SPANDIDOS PUBLICATIONS

Gene expression profiling reveals molecular marker candidates of laryngeal squamous cell carcinoma

JUCIMARA COLOMBO[1], ÂNGELA A. FACHEL[2], MARILIA DE FREITAS CALMON[1], PATRÍCIA MALUF CURY[3], ÉRICA ERINA FUKUYAMA[4], ELOIZA HELENA TAJARA[5], JOSÉ ANTÔNIO CORDEIRO[6], SERGIO VERJOVSKI-ALMEIDA[2], EDUARDO M. REIS[2*] and PAULA RAHAL[1*]

[1]UNESP - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas - IBILCE, Departamento de Biologia, São José do Ri o Preto; [2]USP - Universidade de São Paulo, Instituto de Química, Departamento de Bioquímica, São Paulo; [3]FAMERP - Faculdade de Medicina de São José do Rio Preto, Departamento de Patologia, São José do Rio Preto; [4]Hospital do Câncer Arnaldo Vieira de Carvalho, São Paulo; [5]FAMERP - Faculdade de Medicina de São José do Rio Preto, Departamento de Biologia Molecular; [6]FAMERP - Faculdade de Medicina de São José do Rio Preto, Departamento de Epidemiologia, São José do Rio Preto, São Paulo, Brazil

Received August 14, 2008; Accepted October 6, 2008

DOI: 10.3892/or_00000268

Abstract. Laryngeal squamous cell carcinoma is very common in head and neck cancer, with high mortality rates and poor prognosis. In this study, we compared expression profiles of clinical samples from 13 larynx tumors and 10 non-neoplastic larynx tissues using a custom-built cDNA microarray containing 331 probes for 284 genes previously identified by informatics analysis of EST databases as markers of head and neck tumors. Thirty-five genes showed statistically significant differences (SNR ≥ | 1.0 |, p≤0.001) in the expression between tumor and non-tumor larynx tissue samples. Functional annotation indicated that these genes are involved in cellular processes relevant to the cancer phenotype, such as apoptosis, cell cycle, DNA repair, proteolysis, protease inhibition, signal transduction and transcriptional regulation. Six of the identified transcripts map to intronic regions of protein-coding genes and may comprise non-annotated exons or as yet uncharacterized long ncRNAs with a regulatory role in the gene expression program of larynx tissue. The differential expression of 10 of these genes (*ADCY6*, *AES*, *AL2SCR3*, *CRR9*, *CSTB*, *DUSP1*, *MAP3K5*, *PLAT*, *UBL1* and *ZNF706*) was independently confirmed by quantitative real-time RT-PCR. Among these, the *CSTB* gene product has cysteine protease inhibitor activity that has been associated with an antimetastatic function. Interestingly, *CSTB* showed a low expression in the tumor samples analyzed (p<0.0001). The set of genes identified here contribute to a better understanding of the molecular basis of larynx cancer, and provide candidate markers for improving diagnosis, prognosis and treatment of this carcinoma.

Correspondence to: Dr Eduardo M. Reis, Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, Brazil
E-mail: emreis@iq.usp.br

*Contributed equally

Abbreviations: HNSCC, head and neck squamous cell carcinoma; ORESTES, open reading frame expressed sequence tags; qRT-PCR, quantitative real-time RT-PCR; CSTB, cystatin B

Key words: cDNA microarray, larynx carcinoma, gene expression, real-time RT-PCR, cystatin B

Introduction

Head and neck squamous cell carcinoma (HNSCC) is the fifth most common cancer worldwide, with a global annual incidence of 780,000 new cases (1). More than 90% of these cancer types have a squamous origin, and common sites include the hypopharynx, larynx, oral cavity, nasopharynx, oropharynx, paranasal sinus, nasal cavity, parathyroid and salivary glands (2,3). Laryngeal squamous cell carcinoma is a very common type of head and neck cancer, corresponding to ~25% of HNSCC cases (4). Tobacco and/or alcohol consumption are the two main risk factors involved in the development of HNSCC. Other risk factors include poor oral hygiene, nutritional deficiencies and certain viruses, such as human papilloma (HPV) and Epstein-Barr (2,5,6).

Despite advances in treatment, the long-term survival rate of patients with head and neck squamous cell carcinoma has remained at 50%, with high rates of associated mortality (5). Late presentation of the lesion, lack of suitable markers for early detection and failure of advanced lesions to respond to the available chemotherapy contribute to a poor outcome of

HNSCC (7). However, loco-regional relapse and metastasis after conventional therapy appear to be significant contributing factors for restricted survival of HNSCC patients (1).

The development of HNSCC is a multistep process accompanied by genetic and epigenetic changes, including loss of heterozygosity (LOH), gene inactivation by methylation and gene amplification, which can alter gene expression (8). Various studies have revealed numerous molecular abnormalities in HNSCC, including activation of oncogenes such as *CCND1* (9), *EGFR* (10), *RAS* (11) and *C-MYC* (12); inactivation of tumor suppressor genes such as *CDKN2A (p16)* (13), *TP53* (14), *p27* (15) and *WAF1/C1P1* (16); expression of angiogenic factors and LOH at numerous chromosomal locations (17,18). For instance, the most frequent cytogenetic alterations in HNSCC are gains of 3q, 8q, 20q, 7q, 11q13 and 5p, and losses in 3p, 9p, 21q, 5q, 13q, 18q and 8p (19). The involved loci correspond to genes that encode specific functional classes of proteins, such as cell cycle regulators, tumor suppressors, cell adhesion and protein kinases (20). However, the complete array of molecular changes that occur during oncogenesis and HNSCC progression remains elusive.

Large-scale studies involving microarrays have identified specific gene expression signatures associated with expression changes in HNSCC tissue samples compared to normal tissue, as well as genes involved in clinical outcome and metastasis (19,21,22). However, most microarray studies were performed using tumors from different sites in head and neck.

In the present study, we constructed a cDNA microarray platform containing probes for 284 genes previously identified by the Head and Neck Annotation Consortium (23), and used these arrays to search for genes differentially expressed in clinical samples from squamous cell carcinoma of the larynx. The genes included in the array are candidates for LOH or gene amplification in head and neck tumors identified following an informatics analysis of EST libraries derived from non-tumor and tumor tissues of the oral cavity, larynx, pharynx and thyroid (23). However, their differential expression in malignant and non-tumor head and neck tissues has yet to be evaluated. By analyzing only larynx carcinoma samples, we aimed to minimize the genetic differences that may be present in anatomically diverse tumors of the head and neck region, thus favoring the identification of genes associated with tumors with the same histological origin.

We observed a significant increase or decrease in the mRNA levels of 35 genes in larynx tumor samples. These genes are involved in processes such as cell cycle, apoptosis, DNA repair, protease inhibition, proteolysis and transcriptional regulation. Our findings contribute to the understanding of the molecular basis of larynx cancer and provide a set of genes that may be useful for the development of novel diagnostic markers and/or more effective therapeutic strategies.

Materials and methods

cDNA microarray construction. We selected 135 genes previously identified by the Head and Neck Annotation Consortium (23) to be spotted into a custom-built cDNA microarray. These included genes that map to loss or gain regions involved in head and neck carcinogenesis, according to Knuutila *et al* (24).

Complementary DNA of the human laryngeal carcinoma (Hep2) cell line served as a template for the amplification of probes from microarray spotting. Gene-specific primers for PCR were designed using the Primer 3.0 program (25) to amplify fragments with an average size of 600 bp. In brief, for the PCR, 1 μl cDNA was mixed with 2.5 μl 10X buffer, 1 μl PCR primer (10 pmol/μl), 4 μl dNTPs (1.25 mM), 0.75 μl $MgCl_2$ (50 mM) and 0.2 μl Taq DNA platinum (Invitrogen). PCR conditions were: 94˚C for 4 min, followed by 35 cycles at 94˚C for 30 sec, 60˚C for 1 min and 72˚C for 1 min.

From the products obtained, 5 μl were fractionated by electrophoresis through a 1% agarose gel for size verification, and the remaining 95 μl were purified by filtration on Multiscreen plates (Millipore cat. # MAFB NOB 50). cDNA samples were diluted 1:1 in DMSO and spotted onto silane-coated, reflective Type-7 glass slides (Molecular Dynamics), using a GenIII microarray spotter robot (Molecular Dynamics). Twelve replicates of 4 different plant/bacterial DNA fragments with no similarity to human sequences were evenly distributed along the slide to serve as negative controls of non-specific hybridization (Lucidea ScoreCard, GE Healthcare). Following spotting, cDNAs were fixed to the slide surface by UV cross-linking (500 mJ) and kept in a low-humidity environment until use.

Based on an informatics analysis of the head and neck (HN) transcriptome to identify candidates for LOH in HN tumors (23), an additional set of 196 cDNA clones were included in the array, comprising partial sequences of protein-coding transcripts (123 sequences), as well as of transcripts mapping to intronic (53 sequences) or intergenic regions of the genome (20 sequences). The original sequencing clone collection was stored frozen in bar-coded 96-well plates, and the selected clones were re-arrayed in an automated robotic operation. cDNA fragments were generated from these clones by PCR amplification with universal primers, purified by filtration and spotted as described above.

Each sequence was spotted in 6-replicates in the microarrays. The complete list of the 331 cDNA clones deposited in the head and neck cancer cDNA platform are deposited in the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) under the accession number GPL6426.

Patient samples. This study was approved by the Brazilian National Research Ethics Committee and written informed consent was obtained from the participating individuals prior to sample collection. A total of 15 larynx tumor samples and 10 non-matched, histologically normal adjacent larynx tissue samples were obtained at the time of surgery from the Arnaldo Vieira de Carvalho Hospital. They were immediately snap-frozen in liquid nitrogen.

Medical records were examined to obtain clinical and histopathological information for each patient, including age, gender, histopathological diagnosis, as well as history of tobacco and alcohol consumption. Tumors were staged according to the current TNM classification, as recommended by the UICC (26). Patient clinical data and tumor characteristics are described in Table I.

atient clinical and anatomopathological data.

Case	Gender[a]	Age (years)	Ethnicity[b]	Tobacco[c]	Alcohol[c]	Anatomic location	TNM category and clinical stage	NED or recurrence (months)[d]
T1	M	73	C	P	P	Pyriform sinus and hemilarynx	T4N2bM0 IV	Loc/Rec (9)
T2	M	66	C	P	P	Glottis and subglottis	T4N0M0 IV	Loc/Rec (10)
T3	M	52	C	P	P	Glottis	T3N0M0 III	NED (26)
T4	F	58	C	P	P	Larynx	T4N0M0 IV	Loc/Rec (7)
T5	M	44	C	P	P	Pyriform sinus	T4N2bM0 IV	NED (40)
T6	M	48	N	P	P	Transglottis	T4N2cM0 IV	Loc/Rec (14)
T7	M	53	C	P	P	Transglottis	T2N2M0 III	Loc/Rec (5)
T8	M	58	C	P	P	Transglottis	T4N0M0 IV	NED (30)
T9	M	72	N	P	P	Transglottis	T3N0M0 III	NED (38)
T10	M	72	C	P	P	Vocal cords	T4N2cM0 IV	Loc/Rec (10)
T11	M	44	C	P	N	Vallecula, epiglottis	T4N2cM0 IV	NED (42)
T12	M	60	C	P	P	Pyriform sinus	T4N2cM0 IV	Loc/Rec (7)
T13	M	52	C	P	P	Vallecula	T4N1M0 IV	Loc/Rec (4)
T14	M	67	C	P	P	Vallecula	T4N1M0 IV	Loc/Rec (5)
T15	M	61	C	P	P	Epiglottis, vocal cords	T4N0M0 IV	NED (30)

[a]M, male; F, female; [b]C, caucasian; N, negroid; [c]P, positive; N, negative and [d]NED, no evidence of disease; Loc/Rec, loco-regional recurrence.

Prior to RNA extraction, samples were re-examined, and Giemsa-stained microsections obtained from each side of the frozen block were used to delimit the spatial distribution of the tumor mass or non-tumor tissue. Microdissections were performed to ensure that >70% of the isolated RNA was derived from cancer cells. In the case of surgical margins, only the epithelial tissue was used. Microdissected tumor and non-tumor samples were returned to liquid nitrogen until use.

RNA extraction. Snap-frozen tissue samples were pulverized using a mortar pestle. Total RNA was isolated from tissue specimens and Hep2 cell culture using TRIzol reagent (Gibco BRL, Life Technologies). Total RNA was precipitated by incubating with 0.5 ml of isopropyl alcohol for 10 min, followed by centrifugation at 12,000 x g for 10 min at 4˚C. The pellet was washed with 75% ethanol, solved in RNAse-free water, passed through an RNeasy spin column (Qiagen) for purification and stored at -80˚C until further use.

RNA quantity and quality were evaluated using a spectrophotometer, 1% agarose gel electrophoresis and microelectrophoresis on a Bioanalyzer 2100 (Agilent Technologies). RNA from the samples was of appropriately high quality for cDNA microarray analysis.

Probe labeling and hybridization. The experimental design employed a one-color approach to obtain gene expression measurements (27). Labeled targets for hybridizations were generated from total mRNA in reverse transcription reactions using oligo-dT primers, following the protocol accompanying the CyScribe First-Strand labeling kit strictly (Amersham Biosciences, Piscataway, NJ). Total RNA (15 μg) from each sample was mixed with 4 μl anchored oligo (dT) (500 μg/ml), in a total volume of 11 μl, denatured at 70˚C for 5 min, put on ice for 30 sec, spun down and placed at room temperature for 10 min. Then, 1 μl of dNTP mix, 1 μl of aminoallyl-dUTP, 2 μl of 0.1 M DTT, 1 μl of RNaseOUT (40 U/μl), 4 μl of 5X CyScript buffer and 1 μl of CyScript reverse transcriptase (200 U/μl) were added. The volume was adjusted to 20 μl with water. After incubation for 3 h at 42˚C, RNA was hydrolyzed by adding 2 μl of 2.5 N NaOH for 15 min at 37˚C. Samples were then neutralized with 10 μl of 2 M HEPES-free acid, and reactions were purified using 96-well Millipore multiscreen filter plates as follows: 5 volumes of 5.3 M guanidine-HCl and 150 mM KOAc were added to labeling reactions. The mixture was applied onto the plate and washed 4 times with 80% EtOH by centrifugation at 3500 rpm for 5 min. Residual ethanol was spun out by an additional centrifugation at 3500 rpm for 5 min. Labeled targets were eluted in 50 μl 10 mM Tris pH 8.5, by spinning at 3000 rpm for 5 min, dried on a SpeedVac and kept at -20˚C, protected from light until use. cDNA was re-suspended in 40 μl of 0.1 M $NaHCO_3$ and reacted with monoreactive NHS-ester Cy5 dye. The reaction was incubated for 2 h at room temperature and purified.

Labeled targets were re-suspended in 250 μl of 1X hybridization buffer [25% formamide, 12.5% of proprietary Microarray Hybridization Buffer Version 2 (Amersham Biosciences)], denatured for 2 min at 92˚C and centrifuged at 13,000 rpm for 5 min. The Cy5- (tumor or non-tumor sample) labeled cDNAs were incubated individually with microarrays using an Automated Slide Processor (ASP, GE Healthcare) for 16 h at 42˚C. Following hybridization, slides were washed (1.0X SSC, 0.2% SDS for 10 min at 55˚C; 0.1X SSC, 0.2% SDS for 10 min at 55˚C; 0.1X SSC, 0.2% SDS for 10 min at 55˚C; 0.1X SSC for 1 min at RT; 0.1X SSC for 1 min at RT and dH_2O for 10 sec at RT), and dried with a N2 stream.

Data measurement and normalization. Processed slides were scanned with a 700 V PMT setting (GenIII scanner, Amersham Biosciences), and background-subtracted artifact-removed median intensities of Cy5 emissions were extracted for each spot from raw images, using the ArrayVision V.7.2 software (Imaging Research Inc., Ontario, Canada). An array grid was automatically aligned to locate the position of each spot in the array, and then manually adjusted to obtain the best possible alignment.

To make the experiments comparable, intensity data from each hybridized test sample was normalized by Local weighted scatter-plot smoothing LOWESS (28). Intensity data from a sample with total energy comparable to the average intensity of the samples was used as a reference in the LOWESS normalization procedure. Raw and normalized microarray intensities are deposited in the GEO database (http://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE10288.

Statistical analysis of microarray data. To identify genes with a significantly altered expression in larynx tumors, a signal-to-noise ratio (SNR) metric (29) was used to compare the expression intensity data from tumor and non-tumor samples. The SNR parameter is essentially a measure of signal strength relative to background noise. The distance between the two groups was measured by a signal (expression intensity) to noise (variation) ratio. The signal-to-noise comparison gives an indication of the level of separation for the means of the two distributions defining the gene intensities of the two groups, and was calculated as:

$$SNR = \sqrt{2(\bar{X}_1 - \bar{X}_2)/(S_1 + S_2)}$$

where $\bar{X}_1$ and $\bar{X}_2$, respectively, are the mean intensities of tumor and non-tumor groups, and S_1 and S_2, respectively, the corresponding standard deviations. For each gene, higher absolute SNR values indicate a greater difference of expression between tumor and non-tumor samples, with a lower dispersion within each group. A cut-off SNR $\geq$ | 1 |, $p \leq 0.001$ was used to select differentially expressed genes. Statistical significance of the differential expression (p-values) was ascertained by bootstrap re-sampling, i.e., by re-calculating SNR values following 10,000 random permutations of sample labels and computing the frequency at which each SNR value measured in the original set was observed in the randomly permuted data (29). Expression profiles of selected transcripts were grouped using hierarchical clustering (UPGMA with Euclidean distance) and visualized using the computer software Spotfire Decision Site (Spotfire, Somerville, MA, USA).

Real-time RT-PCR experiments and statistical analysis. For an independent validation of the array data, we selected 14 genes which were down-regulated in tumor tissues as measured by microarray analysis. The cDNA sequence of each gene was selected from the internationally published databases (http://www.ncbi.nlm.ni.gov/) and gene-specific primers for real-time RT-PCR were designed for optimal hybridization kinetics, using the Primer 3.0 program (25). Amplification primers were designed into different exons, to avoid the amplification of any contaminating genomic DNA. Primer sequences are shown in Table III.

For validation, we used total RNA from 15 tumor samples (T1-T15) and a pool of RNA from 7 histologically normal larynx mucosa tissue samples adjacent to tumors, considered here as a normal reference. Double-stranded cDNA was synthesized from total RNA, using high-capacity cDNA archive (Applied Biosystems) according to the manufacturer's instructions.

Real-time PCR was performed using the ABI prism 7300 sequencer detector system and SYBR-Green PCR core reagent

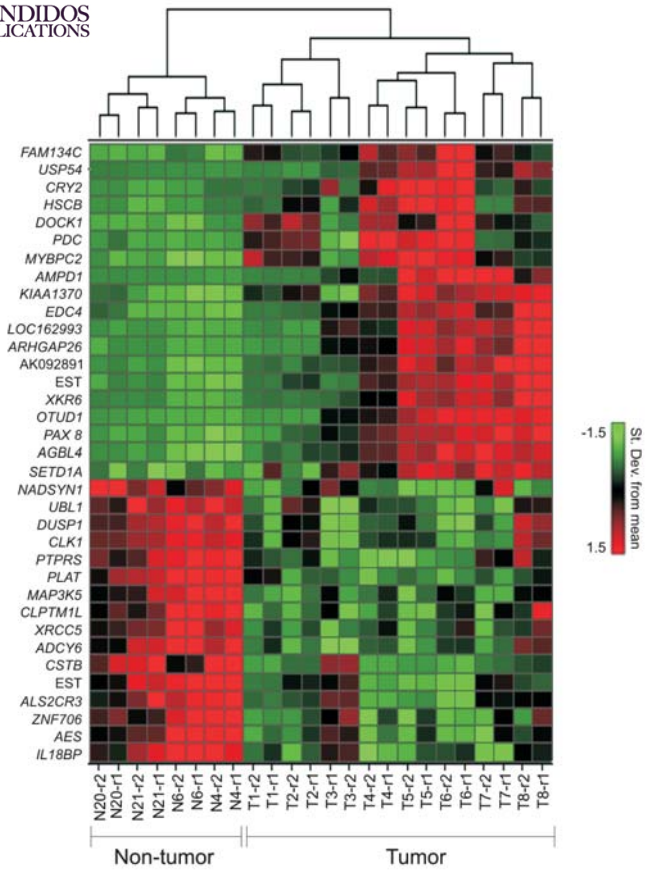

Figure 1. Supervised hierarchical clustering of 8 larynx tumor and 4 non-tumor samples used in the training set. A total of 35 differentially expressed genes was identified, using a signal-to-noise ratio of >|1| and p≤0.001 as cut-offs. Selected transcripts were grouped using hierarchical clustering (Wards method with half square Euclidean distance), and heat maps were constructed using the Spotfire software. For each gene (row), red indicates a higher expression and green a lower one relative to the average level of expression of that gene across the 35 samples (columns).

(Applied Biosystems) following the manufacturer's protocol. In brief, the reaction mixture (20 μl total volume) contained 25 ng of cDNA, gene-specific forward and reverse primers for each gene at a final concentration of 0.4 or 0.5 μM, and 10 μl of 2X quantitative SYBR-Green PCR master mix. The relative quantification was given by the CT values, determined for triplicate reactions for tumor and reference samples from each gene and for the endogenous control (α-tubulin, *TUBA1C*). Thus, the fold-change of each gene was calculated by using the $2^{-\Delta\Delta CT}$ formula, where CT = fluorescence at the defined detection threshold; ΔCT = CT of the target gene - CT of the endogenous control and ΔΔCT = ΔCT (tumor sample) - ΔCT (reference sample). For the tumor sample, evaluation of $2^{-\Delta\Delta CT}$ indicates the fold change in gene expression relative to the reference sample.

Statistical analysis was performed using the Minitab software (version 12.22) and the significance level was defined as α=0.05. Relative expression levels detected by quantitative real-time RT-PCR for the 14 genes in samples from larynx carcinoma were transformed into natural logarithms. The Anderson-Darling normality test was performed to determine whether or not the values presented a normal distribution. As this test did not reject normality for the 14 genes, data were subjected to a one-sample Student's t-test on their expression levels at the logarithmic scale.

Results

Epidemiology. The study population included 14 males and 1 female with larynx carcinoma. Minimum and maximum ages were 44 and 73 years, respectively, with a mean age of 58.7±9.8 years. Patients at the UICC stage were: 1 T2, 2 T3 and 12 T4, while at the neck stage, there were: 7 N0, 2 N1 and 6 N2 patients. Thus, 93% of the patients had T3-T4 tumors and 60% had regional metastases at presentation. None of the patients had distant metastases upon diagnosis of the disease. Regarding the clinical stage, 3 patients had stage III and 12 had stage IV. Six patients presented no clinical evidence of the disease at the last follow-up, and 9 patients showed disease recurrence. Recurrence occurred between 4 and 14 months after surgery. Our data are in agreement with the literature, since 40-50% of patients with advanced disease (stage III and IV) exhibited recurrence, and ~80% of these recurrences occurred within the first two years. Regarding the risk factors, 100% were smokers and 94% used to consume alcohol. Therefore, most of these patients had a history of smoking and drinking, which are the major etiological factors in larynx cancer. Patient data are presented in Table I.

Gene expression analysis of larynx tumors. Expression analysis using the 331-element cDNA microarray was initially performed on 8 larynx tumor (T1-T8) and 4 larynx non-tumor samples.

Differentially expressed genes were identified using the signal-to-noise ratio (SNR) metric as described in Materials and methods. Of the 285 genes analyzed, 35 (12% of the total) showed statistically significant differences in the expression between tumor and adjacent non-tumor larynx tissue (SNR ≥| 1.0 |, p≤0.001). Among these 35 genes, 19 (54%) showed a higher expression in tumor than in non-tumor tissue, and 16 (46%) presented the contrasting pattern.

Supervised hierarchical clustering analysis revealed that the expression patterns of the selected set of 35 differentially expressed genes were able to perfectly distinguish tumors from non-malignant tissues in the set of samples used in this training set (Fig. 1). Two subgroups of tumor samples were distinguishable in the training set based on the expression profile of the 35 genes (Fig. 1), suggesting heterogeneity among larynx cancer cases. Separation of these samples into 2 subgroups does not appear to correlate with the clinical staging or metastatic features of the tumors. In contrast, we observed a statistically significant difference (p=0.030) in the mean age of patients upon disease presentation in each subgroup: in one, the ages of patients ranged between 44 and 58 years (51±5.5), and in the other from 61 to 73 (68±4.7).

To verify the robustness of the 35-gene signature, we investigated whether this gene set was able to distinguish tumor from non-tumor larynx tissues, using data from an additional set of 5 larynx tumor samples (T9, T10, T11, T14 and T15) and 6 additional adjacent non-tumor larynx tissue samples. As shown in Fig. 2, the 35-gene set was able to separate this additional set of larynx samples according to

Table II. The 35 genes showing down- or up-regulation in larynx cancer.

Locus name	EST accession/ HN contig name	Annotation	Genome mapping	Ratio tumor/ non-tumor (log^2)	Biological function/molecular process
NADSYN1	BF825810	NAD synthetase 1	11q13.4	-1.1	Nucleotide binding/NAD biosynthetic process
CSTB	R1_CL240674_CT4	Cystatin B	21q22.3	-1.0	Cysteine protease inhibitor activity/protease inhibition
EST	BI002390	Intergenic region	2q31.2	-0.8	Unknown
UBL1	R1_CL327107_CT1	Ubiquitin-like 1 (sentrin)	2q33.1	-0.8	SUMO-conjugating enzyme activity/DNA repair
PLAT	R1_CL186187_CT2	Plasminogen activator tissue type isoform 2	8p11.21	-0.7	Plasminogen activator activity/proteolysis
CLK1	R1_CL306959_CT1	CDC-like kinase 1	2q33.1	-0.6	Protein serine/threonine kinase activity/cell proliferation
DUSP1	R1_CL117820_CT4	Dual specificity phosphatase 1	5q35.1	-0.6	MAP kinase phosphatase activity/cell cycle
XRCC5	R1_CL315103_CT23	X-ray repair defective repair in Chinese hamster cells	2q35	-0.5	ATP-dependent DNA helicase activity/DNA repair
ALS2CR3	R1_CL303821_CT1	Amyotrophic lateral sclerosis 2 chromosome region, candidate 3	2q33.1	-0.4	GABA receptor binding/neurotransmitter transport
CRR9	R1_CL315539_CT1	Cisplatin resistance-related protein CRR9p	5pter15.33	-0.4	Unknown
ZNF706	R1_CL164403_CT3	Protein zinc finger 706	8q22.3	-0.4	Nucleic acid binding
AES	R1_CL309068_CT2	Amino-terminal enhancer of split isoform b	19p13.3	-0.4	Transcription corepressor activity/transcriptional regulation
MAP3K5	R1_CL237238_CT1	Mitogen-activated protein kinase kinase kinase 5	6q25.1	-0.4	MAP kinase kinase kinase activity/apoptosis
IL18BP	R1_CL305797_CT10	Interleukin 18 binding protein precursor	11q13.4	-0.4	Receptor antagonist activity/T-helper 1 type immune response
ADCY6	R1_CL17269_CT2	Adenylate cyclase 6 isoform a	12q14.1	-0.4	Adenylate cyclase activity/cAMP biosynthesis
PTPRS	R1_CL310075_CT12	Protein tyrosine phosphatase receptor type	19p13.3	-0.3	Protein tyrosine phosphatase activity/cell adhesion
PDC[a]	CV374356	Phosducin isoform a	1q25.2	1.0	Phospholipase inhibitor activity/response to stimulus
SETD1A[a]	BI000562	SET domain containing 1A	16p11.2	1.0	Nucleotide and protein binding/chromatin modification
KIAA1370[a]	BI002185	Hypothetical protein LOC56204	15q21.2	1.1	Unknown

Table II. Continued.

Locus name	EST accession/ HN contig name	Annotation	Genome mapping	Ratio tumor/ non-tumor (log^2)	Biological function/molecular process
DOCK1	BF828101	Dedicator of cytokinesis 1	10q26.2	1.2	GTPase activator activity/apoptosis
MYBPC2	BF826295	Myosin-binding protein C fast type	19q13.33	1.3	Actin binding/cell adhesion
FAM134C	BF377952	Family with sequence similarity 134, member C	17q21.31	1.4	Unknown
EST	AW984383	Intergenic region	6q14.1	1.5	Unknown
AGBL4[a]	AW984388	ATP/GTP binding protein-like 4	1p33	1.6	Carboxypeptidase A activity/proteolysis
PAX 8	BI004213	Paired box gene 8	2q13	1.6	Transcriptional activator activity/transcriptional regulation
HSCB	BI003653	HscB iron-sulfur cluster co-chaperone homolog	22q12.1	1.8	Chaperone binding/protein folding
XKR6[a]	BI002820	Kell blood group complex subunit-related family, member 6	8p23.1	1.9	Unknown
EDC4	BQ359764	Enhancer of mRNA decapping 4	16q22.1	2.0	Protein binding
CRY2	BI004402	Cryptochrome 2 (photolyase-like)	11p11.2	2.0	Nucleotide and protein binding/transcriptional regulation
ARHGAP26	BI002158	Rho GTPase activating protein 26	5q31	2.1	Rho GTPase activator activity/signal transduction
LOC162993[a]	BI000521	LOC162993 hypothetical protein	19p13.2	2.2	Unknown
AK092891	BI001659	Homo sapiens cDNA FLJ35572	17q21.31	2.4	Unknown
OTUD1	BE154599	OTU domain containing 1	10p12.2	3.2	Unknown
USP54	BE153900	Ubiquitin-specific peptidase 54	10q22.2	3.3	Ubiquitin thiolesterase activity/ubiquitin-dependent protein catabolism
AMPD1	BF827836	Adenosine monophosphate deaminase 1	1p13	3.5	AMP deaminase activity/nucleotide metabolic process

[a]EST mapping to intronic region of the annotated gene locus. HN, head and neck.

their histological type, thus confirming the robustness of this gene classifier.

A functional classification of the selected 35-gene set was performed manually by querying the Entrez Gene database (30). We observed that these genes are involved in processes such as apoptosis, cell adhesion, cell cycle, cell motility, DNA repair, metabolism, proteolysis, signal transduction and transcriptional regulation (Table II). Thus, of the 19 genes identified and overexpressed in larynx tumors, 2 were involved in signal transduction, 2 in transcriptional regulation and 1 each in apoptosis, catabolism, cell adhesion, chromatin modification, folding protein, metabolism and response to stimulus, and 8 with unknown functions. Of the 16 genes underexpressed in larynx tumors, 2 were involved in cell cycle, 2 in DNA repair, and 1 each in apoptosis, biosynthesis, cell adhesion, immune response, protease inhibition, proteolysis, transcription regulation and transport, and 4 with unknown functions.

Interestingly, 6 of the 35 transcripts identified map to intronic regions of protein-coding genes (Table II). While these partial transcripts may represent non-annotated exons of these genes, no significant open-reading frames and coding-potential was observed, as determined by the ESTScan software (31), suggesting that these genes represent as yet uncharacterized non-coding RNAs that are deregulated in larynx tumors.

Validation of markers of larynx carcinoma by real-time RT-PCR. Quantitative real-time RT-PCR (qRT-PCR) was employed to validate the larynx tumor gene expression by an independent method. We compared the relative expression levels of the 14 genes *ADCY6*, *AES*, *ALS2CR3*, *CLK1*, *CRR9*, *CSTB*, *DUSP1*, *IL18BP*, *MAP3K5*, *PLAT*, *PTPRS*, *UBL1*, *XRCC5* and *ZNF706* using triplicate measurements and normalization based on the α-tubulin level. Data were analyzed using Student's t-test. Seven genes were statistically under-

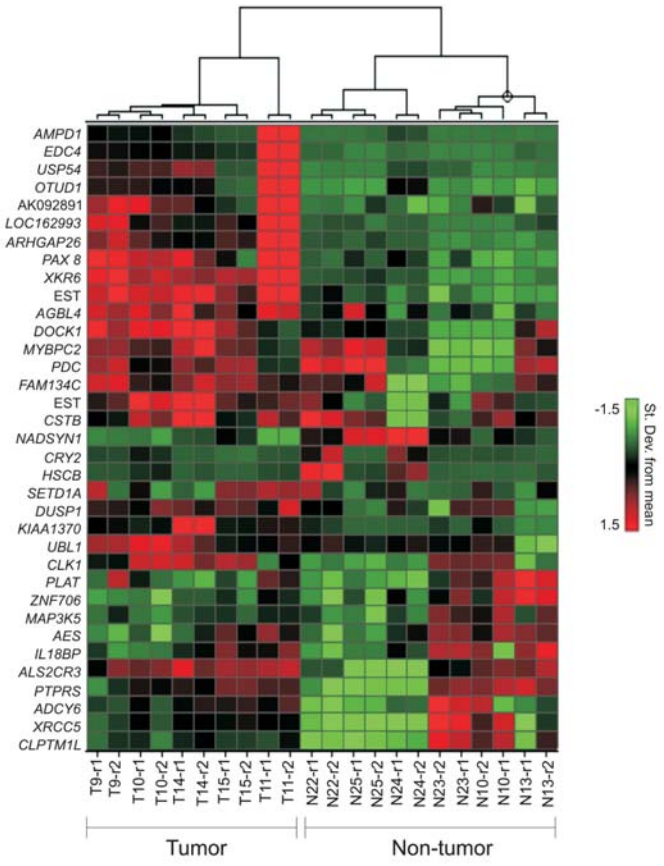

Figure 2. Validation of the 35-gene classifier in an independent set of tumor and non-tumor larynx samples. Expression of the 35 genes identified in the training set was measured in an independent validation set comprised of 5 tumor and 6 non-tumor samples. Selected transcripts were grouped using hierarchical clustering (Wards method with half square Euclidean distance), and heat maps were constructed using the Spotfire software. For each gene (row), red indicates a higher expression and green a lower one, relative to the average level of expression of that gene across the 35 samples (columns).

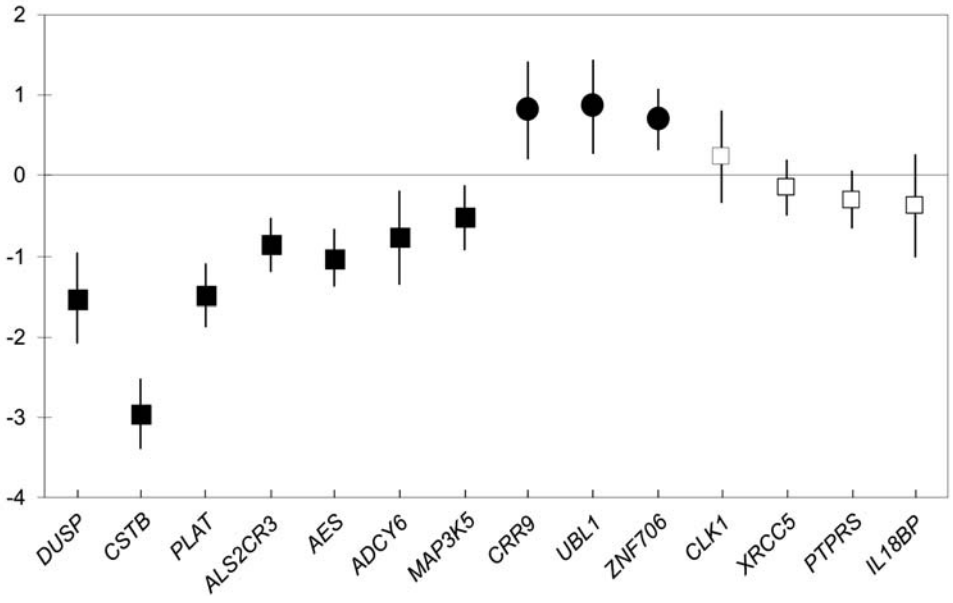

Figure 3. Mean and 95% confidence interval for the natural logarithms of the relative expression values of genes selected for validation by quantitative real-time RT-PCR. Expression of *ADCY6*, *AES*, *ALS2CR3*, *CSTB*, *DUSP1*, *MAP3K5* and *PLAT* (black squares) was significantly lower in tumor tissues than in normal reference. Expression of *CRR9*, *UBL1* and *ZNF706* was significantly higher in tumor tissues than in normal reference (black circles). *CLK1*, *IL18BP*, *PTPRS* and *XRCC5* were not confirmed by real-time RT-PCR as differentially expressed (white squares).

Primer sequences used for validation by quantitative real-time PCR.

Gene	Forward (F) and reverse (R) primers	Amplicon size (bp)
ADCY6	F - 5' - ACACCTGCTGACATCACTGC - 3' R - 5' - GACAGAGCTGGCCAAGAGAC - 3'	150
AES	F - 5' - CCCAGCAACTCAAATTCACC - 3' R - 5' - CTCACTGGCCAACTTGTCAC - 3'	116
ALS2CR3	F - 5' - TTGCCTTCGAGGTTTTATGC - 3' R - 5' - ATGGTTCCCAAGTTTGGTTG - 3'	108
TUBA1C	F - 5' - TCAACACCTTCTTCAGTGAAACG - 3' R - 5' - AGTGCCAGTGCGAACTTCATC - 3'	241
CLK1	F - 5' - AGGAGCATTTAGCAATGATG - 3' R - 5' - CCAGTCTAATCGATCGTGGTG - 3'	104
CRR9	F - 5' - TGGGTGCTGAGAACAACATC - 3' R - 5' - ACAGCGTCCCATTGTTTCTC - 3'	120
CSTB	F - 5' - TGTCATTCAAGAGCCAGGTG - 3' R - 5' - GGGAGAGATTGGAACACTCG - 3'	100
DUSP1	F - 5' - GCGGAATCTGGGTGCAGTT - 3' R - 5' - CAGGTACAGAAAGGGCAGGATTT - 3'	81
IL18BP	F - 5' - TAAGCAGTGTCCAGCATTGG - 3' R - 5' - AGGCCACACAGGATAAGCTC - 3'	83
MAP3K5	F - 5' - ACCGAAGAGAAGGGGAGAAG - 3' R - 5' - CCGACCTGCGTAGACTATCC - 3'	135
PLAT	F - 5' - GAGTGCACCAACTGGAACAG - 3' R - 5' - GGCTTTGAGTCTCGATCTGG - 3'	125
PTPRS	F - 5' - TCCTGCGAGAGTTCAAGGTC - 3' R - 5' - GACTTTGGCACACCCTGTTC - 3'	100
UBL1	F - 5' - TGACAACACATCTCAAGAAACTCAAA - 3' R - 5' - TCTCTGACCCTCAAAGAGAAACCT - 3'	92
XRCC5	F - 5' - ACCAAAGAGGAAGCCTCTGG - 3' R - 5' - CGTCCACATCACCACCTTC - 3'	121
ZNF706	F - 5' - AAATGCCAAAAAGCAAGCTG - 3' R - 5' - TGCTTGCTCTCAAAGTGCTG - 3'	147

expressed in tumor samples compared to the normal reference, while three genes were statistically overexpressed. These data are summarized in Fig. 3. The complete data with the relative expression levels detected by quantitative real-time PCR for the 14 genes in samples from larynx carcinoma are shown in Table IV.

We confirmed by qRT-PCR that genes *ADCY6*, *AES*, *ALS2CR3*, *CSTB*, *DUSP1*, *MAP3K5* and *PLAT* are significantly down-regulated in tumor tissues compared to the normal reference (p=0.012, <0.0001, <0.0001, <0.0001, <0.0001, 0.014 and <0.0001, respectively). These genes, except *ALS2CR3*, are located in regions of chromosomal loss involved in head and neck carcinogenesis. The results of qRT-PCR were in agreement with those of the cDNA microarray analysis, further supporting evidence that these are potential cancer-related genes.

We found by qRT-PCR that genes *CRR9*, *UBL1* and *ZNF706* were significantly up-regulated in tumor tissues when compared to the normal reference (p=0.011, 0.007 and 0.001, respectively). It should be noted that these genes displayed expression patterns which were in contrast to the cDNA microarray analysis results. These genes are located in regions of chromosomal gains involved in head and neck carcinogenesis. The differential expression of *CLK1*, *IL18BP*, *PTPRS* and *XRCC5* was not confirmed by qRT-PCR (p=0.406, 0.222, 0.083 and 0.340, respectively).

Table IV. Natural logarithms of relative expression levels detected by quantitative real-time PCR for the 14 genes in samples from larynx carcinoma.[a]

	Gene													
Sample	*ADCY6*	*AES*	*AL2SCR3*	*CLK1*	*CRR9*	*CSTB*	*DUSP1*	*IL18BP*	*MAP3K5*	*PLAT*	*PTPRS*	*UBL1*	*XRCC5*	*ZNF706*
T1	-0.58	-1.64	-0.93	-0.18	1.07	-3.56	-3.86	0.15	-0.46	-1.06	0.98	1.24	-0.33	0.08
T2	-0.14	-1.57	-0.34	0.37	2.20	-2.61	-1.00	-1.38	0.13	-2.28	0.18	1.38	0.58	1.09
T3	0.24	-0.40	-0.20	1.62	1.80	-2.65	-1.28	-0.30	0.43	-1.41	0.47	1.55	0.16	1.48
T4	-1.96	-1.82	-1.28	1.07	1.20	-2.59	-2.14	-1.62	-0.77	-1.33	-0.62	2.50	-1.59	1.47
T5	-3.26	-1.74	-0.94	0.78	0.06	-2.92	-1.02	-1.68	-0.91	-2.80	-1.30	1.58	0.06	-0.56
T6	-0.72	-0.85	-0.15	1.32	1.38	-1.99	-0.36	0.42	-0.87	-1.84	-0.31	1.05	0.14	0.84
T7	-1.86	-0.20	-1.85	1.77	2.71	-4.47	-0.27	-1.12	-1.98	-2.47	-1.14	1.81	0.12	0.30
T8	0.02	-0.92	-0.68	0.89	1.20	-2.74	-2.43	-0.72	-0.37	-1.07	-0.26	2.15	0.39	0.96
T9	-1.21	-0.81	-1.25	-1.30	-0.27	-1.62	-2.42	-0.72	-0.43	-1.59	-0.25	-1.05	-0.21	0.56
T10	0.76	-0.76	-0.85	-0.52	0.34	-2.75	-0.37	1.10	-0.08	-1.40	0.20	0.20	-0.90	0.30
T11	-0.66	-1.19	-0.76	-0.65	0.05	-2.86	-1.83	-2.01	-0.95	-1.89	-0.21	-0.19	-0.31	0.89
T12	-1.53	-1.95	-1.91	-0.99	-1.27	-4.04	-2.44	-0.65	-0.94	-0.74	-0.23	0.45	0.10	1.80
T13	-0.12	0.30	0.24	0.75	1.76	-2.35	-0.59	0.96	0.50	-0.61	-0.60	1.14	0.54	-0.32
T14	-0.42	-0.93	-1.22	-1.10	-0.43	-3.78	-1.67	0.02	-1.53	-1.55	-1.17	-0.61	-0.95	0.66
T15	-0.12	-0.90	-0.76	-0.42	0.47	-3.53	-1.14	1.92	0.37	-0.24	-0.22	-0.25	-0.10	0.90
Mean ± SD	-0.77± 1.0	-1.02± 0.6	-0.86± 0.6	0.23± 1.0	0.82± 1.1	-2.97± 0.8	-1.528± 1.0	-0.37± 1.1	-0.52± 0.7	-1.49± 0.7	-0.30± 0.6	0.86± 1.1	-0.15± 0.6	0.70± 0.7
P-value	0.012	0.000	0.000	0.406	0.011	0.000	0.000	0.222	0.014	0.000	0.083	0.007	0.340	0.001

[a]A pool of RNA isolated from 7 histologically normal larynx mucosa tissue samples adjacent to tumors were used as a reference. Experiments were performed in triplicate reactions and normalized for an endogenous control (α-tubulin). Statistical significance (P-value) was determined by one-sample Student's t-test as described in Materials and methods.

The current literature includes several studies which included DNA microarray analysis in the study of HNSCC, to determine gene expression changes during disease progression and/or predict disease outcome. However, considerable heterogeneity among these studies exists in terms of study design, number of samples, sites and stage of disease, choice of microarray platform and validation of results by other laboratory methodologies (7). Many studies in HNSCC have been performed using tumors from various anatomical sites, but carcinomas from different upper aerodigestive tract sites may behave differently.Consequently, the results may be misleading when analyzed together. Larynx cancer is one of the most common types of HNSCC, with high mortality rates and poor prognosis (32).

Considering the above, our study examined, by microarray technology, the expression levels of 284 genes in clinical samples of larynx squamous cell carcinoma and non-neoplastic larynx tissue, identifying 35 genes as differentially regulated in tumor as compared to non-tumor tissue. The genes detected are primarily involved in processes such as apoptosis, cell cycle, DNA repair, proteolysis, protease inhibition, signal transduction and transcriptional regulation. Interestingly, six of the transcripts identified map to intronic regions of protein-coding genes. These sequences are unspliced and apparently have no coding potential. Long intronic non-coding RNAs (ncRNAs) are ubiquitously transcribed in the human genome (33,34) and a subset of these ncRNAs were shown to correlate with the degree of tumor cell differentiation in prostate cancer (35). We postulate that the intronic transcripts identified here comprise of as yet uncharacterized long ncRNAs that have a regulatory role in the gene expression program in larynx tissues and are deregulated in the neoplastic tissue.

Microarray data were further confirmed by real-time RT-PCR in a set of 14 genes which were found to be down-regulated in tumor tissues by microarray analysis. Of the genes analyzed by microarray some showed similar patterns of expression by real-time RT-PCR. Thus, of the 14 selected genes tested by microarray and real-time RT-PCR, 7 showed congruent and significant down-regulation in tumor tissue, 3 displayed contrasting expression patterns with the two methods, while 4 did not exhibit any significant change in expression by real-time RT-PCR. One possible source of variation in the measurements of mRNA expression detected by microarray and real-time RT-PCR is the existence of different splicing isoforms, which may affect the transcript levels detected in different regions of the mRNA with these two techniques (36). This result highlights the importance of performing independent real-time RT-PCR validation following the identification of candidates by high-throughput microarray screening before further conclusions on gene expression can be drawn (37).

In the present study, 10 potential biomarkers of larynx carcinoma were validated and showed pronounced differences in expression between tumor and non-neoplastic samples. The validated markers were genes *ADCY6*, *AES*, *AL2SCR3*, *CRR9*, *CSTB*, *DUSP1*, *MAP3K5*, *PLAT*, *UBL1* and *ZNF706*.

The expression of genes *ADCY6*, *AES*, *ALS2CR3*, *CSTB*, *DUSP1*, *MAP3K5* and *PLAT* was significantly lower in tumor tissues than in the non-tumor adjacent tissue. Interestingly, these genes, except *ALS2CR3*, are located in regions of chromosomal loss involved in head and neck carcinogenesis.

Gene *ADCY6*, located at 12q14.1, encodes adenylate cyclase 6, which is a membrane-associated enzyme that catalyzes the formation of the secondary messenger cyclic adenosine monophosphate (cAMP). No information is available in the literature on *ADCY6* expression in head and neck cancer. However, Celano *et al* (38) observed that its expression is significantly lower in hyperfunctioning thyroid tumors than in normal thyroid tissue, evidencing a low expression of this gene in transformed tissues.

The amino-terminal enhancer of the split (*AES*) gene, located at 19p13, encodes a 197-amino acid protein that is homologous to the NH(2)-terminal domain of the *Drosophila* Groucho protein but lacks COOH-terminal WD40 repeats. According to data of Tetsuka *et al* (39), *AES* acts as a co-repressor for NF-κß, which is well recognized as a regulator of genes encoding cytokines, chemokines and cell adhesion molecules important in immune and inflammatory responses, as well as critical genes in the control of cellular proliferation and apoptosis (40).

Additionally, *AES* inhibits NF-κß,-dependent gene expression induced by tumor necrosis factor α, interleukin-1 ß, and mitogen-activated protein kinase/extracellular signal-regulated kinase kinase kinase 1, which is an upstream kinase for NF-κß activation (39). There is evidence that nuclear factor NF-κß modulates a broad program of genes differentially expressed during tumor progression of HNSCC, such as *CCND1* (*PRADI*), *Gro-1* (IL-8 homologue) and *GST* (41). Thus, the *AES* gene may play a role in head and neck carcinogenesis since it acts in the regulation of NF-κß and its target genes. Moreover, gene AES acts on the Wnt receptor signaling pathway. The nuclear output of Wnt signaling is mediated by a complex between DNA-binding proteins of the TCF family and the transcriptional coactivator ß-catenin. The Wnt/ß-catenin/Tcf pathway serves important functions in embryonic development and is constitutively activated in human colorectal cancer (42,43). Groucho proteins, which include AES, act to repress transcriptional activation by the ß-catenin-tcf complex, probably by interacting directly with Tcf transcriptional factors (42). Thus, the low expression of the AES gene may contribute to a higher activation of the Wnt/ß-catenin/Tcf signaling model and consequently to larynx carcinogenesis, as occurs in colorectal cancer.

The *ALS2CR3* gene, located at 2q33.1, is denominated as *GRIF-1* and *TRAK2*. This gene encodes a factor that interacts with the γ-aminobutyric acid (A) receptor and plays a role in the intracellular transport, mainly in the kinesin-mediated transport of mitochondria (44). Alterations of this gene have been associated with neurodegenerative diseases (45). Currently, however, no studies are available in the literature describing alterations of *ALS2CR3* in tumor tissues.

Cystatin B, also called stefin B, is a small protein that is a member of the superfamily of cysteine protease inhibitors. Cystatins have emerged as important players in a multitude of physiological and pathophysiological settings that range from cell survival and proliferation to differentiation, cell signaling and immunomodulation (46). Cysteine proteases,

in turn, have been implicated in multiple steps of tumor progression, including early steps of immortalization and transformation (47), intermediate steps of tumor invasion and angiogenesis (48), and late steps of metastasis and drug resistance (49). The importance of lysosomal cysteine proteases in the progression of tumors from benign to aggressive lesions suggests that cystatins in many ways safeguard against tumor progression (46).

In a review on gene expression profiles, a low expression of the *CSTB* (cystatin B) gene, located at 21q22.3, was reported in 5 of 24 HNSCC microarray studies (7). A decrease of CSTB expression was observed in esophageal carcinoma, breast cancer, prostatic adenocarcinomas and atypical meningiomas (50-53).

The *DUSP1* gene, also referred to as *MSP-1* and located at 5q35.1, encodes a dual-specificity phosphatase for tyrosine and threonin (54). This gene specifically inactivates mitogen-activated protein kinase and suppresses its activation by *ras*. The *DUSP1* gene is also a transcriptional target of tumor suppressor p53, inducing cell cycle arrest or apoptosis. Consequently, *DUSP1* may play an important role in the negative regulation of cell proliferation (54,55).

A low expression of *DUSP1* has been observed in ovarian and prostate cancer (56,57). Moreover, in line with our study, Tomioka *et al* (58) observed a low expression of this gene in head and neck cancer. These authors considered *DUSP1* a candidate for tumor suppression, mediating PTEN signaling pathways. Unoki and Nakamura (59) reported that *DUSP1* gene expression was induced by the introduction of exogenous *PTEN* into endometrial cancer cell lines.

In addition, *DUSP1* plays a vital role in the regulation of innate immune responses via the p38 MAPK and JNK pathways and appears to play a role in the induction of senescence through the inhibition of AP-1 activity and the subsequent transcription of genes involved in DNA replication (60,61).

Gene *MAP3K5*, located at 6q25.1 and denominated as apoptosis signal-regulating kinase 1 (*ASK1*), encodes a multifunctional serine/threonine kinase involved in a broad range of biological activities including cell differentiation and stress-induced apoptosis (62,63). Its catalytic activity can be activated by many stress stimuli such as tumor necrosis factor α (TNF-α), reactive oxygen species (ROS), DNA damage and chemotherapeutic agents such as cisplatin and taxol, and selectively activates JNK and p38 MAPK pathways (64).

The pro-apoptotic role of *MAP3K5* is reinforced by Dasgupta *et al* (65) and Kherrouche *et al* (66). Dasgupta *et al* (65) described a physical and functional interaction between MAP3K5 and the Rb protein in response to apoptotic stimuli. The Rb protein has antiproliferative and antiapoptotic functions. It appears that MAP3K5 has to overcome RB functions to induce apoptosis. Moreover, the MAP3K5-mediated inactivation of Rb correlates with increased levels of the pro-apoptotic protein p73. Kherrouche *et al* (66) suggest that the overexpression of *E2F1* induces the expression of *MAP3K5* and that some of the cellular functions of MAP3K5 are under the control of the E2F transcriptional factor. The authors also suggest that the up-regulation of *MAP3K5* favor the p53-independent E2F1 apoptotic activity. However, no information is available in the literature regarding *MAP3K5* expression in head and neck cancer.

The *PLAT* gene, located at 8p11.21, encodes the tissue-type plasminogen activator, a secreted serine protease which converts the proenzyme plasminogen to plasmin, a fibrinolytic enzyme (67). Numerous studies have provided evidence that plasminogen activators (PA)/plasmin systems represent a key event in angiogenesis. It is well established that tumor-induced neovascularization and angiogenesis are necessary requirements for the growth of tumors and their metastases. The PA/plasmin system comprises two major types of PA, tissue (PLAT) and urokinase (PLAU) (68).

In contrast to *PLAU*, for which considerable evidence indicates that up-regulation of the enzyme correlates with tumor aggression, several observations suggest that high *PLAT* levels correlate with good prognosis in melanoma and breast cancer, whereas lower *PLAT* levels have been associated with malignant tumors (69,70). These findings suggest that an increase in either PLAT activity or expression levels is beneficial, possibly due to the overstimulation of plasmin generation by PLAT that induces the degradation of the pro-angiogenic fibrin matrix, resulting in the inhibition of angiogenesis (71,72). Moreover, Gingras *et al* (72,73) reported that Neovastat, an inhibitor of angiogenesis derived from marine cartilage, specifically stimulates PLAT-dependent plasmin generation through an increase in the affinity of the enzyme towards plasminogen.

Expression of the *CRR9*, *UBL1* and *ZNF706* genes was significantly higher in tumor tissues than in the normal reference. These genes are located in regions of chromosomal gains involved in head and neck carcinogenesis.

The *CRR9* gene, located at 5pter15.33, has an unknown function. A high expression of this gene was observed in a renal carcinoma cell line when compared to normal cells, as well as in a cell line of ovary cancer resistant to treatment with cisplatin (74,75). However, no information is available in the literature on *CRR9* expression in head and neck cancer.

The *UBL1* gene, also referred to as *SUMO-1* and located at 2q33.1, is likely to be involved in many cellular processes, including apoptosis, mitosis regulation, protein translocation, cell proliferation and transcriptional regulation. One function of the UBL1 protein is to conjugate covalently with target proteins and modify their function. Consequently, in the nucleus, the transcriptional activities of UBL1-modified transcriptional factors including p53, c-jun, Sp-3, c-Myc and c/EBP families are reduced (76,77).

The UBL1 protein forms a complex with RAD51 and RAD52 proteins in human cells, which play essential roles in DNA homologous recombination, DNA repair and cell proliferation. *UBL1* overexpression down-regulates DNA double-strand break-induced homologous recombination in CHO cells and reduces cellular resistance to ionizing radiation. Overexpression of *UBL1* reduces the fraction of bidirectional gene conversion tracts, and that of a mutant *UBL1* that is incapable of being conjugated retains the ability to inhibit homologous recombination (78). These results suggest a regulatory role for *UBL1* in homologous recombination.

The UBL1 protein was found to repress *MAPK5* activation through physical interaction (79). Therefore, *UBL1* overexpression may contribute to carcinogenesis by decreasing DNA repair capacity and inhibiting *MAP3K5* activation and, consequently, apoptosis. Currently, no information is available

rature on *UBL1* expression in head and neck cancer., a high *UBL1* expression was observed in a renal carcinoma cell line when compared to normal cells (80).

No information is available in the literature on *ZNF706* expression in cancer. The zinc finger gene family belongs to one of the largest families of transcriptional factors. Most zinc finger proteins bind to specific DNA sequences and are involved in the transcriptional regulation of gene expression. Members of the zinc finger family function as activators or repressors of gene transcription, regulating embryonic development as well as a variety of physiological processes in the adult (81). Zinc-finger-containing transcriptional factors have previously been involved in MAPK signaling pathway regulation. These factors are among the most widespread mechanisms of eukaryotic cell regulation (82).

According to the holistic model of cancer proposed by Hanahan and Weinberg (83), a malignant cell has to acquire 6 biological alterations to dictate pathogenesis: self-sufficiency in proliferative growth signals, insensitivity to growth inhibitor signals, evasion of apoptosis, limitless replication potential, sustained angiogenesis and the induction of invasion. By verifying this model, our study shows a set of differentially expressed genes, especially *AES*, *CSTB*, *DUSP1*, *MAP3K5*, *PLAT* and *UBL1* that can directly or indirectly feed these pathways.

Among the differentially expressed genes, we emphasize *CSTB*. Notably, this gene showed a low expression in the tumor samples analyzed by microarray and real-time RT-PCR. The *CSTB* gene has an antimetastatic function, since it inhibits lysosomal cysteine protease action. Despite mounting evidence showing that the expression of lysosomal cysteine proteases is aberrant in tumor versus normal cells, this class of proteases has yet to be investigated (46). Emphasis has been placed on metalloproteases as potential novel targets for anti-cancer chemotherapy. Clinical trials with cancer patients have provided no indication that metalloprotease inhibitors are successful (84). This has led to a growing interest in members of the cystatin superfamily for potential novel anti-cancer strategies, and pre-clinical studies are promising (85). Several clinical studies have shown that some cathepsins and/or cystatins may have diagnostic and/or prognostic value in a variety of cancer types (52,53,86) Since loco-regional relapse and metastasis appear to be significant contributing factors for the restricted survival of HNSCC patients (1), we suggest that the *CSTB* gene constitutes a good biomarker for larynx cancer and deserves more attention in future studies.

In conclusion, our microarray analysis revealed a gene expression signature of larynx tumors, including several genes whose deregulation is potentially associated with disease progression. Further studies are required to evaluate whether the genes identified in this study are specifically altered in larynx tumors or are deregulated in other head and neck tumors. These findings will contribute to the understanding of the molecular basis of larynx cancer, thus helping to improve diagnosis, treatment and patient outcome.

Acknowledgements

This work was mainly funded by grants and fellowships from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brazil. Partial funding was available as a grant from the Ludwig Institute for Cancer Research under the James R. Kerr Program.

References

1. Lothaire P, de Azambuja E, Dequanter D, *et al*: Molecular markers of head and neck squamous cell carcinoma: promising signs in need of prospective evaluation. Head Neck 28: 256-269, 2006.
2. Döbrossy L: Epidemiolgy of head and neck cancer: magnitude of the problem. Cancer Metastasis Rev 24: 9-17, 2005.
3. Chung CH, Levy S and Yarbrough WG: Clinical applications of genomics in head and neck cancer. Head Neck 28: 360-368, 2006.
4. INCA - Instituto Nacional do Câncer, Ministério da Saúde. Estimativa da incidência e mortalidade por câncer no Brasil, Rio de Janeiro. http://www.inca.org.br, 2008.
5. Hunter KD, Parkinson EK and Harrison PR: Profiling early head and neck cancer. Nat Rev Cancer 5: 127-135, 2005.
6. Ragin CC, Modugno F and Gollin SM: The epidemiology and risk factors of head and neck cancer: a focus on human papillomavirus. J Dent Res 86: 104-114, 2007.
7. Choi P and Chen C: Genetic expression profiles and biology pathway alterations in head and neck squamous cell carcinoma. Cancer 104: 1113-1128, 2005.
8. Kim MM and Califano JA: Molecular pathology of head and neck cancer. Int J Cancer 112: 545-553, 2004.
9. Angadi PV and Krishnapillai R: Cyclin D1 expression in oral squamous cell carcinoma and verrucous carcinoma: correlation with histological differentiation. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 103: 30-35, 2007.
10. Kalyankrishna S and Grandis JR: Epidermal growth factor receptor biology in head and neck. J Clin Oncol 24: 2666-2672, 2006.
11. Ruíz-Godoy RLM, Garcia-Cuellar CM, Herrera Gonzalez NE, *et al*: Mutational analysis of K-ras and Ras protein expression in larynx squamous cell carcinoma. J Exp Clin Cancer Res 25: 73-78, 2006.
12. Papakosta V, Vairaktaris E, Vylliotis A, *et al*: The co-expression of c-myc and p53 increases and reaches a plateau early in oral oncogenesis. Anticancer Res 26: 2957-2962, 2006.
13. Smith EM, Wang D, Kim Y, Rubenstein LM, Lee JH, Haugen TH and Turek LP: p16(INK4a) Expression, human papillomavirus, and survival in head and neck cancer. Oral Oncol 44: 133-149, 2007.
14. Keum KC, Chung EJ, Koom WS, *et al*: Predictive value of p53 and PCNA expression for occult neck metastases in patients with clinically node-negative oral tongue cancer. Otolaryngol Head Neck Surg 135: 858-864, 2006.
15. Hirai T, Hayashi K, Takumida M, Ueda T, Hirakawa K and Yajin K: Reduced expression of p27 is correlated with progression in precancerous lesions of the larynx. Auris Nasus Larynx 30: 163-168, 2003.
16. Nemes JA: p21WAF1/CIP1 expression is a marker of poor prognosis in oral squamous cell carcinoma. J Oral Pathol Med 34: 274-279, 2005.
17. Gleich LL and Salamone FN: Molecular genetics of head and neck cancer. Cancer Control 9: 369-378, 2002.
18. Thomas GR, Nadiminti H and Regalado J: Molecular predictors of clinical outcome in patients with head and neck squamous cell carcinoma. Int J Exp Path 86: 347-363, 2005.
19. Belbin TJ, Singh B, Smith RV, Socci ND, *et al*: Molecular profiling of tumor progression in head and neck cancer. Arch Otolaryngol Head Neck Surg 131: 10-18, 2005.
20. Squire JA, Bayani J, Luk C, *et al*: Molecular cytogenetic analysis of head and neck squamous cell carcinoma: by comparative genomic hybridization, spectral karyotyping, and expression array analysis. Head Neck 24: 874-887, 2002.
21. Al Moustafa AE, Alaoui-Jamali MA, Batist G, *et al*: Identification of genes associated with head and neck carcinogenesis by cDNA microarray comparison between matched primary normal epithelial and squamous carcinoma cells. Oncogene 21: 2634-2640, 2002.
22. Carinci F, Arcelli D, Lo Muzio L, *et al*: Molecular classification of nodal metastasis in primary larynx squamous cell carcinoma. Transl Res 150: 233-245, 2007.
23. Reis EM, Ojopi EP, Alberto FL, *et al*: Large-scale transcriptome analyses reveal new marker candidates of head, neck, and thyroid cancer. Cancer Res 65: 1693-1699, 2005.

24. Knuutila S, Autio K and Aalto Y: Online access to CGH data of DNA sequence copy number changes. Am J Pathol 157: 689, 2000.
25. Rozen S and Skaletsky H: Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132: 365-386, 2000.
26. International Union Against Cancer (UICC): TNM classification of malignant tumours, 6th edition. Wiley-Liss, New York, 2002.
27. Peixoto BR, Vêncio RZ, Egidio CM, Mota-Vieira L, Verjovski-Almeida S and Reis EM: Evaluation of reference-based two-color methods for measurement of gene expression ratios using spotted cDNA microarrays. BMC Genomics 7: 35, 2006.
28. Quackenbush J: Microarray data normalization and transformation. Nat Genet 32: 496-501, 2002.
29. Golub TR, Slonim DK, Tamayo P, *et al*: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286: 531-537, 1999.
30. Maglot D, Ostell J, Pruitt KD and Tatusova T: Entrez Gene: gene-centered information at NCBI. Nucleic Acids Res 35: D26-D31, 2007.
31. Iseli C, Jongeneel CV and Bucher P: ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. Proc Int Conf Intell Syst Mol Biol: 138-148, 1999.
32. Vokes EE and Stenson KM: Therapeutic options for laryngeal cancer. N Engl J Med 349: 2087-2089, 2003.
33. Birney E, Stamatoyannopoulos JA, Dutta A, *et al*: Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447: 799-816, 2007.
34. Nakaya HI, Amaral PP, Louro R, *et al*: Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. Genome Biol 8: R43, 2007.
35. Reis EM, Nakaya HI, Louro R, *et al*: Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer. Oncogene 23: 6684-6692, 2004.
36. Hedge PS, White IR and Debouck C: Interplay of transcriptomics and proteomics. Curr Opin Biotechnol 14: 647-651, 2003.
37. Liew KJL and Chow VTK: Microarray and real-time RT-PCR analysis of a novel set of differentially expressed human genes in ECV304 endothelial-like cells infected with dengue virus type 2. J Virol Meth 131: 47-57, 2006.
38. Celano M, Arturi F, Presta I, *et al*: Expression cyclase types III and VI in human hyperfunctioning thyroid nodules. Mol Cell Endocrinol 203: 129-135, 2003.
39. Tetsuka T, Uranishi H, Imai H, *et al*: Inhibition of nuclear factor-kappaB-mediated transcription by association with the amino-terminal enhancer of split, a Groucho-related protein lacking WD40 repeats. J Biol Chem 275: 4383-4390, 2000.
40. Chang AA and Waes CV: Nuclear factor-kappaß as a common target and activator of oncogenes in head and neck squamous cell carcinoma. Adv Otorhinolaryngol 62: 92-102, 2005.
41. Loercher A, Lee TL and Ricker JL: Nuclear factor-κB is an important modulator of the altered gene expression profile and malignant phenotype in squamous cell carcinoma. Cancer Res 64: 6511-6523, 2004.
42. Lepourcelet M and Shivdasani RA: Characterization of a novel mammalian Groucho isoform and its role in transcriptional regulation. J Biol Chem 277: 47732-47740, 2002.
43. Wong NACS and Pignatelli M: ß-Catenin - A Linchpin in colorectal carcinogenesis. Am J Pathol 160: 389-401, 2002.
44. Brickley K: GRIF-1 and OIP106, members of a novel gene family of coiled-coil domain proteins: association in vivo and in vitro with kinesin. J Biol Chem 280: 14723-14732, 2005.
45. Hadano S, Hand CK, Osuga H, *et al*: A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. Nat Genet 29: 166-173, 2001.
46. Keppler D: Towards novel anti-cancer strategies based on cystatin function. Cancer Lett 235: 159-176, 2006.
47. Fehrenbacher M, Gyrd-Hansen B, Poulsen U, Felbor T, Kallunki T and Boes M: Sensitization to the lysosomal cell death pathway upon immortalization and transformation. Cancer Res 64: 5301-5310, 2004.
48. Joyce JA, Baruch A, Chehade K, *et al*: Cathepsin cysteine proteases are effectors of invasive growth and angiogenesis during multistage tumorigenesis. Cancer Cell 5: 443-453, 2004.
49. Zheng X, Chou PM, Mirkin BL and Rebaa A: Senescence-initiated reversal of drug resistance: specific role of cathepsin L. Cancer Res 64: 1773-1780, 2004.
50. Shiraishi T, Mori M, Tanaka S, Sugimashi K and Akiyoshi T: Identification of cystatin B in human esophageal carcinoma using displays in which the gene expression is related to lymph-node metastasis. Int J Cancer 79:175-178, 1998.
51. Levicar N, Kos J, Blejec A, Golouch R, Vrhovec I, Frkovic-Grazio S and Lah TT: Comparison of potential biological markers cathepsin B, cathepsin L, stefin A and stefin B with urokinase and plasminogen activator inhibitor-1 and clinopathological data of breast carcinoma patients. Cancer Detect Prev 26: 42-49, 2002.
52. Mirtti T, Alanen K, Kallajoki M, Rinne A and Soderstrom KO: Expression of cystatins, high molecular weight cytokeratin, and proliferation markers in prostatic adenocarcinoma and hyperplasia. Prostate 54: 290-298, 2003.
53. Trinkaus M, Vranic A, Dolene VV and Lah TT: Cathepsins B and L and their inhibitors stefin B and Cystatin C as markers for malignant progression of benign meningiomas. Int J Biol Markers 20: 50-59, 2005.
54. Lin YW and Yang JL: Cooperation of ERK and SCFSkp2 for MKP-1 destruction provides a positive feedback regulation of proliferatin signaling. J Biol Chem 281: 915-926, 2006.
55. Li M, Zhou J, Ge Y, Matherly LH and Wu GS: The phosphatase MKP1 is a transcriptional target of p53 involved in cell cycle regulation. J Biol Chem 278: 41059-41068, 2003.
56. Denkert C, Schmitt WD and Berger S: Expression of mitogen-activated protein kinase phosphatase-1 (MKP-1) in primary human ovarian carcinoma. Int J Cancer 102: 507-513, 2002.
57. Rauhala HE, Porkka KP, Tolonen TT, Martikainen PM, Tammela TLJ and Visakorpi T: Dual-specificity phosphatase 1 and serum/glucocorticoid-regulated kinase are downregulated in prostate cancer. Int J Cancer 117: 738-745, 2005.
58. Tomioka H, Morita K, Hasegawa S and Omura K: Gene expression analysis by cDNA microarray in oral squamous cell carcinoma. J Oral Pathol Med 35: 206-211, 2006.
59. Unoki M and Nakamura Y. Growth-suppressive effects of BPOZ and EGR2, two genes involved in the PTEN signaling pathway. Oncogene 20: 4457-4465, 2001.
60. Abraham SM and Clark AR: Dual-specificity phosphatase 1: a critical regulator of innate immune responses. Biochem Soc Trans 34: 1018-1023, 2006.
61. Hardy K, Mansfield L, Mackay A, *et al*: Transcriptional networks and cellular senescence in human mammary fibroblasts. Mol Biol Cell 16: 946-953, 2005.
62. Subramanian RR, Zhang H, Wang H, Ichijo H, Miyashita T and Fu H: Interaction of apoptosis signal-regulating kinase 1 with isoforms of 14-3-3 proteins. Exp Cell Res 294: 581-591, 2004.
63. Mizumura K, Takeda K, Hashimoto S, Horie T and Ichijo H: Identification of Op18/stathmin as a potential target of ASK1-p38 MAP kinase cascade. J Cell Physiol 206: 363-370, 2006.
64. Nagai H, Noguchi T, Takeda K and Ichijo H: Pathophysiological roles of ASK1-MAP kinase signaling pathways. J Biochem Mol Biol 40: 1-16, 2007.
65. Dasgupta P, Betts V, Rastogi S, *et al*: Direct binding of apoptosis signal-regulating kinase 1 to retinoblastoma protein. J Biol Chem 279: 38762-38769, 2004.
66. Kherrouche Z, Blais A, Ferreira E, De Launoit Y and Monté D: ASK-1 (apoptosis signal-regulating kinase 1) is a direct E2F target gene. Biochem J 396: 547-556, 2006.
67. Shim BC, Kang BH, Hong YK, *et al*: The kringle domain of tissue-type plasminogen activator inhibits in vivo tumor growth. Biochem Biophys Res Commun 327:1155-1162, 2005.
68. Pepper MS: Role of matrix metalloproteinase and plasminogen activator-plasmin systems in angiogenesis. Arterioscler Thromb Vasc Biol 21: 1104-1117, 2001.
69. Ferrier CM, Suciu S, van Geloof WL, *et al*: High t-PA-expression in primary melanoma of the limb correlates with good prognosis. Br J Cancer 83: 1351-1359, 2000.
70. Chappuis PO, Dieterich B, Scietta V, Lohse C, Bonnefoi H, Remadi S and Sapino AP: Functional evaluation of plasmin formation in primary breast cancer. J Clin Oncol 19: 2731-2738, 2001.
71. Reijerkerk A, Voest EE and Gebbink MFGB: No grip, no growth: the conceptual basis of excessive proteolysis in the treatment of cancer. Eur J Cancer 36: 16695-16705, 2000.
72. Gingras D, Nyalendo C, Di Tomasso G, Annabi B and Béliveau R: Activation of tissue plasminogen activator gene transcription by Neovastat, a multifunctional antiangiogenic agent. Biochem Biophys Res Commun 16: 205-212, 2004.

as D, Labelle D, Nyalendo C, Boivin D, Demeule M, meuf C and Béliveau R: The antiangiogenic agent Neovastat (AE-941) stimulates tissue plasminogen activator activity. Invest New Drugs 22: 17-26, 2004.
74. Yamamoto, K, Okamoto A, Isonishi S, Ochiai K and Ohtake Y: A novel gene, CRR9, which was up-regulated in CDDP-resistant ovarian tumor cell line, was associated with apoptosis. Biochem Biophys Res Commun 280: 1148-1154, 2001.
75. Asakura T, Imai A, Ohkubo-Uraoka N, *et al*: Relationship between expression of drug-resistance factors and drug sensitivity in normal human renal proximal tubular epithelial cells in comparison with renal cell carcinoma. Oncol Rep 14: 601-607, 2005.
76. Gill G: SUMO and ubiquitin in the nucleus: Different functions, similar mechanisms? Genes Dev 18: 2046-2059, 2004.
77. Chen A, Wang PY, Yang YC, *et al*: SUMO regulates the cytoplasmonuclear transport of its target protein Daxx. J Cell Biochem 98: 895-911, 2006.
78. Li W, Hesabi B, Babbo A, *et al*: Regulation of double-strand break-induced mammalian homologous recombination by UBL1, a RAD51-interacting protein. Nucleic Acids Res 28: 1145-1153, 2000.
79. Lee YS, Jang MS, Lee JS, Choi EJ and Kim E: SUMO-1 represses apoptosis signal-regulating kinase 1 activation through physical interaction and not through covalent modification. EMBO Rep 6: 949-955, 2005.
80. Lichtenfels R, Kellner R, Atkins D, *et al*: Identification of metabolic enzymes in renal cell carcinoma utilizing Proteomex analyses. Biochim Biophys Acta 1646: 21-31, 2003.
81. Xiang, Z: A novel human zinc finger protein ZNF540 interacts with MVP and inhibits transcriptional activities of the ERK signal pathway. Biochem Biophys Res Commun 347: 288-296, 2006.
82. Zhao Y, Zhou L, Liu B, *et al*: ZNF325, a novel human zinc finger protein with a RBaK-like RB-binding domain, inhibits AP-1- and SRE-mediated transcriptional activity. Biochem Biophys Res Commun 346: 1191-1199, 2006.
83. Hanahan D and Weinberg RA: The hallmarks of cancer. Cell 100: 57-70, 2000.
84. Rosenthal EL and Matrisian LM: Matrix metalloproteases in head and neck cancer. Head Neck 28: 639-648, 2006.
85. Gondi CS, Lakka SS, Dinh DH, Olivero WC, Gujrati M and Rao JS: RNAi-mediated inhibition of cathepisin B and uPAR leads to decreased cell invasion, angiogenesis and tumor growth in gliomas. Oncogene 23: 8486-8496, 2004.
86. Kos J, Werle B, Lah T and Brunner N: Cysteine proteinases and their inhibitors in extracellular fluids: markers for diagnosis and prognosis in cancer. Int J Biol Markers 15: 84-89, 2000.