GENES, CHROMOSOMES & CANCER 45:676–691 (2006)

# Analysis of a New Homozygous Deletion in the Tumor Suppressor Region at 3p12.3 Reveals Two Novel Intronic Noncoding RNA Genes

**Debora Angeloni,[1]* Arja ter Elst,[2] Ming Hui Wei,[1] Anneke Y. van der Veen,[2] Eleonora A. Braga,[3] Eugene A. Klimov,[4] Tineke Timmer,[2] Luba Korobeinikova,[4] Michael I. Lerman,[1] and Charles H.C.M. Buys[2]**

[1]Laboratory of Immunobiology, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, MD
[2]Department of Medical Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
[3]The Laboratory of Molecular Diagnostics, Russian State Genetics Center, Moscow, Russia
[4]The Laboratory of Comparative Animal Genetics, Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, Russia

Homozygous deletions or loss of heterozygosity (LOH) at human chromosome band 3p12 are consistent features of lung and other malignancies, suggesting the presence of a tumor suppressor gene(s) (TSG) at this location. Only one gene has been cloned thus far from the overlapping region deleted in lung and breast cancer cell lines U2020, NCI H2198, and HCC38. It is *DUTT1* (*D*eleted in *U* *T*wenty *T*wenty), also known as *ROBO1*, *FLJ21882*, and *SAX3*, according to HUGO. *DUTT1*, the human ortholog of the fly gene *ROBO*, has homology with NCAM proteins. Extensive analyses of *DUTT1* in lung cancer have not revealed any mutations, suggesting that another gene(s) at this location could be of importance in lung cancer initiation and progression. Here, we report the discovery of a new, small, homozygous deletion in the small cell lung cancer (SCLC) cell line GLC20, nested in the overlapping, critical region. The deletion was delineated using several polymorphic markers and three overlapping P1 phage clones. Fiber-FISH experiments revealed the deletion was approximately 130 kb. Comparative genomic sequence analysis uncovered short sequence elements highly conserved among mammalian genomes and the chicken genome. The discovery of two EST clusters within the deleted region led to the isolation of two noncoding RNA (ncRNA) genes. These were subsequently found differentially expressed in various tumors when compared to their normal tissues. The ncRNA and other highly conserved sequence elements in the deleted region may represent miRNA targets of importance in cancer initiation or progression. Published 2006 Wiley-Liss, Inc.†

## INTRODUCTION

Loss of function of tumor suppressor genes (TSGs) is a fundamental genetic change involved in the origin and pathogenesis of human tumors (Knudson, 1971; Marshall, 1991). TSGs have a recessive mode of action; therefore, both copies need to be inactivated to produce a phenotypic effect. This frequently occurs when one allele is hit by a mutation or silenced through promoter region hypermethylation and the other is lost generally because of a large chromosomal deletion. In some instances, both alleles become inactivated by genetic loss as a consequence of a homozygous deletion. Sub-lethal homozygous deletions are frequently smaller than their heterozygous counterparts. This characteristic makes them a useful tool for localizing TSGs.

Cytogenetic and molecular deletion mapping studies have long implicated chromosome bands 3p25-26, 3p21.3, and 3p12-14 as harboring TSGs involved in multiple forms of human cancers, including lung cancers (Whang-Peng et al., 1982; Kok et al., 1987; Zbar, 1989; Kok et al., 1997; Zabarovsky et al., 2002).

The *FHIT* gene, residing in 3p14.2, is subjected to homozygous deletions and alterations of its mRNA in many sporadic cancers (Huebner et al., 1998). Frequent allele loss at the *FHIT* locus has been found in low-grade breast cancer. 3p12 is a particularly significant region (Lerman and Minna, 2000), as demonstrated by functional studies (Lott et al., 1998; Lovell et al., 1999). Rabbitts et al. (1990) reported a homozygous deletion at the *D3S3* locus in the U2020 cell line. It spans about 8 Mb



The content of the publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Supported by: NIH, Dutch Cancer Society; Grant numbers: RUG 94-834, RUG 2000-2317, J.K. de Cock Stichting; Grant number: 99-07, The Russian Foundation for Basic Research; Grant numbers: 04-04-48112, INTAS 03-51-4983, President of the Russian Federation in "Support of young Ph.D."; Grant number: MK-963.2005.4.

*Correspondence to: Dr. Debora Angeloni, Scuola Superiore Sant'Anna and IFC-CNR, Via Moruzzi, 1-56124 Pisa, Italy.
E-mail: angeloni@ifc.cnr.it

Received 27 June 2005; Accepted 5 March 2006
DOI 10.1002/gcc.20332
Published online 10 April 2006 in Wiley InterScience (www.interscience.wiley.com).

and is flanked by the microsatellite markers *D3S1284* and *D3S1276* (Drabkin et al., 1992; Latif et al., 1992). Other nested or overlapping homozygous deletions have also been identified in this region. Todd et al. (1997) reported a homozygous deletion that overlaps the U2020 region and is flanked by the microsatellite markers *D3S1254* and *D3S1776*. Another overlapping homozygous deletion was found in breast cancer cell line HCC38. It spans about 5 Mb and is flanked by the microsatellite markers *D3S2537* and *D3S2527* (Sundaresan et al., 1998a). In small cell lung cancer (SCLC) cell line NCI-H2198, a much smaller deletion was found, which contains the microsatellite markers *D3S1274*, *D3S2498*, and *D3S4492* (telomere to centromere, Sundaresan et al., 1998a).

The *DUTT1* (*D*eleted in *U T*wenty-*T*wenty) gene was also isolated from this low gene-density region of chromosome 3 (Sundaresan et al., 1998b). *DUTT1* gene expression is impaired by hypermethylation of its promoter in primary cancers of kidney and breast, but is hypermethylated to a lesser extent in lung cancers.

In this paper, we describe a new, homozygous deletion at 3p12 identified in the SCLC cell line GLC20. It spans ~130 kb around the *D3S1274* microsatellite marker, and it partly overlaps with the NCI-H2198 deletion. The deletion affects exon 2 and the flanking introns of the *DUTT1* gene. In the second intron, we discovered two novel transcripts with the same orientation as DUTT1. Both are polyadenylated, have small ORFs without any known homologues or orthologues, and likely do not encode proteins. Both are putative noncoding (nc) RNA genes, which we found overexpressed in lung and breast cancers, and underexpressed in kidney cancer as compared to paired normal tissue. The deleted region also contained some sequence elements highly evolutionarily conserved among different genomes (Bejerano et al., 2004; Hillier et al., 2004).

## MATERIALS AND METHODS

### Human Tissues and DNA

Signed, informed consent was obtained from all human donors in the study, according to the NCI institutional review board approved protocol. Tissues were taken from patients not treated by radio- or chemotherapy. Primary tumors were classified by the UICC-TNM (International Union Against Cancer—Tumor-Node-Metastasis) staging system according to Sobin et al. (2002). Histological typing of tumors was performed in accordance with international protocols (Travis et al., 1999; Tavassoli and Devilee, 2003; Eble et al., 2004). Eight renal cell carcinoma cases were represented by clear cell histology type only, ten breast cancers, including mainly ductal breast carcinomas (nine cases and one case of lobular cancer), four nonsmall cell lung cancers (NSCLC), consisting predominantly of adenocarcinomas (three cases and one case of squamous-cell lung cancer).

### Cell Lines and DNA

The lung cancer cell lines NCI-H750, NCI-H2198, and NCI-H1450 were obtained from ATCC (Manassas, VA). U2020 DNA was kindly provided by Dr. Pamela Rabbits (MRC, Cambridge, United Kingdom). SCLC cell line, GLC20, was established from a primary tumor biopsy (De Leij et al., 1985), and is known to harbor a 440-kb homozygous deletion at 3p21.3 (Kok et al., 1994).

For most of our experiments, we focused on GLC20 (established in the Netherlands; De Leij et al., 1985) and NCI-H2198 (established in USA; Phelps et al., 1996). With the exception of the fiber- fluorescence in situ hybridization (FISH) experiment that was performed in the Netherlands on the GLC20 cell line only, all PCR and reverse-transcriptase (RT)-PCR experiments presented here were run separately but in parallel in the Frederick (on H2198) and Groningen (on GLC20) laboratories, for independent mapping of the new deletion.

### EST Clones

EST clones were purchased from the I.M.A.G.E. Consortium (http://image.llnl.gov/).

### PCR

PCR primers were from BioServe Biotechnologies (Laurel, MD). PCR reactions were performed using the PTC-100$^{TM}$ thermalcycler (MJ Research Inc., Bio-Rad Laboratories, Waltham, MA), in a total reaction volume of 12.5 μl, containing 100 ng of genomic DNA, 12.5 pmol of each primer, 200 μM dNTPs, and 1.5 mM $MgCl_2$. The PCR cycles were as follows: 95°C, 5 min; (95°C, 30 sec, [$T_a$ varying between 55 and 64°C], 30 sec, 72°C, 30 sec) for 35 cycles; 72°C, 7 min. PCR products were run on 3% or 4% NuSieve 3:1 agarose gel (Cambrex, Baltimore, MD) and were stained with ethidium bromide (Sigma, St. Louis, MO). Gel analysis was performed with the ViTran software (Biokom, Russia).

### Reverse-Transcriptase and Real-Time Quantitative PCR

RT-PCR was performed with the BD SMART$^{TM}$ RACE cDNA Amplification Kit (Clontech, Palo Alto, CA), according to the manufacturer's instruction.

Real-time quantitative (qRT)-PCR was performed with the GenePak™ Real-Time PCR Core Kit (IsoGen, Russia). The reaction volume was 21 μl (SYBR Green concentration 1×). The PCR thermal profile consisted of an initial denaturation of 3 min at 95°C followed by 70 cycles of 15 sec at 95°C, 15 sec at 52°C (for *W91914* primers F3 and R3) or 55°C (for B2M primers), and 15 sec at 72°C. All signals were detected between 30 and 40 cycles. Amplification, detection, and data analysis were performed with an iCycler IQ real-time detection system (Bio-Rad). Statistical analysis was performed with the REST software (http://www.gene-quantification.de/).

When an RT-PCR was performed for a protein-encoding gene (*RBM5*) with intron-flanking primers yielding different bands for DNA and RNA, similar intensities were observed for both the normal tissues and tumor cell lines (data not shown). Differences in band intensity seen between normal tissues and tumor cell lines were therefore caused by a difference in the amounts of RNA.

qRT-PCR results were analyzed with the comparative DDCt method as described (Senchenko et al., 2003): Delta Delta Ct = (DeltaCt $_{W91914\ in\ tumor}$ − DeltaCt $_{W91914\ in\ norm}$) − (DeltaCt $_{B2M\ in\ tumor}$ − DeltaCt $_{B2M\ in\ norm}$). $2^{-DDCt}$ = RQ, relative expression level in tumor as compared that of normal.

### Southern Blot

Southern blotting was done according to Sambrook et al. (1998). Briefly, genomic DNA samples were digested overnight with *EcoRI* in the presence of 1% spermidine, precipitated with Na acetate, resuspended in TE buffer, loaded on a 1% agarose gel and run overnight at 30 V. The gel was then denatured and DNA transferred overnight in NaOH (0.4 N). The membrane was washed in 2× SSC and dried for 3 hr at 80°C in a vacuum oven. Hybridization was carried out at 60°C overnight. Washes were done in 2× SSC, 1%–0.1% SDS at 65°C.

### Northern Blot

Northern blotting was performed with Clontech Human Multiple Tissue Northern Blots 7760-1 (Clontech). Hybridization was carried out at 42°C overnight. Washes were done first in 2× SSC, 1% SDS and then 2× SSC, 0.1% SDS at 65°C.

### Probe Labeling

*W91914* and *H51703* cDNAs were excised with *EcoRI*/*PacI* from the modified polylinker of the pT7T3D vector (Pharmacia, North Peapack, NJ), and radioactive probes were prepared by $^{32}$P-labeling with random primers (Rediprime DNA Labeling System, Amersham, Arlington Heights, IL).

### P1 Library Screening

We performed a PCR-screening of a P1 genomic library (Genome Systems, St. Louis, MO) and assembled a contig of P1 clones that would represent the entire deleted region (Fig. 5).

The initial screening of the P1 Human Library (Genome Systems, St. Louis, MO) was performed with PCR primers placed within EST *W91914*, and it resulted in the identification of three clones designated P1-97, P1-98, and P1-99. A second round of PCR, with primers placed at the Sp6-end of P1-98, yielded a fourth clone P1-80.

Both ends of each P1 clone were sequenced. PCR reactions performed with primers placed at both ends of each P1 clone (Fig. 3, Table 1) allowed the contig assembly.

### Fiber Fluorescence In Situ Hybridization

Preparations for fiber-FISH analysis were performed essentially according to Giles et al. (1997). GLC20 cells were spun down by centrifugation at 1,200 rpm for 10 min, and resuspended in distilled water to a concentration of 1–5 × $10^5$ cells per ml. Microscope slides were coated by incubating the slides for 30 min with 100 μl of 5% 3-aminopropyl-ethoxy-silane in acetone, washing with distilled water and air drying. The cell suspension was then spread over the coated slide and air dried. Exposure of chromatin threads from the nuclei was obtained by applying two drops of 0.5% SDS, 50 mM EDTA, 0.2M Tris HCl, pH 7.0 lysis solution on 24 mm × 60 mm cover slips. The microscope slides were then placed upside down on top of the coverslips. The slides were then turned over and kept with the coverslip on for up to 30 sec. The coverslips were then slid off gently and preparations were air dried.

Bicolor FISH analysis was performed on these preparations using differentially labeled P1 phages and routine FISH procedures, essentially as described by Driesen et al. (1991).

### Sequencing

Sequencing was performed on an ABI 373 Stretch Automated DNA Sequencer (Applied Biosystems, Foster City, CA).

### Computational Analysis

The following WWW-based servers and databases were used. Global and pairwise sequence alignments: BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) and BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat); mul-

TABLE 1. PCR Primers Used in This Study

| | U2020 | H2198 | H1450 | GLC20 | P1-97 | P1-98 | P1-99 | P1-80 | Forward primer | Reverse primer |
|---|---|---|---|---|---|---|---|---|---|---|
| A: EST clones | | | | | | | | | | |
| D81026 (W118489) | + | + | nd | nd | nd | nd | nd | nd | 5′ ggaatataaagctggagactgtgg | 5′ acgagcttacctcccatgttaact |
| H77734 (SGC35070) | − | + | nd | nd | nd | nd | nd | nd | 5′ atccagtggagagccatctttctt | 5′ cttaaagcctgtattgcaagagcc |
| T02864 (FB13C10) | − | + | nd | nd | nd | nd | nd | nd | 5′ tccttatagcagtgtgggaatg | 5′ caaaatataaacccagatccacc |
| T02956 (FB17G2) | + | + | nd | nd | nd | nd | nd | nd | 5′ caggatgctcaaccagcaggt | 5′ gttttcttcaaagtagccaatccgc |
| T64897 (stSG9120) | − | + | nd | nd | nd | nd | nd | nd | 5′ ttcaatcaagctgtggcatagagg | 5′ agttcaaggtgacagtgaactacg |
| W91914 (stSG8687) | − | − | nd | nd | nd | nd | nd | nd | 5′ gctcttcatcacattttcccatcg | 5′ gttccctgtcttccttatttccc |
| W91914 (nested) | − | − | nd | nd | nd | nd | nd | nd | 5′ tgggccaaatccaaatactc | 5′ gcgattcatttttcccctct |
| Z41019 (WI13003) | − | + | nd | nd | nd | nd | nd | nd | 5′ gtgcaacccctttattcagaatcc | 5′ tatgtctacccctgtttctgctc |
| H51703 | − | − | nd | − | − | − | − | + | 5′ ctgtactacagggaatctctc | 5′ cttcctttgggtctgttcag |
| AW861295 | − | − | nd | − | − | + | − | − | 5′ gggataaagcagtcacacag | 5′ tctcatcacatggcctgttc |
| B: Microsatellite markers | | | | | | | | | | |
| D3S1577 | − | + | + | + | − | − | − | − | 5′ tcaaaagttgcatcgc | 5′ tccattacaatcccctg |
| D3S3681 | − | + | + | + | − | − | − | − | 5′ gtgagaaccatttggggcag | 5′ ggcgagctatctgtcaggg |
| D3S1604 | − | + | + | + | − | − | − | − | 5′ caccattgtaagaggcttca | 5′ aaattgacgcataaaattgtg |
| D3S3 | − | + | + | + | − | − | − | − | 5′ cagaaggacatattcccatttg | 5′ gcagtttcctctagcttttact |
| D3S3049 | − | − | + | − | − | + | − | − | 5′ aaagacacaaggggttttagg | 5′ ttgcacattccatgaacatc |
| D3S1274 | − | − | + | − | + | + | + | − | 5′ ttatacatcagtctctgggaaacac | 5′ tactgtgcatataggttcctgtga |
| D3S3507 | − | + | + | + | − | − | − | − | 5′ tccagcccttatacctactctc | 5′ tggaatcaagacaagactgaac |
| D3S2563 | − | + | + | + | − | − | − | − | 5′ atattttagctggactcggtgctg | 5′ ttgcagactctgctattggcctg |
| D3S2530 | nd | + | + | + | − | − | − | − | 5′ acaggcaattggtgaagcatg Sp6-end | 5′ gatttatccatggcctctgctc Sp6-end |
| C: P1 Clone contig | | | | | | | | | | |
| P1-97 | − | − | + | + | + | − | + | − | 5′ cctgagtttgatttgcatgtgtct | 5′ gagctacagttcaagatgagattg |
| P1-98 | − | − | + | − | − | + | − | + | 5′ aatgaaatcttcgaagttgc | 5′ atagcatatattgacag |
| P1-80 | − | − | + | − | − | + | − | + | 5′ gtgatagcatatattgacag T7-end | 5′ ttttaaggaaaaccatcgcc T7-end |
| P1-97 | − | − | + | − | + | + | + | − | 5′ tggtagcgtgaaacttgcctaccag | 5′ cagtgtggaaagtgggaaggtaga |
| P1-98 | − | − | + | − | + | + | + | − | 5′ tgtcggttgtttcagctctgc | 5′ gggagtcatttttccctcagg |
| P1-80 | − | − | + | + | − | − | − | + | 5′ gaggtggtggcttgaaatgc | 5′ atgagaacccagatgac |
| D: DUTT 1 cDNA (Z95705) | | | | | | | | | | |
| 71-259 | − | nd | nd | + | − | − | − | − | 5′ atcctctctgcccttctctg | 5′ acactcgcacgtcttctggg |
| 601-812 | − | nd | nd | + | − | − | − | − | 5′ cacattgtgagggcgcac | 5′ tcagggcaattactcgtcg |
| 1021-1345 | − | nd | nd | − | − | + | − | + | 5′ tcccgtcttcgtcaggaag | 5′ tggctacttccagcgatgc |
| 1351-1490 | − | nd | nd | + | − | − | − | − | 5′ cttcgggatgacttcagaca | 5′ ttatcatccagtggagagcc |
| 1510-1620 | − | nd | nd | + | − | − | − | − | 5′ cgaggaggaaagctcatgat | 5′ agtcagctcggctacttcac |
| 1623-1776 | − | nd | nd | + | − | − | − | − | 5′ gctcttactgaactcctaaa | 5′ cgtagaaatgctgagttgac |
| 2541-2811 | − | nd | nd | + | − | − | − | − | 5′ tagtgccccatcaaaacctg | 5′ ggatctgatatttggcttgg |
| 5291-5376 | − | nd | nd | + | − | − | − | − | 5′ atcttccaccacctcctgtg | 5′ tagaagggagttttggcacc |
| 6371-6609 | − | nd | nd | + | − | − | − | − | 5′ caaacaattcgaatggggtag | 5′ ggtcattaaaaacatccacttg |
| E: B2M | | | | | | | | | 5′ tgactttgtcacagcccaagatag | 5′ caaatgcggcatcttcaaacctc |
| AP34R | | | | | | | | | 5′ ctgaggagggcgggcttc | 5′ ctgctgcattttagagaccgagtc |
| PDHB | | | | | | | | | 5′ qqtatqqatqaqqaqctqqa | 5′ caqccctcqactaaccttqt |

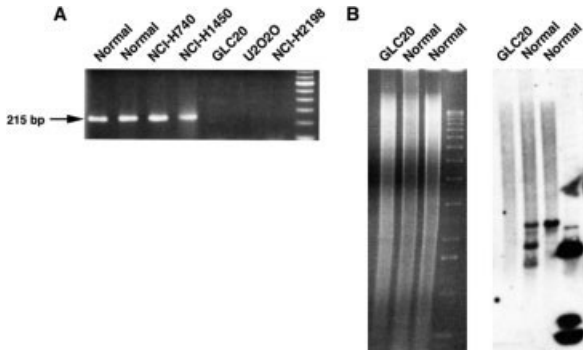



Figure 1. A novel homozygous deletion at 3p12.3 in the SCLC cell lines GLC20 and H2198. (A) Genomic PCR with *W91914* primers on normal (CEPH) DNA and cancer cell lines; 100-bp ladder Invitrogen. (B) Southern blot performed with 10 μg of genomic DNA digested with *EcoRI*. Probe: EST *W91914*. Lane 1: GLC20, lanes 2 and 3: normal individuals, lane 4: molecular marker (Invitrogen). The two normal individuals differ for a *EcoRI* polymorphism.

tiple sequence alignment: ClustalW (http://www.ebi.ac.uk/clustalw/index.html); CpG island search: CpG Plot (http://www.ebi.ac.uk/emboss/cpgplot/index.html?); analysis of nonredundant sets of gene-oriented clusters: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db = unigene; genome annotations: The Human Genome Browser Gateway (http://genome.ucsc.edu/cgi-bin/hgGateway), Acembly (http://www.ncbi.nih.gov/IEB/Research/Acembly/index.html), and Ensembl (http://ensembl.org); prediction of complete gene structures in genomic sequences (exons, introns, promoters, and poly-adenylation signals): GenScan (http://genome.dkfz-heidelberg.de/cgi-bin/GENSCAN/genscan.welcome.pl); analysis of protein features: Psort (http://psort.nibb.ac.jp) and Pfam (http://pfam.wustl.edu); search of possible miRNA target sites: DIANA (http://diana.pcbi.upenn.edu/cgi-bin/micro_t.cgi); analysis of known miRNA sequence: The miRNA Registry (http://www.sanger.ac.uk/Software/Rfam/mirna/index.shtml).

## RESULTS

### Discovery of a New Homozygous Deletion in 3p12.3

We performed PCR experiments on genomic DNA from a panel of lung cancer cell lines with primers located in seven ESTs (*T02956*, *H77734*, *D81026*, *T02864*, *T64897*, *W91914*, and *Z41019*), which map between the genetic markers *D3S1274* (alias *AFM154xa7*, *Z16684*) and *D3S1604* (alias *AFM316vc1*, *Z24325*), Table 1. Among the cell lines analyzed, only U2020 and NCI-H2198 were known to harbor a homozygous deletion at 3p12.

All ESTs, except *T02956* and *D81026*, were found deleted in the U2020 cell line (Table 1 and data not shown). *W91914* was found deleted also in the small cell lung cancer (SCLC) cell lines H2198 and GLC20 (Fig. 1A). The latter was previously not known to harbor deletions in this region of 3p. This serendipitous result was confirmed by Southern blotting experiments (Fig. 1B), and the deletion was designated the "GLC20-3p12 deletion".

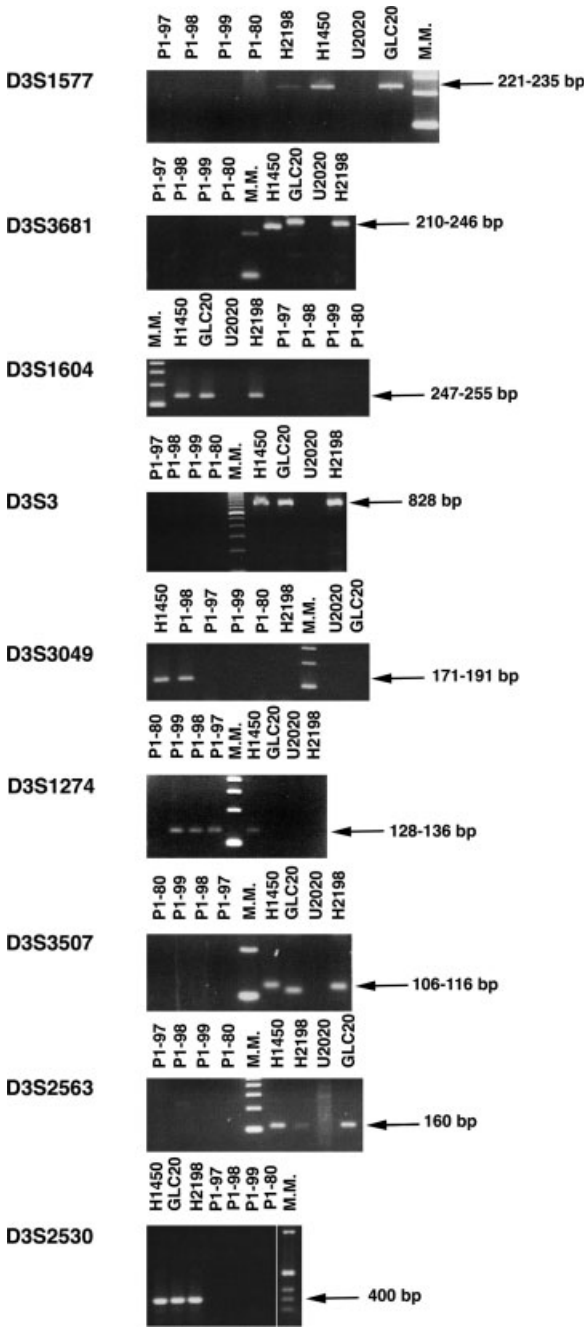



Figure 2. PCR experiments to investigate the position of several microsatellite markers with respect to the GLC20-3p12 homozygous deletion and the four P1 clones that cover the deletion. Microsatellites are ordered (top to bottom) from the more telomeric to the more centromeric. PCR primers are listed in Table 1.

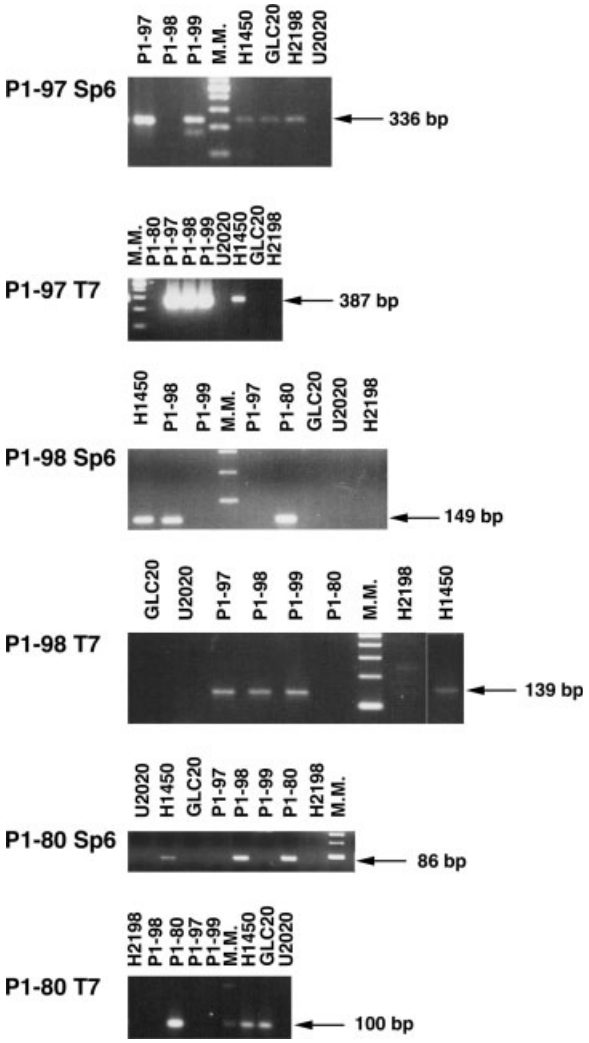



Figure 3. PCR experiments with primers located within the ends of each P1 clone were performed to build a P1 clone contig and to identify the deletion boundaries. The SCLC cell lines GLC20, U2020, and NCI-H2198 were investigated. PCR primers are listed in Table 1.

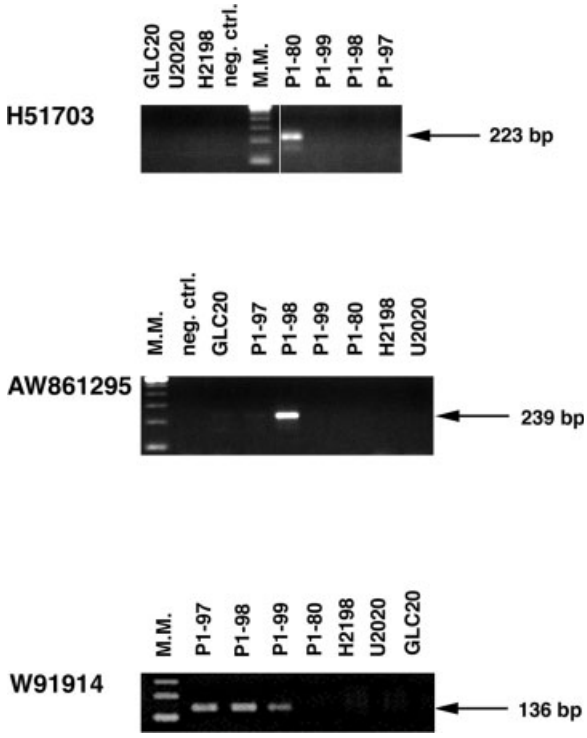



Figure 4. PCR experiments with primers located within the ESTs deleted in the GLC20-3p12 homozygous deletion and represented in the P1 contig. PCR primers are listed in Table 1.

### Characterization of the Deletion Breakpoints

The exact genomic location and extent of the GLC20-3p12 deletion were determined by PCR experiments using several microsatellite genetic markers (Table 1, Fig. 2). The deletion was found to be located around the *D3S1274* marker in both the NCI-H2198 and the GLC20 cell lines (Fig. 2).

PCR-screening of a P1 genomic library (Genome Systems, St. Louis, MO) allowed the assembly of a P1-clone contig representing entirely the deleted region (Fig. 3, Table 1, and Fig. 5). P1-97 contains the proximal boundary of the deletion and is centromeric to P1-98, which overlaps, for a short region, with P1-80, containing the telomeric boundary.

Several ESTs were positioned in the contig by PCR (Figs. 4 and 5). EST *AW861295* was also shown, by PCR, to be homozygously deleted in GLC20-3p12. It mapped to P1-98 (Fig. 4), but was not investigated further, because a BLAST search did not retrieve any other overlapping EST clones.

To corroborate the PCR findings and verify the contig location with respect to the centromeric and telomeric ends of 3p, the P1 clones were used in fiber-FISH experiments. These were conducted on normal DNA and DNA from the GLC20 cell line (Fig. 6), using the referral clones P1-26 and P1-27, previously described by Latif et al. (1992). The fiber-FISH approach was used in addition to the EST and microsatellite PCR experiments, because it allowed physical imaging of the markers' respective position. Using BLAT, the genomic position of the deletion borders, P1-97 Sp6-end and P1-80 T7-end, (deposited under GenBank Accession numbers DQ100613 and DQ100614, respectively) were retrieved. Their positions on chromosome 3 were 78.918,087-78.918,453, and 79.117,734-79.118,091 respectively, according to the Human Genome Browser, May 2004, hg 17 assembly.

The extension of the deletion is hence less than 200.005 bp (i.e., the distance between the two anchoring markers P1-80 T7-end and P1-97 Sp6-end). By means of the fiber-FISH experiments, the

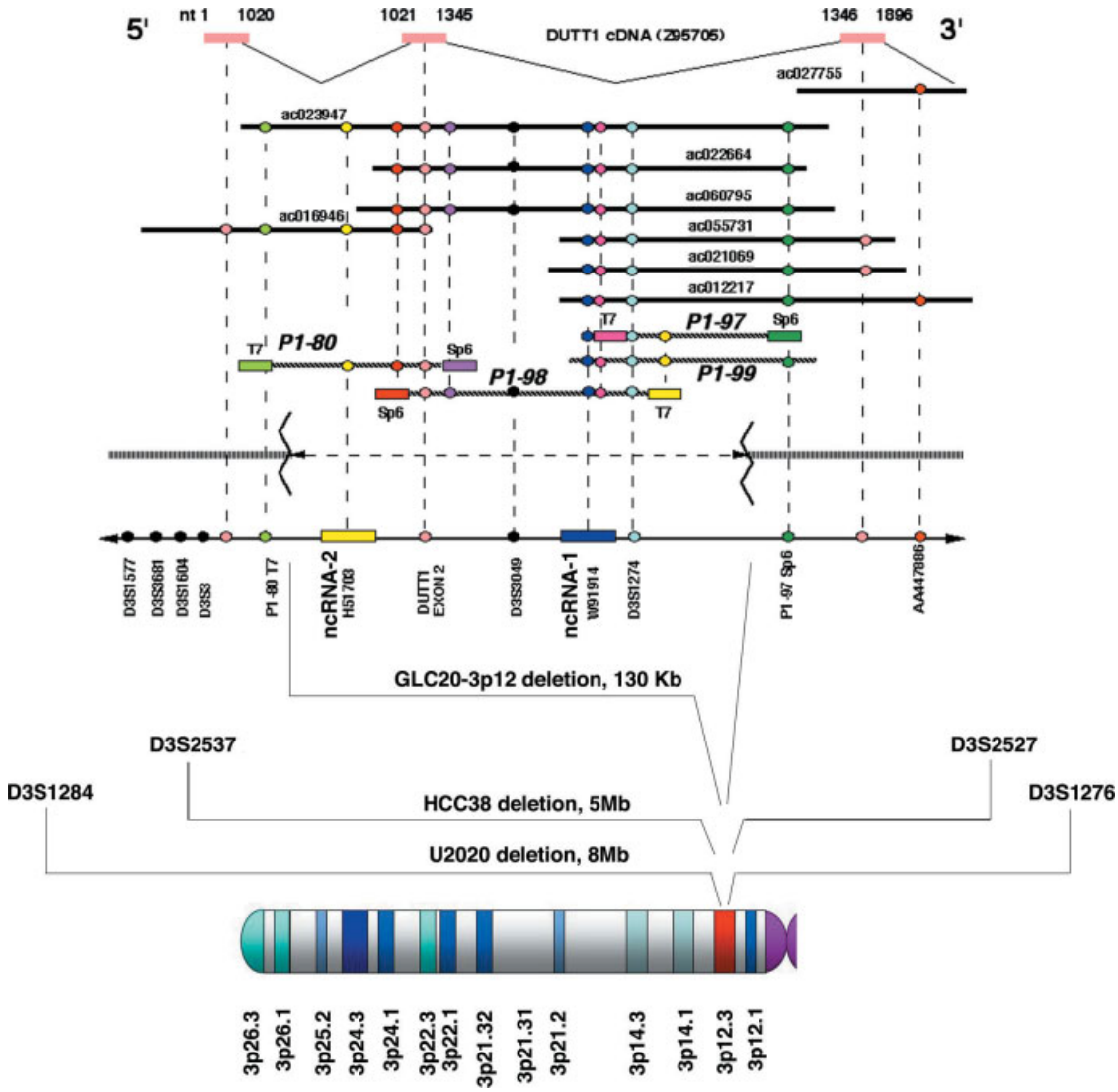



Figure 5. Summary of results obtained by PCR and fiber-FISH experiments, showing the genomic location of the GLC20-3p12 homozygous deletion. Rectangles: sequenced ends of P1 clones. Circles: positive results of specific PCR experiments indicating "anchoring points" of sequences to one another, used to build up the contig.

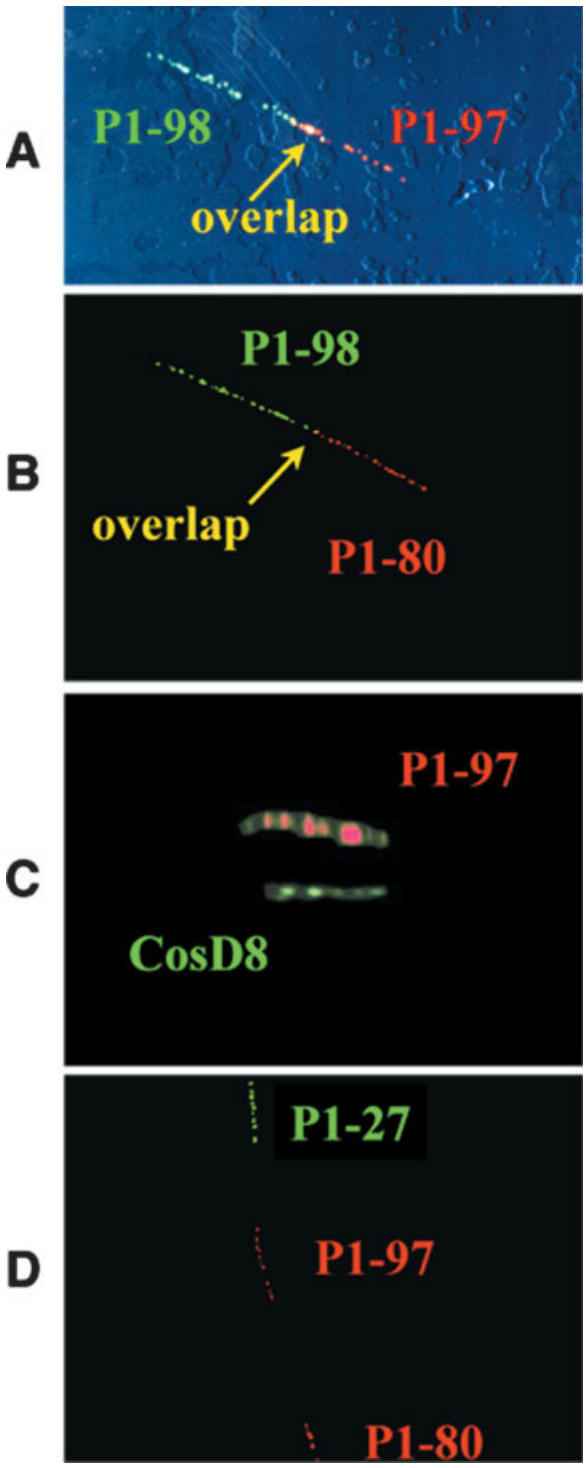
deletion was determined to be ~110–130 kb (see legend of Fig. 6)

### Status of the *DUTT1* Gene in the GLC20-3p12 Homozygous Deletion

Using PCR primers placed within the exonic sequence of the *DUTT1* gene (Table 1), we determined that the GLC20-3p12 homozygous deletion removed one exon of *DUTT1* (exon 2 of *DUTT1* variant 2, GenBank Z95705, NM_133631, corresponding to cDNA residues 1021-1345, Fig. 7), and it comprised entirely the region between introns 1 and 2 of the same gene (Fig. 5). These findings were confirmed by RT-PCR on *DUTT1* cDNA prepared from GLC20 cells. The deletion caused the loss of amino acids 19-128 (as described for the cell line H2198 by Sundaresan et al. (1998b)). These results were also confirmed by PCR on YAC clones 912A11, 15HC9, and 35AH8 (data not shown). The coverage of the new deletion in terms of BAC clones was determined in silico using electronic PCR and the P1-ends sequence as probes (Fig. 5).

Upon comparing our experimental results with the Human Genome Browser database (May 2004, hg 17 assembly, http://genome.ucsc.edu/), we found a perfect match for the location of all markers. The new homozygous deletion in GLC20-3p12 is nested within the 8-Mb deletion of the U2020 cell line and

the 5-Mb deletion of the HCC38 cell line, but it is much smaller. It partly overlaps the deletion in H2198, whose exact size has not been determined, but is known to be internal to *DUTT1* (Sundaresan et al., 1998a). Hence, the same genomic location contains small deletions (H2198 and GLC20-3p12) of comparable size in at least two different SCLC patients.




### The New Deletion Harbors Two Previously Unknown Transcripts—Computational Analysis

Several ESTs (some of which are adenylated) map within the deleted region in 3p12.3. We focused on two clusters and chose one from each, namely *W91914* and *H51703*, for further analysis. Several SAGE entries support these ESTs that originated from different cDNA libraries. BLAST searches with *W91914* and *H51703* against the human EST database made it possible to assemble two separate clusters and analyze their sequence with GenScan looking for predicted exons.

### Cluster of *W91914*

The cluster corresponding to the initial probe *W91914* comprised, at the date of submission of this work, eight ESTs (*W91914*, *W94988*, *H90477*, *H90421*, *AI078492*, *AA639329*, *AA777646*, and *CR740005*) isolated from fetal liver and spleen, head and neck carcinoma, and mammary carcinoma libraries. It is noteworthy that this cluster aligns (chr3: 78,485,247-79,099,496) with the full-length mRNA *BC017743* (isolated by Mammalian Gene Collection Program, Imanishi et al., 2004). A Genscan analysis of this mRNA predicted a putative exon with an ORF of 49 amino acids, but the lack of homologues or orthologues in other species suggests that the transcript is unlikely to code for a protein. Acembly annotated the gene as a single-exon transcript of 3.1 kb. Interestingly, the alignment of the *BC017743* sequence with the chicken genome produced two hits on *Gallus gallus* chromosome 1, one of which with 86% identity over 353

Figure 6. Fiber-FISH with P1 clones. The yellow spots result from merging of green and red labeling and show regions of overlap between two clones. The referral clone P1-27 is centromeric to P1-97. Consequently, the contig is oriented with P1-97 as the most proximal and P1-80 as the most distal (telomeric) clone. (A) P1-97 and P1-98 showing an overlap of an estimated size of about 20 kb when hybridized to fibers from an EBV-transformed lymphoblastoid cell line. (B) P1-98 and P1-80 hardly showing an overlap when hybridized to fibers from the same cell line. (C) Size comparison between the part of P1-97 hybridizing to fibers from the cell line GLC20 and the cohybridized cosmid cosD8 (Kok et al., 1995) showing that the part of P1-97 extending beyond the deletion is somewhat larger than the 40-kb cosmid insert. (D) P1-97 and P1-80 hybridized together with the more centromeric P1-27 to fibers from the lymphoblastoid cell line, showing the orientation of P1-97 and P1-80 with respect to the centromere and the 3p telomere. Since cohybridization of P1-80 and cosD8 to fibers from GLC20 gave a picture very similar to (C) (i.e., both P1 phages extend about 45 kb beyond the deletion), the size of the deletion is about 3 × the P1 insert size of 80 kb (cf. Fig. 3) minus the overlaps of together about 20 kb minus the extending P1 parts of about 2 × 45 kb. Thus the deletion size is ~130 kb.

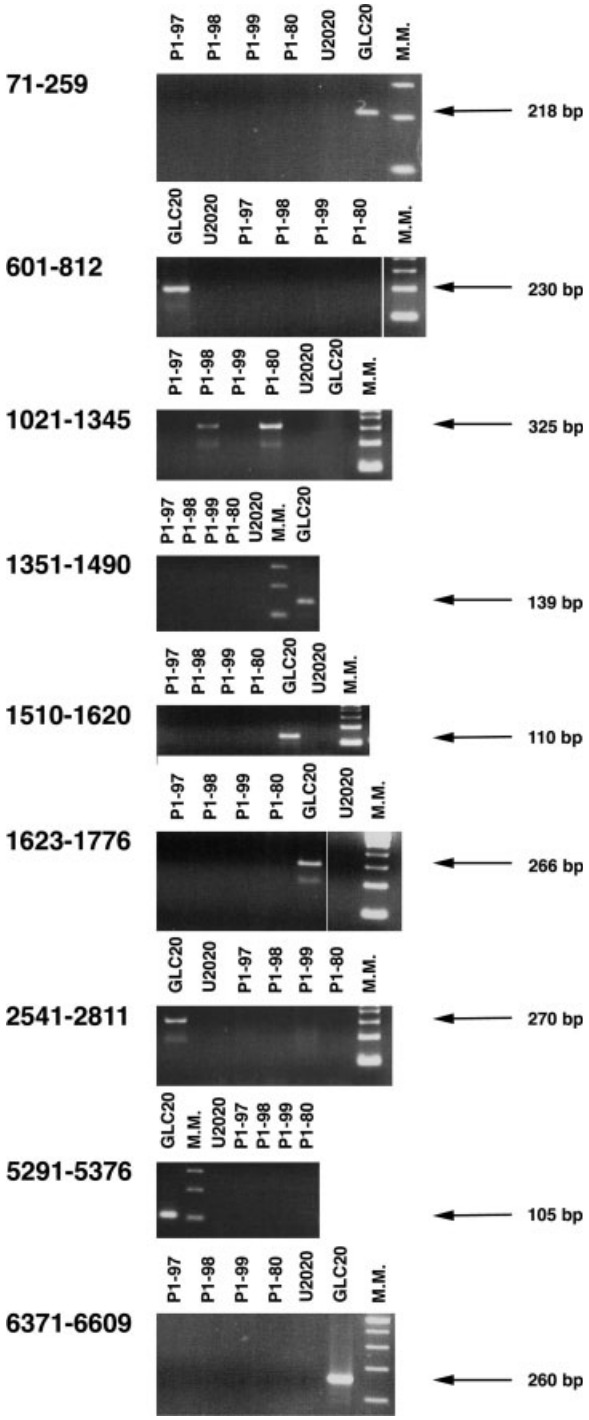



Figure 7. Exon 2 of the *DUTT1* gene (cDNA 1021-1345, as in GenBank Z95705) is homozygously deleted in the GLC20-3p12 deletion. PCR experiments with primers (listed in Table 1) located within exons at the 5′-end of *DUTT1*. The GLC20-3p12 homozygous deletion bridges across exon 2 to introns 1 and 2.

residues. Given the recognized importance of microRNAs (miRNAs) in cancer biology (McManus, 2003; Xu et al., 2004), we looked in the *BC017743* sequence for possible miRNA target sites, which would suggest the existence of a post-transcriptional control mechanism acting on this transcript. The DIANA algorithm predicted two possible target sites in the sequence (Fig. 8A), both scoring in the high-confidence range. Both miRNAs belong to a group of miRNA genes that were experimentally identified and their expression analyzed by Northern blot. They have a similar sequence originating from different loci. The miRNA *has-miR-17-5p* derives from a locus on chromosome 13 and *has-miR-106a* from a locus on the X chromosome (Mourelatos et al., 2002; Dostie et al., 2003; Kasashima et al., 2004; Suh et al., 2004).

### Cluster of *H51703*

When this work was submitted, the cluster assembled around the sequence of *H51703* was represented by 17 ESTs (*H51703*, *T69773*, *T70759*, *T84499*, *H40323*, *H40377*, *R83269*, *AA668381*, *AA669442*, *BI598464*, *BM993003*, *BX105987*, *BE062088*, *BF746150*, *BE061843*, *BF746204*, and *CA440361*) isolated from different libraries.

Again, this cluster aligned (chr3: 78,485,247-79,099,496) with a full-length transcript, namely *BC043430* (The Mammalian Gene Collection Program, Imanishi et al., 2004). Genscan analysis did not predict any exon and the small ORFs did not show similarities with known homologues or orthologues. Acembly annotated this transcript as a putative single-exon gene of about 1.9 kb, with an ORF of 46 amino acids, and no known similarities. However, two bovine ESTs, *AV601957* and *CR454939*, aligned in correspondence of *BC043430* (they showed 92% and 81% identity respectively, in a region free of repetitive DNA). In the *BC043430* sequence, the DIANA algorithm predicted four putative target sites for known miRNAs (Fig. 8B), (Houbaviy et al., 2003; Weber, 2005).

### Conservation Analysis of the Deleted Sequence

Aligning the sequence of the GLC20-3p12 deletion with the chicken genome (Bejerano et al., 2004; Hillier et al., 2004) identified seven blocks of highly conserved sequences. Four blocks contained Alu repeats. Two of the four showed 93% identity between human and chicken, and homology with the 7SL chicken gene on chicken chromosome 5. Two other blocks, almost identical to the previous ones, showed similarities with sequences on chicken chromosome 2, perhaps representing another copy of the 7SL gene. Finally, three highly conserved regions displayed homology with sequences from the chicken chromosome 1 (numbers refer

| | miRNA ID | Energy | Alignment |
|---|---|---|---|
| A | has-miR-17-5p | -33.6 | 5′-------ctgtaatcctagcactttg 3′<br>3′tggacgtgacat-----tcgtgaaac 5′ |
| | has-miR-106a | -35.6 | 5′----------ctgtaatcctagcacttt 3′<br>3′cgatggacgtgacat-----tcgtgaaa 5′ |
| B | has-miR-17-5p | -34.5 | 5′----------ctgtaatcccagcactttg 3′<br>3′tgatggacgtgacat-----tcgtgaaac 5′ |
| | has-miR-106a | -35.6 | 5′----------ctgtaatcctagcacttt 3′<br>3′cgatggacgtgacat-----tcgtgaaa 5′ |
| | has-miR-106ab | -33.6 | 5′ctgc-ctgtaatcccagcacttt 3′<br>3′gacgtgaca-----gtcgtgaaa 5′ |
| | has-miR-34b | -27.8 | 5′-------ctattatgttgactgcct 3′<br>3′gttagtcgat- tac-tgtgacgga 5′ |

Figure 8. Human microRNAs that putatively hit target sequences on (A) *BC017743* and (B) *BC043430* transcripts (adapted from a window of the DIANA algorithm results (http://diana.pcbi.upenn.edu/cgi-bin/micro_t.cgi).

to the human sequence, Human Genome Browser, May 2004 hg 17 assembly):

Seq5: 69 bp (79000588-79000654), showed 84% identity between human and chicken.
Seq6: 353 bp (78954642-78954984), showed 86% identity between human and chicken.
Seq7: 243 bp (79044143-79044377), showed 81% identity between human and chicken.

None of these regions contained repetitive elements. Seq6 aligned inside the noncoding transcript *BC017743*.

### Expression Studies

To independently verify whether the two clusters of ESTs indeed are transcribed, we performed Northern blot experiments using *W91914* and *H51703* cDNA as probes. Under our conditions, the *W91914* probe gave a detectable signal in almost all tissues tested (two bands of about 2.4 and 3 kb, Fig. 9A). The *H51703* probe yielded a band of about 2 kb in pancreas, one of about 3 kb in skeletal muscle, and one of about 1.35 kb in liver (Fig. 9B). RT-PCR experiments performed with primers located in *W91914* and *H51703* confirmed that these sequences were expressed in various normal organs (Fig. 9C). Analysis of *W91914* expression in SCLC cell lines (Fig. 9C) showed that, with the exception of GLC20, the gene was possibly expressed at greater than normal levels. A similar result was seen for *H51703* (Fig. 9C). Next, we investigated with quantitative real-time PCR, the relative expression level of *W91914* in a set of normal/tumor paired samples (Table 2) represented by breast carcinoma, NSCLC, and clear-cell kidney carcinoma

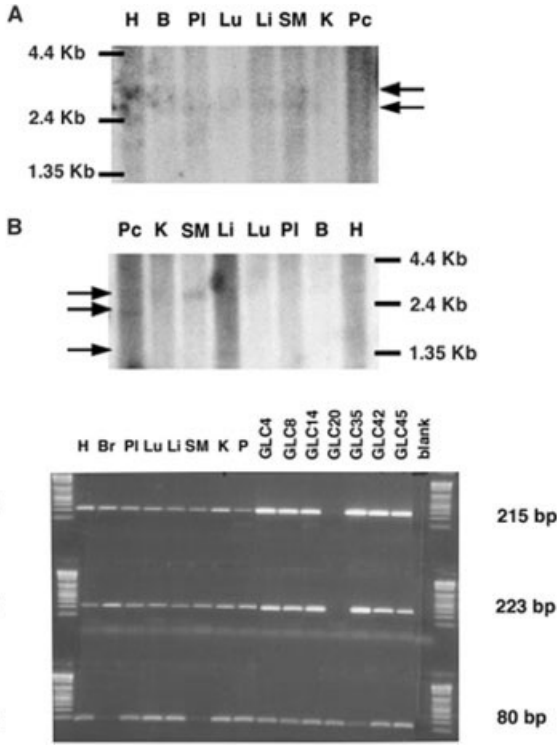



Figure 9. Expression studies of ESTs *W91914* and *H51703*, chosen as representative of two EST clusters located in the region affected by the GCL20-3p12 homozygous deletion. (A) and (B), Northern blots. Probes: *W91914* (A) and *H51703* (B). Arrows indicate bands highlighted by the respective probes. RT-PCR with primers for *BC017743* (*W91914*, 215 bp) in (C), for *BC043430* (*H51703*, 223 bp) in (D), for Beta-2-microglobulin (80 bp) in (E). Template: cDNAs from various organs and SCLC cell lines. H: heart, Br: brain, Pl: placenta, Lu: lung, Li: liver, SM: skeletal muscle, K: kidney, Pc: pancreas, and SLCS GLC cell lines. GLC20: cell line in which *BC017743* and *BC043430* are homozygously deleted. Blank: no-cDNA control.

(RCC). (*H51703* was not analyzed because of technical constraints). *W91914* was found overexpressed in three of four analyzed lung cancer biopsies (75%).

TABLE 2. Description of the Human Tumor Samples Analyzed in This Study

| Sample | Sex | Age | Histology | Stage |
|---|---|---|---|---|
| | | | Breast carcinoma | |
| BC16 | f | 53 | Ductal, md | II a |
| BC17 | f | 53 | Ductal, ld | II b |
| BC19 | f | 49 | Ductal, md | II a |
| BC20 | f | 68 | Ductal, ld | III b |
| BC21 | f | 65 | Lobular, md | I |
| BC22 | f | 55 | Ductal, md | II a |
| BC12 | f | 50 | Ductal, md | II a |
| BC13 | f | 56 | Ductal, md | II a |
| BC14 | f | 56 | Ductal, md | II b |
| BC33 | f | 63 | Ductal, md | II |
| | | | Non-small cell lung carcinoma | |
| LC01 | m | 52 | Adenocarcinoma, hd | I |
| LC02 | m | 56 | Squamous cell, ld | IV |
| LC03 | m | 58 | Adenocarcinoma, ld | IV |
| LC04 | f | 64 | Adenocarcinoma, ld | IV |
| | | | Clear cell renal cell carcinoma | |
| RCC02 | f | 57 | CC-RCC, ld | IV |
| RCC06 | m | 63 | CC-RCC, ld | IV |
| RCC10 | m | 57 | CC-RCC, ld | III |
| RCC11 | f | 59 | CC-RCC, md | II |
| RCC12 | f | 35 | CC-RCC, hd | I |
| RCC13 | f | 54 | CC-RCC, md | II |

However, it was found underexpressed in four of six RCC samples (66.67%) and in six of ten breast cancer biopsies (60%), (Fig. 10). These results show that *BC017743* (represented by *W91914*) and *BC043430* (represented by *H51703*) are differentially expressed in tumor biopsies and cancer cell lines when compared with normal tissues.

By means of RT-PCR experiments using primers located within genomic sequence or across splice sites (*AP43R*, Sulimova et al., 2002; and pyruvate dehydrogenase beta *PDHB*, primer sequence in Table 1), all cDNA samples were tested and found to be free of genomic DNA (data not shown).

## DISCUSSION

Detection of homozygous deletions and subsequent microsatellite mapping to identify allelic loss and deletion boarders remains the most powerful method to localize putative TSGs. Interstitial deletions in the low gene-density chromosome region 3p12 have been reported in lung and other malignancies (Daly et al., 1991; Ganly et al., 1992; Pandis et al., 1993). The U2020 region at chromosome band 3p12.3, in particular, has been strongly suspected to harbor a TSG. Since the mid 1990s, karyotype analysis has shown that deletions in 3p12 are the only evident abnormality in cells cultured from the normal bronchial epithelial cells of lung cancer patients (Sundaresan et al., 1995). Sanchez et al. (1994) showed that the introduction of two centromeric fragments of 3p (encompassing 3p12-q24 and 3p14-q11) into a highly malignant renal cell carcinoma (RCC) cell line resulted in a dramatic suppression of tumor growth in athymic nude mice. This suggested that a locus in this region controlled the growth of RCC cells by inducing rapid cell death in vivo. Lott et al. (1998) and Lovell et al. (1999) demonstrated that a fragment of human chromosome 3, overlapping with the U2020 deletion, mediated rapid cell death and tumor growth suppression of RCC cells in vivo. However, deletion of this region was associated with the immortalization of human uroepithelial cells (Vieten et al., 1998).

Detailed mapping determined that the U2020 deletion was about 8 Mb in size (Drabkin et al., 1992; Latif et al., 1992), and it harbors the smaller deletions HCC38 and H2198 (Sundaresan et al., 1998a). Sundaresan et al. (1998b) cloned a gene disrupted by these deletions: *DUTT1*. *DUTT1* is an integral membrane protein. It is an axon guidance/cell adhesion receptor whose best-characterized function is to regulate the decision by axons to cross the central nervous system midline. *DUTT1* is widely expressed as an about 8-Kb mRNA, and two transcript variants are known that differ in their 5′ terminal end (start site and first exon). The H2198 deletion affects exon 2 (encoding the first Ig domain) of the *DUTT1*-short variant. Introduction of this mutation in the mouse germ line (Xian et al., 2001) generates animals that in the homozygous state frequently die at birth of respiratory failure due to functional immaturity of lungs. Survivors acquire bronchial epithelial abnormalities similar to those involved in the early stages of lung cancer and die in the first year of life (Xian et al., 2001). Heterozygous mice grow normally, but spontaneously develop lymphomas and carcinomas in their second year of life, with a threefold increase in incidence compared with controls (Xian et al., 2004). Invasive lung adenocarcinoma is by far the predominant carcinoma. In addition to the mutant allele, loss of heterozygosity (LOH) analysis indicates that these tumors retain the structurally normal allele but with substantially increased methylation of their promoter region (Xian et al., 2004). Dallol et al. (2002) determined that the *DUTT1* promoter was hypermethylated in 19% of primary invasive breast carcinomas, in 18% of primary clear cell renal cell carcinomas (CC-RCC), and in 4% of primary NSCLC tumors. In addition, 80% of breast carcinomas and 75% of CC-RCC containing *DUTT1* hypermethylation also show allelic loss at 3p12.

| Samples | $2^{-DDCt}$ |
|---|---|
| BC14 | 2E-07 |
| BC21 | 0.073 |
| BC16 | 0.081 |
| BC17 | 0.135 |
| BC12 | 0.157 |
| BC19 | 0.94 |
| BC20 | 8.694 |
| BC22 | 23.1 |
| BC13 | 4482 |
| BC33 | 1E+06 |
| LC2 | 0.712 |
| LC1 | 1.266 |
| LC3 | 15181 |
| LC4 | 5E+05 |
| RCC10 | 4E-07 |
| RCC12 | 6E-06 |
| RCC11 | 9E-06 |
| RCC6 | 0.152 |
| RCC13 | 4.408 |
| RCC2 | 202.3 |

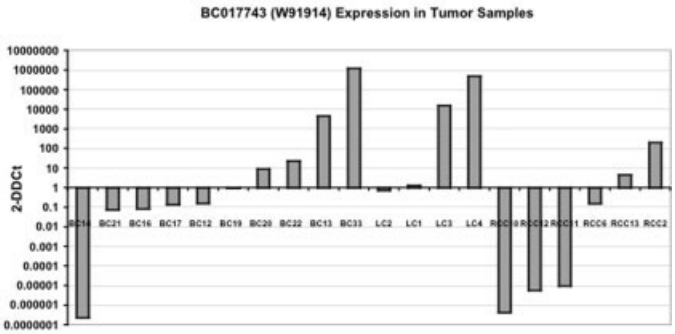



Figure 10. qRT-PCR Expression analysis of *BC017743* (*W91914*) in cancer biopsies and normal related tissues. The *Y* axis indicates the expression value shown in log 10 scale relative to control normal tissues normalized to 1. The *X* axis shows the sample names. The comparative DDCt method was used as described (Senchenko et al., 2003) to analyze the qRT-PCR results.

This suggests that *DUTT1* is a classic TSG requiring inactivation of both alleles to elicit tumorigenesis. However, an extensive mutation analysis of *DUTT1* in lung, breast, and kidney cancers did not reveal inactivating mutations (Dallol et al., 2002).

Here, we report the identification of a new homozygous deletion in the lung cancer cell line GLC20. Located at 3p12.3, it affects the second exon of *DUTT1* and spans about 110–130 kb, perhaps the smallest described thus far in this region. Moreover, using molecular biology and bioinformatics methods, we identified two novel putative genes, *BC017743* and *BC043430*, that reside in the second intron of *DUTT1* and, therefore, are also lost in the deletion. None of the transcripts showed obvious splicing signals; but, in our experiments, each displayed slightly different bands in Northern blot analysis. Both had a polyA tail and small ORFs that are unlikely to encode proteins according to comparative genomic analyses. On the bases of these characteristics, we conclude that they represent possibly mRNA-like ncRNAs (Erdmann et al., 2000; Tupy et al., 2005). Transcribed by RNA polymerase II, in absence of protein products, these types of RNAs serve as riboregulators or regulators of expression of related genes (Erdmann et al., 2000; Numata et al., 2003; Tupy et al., 2005).

Based on their sequence, *BC017743* and *BC043430* seemingly lack homologues or orthologues, a feature that has been reported for other known genes and ncRNAs (Conrad et al., 2002; Weber, 2005). Also, some biologically important ncRNA families show that, inside each group, conservation of secondary structure has a higher significance than conservation of primary sequence (Weinberg and Ruzzo, 2004). *BC017743* and *BC043430* represent the first ncRNA transcripts/genes found in a homozygous cancer deletion affecting the 3p12.3 region.

To date, very few cases of independent transcription units embedded inside introns have been described. The genes *EV12A* and *EV12B* are encoded by one intron of the human *neurofibromatosis type I* (*NF1*) gene and are transcribed in the opposite direction. However, their products are functionally unrelated to *NF1*.

The *Ach transporter* gene, transcribed from the first intron of the rat *ChAT* gene (Bejanin et al., 1994) and the *Saitohin* gene, transcribed from intron 9 of the human *tau* gene (Conrad et al., 2002) are different, as both are transcribed in the same orientation and are functionally related to the longer gene in which they are harbored. The number of such cases is likely to increase given recent reports showing that 233,303 clusters of ESTs are totally contained within intronic regions (Reis et al., 2005).

Both *BC017743* and *BC043430* are transcribed in the same direction as *DUTT1*, suggesting that their product could be co-regulated and possibly related to *DUTT1* function.

It will be interesting to verify whether both transcripts are actually under post-transcriptional control by the miRNAs whose putative target sites were predicted in their sequence. Since the loss of *BC017743* and *BC043430* was found associated with a tumor phenotype, one might argue that both transcripts exert a tumor suppressor action, possibly mediated by the miRNA action. In fact, several cases are now known (McManus, 2003; Xu et al., 2004; Calin et al., 2005), where changes in the expression level of miRNAs may affect the control of cell growth or survival, and be, therefore, involved with cancer onset or progression.

Noteworthy, the miRNA *hsa-mir-17-5p* at 13q31.3 (http://www.ensembl.org/Homo_sapiens/contigview?highlight = &chr = 13&vc_start = 90700860&vc_end = 90900943) is contained within a region frequently lost in breast cancer progression (Eiriksdottir et al., 1998), leading one to speculate that LOH at a miRNA locus (e.g. *hsa-mir-17-5p*) might lead to the same effect as LOH at the miRNA target site (conceivably *BC017743* and *BC043430* in the case of lung, breast, or kidney cancer). Moreover, *hsa-mir-17-5p* was shown to be down-regulated following differentiation induced by TPA treatment of HL-60 promyelocytic leukemia cells (Kasashima et al., 2004). This suggests that miRNA-induced control of cell differentiation might consist of two distinct mechanisms: (i) a mechanism of gene silencing through up-regulation of miRNAs and (ii) a mechanism of gene activation through termination of miRNA-regulated gene silencing. Similarly, it would be interesting to investigate whether LOH at 11q23 in lung cancer involves not only loss of the *TSLC1* gene that resides there (Kuramochi et al., 2001), but also loss of the locus encoding *hsa-mir-34b*, which putatively regulates *BC043430* (as well as other targets in the genome probably; it was shown, for example, that reduced levels of *hsa-mir-15a* and *hsa-mir-16* are a trait shared by different forms of lymphomas and leukemias; Calin et al., 2002; Eis et al., 2005).

A situation diametrically opposed to that described in the GLC20-3p12 deletion is given by the accumulation of the ncRNA *BIC* (Eis et al., 2005). The *BIC* transcript is polyadenylated and has short putative ORFs that are not conserved. Most likely, *BIC* does not encode a protein. However, a phylogenetically conserved region of *BIC* was indeed shown to encode a miRNA, mir-155, whose accumulation strongly correlates with an aggressive B cell neoplasm (Eis et al., 2005).

The GLC20-3p12 deletion might also harbor miRNA loci. Recently, a list of computationally identified human miRNA genes was reported, some of which are encoded at 3p12 (Berezikov et al., 2005). One of them (cand893 HS3, 78.768.573-78.768.661 R, whose closest experimentally identified miRNA is *M. musculus mmu-mir-297*) is harbored inside the deletion (Houbaviy et al., 2003). Further studies are however necessary to investigate whether *DUTT1* itself is a target for this miRNA.

We were also able to analyze the degree of conservation of the noncoding sequence around exon 2 of *DUTT1*. This region contains at least three blocks of extremely conserved, nonrepetitive sequence, spanning 353, 69, and 243 bp respectively, with an identity ranging from 81% to 86% between the human and chicken genome.

In conclusion, it is interesting to note that the two smallest, partly overlapping homozygous deletions described at 3p12.3 (that is, H2198 (Sundaresan et al., 1998a), and the one described here) remove, besides one exon of *DUTT1*, two intronic ncRNA genes. Their presence, along with several short noncoding regions highly conserved between human and chicken genomes (Bejerano et al., 2004; Hillier et al., 2004), suggests that the effect of the 3p12 deletion might be complex.

The expression analysis of *BC017743* (*W91914*) showed that, while it is under-expressed in most breast and kidney tumors analyzed, it is also over-expressed in a variety of SCLC and NSCLC samples (represented by tumor biopsies and cell lines). Because of technical impediments, the expression of *BC043430* (*H51703*) could only be analyzed in lung cancer cell lines only, where it was found to be over-expressed.

Several hypotheses can be suggested regarding their possible role in initiating or promoting cancer growth. Because of their location in a classical tumor suppressor region, it is possible that they represent true candidate TSGs. This hypothesis is supported by their under-expression in kidney and breast carcinomas. Epigenetic inactivation of *DUTT1* due to hypermethylation is a frequent event in breast carcinomas and clear-cell renal cell carcinomas, but rare in NSCLC and not detectable in SCLC (Dallol et al., 2002). Eventhough the ncRNA genes do not have CpG promoters, they are transcribed in the same direction as *DUTT1*. This suggests that the local hypermethylation that affects the *DUTT1* promoter in some cancers (Dallol et al., 2002) might concomitantly be responsible for their decreased expression, whether or not they are functionally correlated to *DUTT1*.

However, both genes were also found to be consistently over-expressed in lung cancer specimens of different histology. Because of their genomic location with respect to *DUTT1*, it is possible that *BC017743* or *BC043430* modulate *DUTT1* function. It has been reported that other ncRNA genes modulate the gene in whose intron they are located (Bejanin et al., 1994; Conrad et al., 2002). If their function is to down-regulate *DUTT1*, their over-expression (caused by a mutation elsewhere in the genome) would have an oncogenic potential. In fact, at least for the GLC20 cell line, it is important to note that another element of complexity is the presence of a second homozygous deletion present at 3p21.3 (Kok et al., 1994). Indeed, with respect to

the short arm of chromosome 3, it has been suggested for several types of tumor, that multiple deletions, coexisting but with different 3p locations, may have a synergistic effect in driving tumorigenesis (Van den Berg et al., 1997; Senchenko et al., 2003). Finally, it cannot be ruled out that *BC017743* and *BC043430* might not have any causative or promoting role in cancer. However, it remains striking how their expression consistently changes in kidney and lung cancers even though in opposite directions.

Further mutation and functional studies of the two ncRNA genes, *BC043430* and *BC017743*, will be critical for understanding their possible role in initiating or promoting cancer development.

## ACKNOWLEDGMENTS

We are grateful to Laura Geil for editing the manuscript and to Andrei Kouranov (Department of Genetics, School of Medicine, University of Pennsylvania, Philadelphia, USA) for his kind help with interpreting the DIANA algorithm results.

Special thanks to each and every LIB member.

## REFERENCES

Bejanin S, Cervini R, Mallet J, Berrard S. 1994. A unique gene organization for two cholinergic markers, choline acetyltransferase and a putative vesicular transporter of acetylcholine. J Biol Chem 269:21944–21947.

Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, Haussler D. 2004. Ultraconserved elements in the human genome. Science 304:1321–1325.

Berezikov E, Guryev V, van de Belt J, Wienholds E, Plasterk RH, Cuppen E. 2005. Phylogenetic shadowing and computational identification of human microRNA genes. Cell 120:21–24.

Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, Aldler H, Rattan S, Keating M, Rai K, Rassenti L, Kipps T, Negrini M, Bullrich F, Croce CM. 2002. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci USA 99:15524–15529.

Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE, Iorio MV, Visone R, Sever NI, Fabbri M, Iuliano R, Palumbo T, Pichiorri F, Roldo C, Garzon R, Sevignani C, Rassenti L, Alder H, Volinia S, Liu CG, Kipps TJ, Negrini M, Croce CM. 2005. A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 353:1793–1801.

Conrad C, Vianna C, Freeman M, Davies P. 2002. A polymorphic gene nested within an intron of the tau gene: Implications for Alzheimer's disease. Proc Natl Acad Sci USA 99:7751–7756.

Dallol A, Forgacs E, Martinez A, Sekido Y, Walker R, Kishida T, Rabbitts P, Maher ER, Minna JD, Latif F. 2002. Tumour specific promoter region methylation of the human homologue of the drosophila roundabout gene *DUTT1* (ROBO1) in human cancers. Oncogene 21:3020–3028.

Daly MC, Douglas JB, Bleehen NM, Hastleton P, Twentyman PR, Sundaresan V, Carritt B, Bergh J, Rabbitts PH. 1991. An unusually proximal deletion on the short arm of chromosome 3 in a patient with small cell lung cancer. Genomics 9:113–119.

de Leij L, Postmus PE, Buys CH, Elema JD, Ramaekers F, Poppema S, Brouwer M, van der Veen AY, Mesander G, The TH. 1985. Characterization of three new variant type cell lines derived from small cell carcinoma of the lung. Cancer Res 45:6024–6033.

Dostie J, Mourelatos Z, Yang M, Sharma A, Dreyfuss G. 2003. Numerous microRNPs in neuronal cells containing novel microRNAs. RNA 9:180–186.

Drabkin HA, Mendez MJ, Rabbitts PH, Varkony T, Bergh J, Schlessinger J, Erickson P, Gemmill RM. 1992. Characterization of the submicroscopic deletion in the small-cell lung carcinoma (SCLC) cell line U2020. Genes Chromosomes Cancer 5:67–74.

Driesen MS, Dauwerse JG, Wapenaar MG, Meerhoek EJ, Mollevanger P, Chen KL, Fishbeck KH, van Ommen GJ. 1991. Generation and fluorescent in situ hybridisation mapping of yeast artificial chromosomes of 1p, 17p, and 19q from a hybrid cell line by high density screening of an amplified library. Genomics 11: 1079–1087.

Eble JN, Sauter GS, Epsten JI, Sisterhenn IA, editors. 2004. World Health Organization: Classification of Tumors. Pathology and Genetics: Tumors of the urinary system and male genital organs. Lyon: IARC Press.

Eiriksdottir G, Johannesdottir G, Ingvarsson S, Bjornsdottir IB, Jonasson JG, Agnarsson BA, Hallgrimsson J, Gudmundsson J, Egilsson V, Sigurdsson H, Barkardottir RB. 1998. Mapping loss of heterozygosity at chromosome 13q: Loss at 13q12-q13 is associated with breast tumour progression and poor prognosis. Eur J Cancer 34:2076–2081.

Eis PS, Tam W, Sun L, Chadburn A, Li Z, Gomez MF, Lund E, Dahlberg JE. 2005. Accumulation of miR-155 and BIC RNA in human B cell lymphomas. Proc Natl Acad Sci USA 102:3627–3632.

Erdmann VA, Szymanski M, Hochberg A, Groot N, Barciszewski J. 2000. Non-coding, mRNA-like RNAs database Y2K. Nucleic Acids Res 28:197–200.

Ganly PS, Jarad N, Rudd RM, Rabbitts PH. 1992. PCR-based RFLP analysis allows genotyping of the short arm of chromosome 3 in small biopsies from patients with lung cancer. Genomics 12:221–228.

Giles RH, Petrij F, Dauwerse HHG, Den Hollander AI, Lushnikova T, van Ommen GJ, Goodman RH, Deaven LL, Doggett NA, Peters DJ, Breuning MH. 1997. Construction of a 1.2-Mb contig surrounding, and molecular analysis of, the human CREB-binding protein (CBP/CREBBP) gene on chromosome 16p13.3. Genomics 42:96–114.

Hillier LW, Miller W, Birney E, Warren W, Hardison RC, Ponting CP, Bork P, Burt DW, Groenen MA, Delany ME, Dodgson JB, Chinwalla AT, Cliften PF, Clifton SW, Delehaunty KD, Fronick C, Fulton RS, Graves TA, Kremitzki C, Layman D, Magrini V, McPherson JD, Miner TL, Minx P, Nash WE, Nhan MN, Nelson JO, Oddy LG, Pohl CS, Randall-Maher J, Smith SM, Wallis JW, Yang SP, Romanov MN, Rondelli CM, Paton B, Smith J, Morrice D, Daniels L, Tempest HG, Robertson L, Masabanda JS, Griffin DK, Vignal A, Fillon V, Jacobbson L, Kerje S, Andersson L, Crooijmans RP, Aerts J, van der Poel JJ, Ellegren H, Caldwell RB, Hubbard SJ, Grafham DV, Kierzek AM, McLaren SR, Overton IM, Arakawa H, Beattie KJ, Bezzubov Y, Boardman PE, Bonfield JK, Croning MD, Davies RM, Francis MD, Humphray SJ, Scott CE, Taylor RG, Tickle C, Brown WR, Rogers J, Buerstedde JM, Wilson SA, Stubbs L, Ovcharenko I, Gordon L, Lucas S, Miller MM, Inoko H, Shiina T, Kaufman J, Salomonsen J, Skjoedt K, Wong GK, Wang J, Liu B, Wang J, Yu J, Yang H, Nefedov M, Koriabine M, Dejong PJ, Goodstadt L, Webber C, Dickens NJ, Letunic I, Suyama M, Torrents D, von Mering C, Zdobnov EM, Makova K, Nekrutenko A, Elnitski L, Eswara P, King DC, Yang S, Tyekucheva S, Radakrishnan A, Harris RS, Chiaromonte F, Taylor J, He J, Rijnkels M, Griffiths-Jones S, Ureta-Vidal A, Hoffman MM, Severin J, Searle SM, Law AS, Speed D, Waddington D, Cheng Z, Tuzun E, Eichler E, Bao Z, Flicek P, Shteynberg DD, Brent MR, Bye JM, Huckle EJ, Chatterji S, Dewey C, Pachter L, Kouranov A, Mourelatos Z, Hatzigeorgiou AG, Paterson AH, Ivarie R, Brandstrom M, Axelsson E, Backstrom N, Berlin S, Webster MT, Pourquie O, Reymond A, Ucla C, Antonarakis SE, Long M, Emerson JJ, Betran E, Dupanloup I, Kaessmann H, Hinrichs AS, Bejerano G, Furey TS, Harte RA, Raney B, Siepel A, Kent WJ, Haussler D, Eyras E, Castelo R, Abril JF, Castellano S, Camara F, Parra G, Guigo R, Bourque G, Tesler G, Pevzner PA, Smit A, Fulton LA, Mardis ER, Wilson RK.(The International Chicken Genome Sequencing Consortium). 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695–716.

Houbaviy HB, Murray MF, Sharp PA. 2003. Embryonic stem cell-specific microRNAs. Dev Cell 5:351–358.

Huebner K, Garrison PN, Barnes LD, Croce CM. 1998. The role of the FHIT/FRA3B locus in cancer. Annu Rev Genet 32:7–31.

Imanishi T, Itoh T, Suzuki Y, O'Donovan C, Fukuchi S, Koyanagi KO, Barrero RA, Tamura T, Yamaguchi-Kabata Y, Tanino M, Yura K, Miyazaki S, Ikeo K, Homma K, Kasprzyk A, Nishikawa T, Hirakawa M, Thierry-Mieg J, Thierry-Mieg D, Ashurst J, Jia L, Nakao M, Thomas MA, Mulder N, Karavidopoulou Y, Jin L, Kim

S, Yasuda T, Lenhard B, Eveno E, Suzuki Y, Yamasaki C, Takeda J, Gough C, Hilton P, Fujii Y, Sakai H, Tanaka S, Amid C, Bellgard M, Bonaldo Mde F, Bono H, Bromberg SK, Brookes AJ, Bruford E, Carninci P, Chelala C, Couillault C, de Souza SJ, Debily MA, Devignes MD, Dubchak I, Endo T, Estreicher A, Eyras E, Fukami-Kobayashi K, Gopinath GR, Graudens E, Hahn Y, Han M, Han ZG, Hanada K, Hanaoka H, Harada E, Hashimoto K, Hinz U, Hirai M, Hishiki T, Hopkinson I, Imbeaud S, Inoko H, Kanapin A, Kaneko Y, Kasukawa T, Kelso J, Kersey P, Kikuno R, Kimura K, Korn B, Kuryshev V, Makalowska I, Makino T, Mano S, Mariage-Samson R, Mashima J, Matsuda H, Mewes HW, Minoshima S, Nagai K, Nagasaki H, Nagata N, Nigam R, Ogasawara O, Ohara O, Ohtsubo M, Okada N, Okido T, Oota S, Ota M, Ota T, Otsuki T, Piatier-Tonneau D, Poustka A, Ren SX, Saitou N, Sakai K, Sakamoto S, Sakate R, Schupp I, Servant F, Sherry S, Shiba R, Shimizu N, Shimoyama M, Simpson AJ, Soares B, Steward C, Suwa M, Suzuki M, Takahashi A, Tamiya G, Tanaka H, Taylor T, Terwilliger JD, Unneberg P, Veeramachaneni V, Watanabe S, Wilming L, Yasuda N, Yoo HS, Stodolsky M, Makalowski W, Go M, Nakai K, Takagi T, Kanehisa M, Sakaki Y, Quackenbush J, Okazaki Y, Hayashizaki Y, Hide W, Chakraborty R, Nishikawa K, Sugawara H, Tateno Y, Chen Z, Oishi M, Tonellato P, Apweiler R, Okubo K, Wagner L, Wiemann S, Strausberg RL, Isogai T, Auffray C, Nomura N, Gojobori T, Sugano S. 2004. Integrative annotation of 21, 037 human genes validated by full-length cDNA clones. PLoS Biol 2:e162.

Kasashima K, Nakamura Y, Kozu T. 2004. Altered expression profiles of microRNAs during TPA-induced differentiation of HL-60 cells. Biochem Biophys Res Commun 322:403–410.

Knudson AG, Jr. 1971. Mutation and cancer: Statistical study of retinoblastoma. Proc Natl Acad Sci USA 68:820–823.

Kok K, Naylor SL, Buys CH. 1997. Deletions of the short arm of chromosome 3 in solid tumors and the search for suppressor genes. Adv Cancer Res 71:27–92.

Kok K, Osinga J, Carritt B, Davis MB, van der Hout AH, van der Veen AY, Landsvater RM, de Leij LFMH, Berendsen HH, Postmus PE, Poppema S, Buys CHCM. 1987. Deletion of a DNA sequence at the chromosomal region 3p21 in all major types of lung cancer. Nature 330:578–581.

Kok K, van den Berg A, Veldhuis PMJF, Franke M, Terpstra P, Buys CHCM. 1995. The genomic structure of the human UBE1L gene. Gene Expr 4:163–175.

Kok K, van den Berg A, Veldhuis PMJF, van der Veen AY, Franke M, Schoenmakers EFPM, Hulsbeek MMF, van der Hout AH, de Leij L, van der Ven W, Buys CHCM. 1994. A homozygous deletion in a small cell lung cancer cell line involving a 3p21 region with a marked instability in yeast artificial chromosomes. Cancer Res 54:4183–4187.

Kuramochi M, Fukuhara H, Nobukuni T, Kanbe T, Maruyama T, Ghosh HP, Pletcher M, Isomura M, Onizuka M, Kitamura T, Sekiya T, Reeves RH, Murakami Y. 2001. TSLC1 is a tumor-suppressor gene in human non-small-cell lung cancer. Nat Genet 27:427–430.

Latif F, Tory K, Modi WS, Graziano SL, Gamble G, Douglas J, Heppell-Parton AC, Rabbitts PH, Zbar B, Lerman MI. 1992. Molecular characterization of a large homozygous deletion in the small cell lung cancer cell line U2020: A strategy for cloning the putative tumor suppressor gene. Genes Chromosomes Cancer 5:119–127.

Lerman MI, Minna JD. 2000. The 630-kb lung cancer homozygous deletion region on human chromosome 3p21.3: Identification and evaluation of the resident candidate tumor suppressor genes. The International Lung Cancer Chromosome 3p21.3 Tumor Suppressor Gene Consortium. Cancer Res 60:6116–6133.

Lott ST, Lovell M, Naylor SL, Killary AM. 1998. Physical and functional mapping of a tumor suppressor locus for renal cell carcinoma within chromosome 3p12. Cancer Res 58:3533–3537.

Lovell M, Lott ST, Wong P, El-Naggar A, Tucker S, Killary AM. 1999. The genetic locus NRC-1 within chromosome 3p12 mediates tumor suppression in renal cell carcinoma independently of histological type, tumor microenvironment, and VHL mutation. Cancer Res 59:2182–2189.

Marshall E. 1991. Breast cancer: Stalemate in the war on cancer. Science 254:1719–1720.

McManus MT. 2003. MicroRNAs and cancer. Semin Cancer Biol 13:253–258.

Mourelatos Z, Dostie J, Paushkin S, Sharma A, Charroux B, Abel L, Rappsilber J, Mann M, Dreyfuss G. 2002. miRNPs: A novel class of ribonucleoproteins containing numerous microRNAs. Genes Dev 16:720–728.

Numata K, Kanai A, Saito R, Kondo S, Adachi J, Wilming LG, Hume DA, Hayashizaki Y, Tomita M;RIKEN GER Group;GSL Members. 2003. Identification of putative noncoding RNAs among the RIKEN mouse full-length cDNA collection. Genome Res 13:1301–1306.

Pandis N, Jin Y, Limon J, Bardi G, Idvall I, Mandahl N, Mitelman F, Heim S. 1993. Interstitial deletion of the short arm of chromosome 3 as a primary chromosome abnormality in carcinomas of the breast. Genes Chromosomes Cancer 6:151–155.

Phelps RM, Johnson BE, Ihde DC, Gazdar AF, Carbone DP, McClintock PR, Linnoila RI, Matthews MJ, Bunn PA, Jr., Carney D, Minna JD, Mulshine JL. 1996. NCI-Navy Medical Oncology Branch cell line data base. J Cell Biochem Suppl 24:32–91.

Rabbitts P, Bergh J, Douglas J, Collins F, Waters J. 1990. A submicroscopic homozygous deletion at the D3S3 locus in a cell line isolated from a small cell lung carcinoma. Genes Chromosomes Cancer 2:231–238.

Reis EM, Louro R, Nakaya HI, Verjovski-Almeida S. 2005. As antisense RNA gets intronic. OMICS 9:2–12.

Sambrook J, Fritsch EF, Maniatis T. 1998. Molecular Cloning, a Laboratory Manual. New York: Cold Spring Harbor Laboratory Press.

Sanchez Y, El-Naggar A, Pathak S, Killary AM. 1994. A tumor suppressor locus within 3p14-p12 mediates rapid cell death of renal cell carcinoma in vivo. Proc Natl Acad Sci USA 91:3383–3387.

Senchenko V, Liu J, Braga E, Mazurenko N, Loginov W, Seryogin Y, Bazov I, Protopopov A, Kisseljov FL, Kashuba V, Lerman MI, Klein G, Zabarovsky ER. 2003. Deletion mapping using quantitative real-time PCR identifies two distinct 3p21.3 regions affected in most cervical carcinomas. Oncogene 22:2984–2992.

Sobin LY, Wittekind CH, editors. 2002. UICC TNM Classification of Malignant Tumors,6th ed. New York: Wiley-Liss.

Suh MR, Lee Y, Kim JY, Kim SK, Moon SH, Lee JY, Cha KY, Chung HM, Yoon HS, Moon SY, Kim VN, Kim KS. 2004. Human embryonic stem cells express a unique set of microRNAs. Dev Biol 270:488–498.

Sulimova GE, Kutsenko AS, Rakhmanaliev ER, Udina IG, Kompaniytsev AA, Protopopov AI, Moisjak EV, Klimov EA, Muravenko OV, Zelenin AV, Braga EA, Kashuba VI, Zabarovsky ER, Kisselev LL. 2002. Human chromosome 3: Integration of 60 NotI clones into a physical and gene map. Cytogenet Genome Res 98:177–183.

Sundaresan V, Chung G, Heppell-Parton A, Xiong J, Grundy C, Roberts I, James L, Cahn A, Bench A, Douglas J, Minna J, Sekido Y, Lerman M, Latif F, Bergh J, Li H, Lowe N, Ogilvie D, Rabbitts P. 1998a. Homozygous deletions at 3p12 in breast and lung cancer. Oncogene 17:1723–1729.

Sundaresan V, Heppell-Parton A, Coleman N, Miozzo M, Sozzi G, Ball R, Cary N, Hasleton P, Fowler W, Rabbitts P. 1995. Somatic genetic changes in lung cancer and precancerous lesions. Ann Oncol 6 (suppl 1):27–31; discussion 31–2.

Sundaresan V, Roberts I, Bateman A, Bankier A, Sheppard M, Hobbs C, Xiong J, Minna J, Latif F, Lerman M, Rabbitts P. 1998b. The *DUTT1* gene, a novel NCAM family member is expressed in developing murine neural tissues and has an unusually broad pattern of expression. Mol Cell Neurosci 11:29–35.

Tavassoli FA, Devilee P, editors. 2003. World Health Organization: Classification of Tumors. Pathology and Genetics of Tumors of the Breast and Female Genital Organs. Lyon: IARC Press.

Todd S, Franklin WA, Varella-Garcia M, Kennedy T, Hilliker CE, Jr., Hahner L, Anderson M, Wiest JS, Drabkin HA, Gemmill RM. 1997. Homozygous deletions of human chromosome 3p in lung tumors. Cancer Res 57:1344–1352.

Travis WD, Colby TV, Corrin B, Shimosato Y, Brambilla E, editors. 1999. Histological Typing of Lung and Pleural Tumors,3rd ed. Berlin: Springer-Verlag.

Tupy JL, Bailey AM, Dailey G, Evans-Holm M, Siebel CW, Misra S, Celniker SE, Rubin GM. 2005. Identification of putative noncoding polyadenylated transcripts in *Drosophila melanogaster*. Proc Natl Acad Sci USA 102:5495–5500.

Van den Berg A, Dijkhuizen T, Draaijers TG, Hulsbeek MMF, Maher ER, Van den Berg E, Störkel S, Buys CHCM. 1997. Analysis of multiple renal cell adenomas and carcinomas suggests allelic loss at 3p21 to be a prerequisite for malignant development. Genes Chromosomes Cancer 19:228–232.

Vieten L, Belair CD, Savelieva L, Julicher K, Brocker F, Bardenheuer W, Schutte J, Opalka B, Reznikoff CA. 1998. Minimal deletion of 3p13→14.2 associated with immortalization of human uroepithelial cells. Genes Chromosomes Cancer 21:39–48.

Weber MJ. 2005. New human and mouse microRNA genes found by homology search. FEBS J 272:59–73.

Weinberg Z, Ruzzo WL. 2004. Exploiting conserved structure for faster annotation of non-coding RNAs without loss of accuracy. Bioinformatics 20 (suppl 1):I334–I341.

Whang-Peng J, Kao-Shan CS, Lee EC, Bunn PA, Carney DN, Gazdar AF, Minna JD. 1982. Specific chromosome defect associated with human small-cell lung cancer; deletion 3p(14–23). Science 215:181–182.

Xian J, Aitchison A, Bobrow L, Corbett G, Pannell R, Rabbitts T, Rabbitts P. 2004. Targeted disruption of the 3p12 gene, DUTT1/Robo1, predisposes mice to lung adenocarcinomas and lymphomas with methylation of the gene promoter. Cancer Res 64:6432–6437.

Xian J, Clark KJ, Fordham R, Pannell R, Rabbitts TH, Rabbitts PH. 2001. Inadequate lung development and bronchial hyperplasia in mice with a targeted deletion in the *DUTT1*/Robo1 gene. Proc Natl Acad Sci USA 98:15062–15066.

Xu P, Guo M, Hay BA. 2004. MicroRNAs and the regulation of cell death. Trends Genet 20:617–624.

Zabarovsky ER, Lerman MI, Minna JD. 2002. Tumor suppressor genes on chromosome 3p involved in the pathogenesis of lung and other cancers. Oncogene 21:6915–6935.

Zbar B. 1989. Chromosomal deletions in lung cancer and renal cancer, In: DeVita V, Hellmann S, Rosenberg S, editors. Important advances in oncology. Philadelphia: Lippincott, p. 41–60.