J Hum Genet (1999) 44:116–120

SHORT COMMUNICATION

Mayumi Tamari · Yataro Daigo · Yusuke Nakamura

Isolation and characterization of a novel serine threonine kinase gene on chromosome 3p22-21.3

Received: October 26, 1998 / Accepted: November 30, 1998

Abstract Through large-scale DNA sequencing of a genomic region on chromosome 3p22-p21.3, we isolated a novel gene encoding a 527-amino-acid protein. Its 18 exons spanned a genomic region of about 90 kilobases, and the 4536-nucleotide cDNA contained an open reading frame of 1581 base pairs. The gene was expressed in all 16 human tissues examined by Northern blotting. The amino acid sequence of the predicted protein was 39% identical to that of human SOK1 (Ste20/oxidant stress response kinase-1), a molecule that is activated by oxidative stress. In view of its significant similarity to SOK1, we suspect that the novel gene, which we named *OSR1*, is a member of the SOK family of kinases in terms of function.

Key words Chromosome 3p · Serine threonine kinase · Oxidant stress · Genomic structure

Introduction

To investigate candidate loci containing putative tumor suppressor genes, we earlier undertook detailed deletion mapping of the short arm of chromosome 3 in a large number of cancerous tissues from lung, uterus, and kidney. That study identified three commonly deleted regions, at 3pter-p25, p22-p21.3, and p21.1-p14 (Yamakawa et al. 1991; Hibi et al. 1992; Yokoyama et al. 1992). Subsequent analysis disclosed homozygous deletions at a locus within one of the commonly deleted regions (3p22-21.3) in several lung cancer cell lines (Yamakawa et al. 1993). Since at least one putative tumor suppressor gene was likely to be present in the homozygously deleted region, we further analyzed one of those cell lines and defined the size of the deletion, which was almost 700 kb. The normal genomic sequence from which these 700 kilobases had been deleted was included in a YAC clone, Y936C1 (Murata et al. 1994). To identify genes that are normally present within that region, we have been performing a large-scale analysis of genomic DNA sequences in the vicinity. DNA sequencing and subsequent screening for cDNAs initially disclosed four genes: *trans-Golgi p230*, *integrin aRLC*, *HYA22*, and a novel gene whose predicted product revealed a high degree of homology to villin (Ishikawa et al. 1997a). Although these four genes were considered to be candidate tumor suppressors for lung and renal cancers, on the basis of their locations within the deleted region, we were unable to find any somatic alterations in any of the primary tumors we examined. However, as the homozygous deletion may have exerted a positional effect by influencing the expression of genes in the near vicinity, we extended the DNA sequencing effort beyond the limits of the 700-kb deleted region. In that effort we located five genes ([*human activin receptor type IIB*, *OCTL1*, *OCTL2*, *XYLB*, and *PLC*δ*1* (phospholipase C delta 1)]) lying distal to the homozygously deleted region (Ishikawa et al. 1997b; 1998; Nishiwaki et al. 1998; Tamari et al. 1998). We now report the identification of a novel serine-threonine kinase gene in the region under investigation.

Materials and methods

Sequencing of cosmid DNAs and computer analysis of genomic DNA sequences

Four cosmid clones (581, 576, 425, and 566) were fragmented by sonication; 1.5- to 6.0-kb fragments were recovered and subcloned into pBluescript IISK(-) cloning vector (Stratagene, La Jolla, CA, USA). Plasmid DNAs were isolated from 300 randomly chosen subclones. To confirm the insert sizes and to eliminate clones that contained cosmid-vector sequences, plasmids were digested with *Eco*RI and *Xho*I, electrophoresed on 0.8% agarose gels, and analyzed

M. Tamari · Y. Daigo · Y. Nakamura (✉)
Laboratory of Molecular Medicine, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
Tel. +81-3-5449-5372; Fax +81-3-54495433
e-mail: yusuke@ims.u-tokyo.ac.jp

Y. Nakamura
Department of Genome Analysis, Cancer Institute, Tokyo, Japan

by Southern blotting with ^{32}P-dCTP-labeled cosmid-vector DNA as the probe. Clones containing human genomic DNA were sequenced with the Dye Terminator Sequencing FS Ready Reaction Kit (Applied Biosystems, Foster City, CA, USA), using T3 or T7 primers and an ABI PRISM 377 DNA sequencer (Perkin-Elmer corporation, Norwalk, CT, USA). DNA sequences were assembled using the ABI Assembler computer software.

cDNA cloning, homology search, and determination of exon-intron junctions

The assembled DNA sequences were analyzed with a computer program (GRAIL 2; Xu et al. 1994) that predicts which sequences might be transcribed. DNA fragments were scored as "excellent", "good", or "marginal" in this context. Using these possible exonic sequences, we performed exon-connection experiments (using reverse transcriptase (RT)-polymerase chain reaction-(PCR)) and screened cDNA libraries. One microgram of poly(A) RNA from human skeletal muscle (Clontech, Palo Alto, CA, USA) was reverse transcribed for single-stranded cDNA, using oligo(dT)15 primer (Boehringer Mannheim, Mannheim, Germany). RT-PCR was carried out in a 20-μl reaction mixture at 94°C for 2 min for initial denaturing, followed by 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s, in a Gene Amp PCR system 9600 (Perkin Elmer Cetus, Norwalk, CT, USA). Each amplified fragment was subcloned in pT7-BlueT (Novagen, Madison, WI, USA); nucleotide sequences were determined with the ABI PRISM 377 DNA sequencer. A total of 1×10^6 clones from a human skeletal-muscle cDNA library (Clontech) were screened with α-^{32}P-dCTP-labeled cDNA fragments that had been isolated by the exon-connection experiments. The amino-acid sequence of the predicted protein was analyzed with the FASTA program (The Genome Net Database service, Japan (http:// www.genome.ad.jp)) to detect homologies with known proteins. Exon-intron junctions of the gene were defined by comparing cDNA sequences with genomic sequences.

Northern-blot analysis with cloned fragments

Northern blot experiments were performed using human multiple tissue Northern blots I and II (Clontech). The cDNA fragments were labeled with α-^{32}P-dCTP by random priming. Membranes were prehybridized and then hybridized according to the manufacturer's protocol. Washed membranes were autoradiographed for 24 h at 80°C.

Single-strand conformation polymorphism (SSCP) analysis

We performed SSCP analysis to detect alterations of the novel gene in DNAs isolated from human lung cancer tissues. Tumors and corresponding non-cancerous tissues were obtained after informed consent was given, from 60 patients with primary lung cancers who had undergone surgery at the Cancer Institute Hospital, Tokyo. RNAs extracted from these carcinomas served as templates for the RT-PCR. The coding region was divided into six segments, each of which was amplified separately, using end-labeled primers. For the SSCP analysis, the amplified DNAs were electrophoresed in 5% acrylamide gels containing 5% glycerol (Orita et al. 1989). DNA sequencing of aberrant bands was performed by the methods described above.

Results

cDNA cloning and homology search

In all, 89,903 base pairs of assembled DNA sequences derived from cosmids 581, 576, 425, and 566 (Fig. 1) were analyzed by GRAIL 2 (Xu et al. 1994). Of the 14 computer-predicted exonic fragments, 11 were scored as "excellent", one as "good", and two as "marginal". Using these possible exonic sequences, we performed exon-connection experiments and screened cDNA libraries. In this manner we obtained a cDNA sequence of 4536 nucleotides, which included 1581 bp of open reading frame encoding 527 amino acids. There was a polyadenylation signal, AATAAA, in the 3′ non-coding region (4502–4507). The nucleotide sequence data will appear in the DDBJ/EMBL/Genbank nucleotide sequence databases with the accession number AB017642. Homology analysis of the predicted amino-acid sequence by the FASTA program revealed significant similarities to human Ste20/oxidant stress response kinase-1 (SOK1; 39% identity in 414 amino acids) and MST1 (35% identity in 384 amino acids) (Fig. 2). Both of these proteins are serine-threonine kinases. Since SOK1 is activated by oxidant stress, we named our gene *OSR1* (*oxidative-stress responsive* gene 1).

Northern-blot analysis

Messenger RNAs isolated from various human tissues revealed a 4.6-kb transcript of OSR1 in all tissues examined (lung, kidney, colon, thymus, heart, liver, spleen, skeletal muscle, ovary, leucocyte, small intestine, testis, prostate, placenta, brain, and pancreas; Fig. 3). An additional, minor, transcript of 7.5 kb was detected in skeletal muscle and heart. However, as we isolated no cDNA clone corresponding to the minor transcript, it remains uncertain whether the slower band reflected a transcript generated from the same gene by alternative splicing or alternative poly(A), or a transcript of some related, cross-hybridizing gene.

Determination of exon-intron junctions

We defined the exon-intron junctions of the *OSR1* gene (Table 1) by comparing cDNA sequences with genomic DNA sequences. The results indicated that it consists of 18 exons and spans a genomic region of about 90 kilobases, as

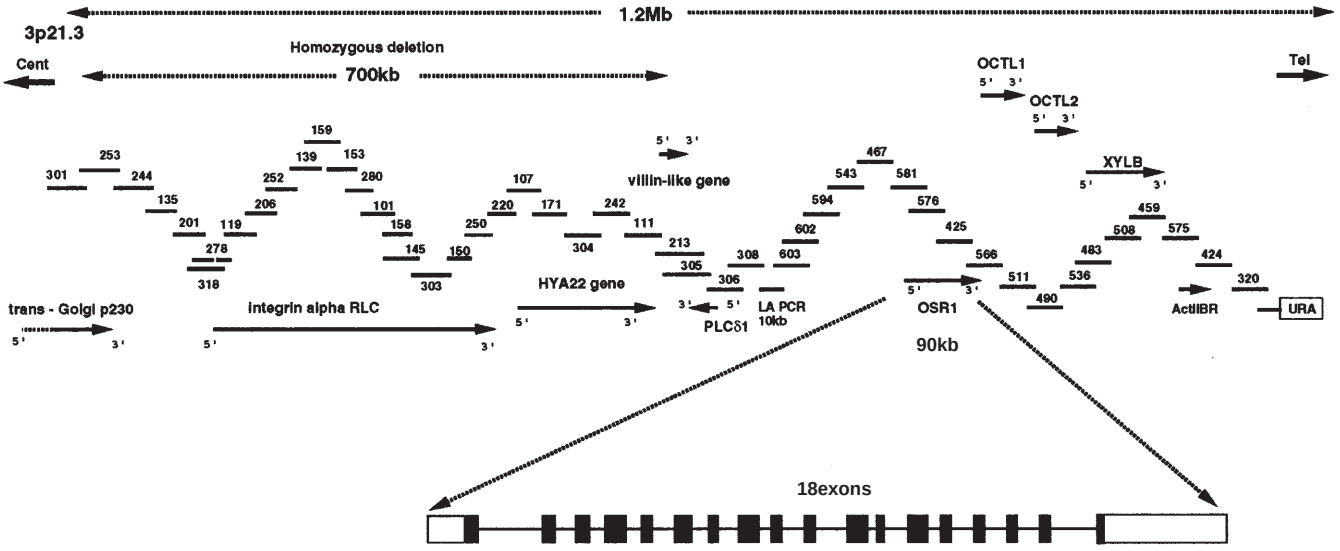

Fig. 1 Location of the novel gene (*oxidative stress-resporsive* gene 1; *OSR1*) and cosmid clones 581, 576, 425, and 566 relative to eight other genes within or near the region at chromosome 3p21.3 that is homozygously deleted in some lung cancer cell lines. The genomic structure of the gene is represented schematically; exons and introns are denoted by *boxes* and *horizontal lines*, respectively. *Arrows* indicate the transcriptional direction of each gene

Fig. 2 Homologies (*black areas*) between the predicted product of the *OSR1* gene and human SOK1 and MST1 proteins

```
OSR1  107 PRAAAAVMSM SEDS-SALPW SINRDDYELQ EVIGSGATAV VQAAYCAPKK EKVAIKRINL
SOK1    1       MAHL RGFA-NQHSR VDPEELFTKL DRIGKGSFGE VYKGIDNHTK EVVAIKIIDL
MST1    8 NPPRRQLKKL DEDSLTKQPE EV----FDVL EKLGEGSYGS VYKAIHKETG QIVAIKQVPV

OSR1  167 EKCQTSMDEL LKEIQAMSQC HHPNIVSYYT SFVVKDELWL VMKLLSGGSV LDIIKHIVAK
SOK1   55 EEAEDEIEDI QQEITVLSQC DSPYITRYFG SYLKSTKLWI IMEYLGGGSA LDLLK-----
MST1   68 E---SDLQEI IKEISIMQQC DSPHVVKYYG SYFKNTDLWI VMEYCGAGSV SDIIRL----

OSR1  227 GEHKSGVLDE STIATILREV LEGLEYLHKN GQIHRDVKAG NILLGEDGSV QIADFGVSAF
SOK1  115 ----PGPLEE TYIATILREI LKGLDYLHSE RKIHRDIKAA NVLLSEQGDV KLADFGVA--
MST1  128 ---RNKTLTE DEIATILQST LKGLEYLHFM RKIHRDIKAG NILLNTEGHA KLADFGVAGQ

OSR1  287 LATGGDITRN KV-RKTFVGT PCWMAPEVME QVRGYDFKAD IWSFGITAIE LATGAAPYHK
SOK1  175 ----GQLTDT QIKRNTFVGT PFWMAPEVIK Q-SAYDFKAD IWSLGITAIE LAKGEPPNSD
MST1  188 LTD----TMA K--RNTVIGT PFWMAPEVIQ EI-GYNCVAD IWSLGITAIE MAEGKRPYAD

OSR1  347 YPPMKVLMLT LQNDPPSLET GVQDKEMLKK YGKSFRKMIS LCLQKDPEKR PTAAELLRHK
SOK1  235 LHPMRVLFLI PKNSPPTLE- G--------Q HSKPFKEFVE ACLNKDPRFR PTAKELLKHK
MST1  248 IHPMRAIFMI PTNPPPTFR- ----KPEL-- WSDNFTDFVK QCLVKSPEQR ATATQLLQHP

OSR1  407 FFQK-AKNKE FLQEKTLQRA PTISERAKKV RRVPGSSGRL HKTEDGGWEW SDDEFDEESE
SOK1  295 FITRYTKKTS FLTE------ --LIDRYKRW K----SEG-- HGEESSS--- EDSDIDGEAE
MST1  308 FVRS-AKGVS ILRD-LINEA MDV-----KL KRQESQQREM DQDDEENSE- -EDEMDSGTM

OSR1  467 EGKAAISQLR SPRVKESISN SELFPTTDPV GTLLQVPEQI SAHLPQPAGQ IATQPTQVSL
SOK1  355 DGEQGPIWTF PPTIRPS-PH SKLHK----- GTALHSSQK- ------PADA VKRQPRSQCL
MST1  368 V-RAVGDEMG TVRVASTMTD GANTMIEHDD TL---PSQ-L GTMVINAEDE EEEGTMKRRD

OSR1  527 PPTAEPA--- -KTAQALSSG S-GSQETKIP ISLVLRLRNS KKELNDIRFE FTPGRDTAEG
SOK1  415 STLVRPVFGE LKEKHKQSGG SVGALEELEN AFSLAEESCP GISDKLMVHL VERVQRFSHN
MST1  428 ETMQPAKPSF LEYFEQKEKE NQINSFGKSV PGPLKNSSDW KIPQDGDYEF -LKSWTVE--
```

indicated in Fig. 1. All of the 14 exons predicted by the GRAIL 2 program were, in fact, transcribed into mRNA.

SSCP

To investigate the candidacy of this gene for involvement in carcinogenesis, we performed SSCP analysis to screen DNAs isolated from human lung cancers for alterations in *OSR1*. An identical aberrant pattern was present in 4 of the 60 primary lung carcinomas examined. DNA sequencing of the PCR products disclosed a C→T transition at the second nucleotide of codon 304, which resulted in a substitution of isoleucine (ATA) for threonine (ACA). However, as this substitution was observed in the corresponding normal tissues and also in 1 of 55 normal control DNAs, we consid-

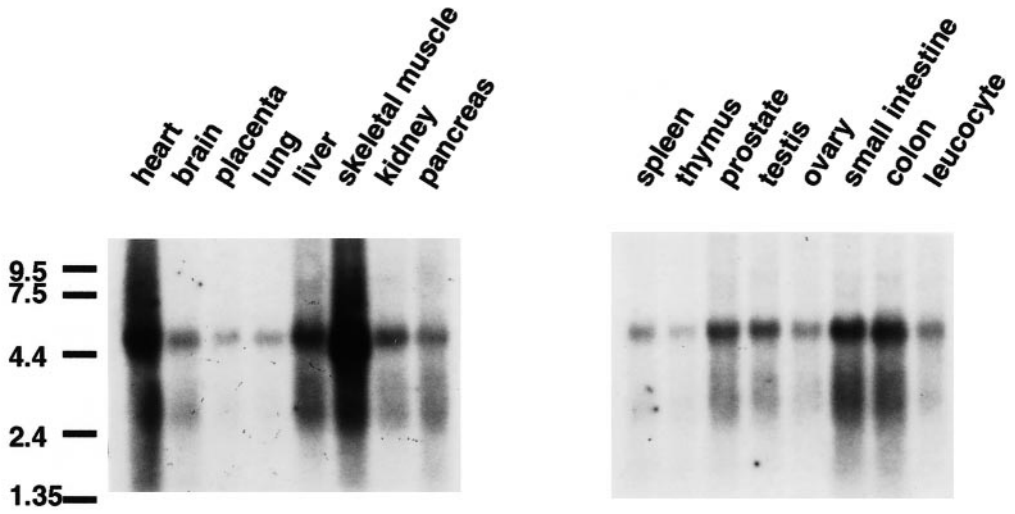

Fig. 3 Multiple-tissue Northern blots of *OSR1* cDNA. The main transcript, 4.5 kb, is observed ubiquitously

Table 1 Exon-Intron boundary sequences of the *OSR1* gene

	Exon length (bp)	Splice acceptor	Splice donor
1	412		GAGGTGATCGgtgagagcag
2	113	ttgtttttagGGAGTGGAGC	TGAACTCCTGgtatgcacat
3	109	ttctttgtagAAAGAAATTC	CTAAGTGGAGgtgagtagag
4	142	ctgttttcagGTTCTGTTCT	AGATCCACAGgtatgtaaaa
5	56	ttttttaaagAGATGTGAAA	CAGATTGCAGgtaatgatta
6	110	tctatttcagACTTTGGGGT	TATGGAACAGgtaccgtgtc
7	102	tgttttgcagGTCCGTGGTT	ACCAATGAAGgtgaggcttg
8	134	attcttgcagGTTTTAATGC	CAGAAAAAAGgtaaaatatg
9	49	ctggttttagACCAACAGCA	GAAAGCAAAGgtaggaaatt
10	66	tcctctccagAATAAAGAAT	AGCAAAAAAGgtaaatcaga
11	123	taatgagcagGTTCGGAGAG	ACAACTCAGGgtaaatttta
12	36	tttcgtttagTCTCCCCGAG	AAATTCTGAGgtaagtaatt
13	147	tgctttttagGCTTTGTCTT	AACAGCTCAGgtaaagccgg
14	65	tttgttacagCTCTTTCCAA	TAAGATTAAGgtaagtagac
15	53	gaattcccaaAAAGAACTAA	CCTGGGAGAGgtgaggcatc
16	69	gtccttgcagATACAGCAGA	TTAGTAATGgtaactccat
17	65	atgattacagTGGCAGCTAA	TTTCAAACTGgtactcatcc
18	2668	tggttcttagGCATCTGGTG	GTAAATTGTCcttagtggtg

The accession number of the OSR1 cDNA sequence is AB017642
OSR1, *Oxidative stress-responsive* gene 1

ered it to reflect a rare polymorphism not associated with cancer.

Discussion

We have reported here the isolation of the cDNA, and characterization of the complete genomic structure, of a novel gene encoding a protein with significant similarity to SOK1 (Ste20/oxidant stress response kinase-1) and MST1. SOK-1, a human Ste20-like kinase from the germinal center (GC) kinase group, functionally resembles the yeast Ste20 proteins that transduce signals in response to environmental stress. However, SOK1 is not involved in the generalized stress response pathway that is stimulated by growth factors, alkylating agents, cytokines, or environmental situations such as heat shock and osmolar stress. It is activated relatively specifically by oxidant stress, although it does not activate any of the known mitogen activated protein (MAP) kinase cascades (Creasy and Sheriff 1995; Pombe et al. 1996). MST1, although it is homologous to a member of a yeast mitogen activated protein kinase (MAPK) cascade, is not involved in regulating any mammalian MAPK pathway; this kinase potentially regulates a novel signaling cascade (Creasy et al. 1996).

We designated the novel gene reported here as *OSR1* (*oxidative-stress responsive* gene 1), because, on the basis of its significant similarity to SOK1, it may be involved specifically in the response to oxidative stress. Characterization of the genomic structure of *OSR1* should contribute to investigations concerned with divergence or evolution among the

SOK1 family of kinases.

Since the *OSR1* gene is located in the close vicinity of the region that is homozygously deleted in several lung cancer cell lines, we consider it to be a candidate tumor suppressor. However, our failure to find mutated forms of this gene in any of 60 primary lung carcinomas examined makes it quite unlikely that *OSR1* plays any carcinogenic role in the lung itself.

Acknowledgments This work was supported in part by a special grant for Strategic Advanced Research on Cancer from the Ministry of Education, Culture, Sports, and Science of Japan; and by a Research for the Future Program Grant of The Japan Society for the Promotion of Science.

References

Creasy CL, Sheriff J (1995) Cloning and characterization of a human protein kinase with homology to Ste20. J Biol Chem 270: 21695–21700

Creasy CL, Ambrose DM, Chernoff J (1996) Ste20-like protein kinase, Mst1, dimerizes and contains an inhibitory domain. J Biol Chem 271: 21049–21053

Hibi K, Takahashi T, Yamakawa K, Ueda R, Sekido Y, Ariyoshi Y, Suyama M, Takagi H, Nakamura Y, Takahashi T (1992) Three distinct regions involved in 3p deletion in human lung cancer. Oncogene 7: 445–449

Ishikawa S, Kai M, Tamari M, Takei Y, Kuwata K, Takeuchi K, Bandou H, Yamane Y, Ogawa M, Nakamura Y (1997a) Sequence analysis of a 685-kb genomic region on chromosome 3p22-p21.3 that is homozygously deleted in a lung carcinoma cell line. DNA Res 4: 35–43

Ishikawa S, Takahashi T, Ogawa M, Nakamura Y (1997b) Genomic structure of the human PLCD1 (phospholipase C delta) locus on 3p22-p21.3. Cytogenet Cell Genet 78: 58–60

Ishikawa S, Kai M, Murata Y, Tamari M, Daigo Y, Murano T, Ogawa M, Nakamura Y (1998) Genomic organization and mapping of the human activin receptor type IIB (hActR-IIB) gene. Jpn J Hum Genet 43: 132–134

Murata Y, Tamari M, Takahashi T, Horio Y, Hibi K, Yokoyama S, Inazawa J, Yamakawa K, Ogawa A, Takahashi T, Nakamura Y (1994) Characterization of an 800-kb region at 3p22-p21.3 that was homozygously deleted in a lung cancer cell line and construction of a cosmid contig. Hum Mol Genet 3: 1341–1344

Nishiwaki T, Daigo Y, Tamari M, Fujii Y, Nakamura Y (1998) Molecular cloning, mapping, and characterization of two novel human genes, *OCTL1* and *OCTL2*, bearing homology to organic-cation transporters. Cytogenet Cell Genet (in press)

Orita M, Suzuki S, Sekiya T, Hayashi K (1989) A rapid and sensitive detection of point mutations and polymorphisms using the polymerase chain reaction. Genomics 5: 874–879

Pombe CM, Bonventre JV, Molnar A, Kyriakis J, Force T (1996) Activation of a human Ste20-like kinase by oxidant stress defines a novel stress response pathway. EMBO J 15: 4537–4546

Tamari M, Daigo Y, Ishikawa S, Nakamura Y (1998) Genomic structure of a novel human gene on chromosome 3p22-21.3 encoding a xylulokinase-like protein. Cytogenet Cell Genet 82: 101–104

Yamakawa K, Morita R, Takahashi E, Hori T, Ishikawa J, Nakamura Y (1991) A detailed deletion mapping of the short arm of chromosome 3 in sporadic renal cell carcinoma. Cancer Res 51: 4707–4711

Yamakawa Y, Takahashi T, Horio Y, Murata Y, Takahashi E, Hibi K, Yokoyama S, Ueda R, Takahashi T, Nakamura Y (1993) Frequent homozygous deletions in lung cancer cell lines detected by a DNA marker located at 3p21.3-p22. Oncogene 8: 327–330

Yokoyama Y, Yamakawa K, Tsuchiya E, Murata M, Sakiyama S, Nakamura Y (1992) Deletion mapping on the short arm of chromosome 3 in squamous cell carcinoma and adenocarcinoma of the lung. Cancer Res 52: 873–877

Xu Y, Mural RJ, Uberbacher EC (1994) Constructing gene models from accurately predicted exons; an application of dynamic programming. Comput Appl Biosci 10: 613–623