Identification of molecular targets in head and neck squamous cell carcinomas based on genome-wide gene expression profiling

SATOYA SHIMIZU[1,2], NAOHIKO SEKI[2], TAKASHI SUGIMOTO[2], SHIGETOSHI HORIGUCHI[1], HIDEKI TANZAWA[3], TOYOYUKI HANAZAWA[1] and YOSHITAKA OKAMOTO[1]

Departments of [1]Otorhinolaryngology, [2]Functional Genomics and [3]Clinical Molecular Biology, Graduate School of Medicine, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8670, Japan

Received May 21, 2007; Accepted June 28, 2007

Abstract. DNA amplifications activate oncogenes and are hallmarks of nearly all advanced cancers including head and neck squamous cell carcinoma (HNSCC). Some oncogenes show both DNA copy number gain and mRNA overexpression. Chromosomal comparative genomic hybridization and oligonucleotide microarrays were used to examine 8 HNSCC cell lines and a plot of gene expression levels relative to their position on the chromosome was produced. Three highly up-regulated genes, *NT5C3, ANLN* and *INHBA*, were identified on chromosome 7p14. These genes were subjected to quantitative real-time RT-PCR on cDNA and genomic DNA derived from 8 HNSCC cell lines. *ANLN* and *INHBA* showed a strong positive correlation between mRNA expression and genomic DNA levels and a similar relationship was shown for the known oncogene, *EGFR*, at 7p11.2. In clinical samples, *ANLN* and *INHBA* showed a significantly higher expression in tumors than in normal tissues. Patients with high expression levels of *INHBA* had a shorter disease-free survival rate. Therefore, *INHBA* may be a promising prognostic marker of HNSCC.

Introduction

Carcinomas of the head and neck represent the sixth most common cancer worldwide and at least 90% of them are squamous cell carcinomas. Despite considerable advances in surgery, radiotherapy and chemotherapy, the 5-year survival rate for HNSCC patients has improved only marginally over the past decade. Local tumor recurrence, second primary tumors and distant metastasis after conventional therapy appear to be major contributing factors for restricted survival of HNSCC patients. Local tumor recurrence affects ~60% of patients and metastases develop in 15-25% of patients (1). Many factors, such as TNM stage, pathological grade and tumor site, influence the prognosis of HNSCC but are not sufficient to predict outcome. In addition, treatment often results in impairment of functions such as speech and swallowing, cosmetic disfiguration and mental pain. These inflictions significantly erode quality of life. To overcome this situation, there is a need to find novel biomarkers that classify patients into prognostic groups, to aid identification of high-risk patients who may benefit from different treatments.

Comparative genomic hybridization (CGH) has facilitated chromosomal characterization of solid tumors as it can provide detailed information on gains and loss of tumor DNA throughout the entire genome. CGH has been widely used to analyze many types of tumors, including HNSCC (2,3). The pattern of aberration, which comprises the numbers and types of aberrations and the regions that are recurrently altered, is characteristic for each tumor type. Most findings based on CGH, however, have shown only slight DNA copy number aberrations because of the large chromosomal areas detected. Also, identification of tumor-related genes is difficult because an extremely large number of genes are within these regions. In this study, gene expression analyses using DNA microarray was performed and combined with the results of CGH. Gene expression profiling has recently been used to promote rational approaches to therapy as well as to improve diagnosis and prognosis in many tumor types (4-6), including HNSCC (1).

Cancers, including HNSCC, are characterized by a complex pattern of cytogenetic and molecular genetic changes. Several chromosomal regions are recurrently amplified or deleted in these tumors, but little is known about specific underlying genes that could be important mediators in tumor formation or progression. The relationship between change in DNA content and gene expression is unknown. Therefore, identification of affected genes in these loci, elucidation of their functions and association of these genes with cancer progression, are required to fully understand HNSCC tumorigenesis and progression.

Oncogenes can be activated by mutation, structural rearrangement or amplification and tumor suppressor genes may be inactivated in some tumors by methylation and in others by mutation or physical deletion. DNA amplification is an important and common mechanism for oncogene overexpression in many cancers. Some oncogenes have been

Correspondence to: Dr Naohiko Seki, Department of Functional Genomics, Graduate School of Medicine, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8670, Japan
E-mail: naoseki@faculty.chiba-u.jp

Key words: head and neck squamous cell carcinoma, comparative genomic hybridization, microarray, 7p14, ANLN, inhibin ß A, activin A

described as being both amplified and overexpressed, for example, *EGFR* (7p11.2) and *CCND1* (11q13) (7). Therefore, to identify prognostic markers of HNSCC, chromosomal CGH and oligonucleotide microarray analysis was performed on 8 HNSCC cell lines. Up-regulated genes at chromosomal gain regions detected by CGH were identified. Quantitative real-time RT-PCR on cDNA and genomic DNA from 8 HNSCC cell lines was performed for 3 selected genes and expression was further analyzed in clinical samples.

Materials and methods

Cell lines and clinical samples. Eight HNSCC cell lines were used. Six cell lines were derived from oral squamous cell carcinomas (OSCC; Ca9-22, H1, HO1N1, HSC2, HSC3 and Sa3) and 2 were from pharyngeal squamous cell carcinomas (PSCC; Detroit562 and FaDu). Six OSCC cell lines were maintained in DMEM/F12 and 2 PSCC cell lines in RPMI-1640. All cell lines were supplemented with 10% fetal bovine serum and were cultured under 5% CO_2 at 37˚C. Eighteen nonmalignant specimens and 22 HNSCC samples were used for quantitative real-time RT-PCR experiments in order to validate candidate genes. HNSCC samples (total 49 samples) were obtained from surgical specimens collected at Chiba University Hospital (Chiba, Japan).

Genomic DNA extraction and chromosomal CGH. Genomic DNA was extracted using the blood and cell culture DNA kit (Qiagen, Valencia, CA) according to the manufacturer's procedures. CGH was performed as described previously (8). DNA from each cell line was directly labeled with Spectrum Green-dUTP (Vysis Inc., Downers Grove, IL) by nick translation using a commercial kit (Vysis). Normal, sex-matched reference DNA was labeled with SpectrumRed-dUTP (Vysis). Labeled cell line and reference DNAs, together with human Cot-1 DNA, were denatured and hybridized to metaphase spreads that were prepared using a standard protocol. The slides were washed and counterstained with DAPI. Threshold values for detection of genomic imbalances were determined as 0.7 for loss and 1.3 for gains.

Oligonucleotide microarray experiment. Total RNA was isolated using TRIzol reagent (Invitrogen, Carlsbad, CA) following the manufacturer's procedures. Total RNA quality was checked by the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). The AceGene® Human Oligo Chip 30K (Hitachi Software Engineering Co. Ltd., Yokohama, Japan) was used for expression profiling. A total of 30,336 spots corresponding to 29,640 independent genes are contained in this array. Oligonucleotide microarray analysis was performed on 8 HNSCC cell lines. aRNA was prepared with the Amino Allyl MessageAmp aRNA kit (Ambion, Austin, TX). Each aRNA from 8 cell lines was labeled with Cyanine 5, whereas aRNA from Universal Reference Total RNA (Clontech, Mountain View, CA) was labeled with Cyanine 3 as a control. The outline of a protocol is available on the web site (http://www.dna-chip.co.jp/thesis/AceGeneProtocol.pdf). Probe purification, hybridization, and washings were performed according to the manufacturer's instructions. Microarray slides were scanned using the Packard GSI Lumonics scanArray 4000 (Perkin Elmer, Boston, MA). Data were analyzed by DNASISarray software (Hitachi Software Engineering Co. Ltd.), which converted the signal intensity of each spot into text format. Log2-ratios of median local background subtracted intensity levels were analyzed. Each microarray data was normalized by Lowess normalization.

Quantitative real-time RT-PCR on cDNA. For cDNA synthesis, the QuantiTect® reverse transcription kit (Qiagen) was used. Reaction mixtures contained 1 μl of a 1/50 dilution of cDNA, 0.5 μM of each primer, 8 μl H_2O and 10 μl of 2X QuantiTect SYBR Green PCR Master mix (Qiagen). Primers were designed to amplify products of 100-200 bp within target and control sequences. Primers used for real-time PCR on cDNA were: *NT5C3*: 5'-aagaatggcagatggagtgg-3' (sense), 5'-acag ttcaattgcacccaca-3' (antisense); *ANLN*: 5'-caatcttgctgcaa ctatttgct-3' (sense), 5'-tgcttaacactgctgctattga-3' (antisense); *INHBA*: 5'-ggagggcagaaatgaatgaa-3' (sense), 5'-ccttggaa atctcgaagtgc-3' (antisense); *EGFR*: 5'-aaggaaatcctcgatgaag cct-3' (sense), 5'-tgtctttgtgttcccggacata-3' (antisense); *ACTB*: 5'-ggcatgggtcagaaggatt-3' (sense), 5'-aggtgtggtgccagattttc-3' (antisense). Reactions were performed under the following conditions: 95˚C for 15 min and 40 cycles of 94˚C, 15 sec; 55˚C, 30 sec; 72˚C, 34 sec. PCR products were assayed using the 7300 Real-Time PCR system (Applied Biosystems, Foster City, CA). Experiments were performed in duplicate and gene expression levels were normalized to *ACTB* expression levels.

Quantitative real-time RT-PCR on genomic DNA. Primers were quoted from the NCBI database. Primers used for real-time PCR on genomic DNA were: *NT5C3*: 5'-ggaga atctgaagctctcctg-3' (sense), 5'-tttccccacaaacttcatcc-3' (antisense); *ANLN*: 5'-gtgtatttaaaagttgagaccctgc-3' (sense), 5'-aacagtgtctagaaacccattttg-3' (antisense); *INHBA*: 5'-tgtc cagagaagacagtggc-3' (sense), 5'-ctgtttcatctgtttcatcagg-3' (antisense); *EGFR*: 5'-aagcaattccagggtgagg-3' (sense), 5'-ctc agacaaacggtcagaagc-3' (antisense); *SUFU* (10q24): 5'-gctgct tcgatggaaac-3' (sense), 5'-aacctgtgtcggctgc-3' (antisense). Reactions were performed with the same conditions as above. Human female genomic DNA (Novagen, Madison, WI) was used as a normal control. Experiments were performed in duplicate. DNA levels were normalized to *SUFU* (10q24) DNA levels, which were designed within a region of euploid copy number as shown by CGH analysis.

Statistical analysis. Microarray expression data and CGH results were combined by arranging gene expression levels on each chromosome. Map positions of the genes were obtained from the NCBI database (http://www.ncbi.nlm.nih.gov/genome/guide/human/). Spearman's rank correlation coefficient (r_s) was used to measure and to test the correlation between two random variables. Relative expression levels obtained by quantitative real-time RT-PCR were subjected to the Mann-Whitney U test to identify differences in gene expression levels between 2 groups. Sample clusters were compared to clinicopathological data of patients using Fisher's exact test. The Kaplan-Meier method was used to plot disease-free survival and the log-rank test was used to compare curves. Statistical significance was defined as P-value of <0.05.

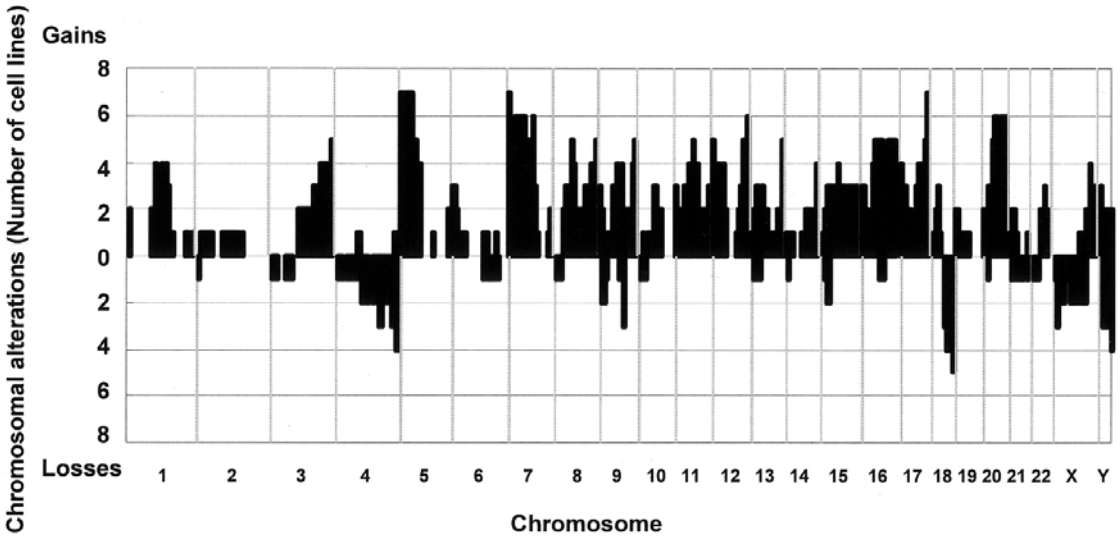

Figure 1. Genomic imbalances detected in 8 HNSCC cell lines. The upper bars indicate the frequency of gains, and the under bars indicate the frequency of losses. The most recurrent regions of DNA copy number gains are on chromosomes 5p, 7p and 20q, while recurrent regions of copy number loss are on chromosomes 4q, 18q and Y.

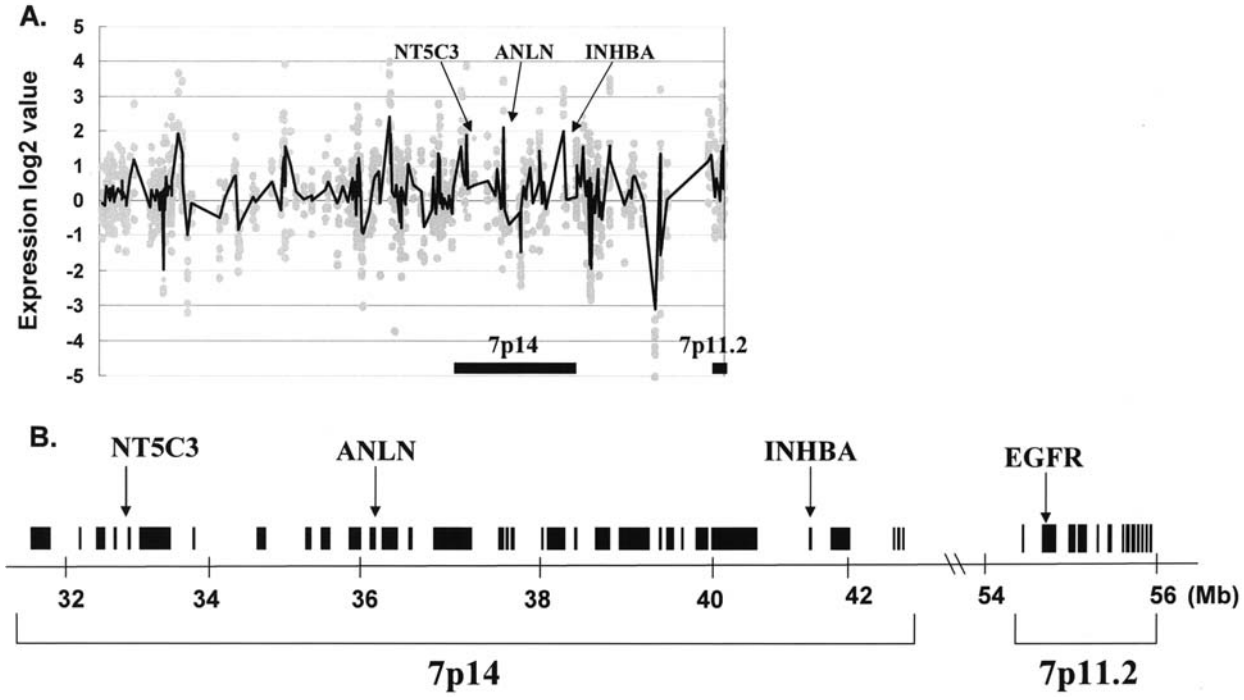

Figure 2. (A) Gene expression levels relative to their position along chromosome arm 7p. The individual values for each gene of 8 HNSCC cell lines in microarray are gray points and the moving average of gene expression values is indicated by a heavy line. Black bars indicate candidate regions. Not all genes included in candidate regions had the tendency to be up-regulated and there were both up- and down-regulated genes in each region. (B) Localization of the genes analyzed on chromosome band 7p14 and 7p11.2. From the 30K microarray, 33 genes were contained in 7p14.

Results

Chromosomal CGH performed on 8 HNSCC cell lines. All cell lines had some chromosomal alterations detected by CGH and many of the copy number alterations were common to most cell lines (Fig. 1). The most frequent chromosomal copy number gains were observed on chromosome arms 5p (7/8 cell lines), 7p (7p22: 7/8; 7p21-p15: 5/8; 7p14-p11.2: 6/8) and 20q (6/8). Chromosomal copy number losses were observed most frequently on chromosome arms 4q (4q27-q33: 3/8; 4q34-q35: 4/8), 18q (18q12: 3/8; 18q21-q22: 4/8; 18q23: 5/8) and Y (Yp: 3/8; Yq: 4/8). Chromosome copy number losses occurred significantly less often than gains.

Gene expression profiles in gain regions detected by chromosomal CGH. Three candidate chromosomal gain regions, altered by CGH, were selected: 5p, 7p and 20q. From the 30K microarray, the number of genes mapped to these regions

Table I. Gene expression profiles of chromosome arms 5p, 7p and 20q.

Chromosomal region	CGH gain	Size (Mb)	No. of transcripts	No. of transcripts on AceGene Human 30K		Average expression level (fold)
				Total	>2-fold	
5p15.3	7/8	8	60	30	6	1.517
5p15.2	7/8	7	25	11	2	1.485
5p15.1	7/8	4	44	5	0	1.479
5p14	7/8	10	25	3	0	1.003
5p13	7/8	13	82	46	10	1.834
5p12	7/8	4	25	12	3	1.739
7p22	7/8	7	104	48	1	1.216
7p21	5/8	13	49	25	7	1.640
7p15	5/8	12	114	63	10	1.431
7p14	6/8	11	87	33	10	1.824
7p13	6/8	4	47	26	4	1.406
7p12	6/8	7	34	11	1	1.273
7p11.2	6/8	3	46	13	6	2.185
20q11.2	6/8	9	156	106	6	1.174
20q12	6/8	4	14	7	2	1.531
20q13.1	6/8	8	135	94	5	1.242
20q13.2	6/8	5	21	12	2	1.394
20q13.3	6/8	8	125	91	6	1.211

was 107 genes in 5p, 219 genes in 7p and 310 genes in 20q. More detailed data of these regions are shown in Table I. The cutoff value of microarray expression levels was set at 2-fold and 81 genes were up-regulated >2-fold in these regions: 21 genes in 5p, 39 genes in 7p and 21 genes in 20q (Table II). These up-regulated genes located with high number (10 genes) at 5p13, 7p15 and 7p14 and located with high probability at 7p14 (10/33) and 7p11.2 (6/13). Average expression levels of the regions were 1.834-fold at 5p13, 1.824-fold at 7p14 and 2.185-fold at 7p11.2. These values were significantly higher than the average of all genes, which was 1.162-fold. These results did not seem to be related to the length of chromosomal region or the percentage of genes located on the chromosome. From these analyses, 5p13, 7p14 and 7p11.2 were likely candidate chromosomal regions. Among these three regions, we focused on 7p14, which is a novel chromosomal gain region of HNSCC.

Candidate genes within 7p14. Gene expression values, relative to their position along the chromosome, were plotted. Fig. 2A shows the gene expression values of chromosome arm 7p obtained from microarray analysis of 8 HNSCC cell lines and the average gene expression values. Although not all genes in 7p14 had the same tendency to be up-regulated, and there were both up- and down-regulated genes in each region, the average expression values of the genes contained in 7p14 (also in 7p11.2) was significantly higher than other 7p regions (P<0.01; Table I).

The 33 genes that were altered on the microarray were contained in 7p14 (Fig. 2B) and among these, the 3 most up-regulated genes (*NT5C3, ANLN* and *INHBA*) and *EGFR* were subjected to quantitative real-time RT-PCR. *EGFR* (epidermal growth factor receptor) is overexpressed in several epithelial malignancies, including HNSCC. The *EGFR* gene plays a critical role in HNSCC growth, invasion, metastasis and angiogenesis so this gene was used as a comparison.

Q*uantitative real-time RT-PCR verification of cell lines*. mRNA expression levels of the cell lines were compared with Human Universal Reference Total RNA, while genomic DNA expression levels of the cell lines were compared with human female genomic DNA (Fig. 3). Among the 3 up-regulated genes, *ANLN* and *INHBA* showed strong positive correlation of mRNA expression levels and genomic DNA expression levels in each cell line (*ANLN*: r_s=0.786, P=0.038; *INHBA*: r_s=0.857, P=0.023). A strong correlation for *EGFR* was also found (r_s=0.905, P=0.017).

Quantitative real-time RT-PCR using clinical samples. Next, mRNA expression levels of the 3 up-regulated genes and *EGFR* in normal squamous epithelial tissues and HNSCC specimens of patients were compared (Fig. 4). Quantitative real-time RT-PCR of 18 normal tissues and 22 tumors demonstrated that *ANLN* and *INHBA* were significantly overexpressed in tumors compared to normal tissues (P<0.001). There was also a statistically significant overexpression of *EGFR* in tumors. Furthermore, when 18 tumors were compared to paired normal tissues, both *ANLN* and *INHBA* expression levels were higher in each of the tumor samples. The average tumor/normal tissue ratios were 5.5 and 28.6, respectively.

Table II. Up-regulated genes >2-fold in 5p, 7p and 20q.

No.	Symbol	Gene name	Accession	Location	Fold in array	Gene function
1	*PDCD6*	Programmed cell death 6	NM_013232	5pter-p15.2	2.11	Apoptosis
2	*Cep72*	Centrosomal protein 72 kDa	NM_018140	5p15.33	2.40	Unknown
3	*KIAA0947*	KIAA0947 protein	XM_029101	5p15.32	4.14	Unknown
4	*NSUN2*	NOL1/NOP2/Sun domain family, member 2	NM_017755	5p15.31	2.84	Unknown
5	*MGC5297*	Hypothetical protein MGC5297	NM_024091	5p15.3-p15.2	3.63	Unknown
6	*MTRR*	5-Methyltetrahydrofolate-homocysteine methyltransferase reductase	NM_024010	5p15.3-p15.2	2.09	Metabolism
7	*CCT5*	Chaperonin containing TCP1, subunit 5ε	NM_012073	5p15.2	4.38	Metabolism
8	*MARCH6*	Membrane-associated ring finger (C3HC4) 6	NM_005885	5p15.2	2.70	Unknown
9	*SUB1*	SUB1 homolog *(S. cerevisiae)*	NM_006713	5p13.3	2.77	Metabolism
10	*TARS*	Threonyl-tRNA synthetase	NM_152295	5p13.2	10.77	Metabolism
11	*RAI14*	Retinoic acid induced 14	NM_015577	5p13.3-p13.2	2.05	Unknown
12	*BXDC2*	Brix domain containing 2	NM_018321	5p13.2	3.94	Ribosome biogenesis
13	*SKP2*	S-phase kinase-associated protein 2 (p45)	NM_032637	5p13	6.66	Cell cycle
14	*SLC1A3*	Solute carrier family 1 (glial high affinity glutamate transporter), member 3	NM_004172	5p13	2.98	Transport
15	*NIPBL*	Nipped-B homolog *(Drosophila)*	NM_015384	5p13.2	5.86	Cell cycle
16	*WDR70*	WD repeat domain 70	NM_018034	5p13.2	2.21	Unknown
17	*OSRF*	Osmosis responsive factor	NM_012382	5p15.2-p12	2.05	Unknown
18	*OXCT1*	3-Oxoacid CoA transferase 1	NM_000436	5p13.1	2.71	Metabolism
19	*PAIP1*	poly(A) binding protein interacting protein 1	NM_006451	5p12	2.54	Metabolism
20	*NNT*	Nicotinamide nucleotide transhydrogenase	NM_012343	5p13.1-5cen	2.73	Transport
21	*FGF10*	Fibroblast growth factor 10	NM_004465	5p13-p12	2.19	Cell cycle
22	*SDK1*	Sidekick homolog 1 (chicken)	NM_152744	7p22.2	2.68	Cell adhesion
23	*C1GALT1*	Core 1 synthase, glycoprotein-N-acetylgalactosamine 3-ß-galactosyltransferase, 1	NM_020156	7p14-p13	5.05	Unknown
24	*FLJ20323*	Hypothetical protein FLJ20323	NM_019005	7p22-p21	3.10	Unknown
25	*RPA3*	Replication protein A3, 14 kDa	NM_002947	7p22	2.24	Metabolism
26	*ANKMY2*	Ankyrin repeat and MYND domain containing 2	NM_020319	7p21	2.23	Unknown
27	*BZW2*	Basic leucine zipper and W2 domains 2	NM_014038	7p21.1	2.87	Unknown
28	*AGR2*	Anterior gradient 2 homolog *(Xenopus laevis)*	NM_006408	7p21.3	3.62	Unknown
29	*AHR*	Aryl hydrocarbon receptor	NM_001621	7p15	2.42	Signal transduction
30	*NUPL2*	Nucleoporin-like 2	NM_007342	7p15	2.38	Transport
31	*GPNMB*	Glycoprotein (transmembrane) nmb	NM_002510	7p15	2.07	Cell proliferation
32	*IGF2BP3*	Insulin-like growth factor 2 mRNA binding protein 3	NM_006547	7p11	2.77	Metabolism
33	*CYCS*	Cytochrome c, somatic	NM_018947	7p15.3	2.03	Apoptosis
34	*NFE2L3*	Nuclear factor (erythroid-derived 2)-like 3	NM_004287	7p15-p14	7.48	Metabolism
35	*HNRPA2B1*	Heterogeneous nuclear ribonucleoprotein A2/B1	NM_031243	7p15	2.41	Metabolism
36	*CBX3*	Chromobox homolog 3 (HP1 γ homolog, *Drosophila*)	NM_007276	7p15.2	4.91	Metabolism
37	*TAX1BP1*	Tax1 (human T-cell leukemia virus type I) binding protein 1	NM_006024	7p15	2.45	Unknown
38	*C7orf24*	Chromosome 7 open reading frame 24 (scraps homolog, *Drosophila*)	NM_024051	7p15-p14	3.15	Unknown
39	*GARS*	Glycyl-tRNA synthetase	NM_002047	7p15	2.67	Metabolism
40	*LSM5*	LSM5 homolog, U6 small nuclear RNA associated *(S. cerevisiae)*	NM_012322	7p14.3	3.65	Metabolism
41	*KIAA0241*	KIAA0241	NM_015060	7p14.3	2.01	Unknown
42	*NT5C3*	5''-nucleotidase, cytosolic III	NM_016489	7p14.3	4.77	Metabolism
43	*SEPT7*	Septin 7	NM_001788	7p14.3-p14.1	2.14	Cell cycle

Table II. Continued.

No.	Symbol	Gene name	Accession	Location	Fold in array	Gene function
44	*ANLN*	Anillin, actin binding protein	NM_018685	7p15-p14	5.92	Cell division
45	*VPS41*	Vacuolar protein sorting 41 (yeast)	NM_014396	7p14-p13	2.20	Transport
46	*RALA*	V-ral simian leukemia viral oncogene homolog A (ras-related)	NM_005402	7p15-p13	3.30	Signal transduction
47	*INHBA*	Inhibin, ß A (activin A, activin AB α polypeptide)	NM_002192	7p15-p13	5.36	Cell differentiation
48	*PSMA2*	Proteasome (prosome, macropain) subunit, α type, 2	NM_002787	7p14.1	2.31	Metabolism
49	*MRPL32*	Mitochondrial ribosomal protein L32	NM_031903	7p14	2.04	Metabolism
50	*STK17A*	Serine/threonine kinase 17a (apoptosis-inducing)	NM_004760	7p12-p14	3.69	Apoptosis
51	*H2AFV*	H2A histone family, member V	NM_012412	7p13	2.13	Metabolism
52	*IGFBP1*	Insulin-like growth factor binding protein 1	NM_000596	7p13-p12	3.81	Cell growth
53	*IGFBP3*	Insulin-like growth factor binding protein 3	NM_000598	7p13-p12	2.68	Cell growth
54	*FIGNL1*	Fidgetin-like 1	NM_022116	7p12.2	3.10	Unknown
55	*SEC61G*	Sec61 γ subunit	NM_014302	7p11.2	2.61	Transport
56	*EGFR*	Epidermal growth factor receptor	NM_005228	7p12	2.99	Cell proliferation
57	*MRPS17*	Mitochondrial ribosomal protein S17	NM_015969	7p11	3.33	Metabolism
58	*PSPH*	Phosphoserine phosphatase	NM_004577	7p15.2-p15.1	3.62	Metabolism
59	*CCT6A*	Chaperonin containing TCP1, subunit 6A (ζ 1)	NM_001762	7p11.2	3.78	Metabolism
60	*CHCHD2*	Coiled-coil-helix-coiled-coil-helix domain containing 2	NM_016139	7p11.2	2.37	Unknown
61	*TPX2*	TPX2, microtubule-associated, homolog *(Xenopus laevis)*	NM_012112	20q11.2	2.73	Cell cycle
62	*TM9SF4*	Transmembrane 9 superfamily protein member 4	NM_014742	20q11.21	2.08	Transport
63	*EIF2S2*	Eukaryotic translation initiation factor 2, subunit 2 ß, 38 kDa	NM_003908	20pter-q12	3.52	Metabolism
64	*RNPC2*	RNA-binding region (RNP1, RRM) containing 2	NM_004902	20q11.22	3.80	Metabolism
65	*RPN2*	Ribophorin II	NM_002951	20q12-q13.1	3.48	Metabolism
66	*FAM83D*	Family with sequence similarity 83, member D	NM_030919	20q11.22-q12	2.33	Metabolism
67	*TOP1*	Topoisomerase (DNA) I	NM_003286	20q12-q13.1	2.65	Metabolism
68	*CHD6*	Chromodomain helicase DNA binding protein 6	NM_032221	20q12	2.05	Unknown
69	*UBE2C*	Ubiquitin-conjugating enzyme E2C	NM_007019	20q13.12	5.53	Cell cycle
70	*ZNF334*	Zinc finger protein 334	NM_018102	20q13.12	2.18	Metabolism
71	*CSE1L*	CSE1 chromosome segregation 1-like (yeast)	NM_001316	20q13	2.36	Cell proliferation
72	*DDX27*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 27	NM_017895	20q13.13	2.05	Unknown
73	*DPM1*	Dolichyl-phosphate mannosyltransferase polypeptide 1, catalytic subunit	NM_003859	20q13.13	3.26	Metabolism
74	*SALL4*	Sal-like 4 *(Drosophila)*	NM_020436	20q13.13-q13.2	2.24	Metabolism
75	*PFDN4*	Prefoldin subunit 4	NM_002623	20q13.2	2.31	Metabolism
76	*AURKA*	Aurora kinase A	NM_003600	20q13.2-q13.3	4.29	Cell cycle
77	*RAB22A*	RAB22A, member RAS oncogene family	NM_020673	20q13.32	2.03	Transport
78	*STX16*	Syntaxin 16	NM_003763	20q13.32	2.88	Transport
79	*C20orf45*	Chromosome 20 open reading frame 45	NM_016045	20q13.32	3.44	Unknown
80	*PSMA7*	Proteasome (prosome, macropain) subunit, α type, 7	NM_002792	20q13.33	2.57	Metabolism
81	*RPS21*	Ribosomal protein S21	NM_001024	20q13.3	2.19	Metabolism

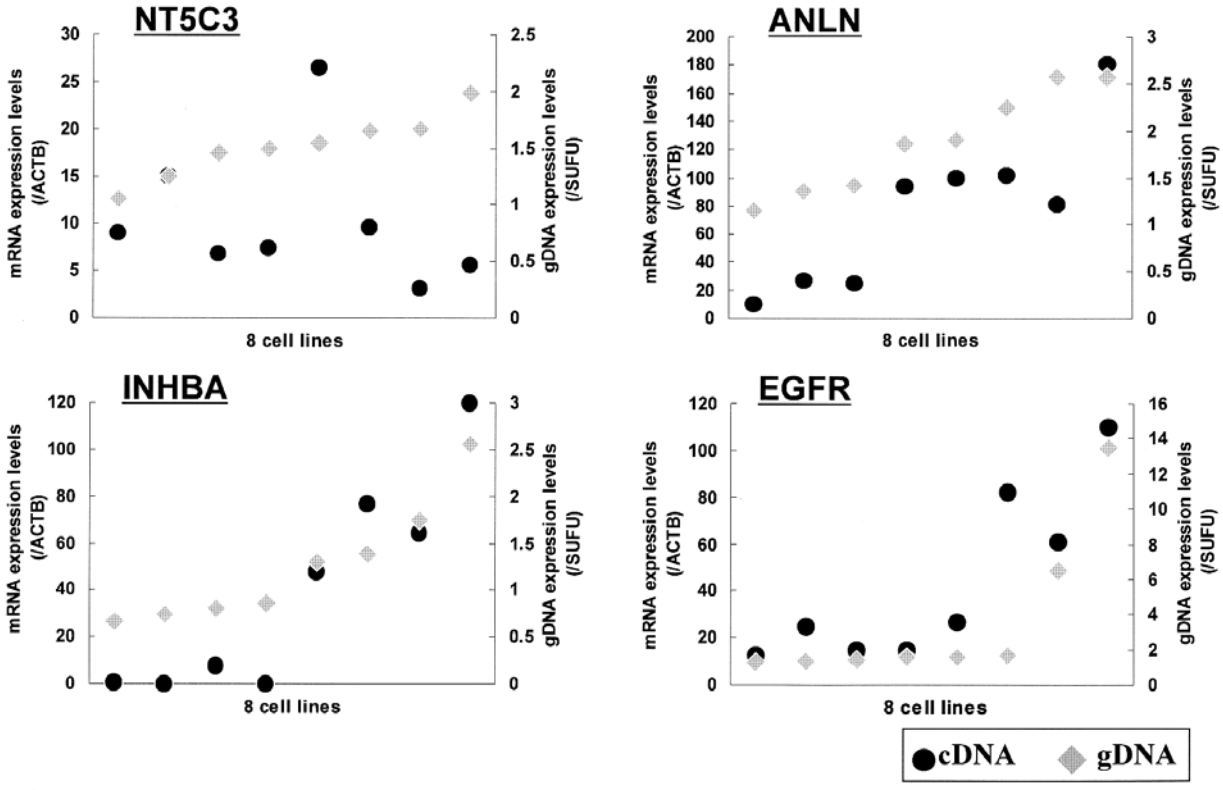

Figure 3. Quantitative real-time RT-PCR analysis on cDNA and genomic DNA (gDNA) of 3 up-regulated genes and *EGFR* gene expression in 8 HNSCC cell lines (random order). Left axis shows mRNA expression levels and right axis shows gDNA expression levels. The relative mRNA and gDNA expression levels were normalized to the amount of *ACTB* and *SUFU*, respectively. mRNA expression levels of cell lines were compared with Human Universal Reference Total RNA and genomic DNA expression levels of cell lines were compared with Human Female Genomic DNA. *ANLN, INHBA* and *EGFR* showed a strong positive correlation between mRNA expression levels and gDNA expression levels in each cell line. Statistical significance was determined by the Spearman's rank correlation coefficient.

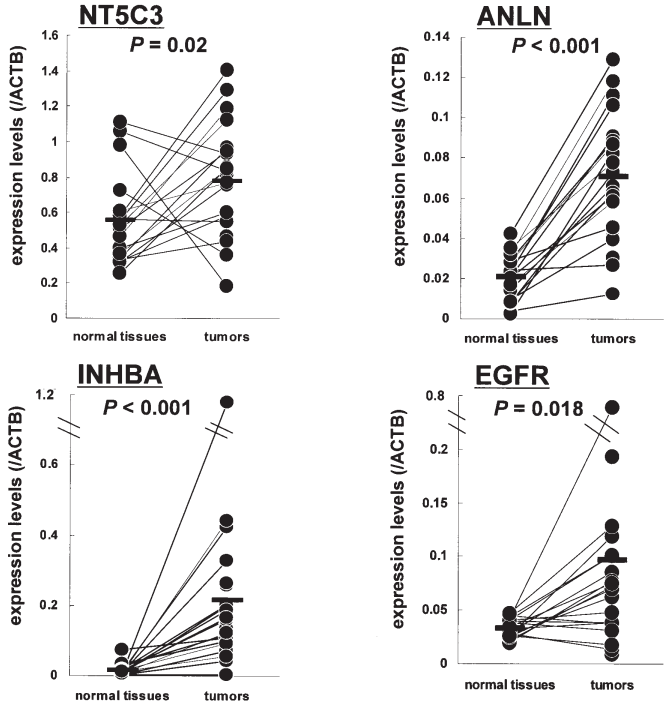

Figure 4. Quantitative real-time RT-PCR analysis of 3 up-regulated genes and *EGFR* expression in clinical samples. Relative mRNA expression examined was normalized to the amount of *ACTB*. Using samples from 18 normal tissues and 22 tumors, *ANLN* and *INHBA* were significantly overexpressed in tumors compared to normal tissues. Statistical significance was determined by the Mann-Whitney U test.

NT5C3 was also overexpressed in tumors compared to normal tissues (P=0.02).

Correlation between ANLN and INHBA expression and clinicopathological characteristics. The expression of *ANLN* and *INHBA* in 49 primary tumors was correlated to clinicopathological characteristics to determine their interrelationship (Table III). Five variables (age, gender, histological grading, T classification, lymph node involvement) were analyzed by Fisher's exact test. Since the average expression levels of *ANLN* and *INHBA* in 49 tumors was 0.057 and 0.188, respectively, samples were separated into two groups: a 'low' expression group, in which expression levels were under the average level and a 'high' expression group, in which expression levels were greater than the average level. Statistical correlations were identified between *ANLN* overexpression and histological grading (P=0.043) and between *INHBA* overexpression and nodal metastasis (P=0.032). For *INHBA*, Kaplan-Meier analysis on 49 primary tumors revealed a significant difference in the time to disease-free survival in the two groups. The 'high' *INHBA* expression group showed a lower survival rate (log-rank test, P=0.014; Fig. 5). A relationship with prognosis was not detected for *ANLN*.

Discussion

Using chromosomal CGH, chromosomal copy number gains were frequently observed on chromosome arms 5p, 7p and

Table III. Correlation between *ANLN* and *INHBA* expression and clinicopathological characteristics.

Factors	*ANLN* expression		P-value[a]	*INHBA* expression		P-value[a]
	Low	High		Low	High	
All cases	27	22		30	19	
Age (years)						
<60	12	4	0.07	11	5	0.54
>60	15	18		19	14	
Gender						
Male	18	18	0.33	21	15	0.53
Female	9	4		9	4	
Histological grading						
G1	10	2	0.043	9	3	0.32
G2 or G3	17	20		21	16	
T classification (pathological)						
T1-2	11	11	0.57	15	7	0.4
T3-4	16	11		15	12	
Lymph node involvement						
Negative	11	7	0.56	15	3	0.032
Positive	16	15		15	16	

[a]P-values were derived from Fisher's exact test.

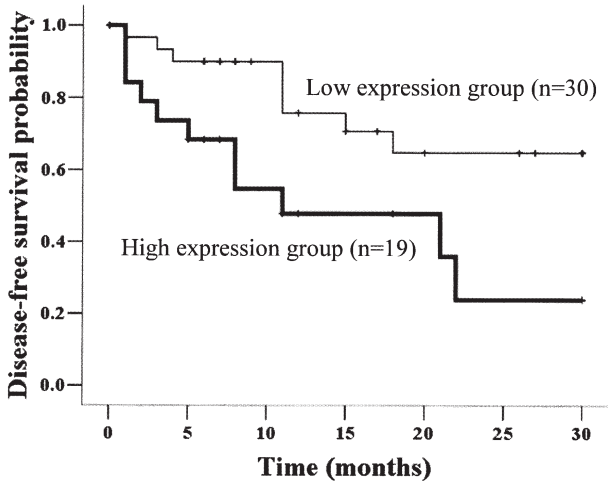

Figure 5. Kaplan-Meier analysis of disease-free survival in patients with HNSCC. Kaplan-Meier analysis of disease-free survival in 49 HNSCC patients is categorized as 'low' and 'high' *INHBA* expression groups. A thin line indicates the 'low' expression group and the heavy line is the 'high' expression group. The 'high' *INHBA* expression group showed a lower survival rate (P=0.014).

20q in 8 HNSCC cell lines. Many regions of copy number alterations matched the regions described in the published literature for HNSCC (2,3) and are also characteristic of other squamous cell carcinomas, such as esophagus, cervix and anal tumors (9). In HNSCC, DNA copy number gains were frequently observed on these chromosome arms as well as the most common gain regions, 3q and 11q13. In this study, 3q26-29 and 11q12-14 were also involved in chromosomal gain regions in >50% of the 8 cell lines. For HNSCC, there are studies that a gain of 5p14 was associated with short-term survival after surgery (10). In addition, the average disease-free survival of OSCC without 7p copy number gain clearly exceeded that of tumors with a gain at 7p (11). No such relationship has been found for 20q.

On chromosome arms 5p, 7p and 20q, 81 genes were identified that were up-regulated >2-fold by microarray analysis. When these microarray results were classified according to their function, 21 of 81 genes were known genes with no functional annotation. Of the remaining 60 genes, 52% (31/60) were involved in metabolism, 13% (8/60) in transport, 12% (7/60) in cell cycle, 5% (3/60) in apoptosis and cell proliferation and 3% (2/60) in cell growth and signal transduction. These functions are essential for normal cellular processes but may be also implicated in development and progression of HNSCC. There are reports that about half of these 81 genes are related to malignant tumors and at least 15 genes are related specifically to squamous cell carcinomas. This gene list will aid in the search for new HNSCC-related genes.

Quantitative real-time RT-PCR on 3 of the up-regulated genes (*NT5C3, ANLN* and *INHBA*) contained in 7p14 was performed. *ANLN* and *INHBA* relative gene copy numbers were significantly correlated with mRNA expression levels, a finding similar to *EGFR*, a gene known to be overexpressed in epithelial malignancies. Also, in the analysis using clinical samples, mRNA expression levels of these two genes were significantly higher in all 18 tumors than matched normal tissues. Furthermore, expression levels of *ANLN* and *INHBA*

in 49 primary tumors were correlated to clinicopathological characteristics. Statistically significant correlations were identified between *ANLN* overexpression and tumor differentiation and between *INHBA* overexpression and nodal metastasis. Nodal metastasis is the most important predictor of outcome in HNSCC. Kaplan-Meier analysis of 49 primary tumors revealed that patients with tumors exhibiting high *INHBA* expression had a poor survival rate.

INHBA (Inhibin ß A, Activin A) functions in the transforming growth factor ß (TGFß) signaling pathway through *DPC4* (12). Pancreatic, prostate and ovarian cancers overexpress Activin A and patients with endometrial and cervical cancers have high serum levels of Activin A (13-16). Furthermore, *INHBA* was markedly overexpressed in stage IV colon cancer (17). Overexpression of *INHBA* was not the result of *SMAD4*/*DPC4* mutations and it remains unclear if loss of expression of *DPC4* up-regulates *INHBA*. *INHBA* also stimulates inflammatory corneal angiogenesis by increasing vascular endothelial growth factor *(VEGF)* levels (18). *VEGF* expression may have prognostic significance for patients with HNSCC (19) but whether *INHBA* is involved in regulating *VEGF* overexpression in HNSCC needs to be established.

This is the first investigation of these genes in HNSCC and our results suggest that *INHBA* may be a promising candidate as an HNSCC prognostic marker. Functional analysis of *INHBA* for HNSCC prognosis may determine whether it could potentially serve as a diagnostic or therapeutic target. Studies to analyze the other chromosomal gain regions identified, 5p and 20q (especially 5p13), are in progress.

In conclusion, combining chromosomal CGH and oligonucleotide microarray analysis, 3 highly up-regulated genes in 7p14 were found. *ANLN* and *INHBA* not only showed strong positive correlation between mRNA expression levels and genomic DNA expression levels in each cell line, but were also significantly overexpressed in tumors in clinical samples. *INHBA* expression may represent a new prognostic factor of HNSCC. Our comprehensive expression profiling data of HNSCC provide new insights into the molecular biology of HNSCC.

Acknowledgements

This study was supported by grants-in-aid for COE research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

1. Choi P and Chen C: Genetic expression profiles and biologic pathway alterations in head and neck squamous cell carcinoma. Cancer 104: 1113-1128, 2005.
2. Speicher MR, Howe C, Crotty P, du-Manoir S, Costa J and Ward DC: Comparative genomic hybridization detects novel deletions and amplifications in head and neck squamous cell carcinomas. Cancer Res 55: 1010-1013, 1995.
3. Bockmuhl U, Schluns K, Schmidt S, Matthias S and Petersen I: Chromosomal alterations during metastasis formation of head and neck squamous cell carcinoma. Gene Chromosomes Cancer 33: 29-35, 2002.
4. Halvorsen OJ, Oyan AM, Bo TH, Olsen S, Rostad K, Haukaas SA, Bakke AM, Marzolf B, Dimitrov K, Stordrange L, Lin B, Jonassen I, Hood L, Akslen LA and Kalland KH: Gene expression profiles in prostate cancer: association with patient subgroups and tumour differentiation. Int J Oncol 26: 329-336, 2005.
5. Nishidate T, Katagiri T, Lin ML, Mano Y, Miki Y, Kasumi F, Yoshimoto M, Tsunoda T, Hirata K and Nakamura Y: Genome-wide gene-expression profiles of breast-cancer cells purified with laser microbeam microdissection: identification of genes associated with progression and metastasis. Int J Oncol 25: 797-819, 2004.
6. Inoue Y, Shirane M, Miki C, Hiro J, Tanaka K, Kobayashi M, Mori K, Yanagi H and Kusunoki M: Gene expression profiles of colorectal carcinoma in neo-adjuvant chemotherapy. Int J Oncol 25: 1641-1649, 2004.
7. Reissmann PT, Koga H, Figlin RA, Holmes EC and Slamon DJ: Amplification and overexpression of the cyclin D1 and epidermal growth factor receptor genes in non-small-cell lung cancer. J Cancer Res Clin Oncol 125: 61-70, 1999.
8. Hashimoto Y, Oga A, Okami K, Imate Y, Yamashita Y and Sasaki K: Relationship between cytogenetic aberrations by CGH coupled with tissue microdissection and DNA ploidy by laser scanning cytometry in head and neck squamous cell carcinoma. Cytometry 40: 161-66, 2000.
9. Myllykangas S, Himberg J, Bohling T, Nagy B, Hollmen J and Knuutila S: DNA copy number amplification profiling of human neoplasms. Oncogene 25: 7324-7332, 2006.
10. Liehr T, Ries J, Wolff E, Fiedler W, Dahse R, Ernst G, Steininger H, Koscielny S, Girod S and Gebhart E: Gain of DNA copy number on chromosomes 3q26-qter and 5p14-pter is a frequent finding in head and neck squamous cell carcinomas. Int J Mol Med 2: 173-179, 1998.
11. Gebhart E, Ries J, Wiltfang J, Liehr T and Efferth T: Genomic gain of the epidermal growth factor receptor harboring band 7p12 is part of a complex pattern of genomic imbalances in oral squamous cell carcinomas. Arch Med Res 35: 385-394, 2004.
12. Zhang Y, Musci T and Derynck R: The tumor suppressor Smad4/DPC4 as a central mediator of Smad function. Curr Biol 7: 270-276, 1997.
13. Kleeff J, Ishiwata T, Friess H, Buchler MW and Korc M: Concomitant over-expression of activin/inhibin beta subunits and their receptors in human pancreatic cancer. Int J Cancer 77: 860-868, 1998.
14. Thomas TZ, Wang H, Niclasen P, O'Bryan MK, Evans LW, Groome NP, Pedersen J and Risbridger GP: Expression and localization of activin subunits and follistatins in tissues from men with high grade prostate cancer. J Clin Endocrinol Metab 82: 3851-3858, 1997.
15. Zheng W, Luo MP, Welt C, Lambert-Messerlian G, Sung CJ, Zhang Z, Ying SY, Schneyer AL, Lauchlan SC and Felix JC: Imbalanced expression of inhibin and activin subunits in primary epithelial ovarian cancer. Gynecol Oncol 69: 23-31, 1998.
16. Petraglia F, Florio P, Luisi S, Gallo R, Gadducci A, Vigano P, Di Blasio AM, Genazzani AR and Vale W: Expression and secretion of inhibin and activin in normal and neoplastic uterine tissues. High levels of serum activin A in women with endometrial and cervical carcinoma. J Clin Endocrinol Metab 83: 1194-1200, 1998.
17. Wildi S, Kleeff J, Maruyama H, Maurer CA, Buchler MW and Korc M: Overexpression of activin A in stage IV colorectal cancer. Gut 49: 409-417, 2001.
18. Poulaki V, Mitsiades N, Kruse FE, Radetzky S, Iliaki E, Kirchhof B and Joussen AM: Activin a in the regulation of corneal neovascularization and vascular endothelial growth factor expression. Am J Pathol 164: 1293-1302, 2004.
19. Kyzas PA, Stefanou D, Batistatou A and Agnantis NJ: Prognostic significance of VEGF immunohistochemical expression and tumor angiogenesis in head and neck squamous cell carcinoma. J Cancer Res Clin Oncol 131: 624-630, 2005.