[CANCER RESEARCH 62, 5273–5279, September 15, 2002]

The Hepatocyte Nuclear Factor 3 α Gene, *HNF3α* (*FOXA1*), on Chromosome Band 14q13 Is Amplified and Overexpressed in Esophageal and Lung Adenocarcinomas[1]

Lin Lin,[2] Charles T. Miller,[2] Jorge I. Contreras, Michael S. Prescott, Susan L. Dagenais, Rong Wu, John Yee, Mark B. Orringer, David E. Misek, Samir M. Hanash, Thomas W. Glover, and David G. Beer[3]

Section of General Thoracic Surgery, Departments of Surgery [L. L., C. T. M., J. I. C., M. S. P., J. Y., M. B. O., D. G. B.], Human Genetics [S. L. D., T. W. G.], Pathology [R. W.], and Pediatrics [D. E. M., S. M. H., T. W. G.], University of Michigan Medical School, Ann Arbor, Michigan 48109

ABSTRACT

Genomic amplification is observed in many, if not all, types of human malignancy and is one of the mechanisms for the activation of dominant-acting oncogenes in tumorigenesis. In the present study, three amplified restriction fragments were identified in an esophageal adenocarcinoma (P16) using the restriction landmark genome scanning two-dimensional gel technique. These fragments were cloned, sequenced, and mapped to chromosome band 14q13. Using the sequence tagged site-amplification mapping approach, we defined the core-amplified domain by screening 75 normal-tumor paired esophageal samples. The frequency of 14q13 amplification is 6.7% in esophageal tumors, and the amplicon spans >6 Mb in 1 tumor but is contained in a region <0.3 Mb in all of the remaining amplified tumors. Quantitative reverse transcription-PCR (RT-PCR) of 8 genes and expressed sequence tags located within the core-amplified domain revealed that the *HNF3α* (*FOXA1*)[4]gene, a forkhead gene family member, was overexpressed in all of the amplified esophageal tumors. *HNF3α* amplification was confirmed by Southern blot and interphase fluorescence *in situ* hybridization analyses, and the results of real-time RT-PCR were consistent with that of the regular quantitative RT-PCR. Increased immunohistochemical nuclear staining of the HNF3α protein was detected in all of the tumors containing 14q13 amplification. Affymetrix oligonucleotide microarrays of 86 lung adenocarcinomas demonstrated that expression of the *HNF3α* mRNA was elevated (≥2.5-fold of mean expression in normal lung) in 37% (32 of 86) of the tumors analyzed. Gene amplification of *HNF3α* was detected in 2 of the 5 overexpressed lung tumors examined. This is the first report of *HNF3α* amplification, and overexpression in esophageal and lung adenocarcinomas. Amplification of *HNF3α* in esophageal and lung tumors may suggest a potential oncogenic role for this gene in tumorigenesis.

INTRODUCTION

The prognosis of patients diagnosed with esophageal adenocarcinoma is poor, with overall 5-year survival rates of 10–15% (1). Additionally, the United States and Western countries have experienced a sharp increase in the incidence of esophageal adenocarcinoma over the past 3 decades (2). Esophageal adenocarcinoma is accompanied frequently by Barrett's esophagus, metaplastic columnar epithelia that replace the normal squamous epithelia of the distal esophagus after chronic gastroesophageal reflux. Progression through a metaplasia-dysplasia-carcinoma sequence is observed (3, 4). However, molecular alterations leading to the development of esophageal adenocarcinoma remain unclear. Therefore, characterization of the genetic lesions involved in the development and progression of this disease are of clinical importance.

Genomic amplification is commonly observed in many types of human malignancies, including esophageal adenocarcinoma. An amplified DNA segment is selected because it contains gene(s), in most cases an oncogene, that when amplified and overexpressed confers a growth advantage for cells. In humans, genomic amplification is exclusively restricted to tumor cells and is a major mechanism for the activation of dominant-acting oncogenes during tumorigenesis. Amplification of the *erbB2*, *EGFR*, and *K-ras* oncogenes has been identified in esophageal adenocarcinomas (5, 6). Genomic amplification on chromosomes 8p, 15q, and 20q has also been reported in this tumor type using comparative genomic hybridization (7). We have previously identified and characterized a novel amplicon on 8p22 in esophageal adenocarcinomas (8). The results suggested that two genes, cysteine protease cathepsin B and transcription factor *GATA-4*, were the selected targets of the 8p22–23 amplification (8, 9). Using the RLGS[5] two-dimensional gel and STS-amplification mapping (9) approaches, we also identified and characterized an amplicon at 19q12, and the *cyclin E* gene was found to be the best candidate gene for selection of this amplicon (10).

In the present study, three *NotI/DpnII* DNA restriction fragments identified in an esophageal adenocarcinoma (P16) demonstrated increased intensity using the RLGS two-dimensional gel approach. STS-amplification mapping was applied to the DNA from 75 normal-tumor paired esophageal samples using STS markers in the chromosomal vicinity of the three cloned restriction fragments to define the frequency and extent of amplification. We found that the core-amplified domain includes the *HNF3α* gene and the EST stSG46936. The transcription factor *HNF3α* is a member of the *forkhead* gene family, and the involvement of the *HNF3α* gene in liver differentiation and regeneration, embryonic gut development, and pancreatic development has been reported in rodent studies (11–14). The pathogenic significance of overexpressed *HNF3α* is currently unclear. However, the evidence from our study suggests potential oncogenic involvement for *HNF3α* in the development and/or progression of gastroesophageal and lung adenocarcinomas.

MATERIALS AND METHODS

Tissue Collection. After obtaining written consent and University of Michigan Institutional Review Board approval, 75 esophageal and gastric cardia adenocarcinomas, and their corresponding normal esophagus or gastric mucosa, 20 specimens of Barrett's metaplasia, and 114 lung adenocarcinomas were obtained from patients undergoing esophagectomy or pulmonary resection at the University of Michigan Medical Center between 1992 and 2000. Patients in this study had no preoperative radiotherapy or chemotherapy. A small portion of each tissue was embedded in OCT compound (Miles Scien-

Received 11/15/01; accepted 7/19/02.

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.

[1] Supported by National Cancer Institute Grant CA71606 and the Weber Research Fund.

[2] These authors contributed equally to this work.

[3] To whom requests for reprints should be addressed, at MSRB II B560, Box 0686, Department of Surgery, Section of General Thoracic Surgery, University of Michigan Medical School, Ann Arbor, MI 48109. Phone: (734) 763-0325; Fax: (734) 763-0323; E-mail: dgbeer@umich.edu.

[4] According to the new nomenclature, all vertebrate genes encoding winged helix proteins will be termed *Fox*, for *forkhead box* (Kaestner, K. H., Knochel, W., and Martinez, D. E. Unified nomenclature for the winged helix/forkhead transcription factors. Genes Dev., *14*: 142–146, 2000).

[5] The abbreviations used are: RLGS, restriction landmark genome scanning; FISH, florescent *in situ* hybridization; STS, sequence tagged site; EST, expressed sequence tag; QG-PCR, quantitative genomic-PCR; NCBI, the National Center for Biotechnology Information; BAC, bacterial artificial chromosome; Tsc:Ns/c, the intensity ratio of tumor (Ts/c) or normal (Ns/c) samples versus *GAPDH* control from QG-PCR.

tific, Naperville, IL) and stored at −80°C; the remainder was frozen in liquid nitrogen and stored at −80°C until use.

DNA Isolation and RLGS Two-dimensional Gel Electrophoresis. High molecular weight DNA was isolated as described previously (15). All of the tumor portions used for DNA isolation in this study contained >70% tumor cells as determined by frozen tissue sections. RLGS two-dimensional gel electrophoresis was performed as described previously (16). In brief, DNA from normal and tumor tissues was double-digested using the restriction enzymes *Not*I and *Eco*RV (New England Biolabs Inc., Beverly, MA), and the *Not*I ends were filled with [α-^{32}P]dCTP and [α-^{32}P]dGTP (NEN Life Science Products, Boston, MA). First-dimensional size fractionation was performed in a 32-cm 0.9% disk-agarose gel. The resulting DNA fragments were then in-gel digested with *Hinf*I (or *Dpn*II for cloning purposes) and separated in the second dimension on a 5.25% polyacrylamide gel. Gels were then dried and autoradiographic images obtained (Molecular Dynamics, Sunnyvale, CA). Digital images were analyzed using ImageQuant v1.2 software (Molecular Dynamics). Amplified DNA fragments were quantified by densitometry using the two-dimensional images from the corresponding normal tissue DNA as standards. For cloning purposes, preparative gels were run and the amplified spots were collected (8).

Isolation and Cloning of Two-dimensional DNA Fragments. The gel pieces containing the amplified DNA fragments were isolated and electrophoresed onto a DEAE membrane (Schleicher & Schuell, Keene, NH). The DNA was then eluted and purified as described previously (8), and dissolved in 7 μl of distilled water. A pBC vector (Stratagene, La Jolla, CA) was cleaved with the enzymes *Not*I and *Bam*HI (*Dpn*II compatible end), and was gel-purified using SeaKem GTG agarose (FMC BioProducts, Rockland, ME). Purified two-dimensional gel fragments were then incubated with 1 μl of vector (1 ng/μl) and 1 μl of 10× ligation buffer (Boehringer Mannheim, Indianapolis, IN) at room temperature for 10 min and at 65°C for 7 min, and 1 μl T4 ligase (5 ng/μl; Boehringer Mannheim) was added. The ligation reaction was carried at 16°C for 25 h. The ligation mixture was purified using a standard phenol:chloroform extraction protocol (17). The XL1 Blue competent cells were then transformed with the purified plasmid DNA through electroporation using *Escherichia coli* Pulser (Bio-Rad Laboratories, Hercules, CA). Individual colonies were collected for DNA isolation (mini-preps) via the alkaline lysis method as described by Sambrook *et al.* (17).

Electronic-PCR Analysis. The cloned fragments were sequenced and analyzed by the BLAST tools[6] for similarity matches in the databases. Resulting sequences were used as templates to perform electronic-PCR analysis.[6] Chromosomal location of the template was determined based on an alignment to sequences in the databases.

STS-Amplification Mapping Using QG-PCR. STS markers located in the 14q13 region were selected for quantitative PCR, as described previously (9). In brief, PCR primers were designed to ensure that the melting temperature (*Tm*) of each primer set was compatible with that of the internal control (*GAPDH* or same chromosome control D14S72). The quantity of the normal and tumor genomic DNA was measured by a fluorometer (TKO100; Hoefer Scientific Instruments, San Francisco, CA) to ensure equity of the starting materials. Forward primers for *GAPDH*, D14S72 and STS markers, were end-labeled with [γ-^{32}P]ATP (NEN Life Science Products) using T4 polynucleotide kinase (New England Biolabs). PCR was performed using Taq polymerase (Promega, Madison, WI), and the PCR products were resolved on 8% denaturing polyacrylamide gels. PCR product signal ratios [the intensity ratio of tumor (Ts/c) or normal (Ns/c) samples versus *GAPDH* control from QG-PCR] for both the tumor (Ts/c, tumor STS fragment/tumor *GAPDH* fragment) and normal DNA samples (Ns/c, normal STS fragment/normal *GAPDH* fragment) were quantified using ImageQuant software (Molecular Dynamics). Values ≥2.0 were considered indicative of DNA amplification. All of the results were repeated and verified.

Southern Blot Analysis and Interphase FISH Assay. Ten μg of high molecular weight DNA from paired normal-tumor tissues were loaded into a 1.0% agarose gel. DNA was blotted, and a labeled *HNF3α* probe was hybridized to the Southern filter using the protocol described previously (10). The 459-bp probe spans the sequence of the *HNF3α* gene (NM_004496, a curated mRNA from National Center for Biotechnology Information databases) from 1485 bp to 1943 bp in the 3′ untranslated region and was selected to avoid cross-hybridization to other members of *HNF3* family. All of the reagents for interphase FISH were purchased from Vysis, Inc. (Downers Grove, IL) except pepsin (Sigma, St. Louis, MO) and Cot I DNA (Life Technologies, Inc., Gaithersburg, MD). Interphase slides were made using a touch approach, and nick translation and hybridization were conducted according to the Vysis protocol with modification. A prelabeled chromosome 14q telomeric probe (Part number, 33-260014; Vysis) was cohybridized with the target probe. The BAC clone (R-356O9) that includes the *HNF3α* sequence was purchased from the BACPAC Resources (Children's Hospital Oakland Research Institute, Oakland, CA).[7] BAC DNA was prepared using Concert High Purity Maxiprep System (Life Technologies, Inc.).

RNA Isolation and Real-Time Quantitative RT-PCR. Total RNA was isolated using TriZol reagent (Life Technologies, Inc.). Agarose gel electrophoresis and the A260:280 ratio were used to assess RNA quality. RNA samples were stored at −80°C until use. All of the RNA samples were treated with DNase I (Promega) before performing reverse transcription. Two μg of total RNA was reverse transcribed using reverse transcriptase (Life Technologies, Inc.) and primed by both $(dT)_{18}$ and random hexamers in a total of 20 μl of reaction volume. Two μl of the cDNA products underwent RT-PCR with *GAPDH* coamplified as an internal control. RT-PCR products were then resolved on 8% denaturing PAGE gels, and gel data were analyzed using ImageQuant software as in QG-PCR. Real-time quantitative RT-PCR for *HNF3α* was performed to verify the expression levels in the tumors. *HNF3α* and *GAPDH* were amplified using TaqMan probes labeled with 5′ fluorescent label, 6FAM or VIC, respectively, and 3′ Quencher 6-carboxytetramethylrhodamine (PE Applied Biosystems, Foster City, CA). The forward and reverse primers flanking a 78-bp fragment of *HNF3α*, which maps to the 3′ untranslated region of the gene and ensures the specificity of the fragment to *HNF3α*, are (5′-3′) CATTGCCATCGTGTGCTTGT and CCCGTCTGGCTATACTAACACCAT, respectively. The sequence of 6FAM-labeled probe is (5′-3′) TCATCCAGTGTTATGCACTTTCCACAGTTGG. Human *GAPDH* primers and VIC-prelabeled probe were purchased from PE Applied Biosystems (catalogue no. 4310884E). Real-time quantitative PCR was performed and analyzed using the PRISM 7700 Sequence Detector according to instructions of the manufacturer (PE Applied Biosystems).

Immunohistochemical Analysis. Frozen specimens were sectioned at 5 μm, placed on 0.1% poly-L-lysine-coated slides, and fixed in 100% acetone at −20°C for 10 min. Endogenous peroxidase activity was quenched, and nonspecific binding was blocked using horse serum in PBS-1% BSA. The HNF3α protein was detected using an anti-HNF3α antibody (Santa Cruz Biotechnology, Santa Cruz, CA) at a 1:100 dilution in PBS-1% BSA. Immunoreactivity was detected using the Vectastain avidin/biotin complex kit (Vector Laboratories, Burlingame, CA) as described previously (18).

Expression Analysis Using Oligonucleotide (Affymetrix) Microarrays. Total cellular RNA was isolated from 10 normal and 86 lung adenocarcinoma samples using TriZol reagent (Life Technologies, Inc.). Total RNA was additionally purified using RNeasy spin column (Qiagen, Valencia, CA) and used to generate cRNA probes. Preparation of cRNA and hybridization of the HuGeneFL Arrays were performed according to the manufacturer's protocol (Affymetrix, Santa Clara, CA). The arrays were scanned using the GeneArray scanner (Affymetrix), and data analysis was performed using GeneChip 4.0 software as described previously (19).

RESULTS

Restriction Landmark Genomic Scanning Two-dimensional Gel Analysis of Esophageal Adenocarcinomas. DNA samples from 44 primary esophageal and gastric cardia adenocarcinomas, and the corresponding normal tissues were analyzed using the RLGS two-dimensional gel technique. More than 2000 individual *Not*I/*Hinf*I restriction fragments were visualized and used to compare the two-dimensional images of normal and tumor DNA (Fig. 1). Comparison of one esophageal adenocarcinoma, P16 (Fig. 1*B*), with its corresponding normal tissue (Fig. 1*A*) revealed a *Not*I/*Hinf*I fragment with increased intensity in the tumor that was 4400 bp in the first dimen-

[6] Internet address: http://www.ncbi.nlm.nih.gov.

[7] E-mail address: BACPACorders@chori.org.

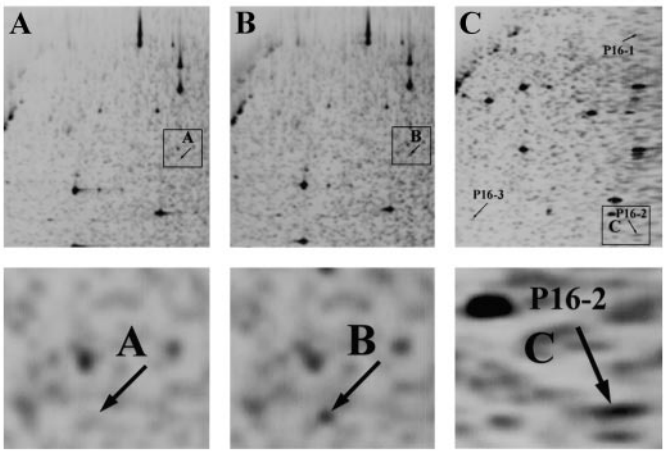

Fig. 1. RLGS two-dimensional gel analysis of DNA from normal and tumor tissues from patient P16. Two-dimensional PAGE gel was used to separate the *Not*I/*Hinf*I DNA fragments from normal tissue from patient P16 (*A*) and the corresponding adenocarcinoma (*B*). *Not*I/*Dpn*II digestion was performed for cloning purposes (*C*). Cloned *Not*I/*Dpn*II fragments P16–1, P16–2, and P16–3 are shown (*arrows; C*). The enlarged thumb view of fragment P16–2 is shown in normal (*A, bottom panel*) and tumor (*B* and *C, bottom panel*) two-dimensional gel images.

sion and 420 bp in the second dimension. The tumor DNA was digested with *Not*I/*Dpn*II restriction enzymes for cloning purposes (Fig. 1*C*). Two *Not*I/*Dpn*II fragments, P16–1 and P16–2, were found to share the sequences near the *NotI* cut-site (Table 1; Fig. 1*C*), indicating that two *Dpn*II restriction sites are located within their respective *Not*I/*EcoR*V fragments separated on one-dimensional gel. Three distinct fragments (P16–1, P16–2, and P16–3) with increased intensity were collected from the two-dimensional gels, and the DNA was extracted and purified. Shown in the enlarged thumb view of Fig. 1*C* is fragment P16–2 from tumor P16. The very intense fragments visible on these gels represent ribosomal DNA, which are present in multiple copies in the genome.

Chromosome Localization of Amplified Two-dimensional Fragments. Amplified fragments were cloned, sequenced, and subject to a BLAST search for similarity match. Both P16–1 and P16–2 were matched to a sequence of the same BAC clone R-356O9 (GenBank accession no. AL121790), which is mapped to the chromosomal band 14q13. Fragment P16–3 was matched to BAC clone R-964E11 (GenBank accession no. AL079303), which is also located at 14q13 (Table 1). Fragments P16–1, P16–2, and P16–3 are all contained within the same contig, NT_010164, indicating a close genetic linkage of all three of the two-dimensional fragments. Fragment P16–2 was confirmed as part of the sequence of the *HNF3α* gene, and the sequence of the P16–3 fragment was found to overlap the sequence of gene *PAX9* (Table 1).

Defining and Characterizing the Core-amplified Domain of the 14q13 Amplicon Using the STS-Amplification Mapping Approach. To define the core-amplified domain and determine the frequency of the 14q13 amplification present in esophageal adenocarcinomas, >13 STS or gene markers, spanning a distance >6 Mb in the vicinity of the cloned genomic fragments, were selected by searching the NCBI, Whitehead Institute Center for Genome Research, and Stanford Human Genome Center databases. QG-PCR analysis was applied to DNA from 75 normal-tumor paired esophageal samples with these markers (Fig. 2*A;* Fig. 3). The housekeeping gene *GAPDH*, located on 12p13, was coamplified in every reaction as an internal control. D14S72, which maps to centromeric chromosome band 14q11, was coamplified with target genes or STS markers in additional reactions as a chromosome 14 control to exclude aneuploidy as the cause of increased DNA dosage in the tumors containing the 14q13 amplicon. *GAPDH* and D14S72 were PCR coamplified before the targeted assays for each normal-tumor pair to ensure that D14S72 is not increased in copy number in the tumors examined. DNA copy number of D14S72 was found to parallel that of *GAPDH* between normal and tumor DNAs in all of the tumors tested (data not shown). This indicates that the 14q13 amplicon represents an intrachromosomal amplification and that aneuploidy is not the cause of increased DNA content in amplified tumors. This is consistent with the results of interphase FISH, where a telomeric 14q (14q32) probe and a *HNF3α* containing BAC probe (R-356O9) were cohybridized to the tumor sections (Fig. 6, *B* and *C*). Therefore, D14S72 can be used as a reference marker to more accurately reflect copy number of amplified genomic fragments at 14q13 in a QG-PCR assay. Increased DNA dosage (value ≥2.0) was observed in amplified tumors (Fig. 2*A*). The frequency of the 14q13 amplification is 6.7%, and the amplicon spans ~6 Mb in tumor P16 and is contained in a region <0.3 Mb in all of the remaining amplified tumors (Fig. 2*A* and Fig. 3). The core-amplified domain of this amplicon maps between markers AA029430 and Stanford Human Genome Center-103312 (Fig. 3), a distance of ~254 kb based on the physical maps from the NCBI. The transcription factor, *HNF3α*, resides in the core-amplified domain (Fig. 3). Genomic coamplification of the *PAX9* (cloned two-dimensional fragment P16–3) and the *NKX-2.8* genes, an oncogene 80 kb centromeric to *PAX9*, with *HNF3α* was only observed in tumor P16 (Fig. 3). The copy number of the *NFKBIA* gene (*IKBA* or *MAD-3*) was not found to be increased in any tumors examined.

Verification of Amplification on Chromosome Band 14q13. Southern blot analysis was performed on three pairs of esophageal normal-tumor samples using a *HNF3α*-specific probe, and gene amplification of *HNF3α* was verified in tumor P16 (Fig. 2*B*). Interphase FISH was conducted to examine 14q13 regional changes in sectioned esophageal tumors using a 200-kb *HNF3α* containing BAC probe. FISH images demonstrated an apparent increase in copy number of the 14q13 region in the tumors containing 14q13 amplicon, whereas

Table 1 *Cloned RLGS two-dimensional fragments from NotI/DpnII preparative gels*

Name (estimated size)	Chromosome location	Accession ID (BAC clone)	Sequence[a] (represented in BAC clone)	EST hit	Gene hit (name)	Representative STS-marker
P16-1 (1545 bp)	14q-12-13	AL121790 (R356O9)	128434−	BE550975 AW295342	None	U39840 STS68217
		AL135879 (R-356O9)	128425−	BE218603		
P16-2 (243 bp)	14q12-13	AL121790	124035−	None	U86584 (*HNF-3α*,[b] rat) AF056705 (*HNF-3α*,[b] mouse)	
		AL135879	124027−			
P16-3 (278 bp)	14q12-13	AL079303 (R-964E11)	101020−	BF514698 BF514299	L09745 (*PAX9*) AJ238381 (*PAX9*)	STS40888

[a] 5′ starting point of the cloned fragments within the BAC clone.
[b] High homology to the sequences of the promoter and 5′ untranslated region in rat and mouse.

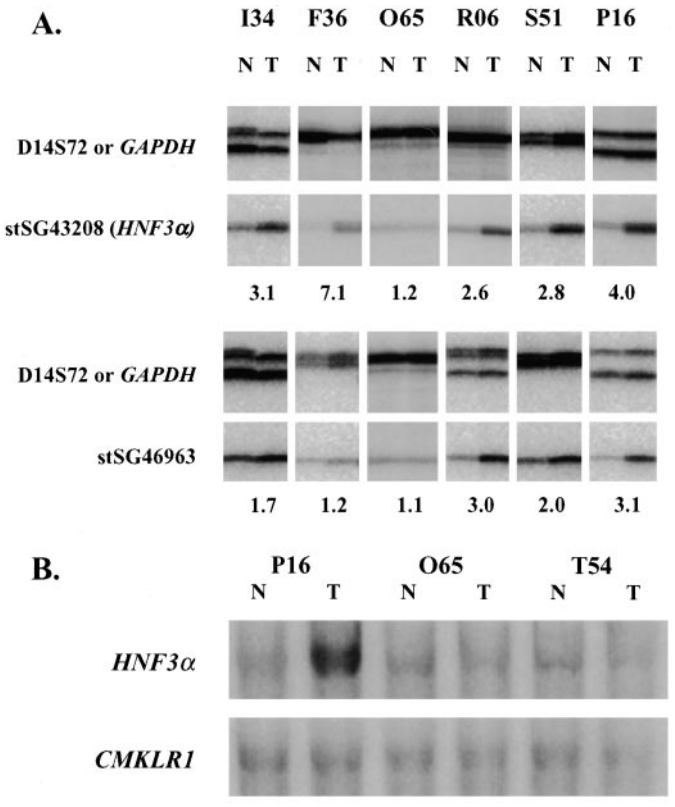

Fig. 2. *A*, the QG-PCR assay was used to map the core-amplified domain and to characterize the 14q13 amplicon in 75 esophageal and gastric cardia adenocarcinomas. The D14S72 marker or housekeeping gene *GAPDH* (in sample F36, O65, and R06) was coamplified as an internal control. The D14S72 marker was selected as a syntenic control to more accurately distinguish the difference in DNA copy number between normal and tumor tissues. STS markers neighboring the cloned fragments were selected from the databases (Table 1). Shown here is a representative QG-PCR gel image from the 5 tumors containing the 14q13 amplicon and 2 not possessing the 14q13 amplicon (tumor O65). Three tumors, P16, R06, and S51, were amplified at stSG46963. Marker stSG43208 (gene *HNF3α*), which is closely linked to stSG46963, was amplified in 5 tumors shown. *B*, gene amplification of *HNF3α* is verified by Southern blot analysis using *HNF3α* gene-specific probe.

the telomeric region of 14q remained unchanged in the same tumors, indicating a nonaneuploid status of chromosome 14 for these tumors and that the gene amplification of *HNF3α* is not a result of increased numbers of chromosome 14 (Fig. 6, *B* and *C*). Tumor P16 demonstrated a highly increased DNA dosage of the 14q13 band with hybridization of the *HNF3α* containing BAC probe as compared with its cohybridized telomeric 14q probe (Fig. 6*B*). However, the FISH images might not ultimately discriminate whether the increased genomic dosage at region 14q13 constitutes a homogenously staining region, double minute, or other chromosomal rearrangements, *e.g.*, chromosomal translocation. Consistent with the QG-PCR results (Fig. 2*A*), a low increase in copy number of *HNF3α* was confirmed in tumor I34, where three or more green spots (*HNF3α* containing BAC probe) were observed (Fig. 6*C*). In contrast, the amount of genomic DNA at region 14q13 in tumor O65 is paralleled comparatively with the amount at 14q32 (Fig. 6*D*), which is consistent with the QG-PCR and Southern blot results.

Real-Time Quantitative PCR Reveals Increased Expression of the *HNF3α* mRNA. Quantitative RT-PCR was applied to amplified tumors and a few tumors not containing the 14q13 amplicon to investigate the expression level of 8 genes and ESTs within or near the core-amplified region (Table 2; Fig. 4). *HNF3α* was found to be overexpressed in all of the amplified tumors. *HNF3α* exhibits the most frequent rate of both overexpression and amplification among the examined 8 genes and ESTs (Table 2) located within or near the core amplified domain (~254 kb). This suggests that *HNF3α* is the best candidate for selection in the 14q13 amplicon. Moreover, *HNF3α* mRNA expression data, analyzed using real-time quantitative RT-PCR of an *HNF3α*-specific sequence, are consistent with the results from conventional quantitative RT-PCR (Table 2).

Affymetrix Oligonucleotide Microarrays Revealed That Overexpression of the *HNF3α* mRNA Is a Frequent Event in Lung Adenocarcinomas. Affymetrix oligonucleotide microarrays of 86 lung adenocarcinomas demonstrated an elevated expression of the *HNF3α* mRNA (≥2.5-fold of mean expression in normal lungs) in 32 of the 86 (37%) lung tumors analyzed (Table 3). To determine whether gene amplification was present in these overexpressed lung adenocarcinomas, 5 tumors demonstrating overexpression of *HNF3α* mRNA (all ≥4.8-fold of mean expression in normal lungs) in microarrays were analyzed for *HNF3α* amplification by QG-PCR (Fig. 5; Table 3). Two tumors, S53 and C66, were confirmed to be amplified (Fig. 5). Gene amplification of *HNF3α* was also found in 1 lung adenocarcinoma, H01, in an independent series of 27 lung tumors examined (Fig. 5).

Increased Expression of the *HNF3α* Protein in Esophageal Tumors Containing the 14q13 Amplicon. Immunohistochemistry was used to determine the expression level and cellular localization of the HNF3α protein in esophageal tumors with or without *HNF3α* amplification. Intense nuclear staining was observed in tumor P16 (Fig. 6*F*) as compared with the section from the same tumor receiving no HNF3α primary antibody (negative control; Fig. 6*E*). Tumor I34 demonstrated light nuclear staining in the majority of tumor cells (Fig. 6*G*). Nuclear staining of HNF3α was sporadic in tumor O65, which does not contain the 14q13 amplicon (Fig. 6*H*) and was not detected in any of the normal esophagus or Barrett's metaplasia examined (data not shown). Increased nuclear staining of the HNF3α protein was also found in lung adenocarcinomas containing the 14q13 amplification (data not shown).

DISCUSSION

Genomic amplification is associated frequently with an increased copy number of an oncogene or other cancer-related genes in human malignancy. Using the RLGS two-dimensional gel approach, we

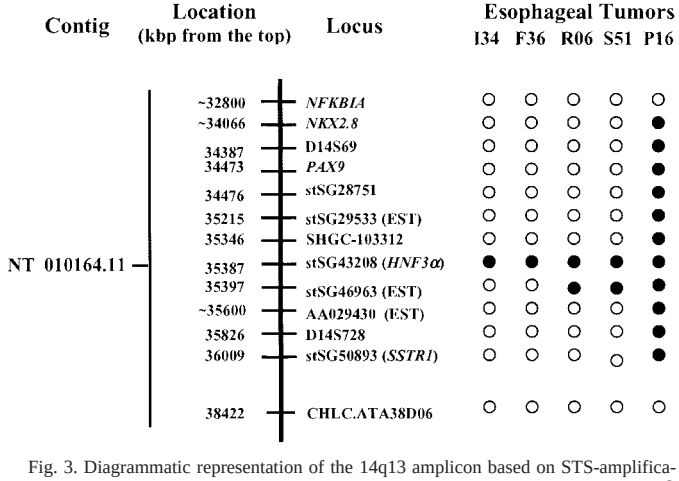

Fig. 3. Diagrammatic representation of the 14q13 amplicon based on STS-amplification mapping using QG-PCR. The map order is drawn based on NCBI databases. ○ indicate the absence of genomic amplification and ● represent the presence of genomic amplification, which was determined by a ratio of tumor:normal ≥2.0 (see "Materials and Methods"). Tumor P16 possesses the largest amplicon, spanning >6 Mb. Marker stSG43208 (gene *HNF3α*), demonstrates the highest frequency of amplification in 75 tumors examined.

Table 2 *RT-PCR analysis of genes and ESTs in tumors containing the 14q13 amplicon*

EST[a]	Gene	Tumor I34	F36	R06	S51	P16
stSG15955	Unknown	−[b]	−	−	−	+[c]
stSG29533	Unknown	−	+	+	−	+
SHGC-103312	Unknown	−	−	+	−	+
stSG43208	*HNF3α*	+	+	+	+	+
SGC35681	*HNF3α*	+	+	+	+	+
stSG46963	Unknown	+	+	+	−	+
AA029430	Unknown	+	+	−	−	+
stSG50893	*SSTR1*	−	−	−	−	+
Real-time RT-PCR of *HNF3α* (-fold increase)		2.0	4.3	3.1	7.3	12.8

[a] ESTs located within the core-amplified domain showing increased expression in the tumors demonstrating *HNF3α* amplification. The order from top to bottom represents ESTs oriented from the centromeric to the telomeric direction.
[b] "−" (Ts/c:Ns/c) intensity ratio <2.0.
[c] "+" (Ts/c:Ns/c) intensity ratio ≥2.0.

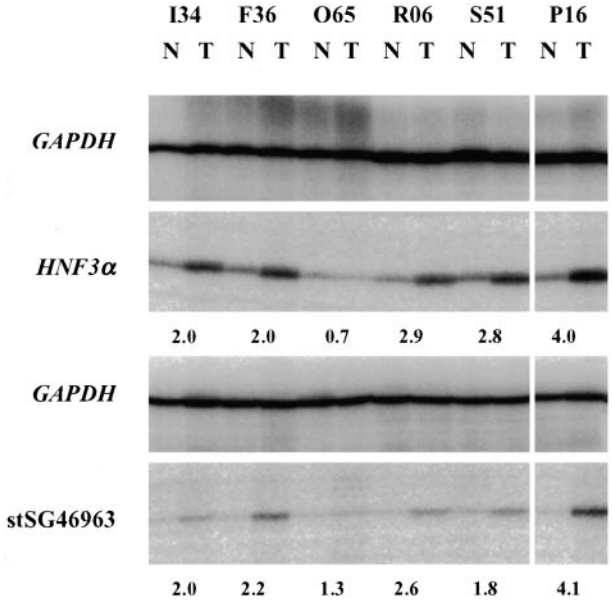

Fig. 4. Quantitative RT-PCR analysis of genes and ESTs in the core-amplified region. Overexpression of the *HNF3α* mRNA was found in all tumors amplified (*P16, R06, F36, S51,* and *I34*) and was the most frequently overexpressed gene of the 8 genes/ESTs tested. It was verified by real-time RT-PCR approach (see Table 2).

identified three amplified *Not*I/*Dpn*II fragments in an esophageal adenocarcinoma (P16). The fragments were cloned, sequenced, and mapped to chromosome band 14q13. The core-amplified domain of the 14q13 amplicon was subsequently defined and characterized. The amplicon extends >6 Mb, and the core-amplified domain is contained in a region <0.3 Mb, which includes the *HNF3α* gene (Table 1). Real-time quantitative RT-PCR analysis revealed that *HNF3α* is overexpressed in all 5 of the amplified tumors (Table 2). Nuclear staining of the *HNF3α* protein was observed in the esophageal tumors containing the *HNF3α* amplicon. Amplification and overexpression of *HNF3α* were also found in lung adenocarcinomas. Overexpression of the *HNF3α* mRNA can be gene amplification dependent and independent, because elevated expression of this gene was observed not only in the tumor RNA, which possesses *HNF3α* gene amplification, but also in lung tumor RNAs not containing gene amplification. This is the first study to report and characterize the 14q13 amplification in esophageal and gastric cardia adenocarcinomas as well as in lung adenocarcinomas.

During the preparation of this manuscript, amplification of chromosome band 14q13 was reported in cell lines derived from esophageal squamous cell carcinomas (20). The minimal common region of amplification was 4 Mb with 7 genes named as the candidates for the amplicon. One of the most frequently amplified genes in that study, *NFKBIA*, which is 2.6 Mb centromeric to *HNF3α*, was not found amplified in any of the 75 primary esophageal adenocarcinomas examined in the present study. The analyses of genomic DNA amplification, and mRNA and protein expression in these tumors indicate that *HNF3α* is a strong candidate for the 14q13 amplicon in our series of esophageal adenocarcinomas. Analyses of virtual Northern and serial analysis of gene expression (SAGE) data indicated that increased expression of *HNF3α* is observed in cell lines derived from prostate, breast, and ovarian cancers.[6] High expression of the *HNF3α* mRNA has also been observed in a luminal subtype of primary breast carcinomas (21). Increased *HNF3α* mRNA expression in many tumor types indicates a possible role for *HNF3α* in the development or progression of many cancers.

HNF3α (*FOXA1*) belongs to the hepatocyte nuclear factor 3 (*HNF3*) gene family, which includes *HNF3β* (*FOXA2*) and *HNF3γ* (*FOXA3*). *HNF3α* encodes a polypeptide of 473 amino acids in humans (22). The *HNF3* genes are members of the forkhead class of DNA-binding proteins (23), all of which contain a highly conserved 110-amino acid forkhead motif, a variant of the helix-turn-helix motif, first identified in the *Drosophila* gene fork head (*fkh*; Ref. 11). *HNF3α* is expressed in embryonic endoderm and adult tissues of endodermal origin, including stomach, intestines, liver, and lung (24). Recently, expression of the secreted signaling factor sonic hedgehog (*Shh*) in mice has been shown to be regulated by an *HNF3*-dependent mechanism (25). Overexpression of *Shh* in mice has been associated with the development of basal cell carcinoma (26).

Forkhead-family genes have been associated previously with cancer and leukemia, but our report is the first example of amplification and overexpression in primary tumors of the esophagus and lung.

Table 3 *HNF3α amplification and mRNA expression analysis of lung adenocarcinomas using QG-PCR and oligonucleotide microarrays*

Tumor	Microarray expression value[a]	Fold expression increase[b]	*HNF3α* gene amplification
S53	1980	9.4	+[c]
B31	1014	4.8	−[d]
B32	1097	5.2	−
S82	2111	10.0	−
C66	1193	5.6	+

[a] Expression value based on oligonucleotide microarray analysis (*HNF3α* mean expression = 211 in 10 normal lung samples).
[b] Fold expression increase is compared with the mean expression in normal lungs.
[c] "+" (Ts/c:Ns/c) intensity ratio ≥2.5.
[d] "−" (Ts/c:Ns/c) intensity ratio <2.5.

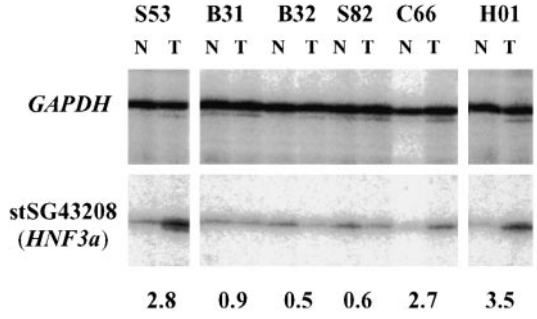

Fig. 5. The *HNF3α* gene is amplified in the lung adenocarcinomas demonstrating overexpression of *HNF3α* mRNA in Affymetrix oligonucleotide microarray analysis. Five of 86 lung adenocarcinomas, all of which demonstrated a >4.8-fold expression increase of *HNF3α* (Table 3), were analyzed to determine whether *HNF3α* amplification is present in these tumors. *HNF3α* gene amplification was observed in 2 tumors (*S53* and *C66*) of these 5 tested. *HNF3α* amplification was also found in another lung tumor, H01, from an independent series of 27 lung adenocarcinomas.

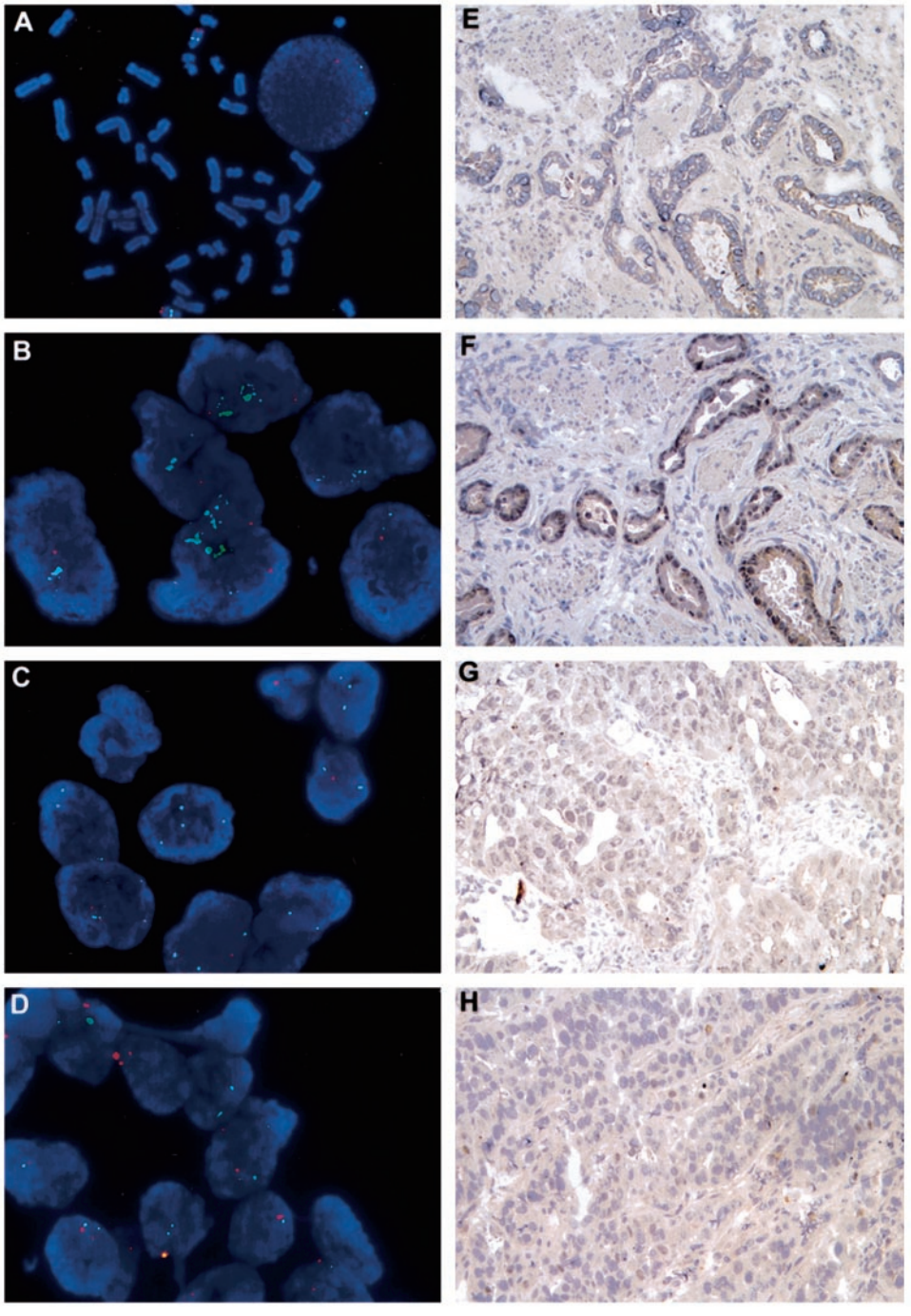

Fig. 6. Left panel (*A–D*) shows images of an interphase FISH study with the tumors tested coordinated to the tumors examined at *right panel* (*E–H*) of immunohistochemical assays. A chromosome telomeric probe (at 14q32, *red*) was cohybridized with the *HNF3α* containing BAC probe (R-356O9, *green*) to a metaphase section as the control (*A*) and the sections from tumors P16 (*B*, highly amplified), I34 (*C*, low increase in copy number), and O65 (*D*, not containing the 14q13 amplicon). Tumor P16 demonstrates intense *HNF3α* nuclear staining in the immunohistochemical assay (*F*) as compared with the section of the same tumor not receiving the primary antibody (negative control; *E*). An esophageal tumor (*I34*), observed low increase in DNA copy number, demonstrates low-level nuclear staining in a majority of tumor cells (*G*). A tumor not containing *HNF3α* amplification demonstrates a sporadic nuclear staining of *HNF3α* (*H*).

Breakpoints within the forkhead domain of two forkhead-family genes, *FOXO4* and *FOXO2*, have been found to be associated with chromosomal translocations in acute lymphocyte leukemia and secondary acute leukemia (27, 28). Translocation involving another forkhead family gene, *FOXO1a*, has also been associated with pediatric solid tumor alveolar rhabdomyosarcoma (29). Oncogenic transformation was observed when the fused gene of *PAX3* and *FOXO1a* was transfected into chicken embryo fibroblasts (30). In addition, the *Foxg1b* (*qin*) oncogene, a member of the forkhead/*HNF3* family of transcriptional regulators, has been found to transform chicken embryo fibroblasts and induce fibrosarcomas in chickens (31, 32). A deletion of the winged helix forkhead domain of the *qin* gene was shown to completely abrogate the oncogenic capacity of *qin/foxg1b* in chicken embryo fibroblasts, indicating the necessity of the forkhead domain for the transforming activity of the gene (32). Interestingly, several forkhead family genes, *e.g.*, FOXA1(*HNF3α*), *FOXG1A*, and *FOXG1B* (*QIN*) are clustered and located at 14q13,[6] but both *FOXG1A* and *FOXG1B* are outside our 14q13 amplicon. Evidence from our study and oncogenic properties of other members of the forkhead gene family suggest a potential oncogenic role for *HNF3α* (*FOXA1*) in the development and/or progression of esophageal and lung adenocarcinomas.

ACKNOWLEDGMENTS

We thank Caitlin Biedron for excellent technical assistance with these studies.

REFERENCES

1. Farrow, D. C., and Vaughan, T. L. Determinants of survival following the diagnosis of esophageal adenocarcinoma (United States). Cancer Causes Control, *7:* 322–327, 1996.
2. Devesa, S. S., Blot, W. J., and Fraumeni, J. F. Changing patterns in the incidence of esophageal and gastric carcinoma in the United States. Cancer (Phila.), *83:* 2049–2053, 1998.
3. Neshat, K., Sanchez, C. A., Galipeau, P. C., Cowan, D. S., Ramel, S., Levine, D. S., and Reid, B. J. Barrett's esophagus: a model of human neoplastic progression. Cold Spring Harbor Symp. Quant. Biol., *59:* 577–583, 1994.
4. Jankowski, J. A., Wright, N. A., Meltzer, S. J., Triafilopoulos, G., Geboes, K., Casson, A. G., Kerr, D., and Young, L. S. Molecular evolution of the metaplasia-dysplasia-adenocarcinoma sequence in the esophagus. Am. J. Pathol., *152:* 965–973, 1999.
5. Houldsworth, J., Cordon-Cardo, C., Ladanyi, M., Kelsen, D. P., and Chaganti, R. S. Gene amplification in gastric and esophageal adenocarcinomas. Cancer Res., *50:* 6417–6422, 1990.
6. Brien, T. P., Odze, R. D., Sheehan, C. E., McKenna, B. J., and Ross, J. S. HER-2/neu gene amplification by FISH predicts poor survival in Barrett's esophagus-associated adenocarcinoma. Hum. Pathol., *31:* 35–39, 2000.
7. Riegman, P. H. J., Vissers, K. J., Alers, J. C., Geelen, E., Hop, W. C. J., Tilanus, H. W., and van Dekken, H. Genomic alterations in malignant progression of Barrett's esophagus. Cancer Res., *61:* 3164–3170, 2001.
8. Hughes, S. J., Glover, T. W., Zhu, X. X., Kuick, R., Thoraval, D., Orringer, M. B., Beer, D. G., and Hanash, S. A novel amplicon at 8p22–23 results in overexpression of cathepsin B in esophageal adenocarcinoma. Proc. Natl. Acad. Sci. USA, *95:* 12410–12415, 1998.
9. Lin, L., Aggarwal, S., Glover, T. W., Orringer, M. B., Hanash, S., and Beer, D. G. A minimal critical region of the 8p22–23 amplicon in esophageal adenocarcinomas defined using sequence tagged site-amplification mapping and quantitative polymerase chain reaction includes the GATA-4 gene. Cancer Res., *60:* 1341–1347, 2000.
10. Lin, L., Prescott, M. S., Zhu, Z., Singh, P., Chun, S. Y., Kuick, R. D., Hanash, S. M., Orringer, M. B., Glover, T. W., and Beer, D. G. Identification and characterization of a 19q12 amplicon in esophageal adenocarcinoma reveals *cyclin E* as the best candidate for this amplicon. Cancer Res., *60:* 7021–7027, 2000.
11. Lai, E., Prezioso, V. R., Tao, W. F., Chen, W. S., and Darnell, J. E. Jr. Hepatocyte nuclear factor 3 α belongs to a gene family in mammals that is homologous to the Drosophila homeotic gene fork head. Genes Dev., *5:* 416–427, 1991.
12. Ang, S. L., Wierda, A., Wong, D., Stevens, K. A., Cascio, S., Rossant, J., and Zaret, K. S. The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead proteins. Development (Camb.), *119:* 1301–1315, 1993.
13. Sasaki, H., and Hogan, B. L. Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. Development (Camb.), *118:* 47–59, 1993.
14. Kaestner, K. H., Katz, J., Liu, Y., Drucker, D. J., and Schutz, G. Inactivation of the winged helix transcription factor HNF3α affects glucose homeostasis and islet glucagon gene expression *in vivo*. Genes Dev., *13:* 495–504, 1999.
15. Blin, N., and Stafford, D. W. A general method for isolation of high molecular weight DNA from eukaryotes. Nucleic Acids Res., *3:* 2303–2308, 1976.
16. Kuick, R., Asakawa, J., Neel, J. V., Satoh, C., and Hanash, S. M. High yield RFLPs in two-dimensional separations of human genomic DNA. Genomics, *25:* 345–353, 1995.
17. Sambrook, J., Fritsch, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual, Section I, pp. 1.25–1.28. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1989.
18. Hughes, S. J., Nambu, Y., Soldes, O. S., Hamstra, D., Rehemtulla, A., Iannettoni, M. D., Orringer, M. B., and Beer, D. G. Fas/APO-1 (CD95) is not translocated to the cell membrane in esophageal adenocarcinoma. Cancer Res., *57:* 5571–5578, 1997.
19. Giordano, T. J., Shedden, K. A., Schwartz, D. R., Kuick, R., Taylor, J. M., Lee, N., Misek, D. E., Greenson, J. K., Kardia, S. L., Beer, D. G., Rennert, G., Cho, K. R., Gruber, S. B., Fearon, E. R., and Hanash, S. Organ-specific molecular classification of primary lung, colon, and ovarian adenocarcinomas using gene expression profiles. Am. J. Pathol., *159:* 1231–1238, 2001.
20. Yasui, K., Imoto, I., Fokuda, Y., Pimkhaokham, A., Yang, Z., Naruto, T., Shimada, Y., Nakamura, Y., and Inazawa, J. Identification of target genes within an amplicon at 14q12–q13 in esophageal squamous cell carcinoma. Genes Chromosomes Cancer, *32:* 112–118, 2001.
21. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Thorsen, T., Quist, H., Matese, J. C., Brown, P. O., Botstein, D., Eystein Lonning, P., and Borresen-Dale, A. L. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc. Natl. Acad. Sci. USA, *98:* 10869–10874, 2001.
22. Bingle, C. D., and Gowan, S. Molecular cloning of the forkhead transcription factor HNF-3 α from a human pulmonary adenocarcinoma cell line. Biochim. Biophys. Acta, *1307:* 17–20, 1996.
23. Kaestner, K. H., Hiemisch, H., Luckow, B., and Schutz, G. Hepatocyte nuclear factor 3 α belongs to a gene family in mammals that is homologous to the Drosophila homeotic gene fork head. Genomics, *20:* 377–385, 1994.
24. Monaghan, A. P., Kaestner, K. H., Grau, E., and Schutz, G. Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 α, β and γ genes in the determination of the definitive endoderm, chordamesoderm and neuroectoderm. Development (Camb.), *119:* 567–578, 1993.
25. Epstein, D. J., McMahon, A. P., and Joyner, A. L. Regionalization of Sonic hedgehog transcription along the anteroposterior axis of the mouse central nervous system is regulated by Hnf3-dependent and -independent mechanisms. Development (Camb.), *126:* 281–292, 1999.
26. Oro, A. E., Higgins, K. M., Hu, Z., Bonifas, J. M., Epstein, E. H., Jr., and Scott. M. P. Basal cell carcinomas in mice overexpressing sonic hedgehog. Science (Wash. DC), *276:* 817–821, 1997.
27. Parry, P., Wei, Y., and Evans, G. Cloning and characterization of the t(X;11) breakpoint from a leukemic cell line identify a new member of the forkhead gene family. Genes Chromosomes Cancer, *11:* 79–84, 1994.
28. Hillion, J., Le Coniat, M., Jonveaux, P., Berger, R., and Bernard, O. A. AF6q21, a novel partner of the MLL gene in t(6;11)(q21;q23), defines a forkhead transcriptional factor subfamily. Blood, *90:* 3714–3719, 1997.
29. Galili, N., Davis, R. J., Fredericks, W. J., Mukhopadhyay, S., Rauscher, F. J., III, Emanuel, B. S., Rovera, G., and Barr, F. G. Fusion of a fork head domain gene to PAX3 in the solid tumour alveolar rhabdomyosarcoma. Nat. Genet., *5:* 230–235, 1993.
30. Scheidler, S., Fredericks, W. J., Rauscher, F. J., III, Barr, F. G., and Vogt, P. K. The hybrid PAX3-FKHR fusion protein of alveolar rhabdomyosarcoma transforms fibroblasts in culture. Proc. Natl. Acad. Sci. USA, *93:* 9805–9809, 1996.
31. Li, J., and Vogt, P. K. The retroviral oncogene qin belongs to the transcription factor family that includes the homeotic gene fork head. Proc. Natl. Acad. Sci. USA, *90:* 4490–4494, 1993.
32. Chang, H. W., Li, J., and Vogt, P. K. Domains of the qin protein required for oncogenic transformation. Oncogene, *13:* 441–444, 1996.