP (27) and binary segmentation analysis (32), were used to infer copy number and thereby to identify genomic amplifications and deletions.

The genomes of lung carcinomas are often complex with numerous chromosomal alterations. For example, the lung adenocarcinoma cell line H2122 shows homozygous deletions on chromosome arms 2q, 8q, and 9p, and a significant amplification on chromosome arm 8q (Fig. 1*A*). Analyses of all tumor samples, throughout the genome, identifies recurrent regions of copy number gain and loss in lung carcinomas (Fig. 1*B* and *C*; Supplementary Fig. S1 shows copy number estimates for each sample across the entire genome). In both NSCLC and SCLC, the most frequent copy number gains were found in chromosome arm 5p, with ≥3 copies found in 25% and 43% of samples, respectively (Fig. 1*B* and *C*). However, the region of 5p copy number gain is usually large, and we were not able to identify any area of focal amplification from our data set.

Maximum degrees of copy number loss at a given locus, with an inferred copy number <1.5 based on HMM analysis, were found most often in chromosome arms 8p (33%) and 9p (26%) in NSCLC (Fig. 1*B*) and in chromosome arms 3p (68%) and 4q (58%) in SCLC (Fig. 1*C*). These findings are broadly similar to results reported by CGH (33); however, overall, we are reporting a somewhat lower frequency of these chromosome alterations. Possible reasons include the following: (*a*) the presence of stromal admixture in primary tumor samples; (*b*) copy number signal is attenuated either at the level of primary hybridization intensity or due to the application of the HMM; or (*c*) we are analyzing maximal regions of loss rather than tallying the presence of loss at any site within a whole chromosome arm. For example, the SCLC cell lines NCI-H524 and NCI-H526 only show partial loss of chromosome 3p in SNP array analysis (Supplementary Fig. S1); similar results are seen for the same cell lines upon analysis with cDNA array CGH.[15] The large regions of modest copy gain and loss shown in Fig. 1*B* and *C* do not determine the presence or absence of alterations in specific genes within these regions.

Homozygous deletions and amplifications are of particular interest because they may indicate tumor suppressor gene and oncogene loci, respectively. Regions of homozygous deletion were

[15] L. Girard and J. Minna, unpublished data.

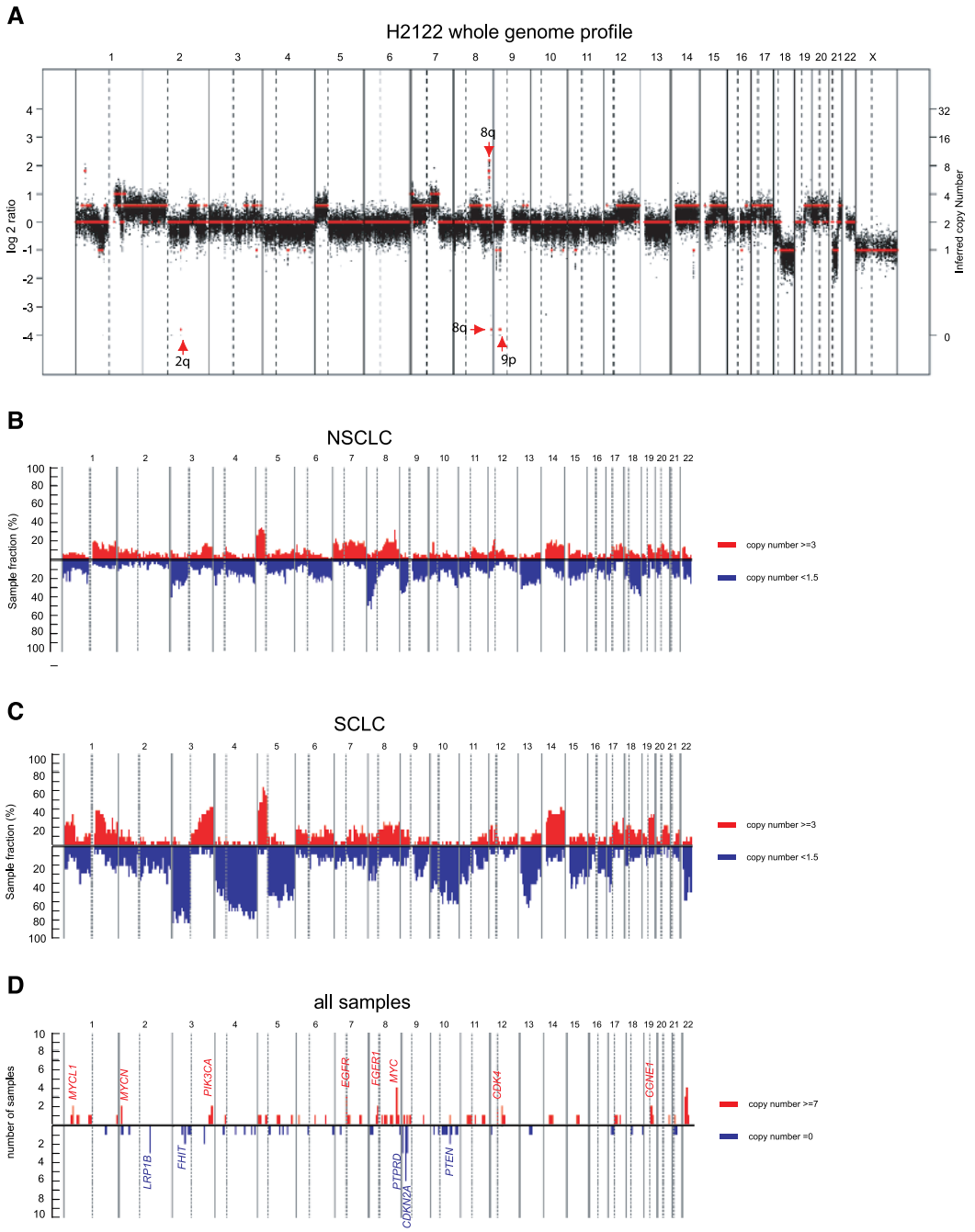

Figure 1. DNA copy number alterations by SNP array analysis. *A,* scatter plot of $\log_2$ copy number ratios (*black dots, left axis*) and inferred copy number derived by HMM (*red dots, right axis*) versus genomic position of all SNPs, for the NSCLC cell line H2122. The ratios are displayed as a median of each SNP and its two nearest neighbors. *Red arrows,* high-level amplification within chromosome arm 8q (inferred copy number ≥7) and homozygous deletions within chromosome arms 2q, 8q, and 9p (inferred copy number = 0). *B* and *C,* the fraction of samples with copy number amplification of at least three copies (*red*), and copy number reduction to <1.5 copies (*blue*) across all autosomal SNPs is shown for all NSCLC samples (*B*) and all SCLC samples (*C*). SNP markers are ordered according to their mapped positions from chromosome 1 to 22 with vertical solid and dashed lines (*gray*) indicating chromosome boundaries and centromeres, respectively. *D,* the recurrence of high-level amplification and homozygous deletions across all autosomal SNPs. *Red bars,* the number of samples with amplifications of at least copy number 7; *blue bars,* the number of samples with homozygous deletions. The recurrent regions (≥2) that harbor known or candidate oncogene (*red*) and tumor suppressor genes (*blue*) are highlighted. Note that the indicated genes are representative and do not necessarily represent that functional target of the chromosomal alteration.

defined as segments of at least four SNP loci covering >5 kb with an inferred copy number of 0 by HMM. Similarly, regions of high-level amplification were defined as segments at least four SNP loci covering >5 kb with an inferred copy number ≥7.

Copy number alterations that recurred in multiple samples were verified by real-time PCR (Table 1). In general, copy number estimation was consistent between the HMM and binary segmentation approaches. Note that whereas quantitative real-time PCR results

Table 1. Predicted regions of recurrent amplification and homozygous deletion

Cytoband*	Start (Mb)*	Stop (Mb)*	Size (Mb)*	Sample	Mean dCHIP copy number†	Binary segmentation copy number	Real-time PCR copy number	Representative gene within interval	Total no. genes*
1p34.2	39.51	40.55	1.04	H1963	10.6	11	147.8 ± 32.3	*MYCL1*	23
	39.55	40.91	1.36	S0173T	10.7	8	20.7 ± 4.9		24
2p24.3-p24.2	14.20	16.38	2.18	S0172T	14.2	12	56.2 ± 13.7	*MYCN*	9
	15.25	17.06	1.81	H526	7.0	5	42.3 ± 6.4		9
2q22.1	141.71	142.45	0.73	H2122	0	0	0.0 ± 0.0	***LRP1B***‡	1
	141.79	142.78	0.99	HCC95	0	1	0.0 ± 0.0		1
	141.94	142.20	0.26	H2126	0	0	0.0 ± 0.0		1
	142.00	142.20	0.20	H157	0	0	0.1 ± 0.0		1
3p14.2	60.29	60.78	0.49	HCC95	0	0	0.0 ± 0.0	*FHIT*	4
	60.32	60.40	0.08	H2887	0	0	0.0 ± 0.0		1
3q25.1	152.82	152.95	0.12	H2882	0	0	0.0 ± 0.0§	*AADAC, SUCNR1*	2
	152.82	152.95	0.12	S0177T	0	0	0.0 ± 0.0§		2
3q26.31-q27.1	174.86	184.52	9.66	S0465T	7.8	5	10.3 ± 1.7	*PIK3CA*	41
	182.50	184.47	1.98	S0515T	13.2	11	3.9 ± 0.4		12
7p12.1-q11.22	53.16	61.49	11.34	HCC827	11.3	9	41.7 ± 8.6	*EGFR*	50
	54.24	69.62	9.73	AD347T	9.8	9	18.3 ± 4.8		123
	54.37	55.63	13.65	S0480T	13.7	10	67.1 ± 10.9		12
8p12-p11.22	38.05	39.97	1.93	MGH1622T	10.4	11	14.9 ± 5.9	*FGFR1*	22
	38.73	39.84	1.11	S0449T	6.4	5	6.07 ± 1.7		11
8q24.13-q24.21	126.60	128.89	2.3	H524	6.6	5	174.5 ± 39.3	*MYC*	6
	127.46	128.89	1.43	HCC827	6.9	6	8.6 ± 1.4		6
	127.59	130.83	3.24	NCI-H23	8.0	7	11.1 ± 2.8		11
	127.90	129.62	1.72	H2122	3.6	5	14.5 ± 4.3		6
	128.44	129.60	1.16	H2087	7.9	8	16.0 ± 5.8		4
9p23	8.61	9.12	0.51	S0177T	0	0	0.0 ± 0.0	*PTPRD*	2
	8.79	9.55	0.77	H358	0	0	0.1 ± 0.0§		2
	9.41	9.61	0.20	HCC1171	0	0	0.2 ± 0.0§		1
	9.50	9.75	0.25	H2347	0	0	0.0 ± 0.0§		1
9p21.3	20.90	22.94	2.03	H2126	0	0	0.0 ± 0.0	*CDKN2A*	35
	21.20	22.19	0.98	HCC1359	0	0	0.0 ± 0.0		21
	21.58	25.10	3.52	HCC1171	0	0	0.0 ± 0.0		11
	21.70	23.39	1.69	H2882	0	1	0.0 ± 0.0		6
	21.84	26.83	4.99	HCC95	0	0	0.0 ± 0.0		11
	21.95	22.09	0.14	H2122	0	0	0.0 ± 0.0		3
	24.34	24.70	0.36	H157	0	0	0.0 ± 0.0§		1
10q23.31	89.03	89.40	0.37	H1607	0	0	0.0 ± 0.0	*PTEN*	4
	89.18	89.88	0.69	S0187T	0	0	0.1 ± 0.0		4
	89.35	91.16	1.80	S0189T	0	0	0.1 ± 0.0§		22
12p11.21	32.17	33.02	0.85	S0515T	8.8	7	10.8 ± 5.4	*PKP2*	6
	32.69	36.59	3.90	H2087	7.8	8	11.4 ± 5.7		12
12q13.3-q14.1	56.26	57.37	1.10	H2087	8.8	9	23.4 ± 11.3	*CDK4*	20
	55.82	56.67	0.85	HCC827‖	13.0	13	30.3 ± 11.8		34
19q12	34.02	35.55	1.53	S0524T	6.7	7	6.8 ± 1.93	*CCNE1*	12
	34.79	37.09	2.30	S0188T	7.9	7	10.9 ± 3.2		12
22q11.21-q11.22	16.99	20.31	3.32	H1819	6.8	7	12.6 ± 4.8	*CRKL*	92
	17.51	21.44	3.93	HCC515	7.4	7	14.0 ± 3.6		169
	18.47	20.61	2.14	S0380T	6.4	5	8.4 ± 1.2		68
	19.45	20.75	1.29	HCC1359	6.5	7	8.1 ± 1.7		48

NOTE: Predicted regions of recurrent amplification contain at least four SNPs at least 5 kb in size and with an inferred copy number of ≥7, which occur in two or more samples. Amplified regions separated by <2 Mb of unamplified sequenced have been combined. Predicted regions of homozygous deletion, at least 5 kb in size, contain at least four consecutive SNPs with inferred copy number = 0, which occur in two or more samples. Deleted regions separated by <2 Mb of undeleted sequence have been combined.

*Based on hg16 genome assembly.

† Mean of amplified segments include sequences with copy number <7 if regions were combined.

‡ Bold indicates only gene in region.

§These real-time PCR values denote targets that are not in an exon of the representative gene.

‖Less than 4 SNPs, but validated with real-time PCR.

have been highly reproducible between duplicate experiments, the copy number increases detected by the arrays seem to saturate at a much lower level. Thus, an inferred copy number of ~7 may correspond to a quantitative PCR copy number of as high as 175 or as low as 9 (*MYC* amplification; Table 1).

On average, the number of annotated genes is greater for regions of recurrent, high-level amplification (copy number ≥7) than for recurrent homozygously deleted regions (14.6 versus 7.7 genes/Mb, respectively). Given the parameters used in these experiments, the HMM algorithm was able to identify several amplified regions that were not found by binary segmentation but were verified by real-time PCR analysis. (All homozygous deletions and high-level amplifications identified in this study are shown in Supplementary Tables S1 and S2, respectively).

The frequencies of high-level amplification (copy number ≥7) and homozygous deletion (copy number = 0) across the genome in NSCLC and SCLC are displayed in Fig. 1*D*. Note that gene names shown are merely representative of the locus. They do not imply that the indicated gene is the only or key target of chromosomal alteration or that it is involved in cancer pathogenesis.

Six distinct recurrent homozygous deletions were identified (Table 1; Fig. 1*D*, *blue bars below the line*). The most common homozygously deleted locus (7 of 26 NSCLC lines) includes the cyclin-dependent kinase inhibitor gene, *CDKN2A*, on chromosome 9p21, well-known to be deleted in NSCLC. Other deletions comprise the phosphatase and tensin homologue *PTEN* tumor suppressor gene on chromosome 10q23, in 1 of 5 SCLC lines and 2 of 19 primary tumors, and the *FHIT* gene on chromosome 3p, also primarily in SCLC. A homozygous deletion of chromosome 2q22.1 was found in 4 of 26 NSCLC cell lines (Table 1). Each of these four deletions fall within a single known gene, the low-density lipoprotein-related protein 1B gene (*LRP1B*), but in no case is the entire *LRP1B* gene deleted; it is unknown whether *LRP1B* or some undescribed transcript is the target of these deletions. Homozygous deletions in *LRP1B* have previously been identified in ~17% of NSCLC cell lines (34). Interestingly, every cell line with interstitial *LRP1B* deletion also has undergone complete deletion of *CDKN2A* (Table 1). Finally, we identified two recurrent homozygous deletions on chromosome segments 3q25 and 9p23 (see below).

Genes within loci that most frequently undergo copy number gain (copy number ≥4) include the Myc family members *MYC, MYCL1*, and *MYCN* (35–37); regions encompassing the *EGFR* (38), *FGFR1*, and *CDK4* (39) kinase genes and the *CCNE1* cyclin gene; and the *PIK3CA* gene (18). All of these amplifications have been previously reported in lung carcinoma except the loci containing *FGFR1* and *CCNE1*, which have been reported in other tumor types (40–44).

The 8q12-13 locus, where we recently identified amplification (27), is amplified in a second SCLC sample in this study. We have also identified two novel amplicons on chromosome 12p11 and 22q11.

Homozygous deletions within chromosome segments 3q25 and 9p23. There were two recurrent regions of homozygous deletion within our data set that had not been previously reported, on chromosome segments 3q25 and 9p23. Homozygous deletions within 3q25 were found in two samples, the H2882 NSCLC cell line and the S0177T primary SCLC tumor, by array analysis and quantitative PCR (Table 1; Supplementary Fig. S2*A*). This 120 kb region includes only two annotated genes, *AADAC* and *SUCNR1*. *AADAC* encodes arylacetamide deacetylase, which is predicted to catalyze biotransformation pathways for arylamine and heterocyclic amine carcinogens. *SUCNR1* is a G-protein–coupled receptor for the citric acid cycle intermediate, succinate, and may be involved in succinate-induced hypertension (45).

Chromosome 9p undergoes frequent LOH in lung and other cancers, typically associated with homozygous deletion or other inactivation of *CDKN2A*. Our data have identified an additional region of homozygous deletion on chromosome 9p23-24.1, telomeric to *CDKN2A,* which includes sequence upstream of and in the 5′-most portion of the protein tyrosine phosphatase, receptor type, D (*PTPRD*) gene (Fig. 2*A* and *B*; Table 1). One SCLC primary tumor, S0177T, and one NSCLC cell line, H358, contain homozygous deletions confirmed by real-time PCR upstream of *PTPRD* and in the 5′ untranslated region, exon 1, and intron 1 of this gene (Fig. 2*C*). Two additional cell lines, H2347 and HCC1171, as well as H358, also contain homozygous deletions further upstream of the *PTPRD* coding region and more centromeric on chromosome 9 (Fig. 2*C*; Table 1). Whereas not all of the deletions remove exons of *PTPRD*, all of the homozygous deletions eliminate exons from an uncharacterized spliced transcript whose sequence and exonic structure are conserved between human and mouse (BC028038; corresponds to the position ~8.52 to 10.60 Mb on human chromosome 9). This transcript contains a unique 5′ end, several central exons shared with *PTPRD*, and a unique 3′ end. An alignment of *PTPRD*, *BC028038*, and the similar mouse transcript is shown in Supplementary Fig. S3. No role for *PTPRD* or the BC028038 transcript has been described as yet in lung tumorigenesis; we seek to determine the frequency of alterations in these genes in lung cancer by further SNP array analysis, by quantitative PCR, and by sequencing.

8q amplification in small cell lung carcinoma. A high-level amplicon of chromosome 8q12.1-q13.11, 1.7 to 2.6 Mb in size in the SCLC cell line H2171, was identified in our previous study using SNP arrays representing ~10,000 SNP loci (27). Interphase FISH analysis on the H2171 line confirmed the amplification of the 8q12-13 locus, with an estimated copy number of at least 12 to 20 (Fig. 3*A*). In this study, SNP array analysis of one SCLC primary tumor sample, S0177T, revealed high-level amplification of the 8q12-13 region near the *ASPH* locus (Fig. 3*B*). Quantitative real-time PCR revealed a copy number of 89.9, a 45-fold increase over normal genomic DNA. SNP array analysis revealed lower level copy number gains (≥4) of the 8q12-13 region in additional SCLC primary tumors, which were verified by quantitative real-time PCR (Fig. 3*C*) and in 3 of 22 NSCLC cell lines (not shown).

Whereas we have identified other novel regions of chromosome amplification at similar frequencies, this 8q12-13 region is of interest because of the small number of genes involved. The primary SCLC tumor sample, S0177T, contains an amplicon of 670 to 750 kb in size that does not contain the entire *ASPH* gene, but does include its catalytic domain and one additional open reading frame, *MGC34646*, encoding a protein containing a Sec14p-like lipid-binding domain (Fig. 3*D*). Real-time reverse transcription-PCR (RT-PCR) analysis showed the relative expression of *ASPH* was 12-fold higher in H2171 cell line than in normal lung. Quantitative PCR analysis of *MGC34646* expression revealed it to be >100-fold higher in H2171 than in control SCLC cell lines that do not have 8q12-13 amplification (not shown); *MGC34646* expression was not detectable from normal lung tissue. These data suggest but do not prove that *MGC34646* may be a target of chromosomal amplification resulting in significantly increased gene expression.

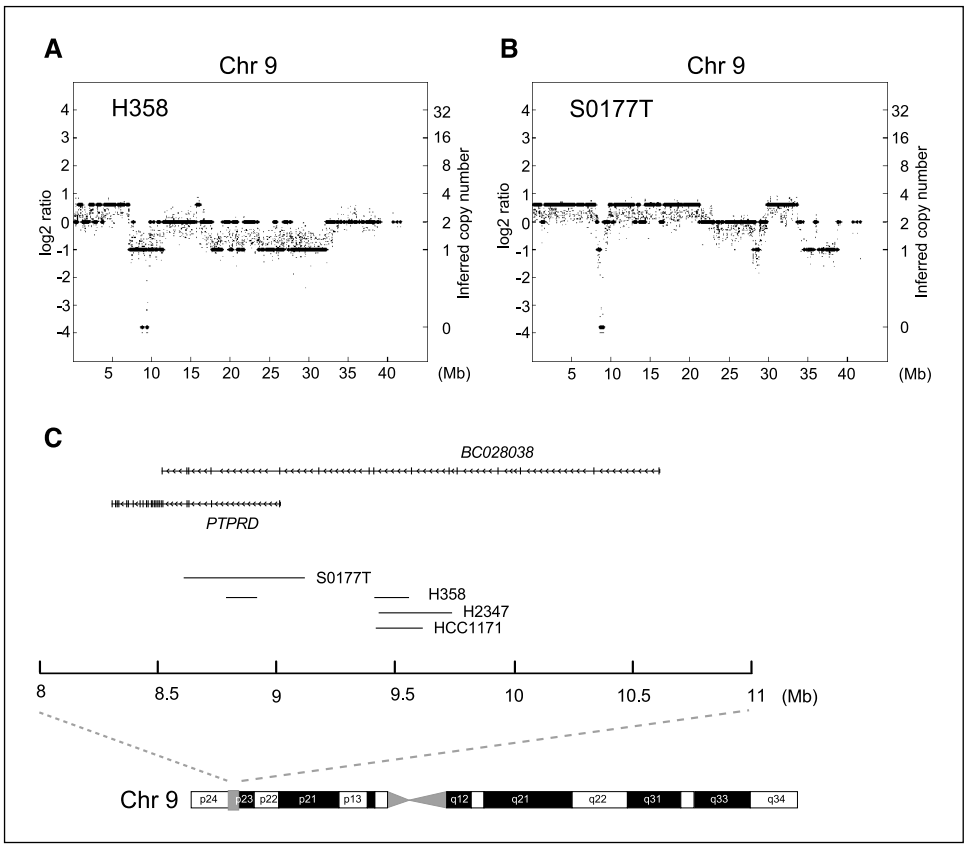

Figure 2. Homozygous deletions within chromosome 9q23. Scatter plots of $\log_2$ copy number ratios (*small dots, left axis*) and inferred copy number (*large dots, right axis*) against the position of all autosomal SNPs show homozygous deletions at chromosome 9p23 in the BAC cell line H358 (*A*) and the SCLC primary tumor sample S0177T (*B*). *C,* the positions of the four homozygous deletions (S0177T, H358, HCC1171, and H2347) are plotted along chromosome 9 (Mb). The positions of the *PTPRD* gene and the BC028038 spliced transcript are shown in relation to the deletions.

Recurrent amplification on 12p11 and 22q11 in non–small cell lung carcinoma. Amplification of 12p11 was found in two NSCLC samples (Table 1; Supplementary Fig. S2*B*), with an overlapping region from 32.7 to 33 Mb. The minimally amplified region contains four genes, *LOC283343*, a pseudogene similar to argininosuccinate synthetase, and *CGI-04*, a provisional protein coding gene, as well as *DNM1L* and *PKP2*. *DNM1L* encodes dynamin like protein 1, which is a member of the dynamin family of GTPases, involved in the fission of organelles, such as mitochondria and peroxisomes (46, 47). *PKP2* encodes plakaphilin 2, which may be involved in β-catenin signaling (48).

High-level amplification of chromosome segment 22q11 was found in two adenocarcinoma cell lines, HCC515 and H1819, one primary adenocarcinoma tumor, S0380T, and one large cell carcinoma cell line, HCC1359, and confirmed by quantitative PCR (Table 1), with a minimal region of overlap from 19.45 to 20.31 Mb (Supplementary Table S3). Examples from the HCC515 cell line and S0380T primary tumor are shown in Supplementary Fig. S2*C*. High-level amplification on 22q11 has not been previously described in lung cancer. Although we have not yet identified the target gene within the 22q11 amplicon, genes of interest that map to this region include *CRKL* and *PIK4CA*. *CRKL* is a member of the CRK adapter protein family, which include homologues of the *v-CRK* oncogene (49). *PIK4CA* is the catalytic subunit of phosphatidylinositol 4-kinase α, which is responsible for the downstream production of certain cell signaling molecules. Real-time quantitative RT-PCR analysis of cell line–derived mRNA showed the relative expression of *CRKL* in cells with 22q11 amplification to be higher (5.32-fold in HCC515 and 3.75-fold in HCC1359) than *PIK4CA* expression (2.61-fold in HCC515 and to be 0.4-fold higher in HCC1359), compared with both normal human lung and lung cancer lines without 22q11 amplification. However, a significant increase in CRKL protein expression was not found in cell lines containing the amplicon (data not shown). Thus, the target of the amplification remains unknown.

Copy number aberrations in the *CDK4/CDK6* pathway in non–small cell lung carcinoma. The *RB/p16/CDK4/CDK6* pathway is often disrupted in tumorigenesis. Copy number alterations of *CCND1, CDK4, CDK6, p16, RB1*, as well as *CCNE1*, were present in our data and were for the most part nonoverlapping, as expected with genes in a pathway. However, clear target genes have yet to be identified, as these regions with copy number alterations contain several genes in addition to these candidate ones. The *CDK2NA* locus on chromosome 9p21 is frequently subject to homozygous deletion in NSCLC (50) as in many other cancers. Seven NSCLC cell lines in this study were found to undergo loss of both copies of the *CDKN2A* locus (Table 1), confirming frequent deletion of this region; we suspect that there were also primary tumors with homozygous deletion but cannot confirm this finding in the face of stromal admixture.

Cyclin D1 (*CCND1*) amplification and overexpression has been previously described in NSCLC (51). The region containing the cyclin D1 gene *CCND1* was amplified in five NSCLC cell lines (Supplementary Table S3). High-level amplification (3- to 4-fold) of the region surrounding the cyclin E (*CCNE1*) gene (19q12) was also present in two primary tumor samples, one SCLC and one adenocarcinoma (Table 1). High-level amplification on chromosome 12q13-12q14, encompassing the *CDK4* locus, was found in two samples (Table 1; Supplementary Fig. S2*B, left*). High-level amplification of this region has recently been reported in lung

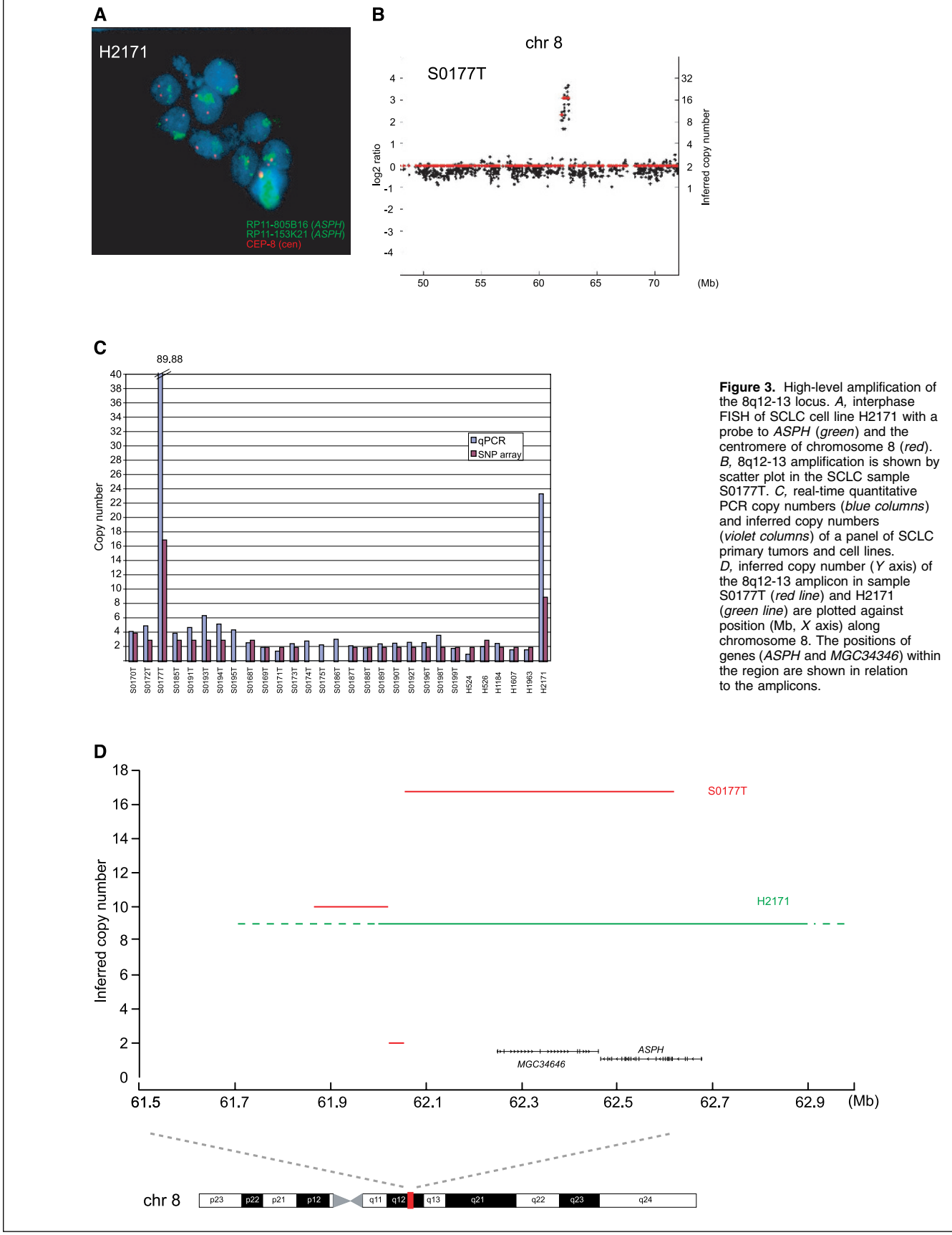

Figure 3. High-level amplification of the 8q12-13 locus. *A,* interphase FISH of SCLC cell line H2171 with a probe to *ASPH* (*green*) and the centromere of chromosome 8 (*red*). *B,* 8q12-13 amplification is shown by scatter plot in the SCLC sample S0177T. *C,* real-time quantitative PCR copy numbers (*blue columns*) and inferred copy numbers (*violet columns*) of a panel of SCLC primary tumors and cell lines. *D,* inferred copy number (*Y* axis) of the 8q12-13 amplicon in sample S0177T (*red line*) and H2171 (*green line*) are plotted against position (Mb, *X* axis) along chromosome 8. The positions of genes (*ASPH* and *MGC34346*) within the region are shown in relation to the amplicons.

cancer (39). The amplified region contains from 20 to 34 genes, among which *CDK4* is an intriguing candidate oncogene (Supplementary Table S3). Two- to three-fold amplification of *CDK6* was also found in four adenocarcinoma samples (Supplementary Table S3).

Tyrosine kinase copy number alterations. A survey of tyrosine kinase gene copy number identified four receptor tyrosine kinase (RTK) genes, *FGFR1, ERBB2* (*Her-2/neu*), *MET* (*HGFR*), and *EGFR,* as being highly amplified (copy number ≥8) in at least one sample. These kinase genes were found in regions containing several genes and, therefore, the targets of the amplicons are still unknown. Whereas some of these amplifications are rare, we mention them because of the possibility of targeted therapy directed against aberrantly expressed protein tyrosine kinases.

A novel amplification of chromosome 8p11.2-p11.1 was identified, which includes the *FGFR1* gene. High-level amplification (copy number ≥8) was found in two NSCLC samples, S0449T and MGH1622T (Fig. 4*A*) and was confirmed by real time PCR (Table 1); lower level copy number gains (copy number ≥4) were found in four additional NSCLC samples. Amplification of the *FGFR1* locus has been found in other tumor types, including breast and urinary carcinoma (41, 42, 52). However, the role of *FGFR1* in lung tumorigenesis remains unknown and we have no evidence to implicate *FGFR1* as the target of this amplicon.

The identification of *ERBB2*, an RTK shown to be overexpressed in many human tumors, including breast, colorectal, ovarian, and NSCLC (53), has led to the development of the targeted breast cancer therapy, trastuzumab (Herceptin; ref. 54). Over 4-fold amplification (copy number ≥8) of the region surrounding *ERBB2* was found in one adenocarcinoma cell line, H1819 (Fig. 4*B*). Additional primary tumors, including two adenocarcinoma samples and one SCLC sample, had moderate copy number gains (copy number ≥4) of the *ERBB2* locus. Overexpression and activating mutations of another RTK, *MET*, have been found in a variety of tumor types (53). A high-level amplification (copy number = 10) of chromosome 7q31.2 resolved into three peaks in the H1993 sample. One of these peaks was a 390 kb amplicon that included *MET* (Fig. 4*C*); lower-level gain of this region was found in two additional NSCLC samples.

***EGFR* mutation compared with *EGFR* amplification.** We sought to analyze amplification of the *EGFR* region in an unbiased fashion and to determine the degree or correlation between amplification, mutation (12), and expression (55) of *EGFR* within the same lung carcinoma samples. Our analysis revealed amplification of the *EGFR* region on chromosome 7p11.2 to copy number ≥8 in one primary squamous cell carcinoma (S0480T), one adenocarcinoma cell line (HCC827; Fig. 4*D*), and one primary adenocarcinoma (AD347T; Tables 1 and 2). Interestingly, HCC827 also contains an *EGFR* mutation (Table 2), whereas AD347T and S0480T do not.

Comparisons of *EGFR* amplification, mutation, and expression data indicate that high and moderate level amplifications are not always associated with *EGFR* gene mutation (Table 2). In addition, 17 samples run on SNP arrays, with known *EGFR* mutation status, also had expression information available (Table 2). Samples with *EGFR* mutation and/or copy number gain (copy number ≥4) of the *EGFR* gene were shown to have higher average *EGFR* expression than samples with wild type, unamplified *EGFR* on average, as measured by two probe sets on Affymetrix U95AV2 arrays [1537_at (*EGFR*): $P = 0.035$; 37327_at (*EGFR*): $P = 0.004$]. Our study shows that mutations in the *EGFR* gene are at least partially independent of *EGFR* gene amplification and that *EGFR* expression and amplification are correlated.

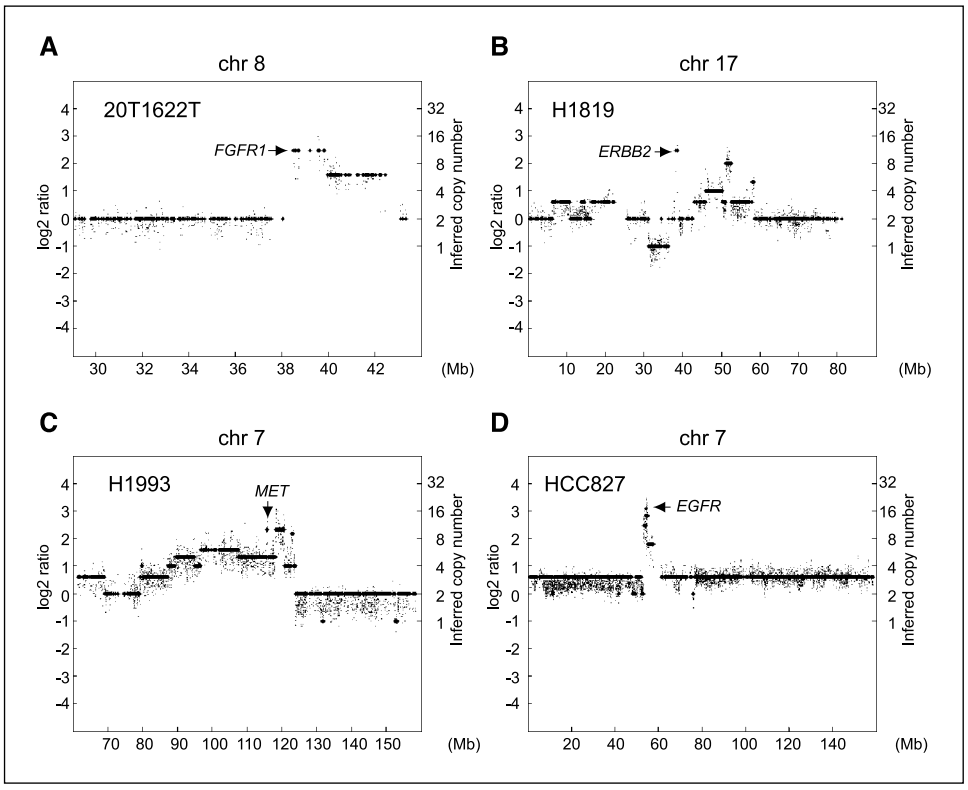

Figure 4. Tyrosine kinase amplifications. Scatter plots of $\log_2$ copy number ratios (*small dots, left axis*) and inferred copy number (*large dots, right axis*). *A,* over 5-fold *FGFR1* amplification in the primary adenocarcinoma sample MGH1622T on chromosome 8. *B,* 4-fold *ERBB2* amplification in the adenocarcinoma cell line H1819 on chromosome 17. *C,* 5-fold *MET* amplification in the adenocarcinoma cell line H1993 on chromosome 7. *D,* high-level *EGFR* amplification in the adenocarcinoma cell line HCC827, containing an exon 19 deletion (E746-A750del) in *EGFR* (Table 2).

Table 2. Comparison of EGFR mutation, amplification, and expression

Sample	Histology	SNP data set*	EGFR copy number (dCHIP)	EGFR kinase domain mutation†	EGFR expression 1537_at‡	EGFR expression 37327_at‡
S0480T	Squamous	120K	14	None	N/A	N/A
HCC827	Adenocarcinoma	120K	14	E746_A750del	N/A	N/A
AD347T	Adenocarcinoma	120K	11	None	1,534.43	699.58
H3255	Adenocarcinoma	60K	8	L858R	4,059.73	1,913.25
H1819	Adenocarcinoma	120K	4	N/A	N/A	N/A
H1993	Adenocarcinoma	120K	4	N/A	N/A	N/A
S0405T	Adenocarcinoma	120K	4	E746_A750del	N/A	N/A
DFCI-LU-01	Adenocarcinoma	60K	4	L747_E749del, A750P	N/A	N/A
S0412T	Adenocarcinoma	120K	3	E746_A750del	N/A	N/A
S0514T	Adenocarcinoma	120K	3	G719S	N/A	N/A
S0380T	Adenocarcinoma	120K	3	E746_A750del	N/A	N/A
AD309T	BAC	120K	3	L747_E752del, P753S	298.48	31.54
AD157T	Adenocarcinoma	120K	3	None	119.39	104.56
AD337T	Adenocarcinoma	120K	3	None	338.97	221.75
H1975	Adenocarcinoma	60K	3	L858R	N/A	N/A
S0377T	Adenocarcinoma	120K	2	G719S	N/A	N/A
H358	BAC	120K	2	None	183.33	75.11
AD327T	Adenocarcinoma	120K	2	None	94.58	43.09
AD338T	Adenocarcinoma	120K	2	None	127.80	39.51
AD311T	Adenocarcinoma	120K	2	None	33.51	35.80
AD362T	Adenocarcinoma	120K	2	None	48.22	36.63
AD330T	Adenocarcinoma	120K	2	None	44.88	14.70
AD334T	Adenocarcinoma	120K	2	None	31.65	63.74
AD335T	Adenocarcinoma	120K	2	None	24.27	35.41
AD336T	Adenocarcinoma	120K	2	None	132.19	82.49
AD163T	Adenocarcinoma	120K	2	None	29.66	33.82
H1650	Adenocarcinoma	60K	2	None	2,125.20	640.89
H1666	Adenocarcinoma	60K	2	None	600.01	205.29

NOTE: Samples are those with *EGFR* amplification (copy number ≥ 4) or *EGFR* gene mutation or those with *EGFR* expression, *EGFR* copy number, and *EGFR* mutation status available.
Abbreviation: N/A, not available.
*120K set was from both *Xba*I and *Hin*dIII chips; 60K set was from only *Xba*I chip.
† Ref. 12; Naoki et al., unpublished data.
‡ Ref. 55; Naoki et al., unpublished data.

Discussion

The present study represents the first application of genome-wide copy number analysis in lung cancer by SNP array. Our group and others recently showed that this technology provided us with a unique opportunity to assess DNA copy number changes and LOH simultaneously throughout the entire genome (27, 28). Our results illustrate the application of SNP arrays to the analysis of chromosomal alterations across lung cancer genomes. The ability to identify aberrations in the chromosomal structure of cancer cells, and the genes affected by them, has been improved through the present analysis with high resolution. The power of SNP arrays to detect small homozygous deletions with high reliability has been shown in this study with the identification of the novel chromosome 9p23 deletion. As used on unpaired samples in this study, SNP array analysis is comparable with oligonucleotide array CGH, although the SNP arrays also offer the ability to identify copy-neutral LOH when coupled with paired normal samples.

Targeting therapies based on the identification of specific chromosomal abnormalities in tumor cells have already offered great potential; inhibitors to kinases, such as *BCR/ABL, KIT, ERBB2*, and *EGFR*, have proven to be valuable as cancer therapies. Thus, kinases, such as *CDK4, FGFR1*, and *MET*, identified in this study may be future therapeutic targets in lung cancers upon further validation.

However, it is important to note that the presence of DNA copy number changes does not provide absolute proof of an involvement of these genes within the regions. Therefore, a detailed characterization should be done to elucidate the significance of the regions we have detected in this study, many of which may contain candidate tumor suppressor genes and oncogenes. Novel loci including the recurrent deletion within chromosome segments 3q25 and 9p23 and amplifications within 8q12-13, 12p11, and 22q11 will be of particular interest.

Acknowledgments

Received 12/29/2004; revised 4/4/2005; accepted 4/14/2005.
Grant support: NIH grant 2P30 CA06516-39 (C. Li), American Cancer Society grant RSG-03-240-01-MGO (M. Meyerson), Flight Attendant Medical Research Institute

(M. Meyerson), Arthur and Rochelle Belfer Foundation (C. Li and M. Meyerson), and Lung Cancer Specialized Programs of Research Excellence grant P50CA70907 (J. Minna and L. Girard).

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.

Raw data will be available at the web site http://research2.dfci.harvard.edu/dfci/snp/ and the latest version of the dChip program is available at http://biosun1.harvard.edu/~cli/dchip2004.exe.

We thank Pamela Mole, Maura Berkeley, and Dr. Ed Fox at Dana-Farber Cancer Institute microarray facility for valuable assistance with the SNP array hybridization and Dr. Adi Gazdar for discussions and tumor cell lines.

References

1. Friend SH, Bernards R, Rogelj S, et al. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature 1986;323:643–6.
2. Kamb A, Gruis NA, Weaver-Feldhaus J, et al. A cell cycle regulator potentially involved in genesis of many tumor types. Science 1994;264:436–40.
3. Nobori T, Miura K, Wu DJ, Lois A, Takabayashi K, Carson DA. Deletions of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers. Nature 1994;368:753–6.
4. Steck PA, Pershouse MA, Jasser SA, et al. Identification of a candidate tumour suppressor gene, *MMAC1*, at chromosome 10q23.3 that is mutated in multiple advanced cancers. Nat Genet 1997;15:356–62.
5. Li J, Yen C, Liaw D, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. Science 1997;275:1943–7.
6. Hahn SA, Hoque AT, Moskaluk CA, et al. Homozygous deletion map at 18q21.1 in pancreatic cancer. Cancer Res 1996;56:490–4.
7. Fleming TP, Saxena A, Clark WC, et al. Amplification and/or overexpression of platelet-derived growth factor receptors and epidermal growth factor receptor in human glial tumors. Cancer Res 1992;52:4550–3.
8. Heinrich MC, Corless CL, Duensing A, et al. PDGFRA activating mutations in gastrointestinal stromal tumors. Science 2003;299:708–10.
9. Lin CR, Chen WS, Kruiger W, et al. Expression cloning of human EGF receptor complementary DNA: gene amplification and three related messenger RNA products in A431 cells. Science 1984;224:843–8.
10. Merlino GT, Xu YH, Ishii S, et al. Amplification and enhanced expression of the epidermal growth factor receptor gene in A431 human carcinoma cells. Science 1984;224:417–9.
11. 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.
12. 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.
13. 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.
14. Ullrich A, Coussens L, Hayflick JS, et al. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature 1984;309:418–25.
15. Semba K, Kamata N, Toyoshima K, Yamamoto T. A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epidermal growth factor-receptor gene and is amplified in a human salivary gland adenocarcinoma. Proc Natl Acad Sci U S A 1985;82:6497–501.
16. Stephens P, Hunter C, Bignell G, et al. Lung cancer: intragenic ERBB2 kinase mutations in tumours. Nature 2004;431:525–6.
17. Shayesteh L, Lu Y, Kuo WL, et al. PIK3CA is implicated as an oncogene in ovarian cancer. Nat Genet 1999;21:99–102.
18. Massion PP, Kuo WL, Stokoe D, et al. Genomic copy number analysis of non-small cell lung cancer using array comparative genomic hybridization: implications of the phosphatidylinositol 3-kinase pathway. Cancer Res 2002;62:3636–40.
19. Samuels Y, Wang Z, Bardelli A, et al. High frequency of mutations of the *PIK3CA* gene in human cancers. Science 2004;304:554.
20. Pinkel D, Segraves R, Sudar D, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 1998;20:207–11.
21. Pollack JR, Perou CM, Alizadeh AA, et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 1999;23:41–6.
22. Mei R, Galipeau PC, Prass C, et al. Genome-wide detection of allelic imbalance using human SNPs and high-density DNA arrays. Genome Res 2000;10:1126–37.
23. Lindblad-Toh K, Tanenbaum DM, Daly MJ, et al. Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. Nat Biotechnol 2000;18:1001–5.
24. Lieberfarb ME, Lin M, Lechpammer M, et al. Genome-wide loss of heterozygosity analysis from laser capture microdissected prostate cancer using single nucleotide polymorphic allele (SNP) arrays and a novel bioinformatics platform dChipSNP. Cancer Res 2003;63:4781–5.
25. Hoque MO, Lee CC, Cairns P, Schoenberg M, Sidransky D. Genome-wide genetic characterization of bladder cancer: a comparison of high-density single-nucleotide polymorphism arrays and PCR-based microsatellite analysis. Cancer Res 2003;63:2216–22.
26. Janne PA, Li C, Zhao X, et al. High-resolution single-nucleotide polymorphism array and clustering analysis of loss of heterozygosity in human lung cancer cell lines. Oncogene 2004;23:2716–26.
27. Zhao X, Li C, Paez JG, et al. An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. Cancer Res 2004;64:3060–71.
28. Bignell GR, Huang J, Greshock J, et al. High-resolution analysis of DNA copy number using oligonucleotide microarrays. Genome Res 2004;14:287–95.
29. Jemal A, Murray T, Ward E, et al. Cancer statistics, 2005. CA Cancer J Clin 2005;55:10–30.
30. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. Genome Biol 2001;2:RESEARCH0032.
31. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc Natl Acad Sci U S A 2001;98:31–6.
32. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 2004;5:557–72.
33. Balsara BR, Testa JR. Chromosomal imbalances in human lung cancer. Oncogene 2002;21:6877–83.
34. Liu CX, Musco S, Lisitsina NM, Forgacs E, Minna JD, Lisitsyn NA. *LRP-DIT*, a putative endocytic receptor gene, is frequently inactivated in non-small cell lung cancer cell lines. Cancer Res 2000;60:1961–7.
35. Little CD, Nau MM, Carney DN, Gazdar AF, Minna JD. Amplification and expression of the c-*myc* oncogene in human lung cancer cell lines. Nature 1983;306:194–6.
36. Nau MM, Brooks BJ, Battey J, et al. L-*myc*, a new *myc*-related gene amplified and expressed in human small cell lung cancer. Nature 1985;318:69–73.
37. Nau MM, Brooks BJ Jr, Carney DN, et al. Human small-cell lung cancers show amplification and expression of the N-*myc* gene. Proc Natl Acad Sci U S A 1986;83:1092–6.
38. Reissmann PT, Koga H, Figlin RA, Holmes EC, Slamon DJ. Amplification and overexpression of the cyclin D1 and epidermal growth factor receptor genes in non-small-cell lung cancer. Lung Cancer Study Group. J Cancer Res Clin Oncol 1999;125:61–70.
39. Wikman H, Nymark P, Vayrynen A, et al. CDK4 is a probable target gene in a novel amplicon at 12q13.3-q14.1 in lung cancer. Genes Chromosomes Cancer 2005;42:193–9.
40. Marone M, Scambia G, Giannitelli C, et al. Analysis of cyclin E and CDK2 in ovarian cancer: gene amplification and RNA overexpression. Int J Cancer 1998;75:34–9.
41. Dib A, Adelaide J, Chaffanet M, et al. Characterization of the region of the short arm of chromosome 8 amplified in breast carcinoma. Oncogene 1995;10:995–1001.
42. Simon R, Richter J, Wagner U, et al. High-throughput tissue microarray analysis of 3p25 (RAF1) and 8p12 (FGFR1) copy number alterations in urinary bladder cancer. Cancer Res 2001;61:4514–9.
43. Khatib ZA, Matsushime H, Valentine M, Shapiro DN, Sherr CJ, Look AT. Coamplification of the *CDK4* gene with MDM2 and GLI in human sarcomas. Cancer Res 1993;53:5535–41.
44. Maelandsmo GM, Berner JM, Florenes VA, et al. Homozygous deletion frequency and expression levels of the CDKN2 gene in human sarcomas—relationship to amplification and mRNA levels of CDK4 and CCND1. Br J Cancer 1995;72:393–8.
45. He W, Miao FJ, Lin DC, et al. Citric acid cycle intermediates as ligands for orphan G-protein-coupled receptors. Nature 2004;429:188–93.
46. Koch A, Thiemann M, Grabenbauer M, Yoon Y, McNiven MA, Schrader M. Dynamin-like protein 1 is involved in peroxisomal fission. J Biol Chem 2003;278:8597–605.
47. Smirnova E, Griparic L, Shurland DL, van der Bliek AM. Dynamin-related protein Drp1 is required for mitochondrial division in mammalian cells. Mol Biol Cell 2001;12:2245–56.
48. Chen X, Bonne S, Hatzfeld M, van Roy F, Green KJ. Protein binding and functional characterization of plakophilin 2. Evidence for its diverse roles in desmosomes and β-catenin signaling. J Biol Chem 2002;277:10512–22.
49. ten Hoeve J, Morris C, Heisterkamp N, Groffen J. Isolation and chromosomal localization of *CRKL*, a human crk-like gene. Oncogene 1993;8:2469–74.
50. Hayashi N, Sugimoto Y, Tsuchiya E, Ogawa M, Nakamura Y. Somatic mutations of the MTS (multiple tumor suppressor) 1/CDK4I (cyclin-dependent kinase-4 inhibitor) gene in human primary non-small cell lung carcinomas. Biochem Biophys Res Commun 1994;202:1426–30.
51. Betticher DC, Heighway J, Hasleton PS, et al. Prognostic significance of CCND1 (cyclin D1) overexpression in primary resected non-small-cell lung cancer. Br J Cancer 1996;73:294–300.
52. Cuny M, Kramar A, Courjal F, et al. Relating genotype and phenotype in breast cancer: an analysis of the prognostic significance of amplification at eight different genes or loci and of p53 mutations. Cancer Res 2000;60:1077–83.
53. Blume-Jensen P, Hunter T. Oncogenic kinase signalling. Nature 2001;411:355–65.
54. Slamon DJ, Leyland-Jones B, Shak S, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 2001;344:783–92.
55. Bhattacharjee A, Richards WG, Staunton J, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc Natl Acad Sci U S A 2001;98:13790–5.

Homozygous Deletions and Chromosome Amplifications in Human Lung Carcinomas Revealed by Single Nucleotide Polymorphism Array Analysis

Xiaojun Zhao, Barbara A. Weir, Thomas LaFramboise, et al.

Cancer Res 2005;65:5561-5570.