Cancer Research and Treatment 2003;35(4):304-313

New Approaches to Functional Process Discovery in HPV 16-Associated Cervical Cancer Cells by Gene Ontology

Yong-Wan Kim, Ph.D.[1], **Min-Je Suh**, M.S.[1], **Jin-Sik Bae**, M.S.[1], **Su Mi Bae**, M.S.[1], **Joo Hee Yoon**, M.D.[2], **Soo Young Hur**, M.D.[2], **Jae Hoon Kim**, M.D.[2], **Duck Young Ro**, M.D.[2], **Joon Mo Lee**, M.D.[2], **Sung Eun Namkoong**, M.D.[2], **Chong Kook Kim**, Ph.D.[3] and **Woong Shick Ahn**, M.D.[2]

[1]Catholic Research Institutes of Medical Science, [2]Department of Obstetrics and Gynecology, College of Medicine, The Catholic University of Korea, Seoul; [3]College of Pharmacy, Seoul National University, Seoul, Korea

Purpose: **This study utilized both mRNA differential display and the Gene Ontology (GO) analysis to characterize the multiple interactions of a number of genes with gene expression profiles involved in the HPV-16-induced cervical carcinogenesis.**

Materials and Methods: **mRNA differential displays, with HPV-16 positive cervical cancer cell line (SiHa), and normal human keratinocyte cell line (HaCaT) as a control, were used. Each human gene has several biological functions in the Gene Ontology; therefore, several functions of each gene were chosen to establish a powerful cervical carcinogenesis pathway. The specific functions assigned to these genes were then correlated with the gene expression patterns.**

Results: **The results showed that 157 genes were up- or down-regulated at least 2-fold and organized into reciprocally dependent sub-function sets, depending on their cervical cancer pathway, suggesting the potentially significant genes of unknown function affected by the HPV-16-derived pathway. The GO analysis suggested that the cervical cancer cells underwent repression of the cancer-specific cell adhesive properties. Also, genes belonging to DNA metabolism, such as DNA repair and replication, were strongly down-regulated, whereas significant increases were shown in the protein degradation and synthesis.**

Conclusion: **The GO analysis can overcome the complexity of the gene expression profile of the HPV-16-associated pathway, identify several cancer-specific cellular processes and genes of unknown function. It could also become a major competing platform for the genome-wide characterization of carcinogenesis.** ***(Cancer Research and Treatment 2003;35:304-313)***

Key Words: **Cervix neoplasia, mRNA differential display, Gene ontology**

INTRODUCTION

Human papillomaviruses (HPV-16 and -18) infection has been commonly identified in cervical carcinomas (1). Generally, after a high-risk HPV infection, the E6 and E7 oncoproteins are consistently expressed, and essential for the immortalization and transformation of human squamous epithelial cells (2,3). Although a number of cervical cancer-related genes and cellular processes have been studied, many of the molecular events involved in the cervical cancer pathway are still unclear, as a gene can be involved in multiple independently regulated cancer-specific pathways. Moreover, many studies on cervical cancer lack physiological relevance because those were performed by using the small number of genes and established only in cell lines.

While it is becoming increasingly clear that there are wide variations in the efficiency of cancer therapy among different cell types, there is still relatively little known regarding the mechanism by which genes or gene complexes are directly cancer-specific. Here, to quantitatively understand the possible multiple relationships between differentially expressed profiles of a gene and the cervical cancer-specific pathway, the annotation project, directed by the Gene Ontology (GO) Consortium (http://www.geneontology.org), was used (4). Despite the significance of functional analysis in cancer research, the GO analysis has not been widely used in carcinogenesis, mainly due to its complexity and rapidly evolving property. With the GO analysis, the regulated genes are organized into three separated ontologies comprised of; biological process, cellular component and molecular function, and defines a set of well-defined terms and relationships by which the role of a particular gene, gene product or gene-product group can be interpreted. Thus, an advanced strategy for the identification of preferential tumor-specific pathways would be needed by using the GO analysis. As of March 12, 2003, GO contained 6959 processes, 5339 functions and 1199 component terms, with a

Correspondence: Woong Shick Ahn, Department of Obstetrics and Gynecology, College of Medicine, The Catholic University of Korea, 505 Banpo-dong, Seocho-gu, Seoul 137-701, Korea. (Tel) 02-590-2409, (Fax) 02-599-4120, (Email) ahnws@cmc.cuk.ac.kr

Received April 15, 2003, Accepted June 2, 2003
This research was supported by the grant from Ministry of Health and Welfare (02-PJ1-PG3-20801-0010).

total of 9791 GO term definitions.

In this study, mRNA differential displays were used to investigate the expression patterns of genes related to the cervical cancer pathway, where the biological functions were primarily affected by the HPV-16-derived gene regulation. These functional profiles are cervical cancer-dependent, clearly differentiated and resulted in finding a large subset of cellular functional changes that could be described as tending to increase, or decrease, with the number of cervical cancer cells. The GO analysis was identified as being descriptive of cervical carcinogenesis, suggesting that several previously unreported functions were cervical cancer-dependent. This report has shown that the GO analysis is a valuable tool for diagnostic and therapeutic interventions in cancer research.

MATERIALS AND METHODS

1) Cell lines and cell culture

The human cervical cancer cell line, HPV-16 positive SiHa, and human keratinocyte cell line, HaCaT (5), which has similar properties to normal keratinocytes, were used for the experiment. The cells were cultured in a 5% CO_2 incubator at 37°C.

2) mRNA differential display

The total RNA was isolated from cultured cells using TRIzol (Gibco-BRL). The differential display was carried out as previously described (6). Briefly, RT-PCR was performed using RNAimage kits (GenHunter, Brookline, MA), according to the manufacturer's instructions. 0.2μg of total RNA was used for the PCR amplification, and the amplified cDNA fragments were cloned into the pGEM-T easy vector (Promega, Madison, WI). Plasmid DNA was prepared using the Wizard Miniprep Purification System (Promega). The sequencing reactions were performed with an Applied Biosystems sequencer, model 3,100, with T7 and SP6 primers. The homology of the cDNA sequences to other nucleic acids was determined using the National Center for Biotechnology Information BLAST database search algorithms.

3) Northern blotting

20μg of total RNA was electrophoresed, and transferred to nylon membranes (Hybond-N+; Amersham, Uppsala, Sweden). The insert into the pGEM-T easy vector was amplified using T7 and SP6 primers, and then gel-eluted. The probes were then [α-^{32}P]-dCTP-labeled with a Random primer labeling kit (TaKaRa, Shiga, Japan). The hybridization was carried out overnight, at 68°C, in Rapid Hyb-buffer (Amersham). After hybridization, the membrane was radiophotographed at -70°C for 24～48 hr.

3) Analysis

The ImaGene v5.0 (Biodiscovery, Marina Del Rey, CA) was used to analyze the image data. Genes were excluded from the analyses if their expressions were negative or too smeared, with those showing differences of at least 2 fold in their expression levels being selected for the function analysis. To classify the gene expression profiles, functional analyses were carried out as previously described (7). Each gene was annotated by integrating the information (as of March 12, 2003) on the Gene Ontology website (http://www.geneontology.org). All the files, including the BLAST database search results for the transcripts from differential displays, and those of the Gene Ontology analysis were downloaded from our anonymous FTP site: ftp://160.1.9.42/work/ddpcr.

RESULTS

As shown in Fig. 1, after the mRNA differential display and gel electrophoresis, approximately 640 bands were identified. After the exclusion of the poor bands, 349 of the 640 bands remained. From the sequencing and BLAST database search, 157 transcripts showing at least a 2-fold increased change were identified.

Table 1 shows a description of the transcripts, with new cervical cancer-related genes differentially up- or down-regulated in SiHa cells. The expression profiling for the down-regulation was different from that for the up-regulation, indicating that repression of the gene expression may have an important impact on the HPV-16-derived cellular processes. Northern

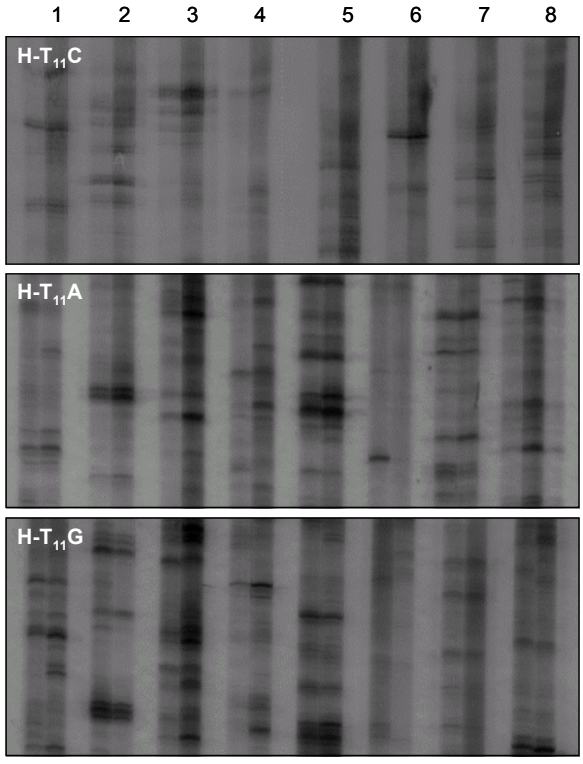

Fig. 1. Differential display of mRNA expression between HaCAT and SiHa cells. Representative mRNA DD results were obtained using different H-T primers for reverse transcription and H-AP random primer pairs for PCR. 1. H-AP17; 2. H-AP18; 3. H-AP19; 4. H-AP20; 5. H-AP21; 6. H-AP22; 7. H-AP24. Each number consists of two lanes of HaCaT and SiHa cell lines.

Table 1. Summary of gene expression changes. The genes are adversely ranked by fold change, i.e., from down-regulation to upregulation

Increase								
Unigene number	Gene symbol	Fold change	Unigene number	Gene symbol	Fold change	Unigene number	Gene symbol	Fold change
Hs.83916	NDUFA5	-8.9	Hs.301434	**KIAA1387**	-2.8	Hs.2726	HMGA2	2.9
Hs.181369	UFD1L	-8.9	Hs.211577	KTN1	-2.8	Hs.40342	HRIHFB2122	3.4
Hs.77171	MCM5	-8.7	Hs.285976	**LASS2**	-2.6	Hs.127376	**KIAA0266**	3.7
Hs.75355	UBE2N	-8.6	Hs.326766	**LOC220070**	-2.6	Hs.32511	KIAA0441	3.9
Hs.227730	ITGA6	-6.9	Hs.75639	LOC91137	-2.6	Hs.6834	**KIAA1014**	3.9
Hs.151134	OXA1L	-6.8	Hs.3688	LUC7A	-2.6	Hs.288697	**MGC11349**	4.0
Hs.153834	PUM1	-6.8	Hs.45105	**MCJ**	-2.6	Hs.250705	MGC17921	4.3
Hs.184014	RPL31	-6.8	Hs.179565	MCM3	-2.6	Hs.66521	MJD	4.3
Hs.334612	SNRPE	-6.7	Hs.39132	**MGC11115**	-2.5	Hs.111039	NMT1	4.3
Hs.25197	STUB1	-6.7	Hs.40094	**MGC8974**	-2.5	Hs.89545	PSMB4	4.5
Hs.30213	CLN5	-6.7	Hs.108931	MPP6	-2.5	Hs.196209	RAE1	4.5
Hs.306237	**KIAA1680**	-6.5	Hs.107474	NAB1	-2.5	Hs.241567	RBMS1	4.5
Hs.111632	LSM3	-6.5	Hs.5273	NDUFS3	-2.5	Hs.356593	RPL7	4.7
Hs.184352	**PCNT1**	-5.8	Hs.64056	PAK1	-2.5	Hs.182740	RPS11	4.8
Mm.6332	ALOX15B	-5.7	Hs.267289	POLA	-2.5	Hs.151604	RPS8	4.8
Hs.72160	AND-1	-5.6	Hs.110857	POLR3K	-2.5	Hs.1027	RRAD	4.8
Hs.155172	AP3B1	-5.6	Hs.184011	PP	-2.5	Hs.271926	SDCCAG16	4.9
Hs.154149	APEX2	-5.4	Hs.106415	PPARD	-2.5	Hs.75231	SLC16A1	5.0
Hs.227913	API5	-5.4	Hs.279554	PSMD13	-2.3	Hs.77496	SNRPG	5.0
Hs.7351	ARPP-19	-5.1	Hs.366	PTS	-2.3	Hs.128645	SNX16	5.4
Hs.81886	AUH	-4.9	Hs.168095	**RNF20**	-2.3	Hs.7766	UBE2E1	5.4
Hs.75415	B2M	-4.9	Hs.278526	**RNTRE**	-2.3	Hs.78563	UBE2G1	5.4
Hs.15259	BAG3	-4.7	Hs.110165	RPL26L1	-2.3	Hs.75875	UBE2V1	5.5
Hs.202	BZRP	-4.7	Hs.301547	RPS7	-2.2	Hs.78902	VDAC2	5.5
Hs.3402	C17orf26	-4.7	Hs.21858	SERPINE2	-2.2	Hs.206594	VPS33A	5.5
Hs.155560	CANX	-4.8	Hs.78713	SLC25A3	-2.2	Hs.149923	XBP1	5.9
Hs.1592	CDC16	-4.5	Hs.2043	SLC25A4	-2.2	Hs.173518	**ZFR**	6.1
Hs.179902	**CDW92**	-4.5	Hs.878	SORD	-2.2	Hs.172673	AHCY	6.2
Hs.278562	CLDN7	-4.1	Hs.75975	SRP9	-2.1	Hs.7041	**MGC4170**	6.4
Hs.106552	CNTNAP2	-4.0	Hs.380938	STX8	-2.1	Hs.159322	SLC35A3	6.4
Hs.79194	CREB1	-3.9	Hs.380364	TARS	-2.1	Hs.11411	**DKFZp566O084**	6.7
Hs.330208	CRYZL1	-3.8	Hs.173824	TDG	-2.1	Hs.75334	EXT2	6.7
Hs.349699	DSP	-3.8	Hs.1287	TRIM26	-2.1	Hs.63304	PME-1	6.8
Hs.177534	DUSP10	-3.8	Hs.100090	TSPAN-3	-2.1	Hs.164915	SNAPC3	6.8
Hs.346868	EBNA1BP2	-3.7	Hs.28423	UBP1	-2.1	Hs.12163	EIF2S2	7.3
Hs.275959	EEF1B2	-3.7	Hs.234282	VPS11	-2.0	Hs.51043	HEXB	7.3
Hs.348389	SFXN1	-3.7	Hs.173081	ZNF292	-2.0	Hs.210850	KIAA1131	7.4
Hs.331318	**FLJ13213**	-3.7	Hs.189829	**AGO3**	2.0	Hs.334368	**MGC11257**	7.4
Hs.47125	**FLJ13912**	-3.7	Hs.1227	ALAD	2.0	Hs.44143	**PB1**	7.7
Hs.71616	**FLJ14431**	-3.5	Hs.168625	AS3	2.0	Hs.333139	RAB6C	7.7
Hs.79828	FLJ20333	-3.5	Hs.78629	ATP1B1	2.2	Hs.178391	**RPL36A**	7.8
Hs.349887	**FLJ30656**	-3.5	Hs.73851	ATP5J	2.2	Hs.113029	RPS25	9.0
Hs.374897	**FLJ36874**	-3.3	Hs.76572	ATP5O	2.2	Hs.79968	SPF30	9.4
Hs.180296	FOP	-3.3	Hs.77204	CENPF	2.2	Hs.9614	NPM1	9.5
Hs.191356	GTF2H2	-3.2	Hs.289097	DDX19	2.3	Hs.93659	ERP70	9.8
Hs.55468	**H4F2**	-3.2	Hs.5364	**DKFZP564I052**	2.4	Hs.29797	RPL10	10.2
Hs.13340	HAT1	-3.2	Hs.334798	EEF1D	2.4	Hs.106826	**BHC80**	10.3
Hs.15318	HAX1	-3.2	Hs.100914	**FLJ10352**	2.5			
Hs.372673	**HNRPDL**	-3.0	Hs.168232	FLJ13855	2.5			
Hs.282804	**HPS3**	-3.0	Hs.300700	**FLJ20727**	2.5			
Hs.154762	**HRB2**	-2.9	Hs.42194	**FLJ22649**	2.7			
Hs.31323	IKBKAP	-2.9	Hs.78915	GABPB1	2.7			
Hs.141296	**KIAA0226**	-2.9	Hs.111676	H11	2.7			
Hs.71109	**KIAA1229**	-2.9	Hs.180877	**H3F3B**	2.7			
Hs.106204	KIAA1327	-2.9	Hs.343261	HM13	2.8			

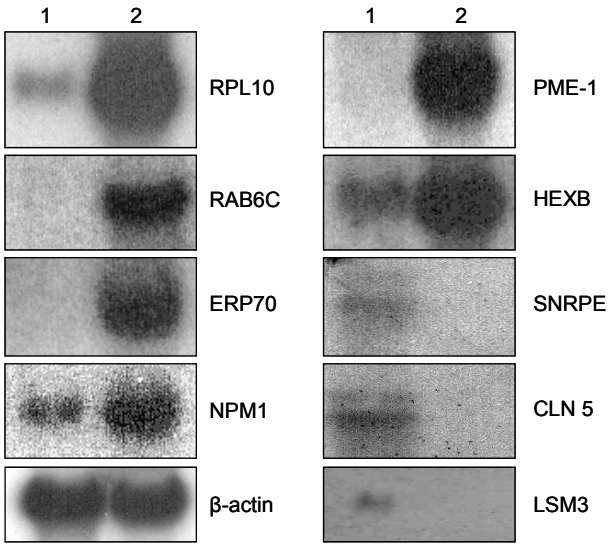

Fig. 2. Results of Northern blot. Total RNA obtained from HaCAT cells (lane 1) and SiHa cells (lane 2) was subjected to Northern analysis as described in materials and methods.

blots were performed to confirm the results of the differential displays. As shown in Fig. 2, several up- and down-regulated transcripts confirmed the patterns obtained from the differential displays, showing the consistency of the experimental repeatability.

As shown in Tables 2, 3 and 4, the significantly up- and down-regulated functional activities, as diverse as cell communication, cell growth and maintenance, cell death, development, apoptosis regulation, nucleic acid binding, chaperon and enzymes, etc., were analyzed according to the biological processes, cellular components and molecular function ontologies. Of the HPV-16-derived biological functions, over 50% of the functions were included in the biological processes, with half of these being in the cell growth and maintenance.

As shown in Table 2 and Fig. 3A, the transcripts involved in the cell communication (6/10, i.e., up-regulation of RAB6C, EXT2, XBP1, SNX16, RRAD and MJD, or the down-regulation of ITGA6, OXA1L, B2M, CNTNAP2, CREB1, DUSP10, IKBKAP, PAK1, MPP6 and SORD) were relatively repressed in the SiHa cells as compared to the normal cells. The transcripts in the signal transduction (4/6: RAB6C, EXT2, SNX16, RRAD/ITGA6, OXA1L, CREB1, DUSP10, PAK1, MPP6) showed relatively repressed expression profiles. The cell adhesion function (0/2) was likely to be affected by the HPV 16-associated cellular process. For instance, alpha-6 integrin (ITGA6) and contactin-associated protein 2 (CNTNAP2), another cell-to-matrix attachment-related components, were highly down-regulated.

As shown in Table 2, the transcripts involved in the cell cycle (2/3: UBE2V1, CENPF/MCM5, MCM3, CDC16) were repressed in the SiHa cells. As shown in Fig. 3B, the transcripts in the transport (10/9: RAB6C, SLC35A3, SNX16, SLC16A1, RAE1, ATP5O, ATP5J, VDAC2, ATP1B1, DDX19/LOC91137, STX8, SLC25A4, SLC25A3, SFXN1, AP3B1, C17orf26, KTN1, VPS11) were relatively balanced.

As shown in the metabolism process, a number of transcripts were repressed in the SiHa cells. The transcripts involved in the DNA metabolism, such as DNA repair (1/4: UBE2V1/GTF2H2, TDG, UBE2N, APEX2) and DNA replication (1/3: RBMS1/POLA, MCM3, MCM5), were down-regulated. In the lipid and carbohydrate metabolism, the HEXB, associated with HPV 16-derived cervical cancer, was highly up-regulated. Also, the NADH dehydrogenase subunit 4, as a mitochondrial enzyme, was the highly over-expressed clone involved in the metabolic pathway.

In the development function, several transcripts involved in the morphogenesis (0/2: TSPAN-3, DSP) were down-regulated, whereas those in the organogenesis (2/2: MJD, EXT2/PTS, UFD1L) were maintained at relatively constant levels, as shown in Table 2 and Fig. 3D.

As shown in Table 3, the apoptosis function, the transcripts involved in heat-shock proteins, such as H11, the eighth known human small Hsp, named HspB8 described as a serine-threonine protein kinase, were up-regulated in the SiHa cells. It is well known that cervical carcinomas produce DNA damage in cancer cells, where tumor suppressor-related genes, such as EXT2, were highly induced, as shown in this study. In contrast, down-regulation of the genes related to DNA repair, such as UBE2N and TDG, were also apparent in the SiHa cells. Conversely, there was a down-regulated change in the antiapoptotic transcripts, BAG3 and API5. The molecular function ontology includes transcripts with an apoptosis regulator and chaperone. The transcripts in chaperone (1/2: H11/BAG3, CANX) were shown in the SiHa cells.

The DNA binding function includes several transcripts, which were down-regulated (6/12: SNAPC3, XBP1, KIAA0441, HMGA2, GABPB1, AS3/MCM5, PPARD, CREB1, GTF2H2, MCM3, NAB1, POLR3K, UBP1, ZNF292, KIAA1327, FLJ20333, AND-1), as shown in Table 3 and Fig. 3E. Note that the GABPB1, involved in the expression of the adenovirus E4 gene, was up-regulated. Several transcripts were involved in protein translation, where the generally down-regulated genes were translation elongation factors, such as EEF1B2 and SRP9, compared to highly up-regulated genes, such as EEF1D, EIF2S2 and RBMS1. The E2F family of transcription factors, which have a central role in the regulation of cell proliferation, were not shown in this study.

As shown in Table 4, the transcripts in the cytoskeleton (4/3: HRIHFB212, RAE1, MGC17921, CENPF/DSP, CLDN7, CDC16) were balanced, whereas a clone on chromosome Xq25-26.3, which is remarkably similar to that of the human molecule beta-tubulin, with 98% homology, was highly down-regulated in the SiHa cells. No nuclear structural proteins were expressed in this assay.

Transcripts in the proteasome and ubiquitin degradation pathways, such as UBE2E1, UBE2G1, KIAA1131, FLJ13855 and PSMB4, were up-regulated 2-fold, but in contrast, the UBE2N, FLJ20333 and UFD1L, involved in the degradation of ubiquitin fusion proteins, were down-regulated. On the other hand, as shown in Table 2 and Fig. 3F, the protein biosynthesis activity in the metabolism was relatively up-regulated (8/6: RPS25, EIF2S2, RPS8, RPS11, RPL7, EEF1D, RBMS1, ATP5O/RPL31 EEF1B2, RPS7, TARS, RPL26L1, SRP9).

Table 2. Summarized biological process ontology of up- and down-regulated transcripts. The GO code and the number of up-/down-regulated genes were shown

Biological process ontology	GO code	Up	Down
Cell communication	GO:0007154	6	10
Cell adhesion	GO:0007155	0	2
Cell-matrix adhesion	GO:0007160	0	1
Cell-cell signaling	GO:0007267		
Transmission of nerve impulse	GO:0019226	0	1
Synaptic transmission	GO:0007268	1	0
Response to external stimulus	GO:0009605		
Vision	GO:0007601	0	1
Antigen presentation, endogenous antigen	GO:0019833	0	1
Antigen processing, endogenous antigen via MHC class 1	GO:0019885	0	1
Immune response	GO:0006955	1	1
Signal transduction	GO:0007165	4	6
Integrin-mediated signaling pathway	GO:0007229	0	1
Intracellular signaling cascade	GO:0007242	1	0
JNK cascade	GO:0007254	0	2
Small GTPase mediated signal transduction	GO:0007264	1	0
Cell growth and maintenance	GO:0008151	3	1
Cell cycle	GO:0007049	2	3
DNA replication and chromosome cycle	GO:0000067	1	0
Mitosis	GO:0007067	1	0
Regulation of cell cycle	GO:0000074	0	1
Cell motility	GO:0006928	0	1
Regulation of cell migration	GO:0030334	0	1
Cell organization and biogenesis	GO:0016043		
Cell-substrate junction assembly	GO:0007044	0	1
Cell proliferation	GO:0008283	0	1
Metabolism	GO:0008152		
Amino acid metabolism	GO:0006520	0	1
Protein biosynthesis	GO:0006412	8	6
ATP biosynthesis	GO:0006754	1	0
Carbohydrate metabolism	GO:0005975	1	1
Regulation of proteolysis and peptidolysis	GO:0030162	0	1
Electron transport	GO:0006118	1	2
Energy pathways	GO:0006091	1	3
Lipid metabolism	GO:0006629	0	1
Fatty acid metabolism	GO:0006631	0	1
Nucleobase, nucleoside, nucleotide and nucleic acid metabolism	GO:0006139		
DNA metabolism	GO:0006259		
DNA repair	GO:0006281	1	4
DNA replication	GO:0006260	1	3
RNA metabolism	GO:0016070		
RNA processing	GO:0006396	1	0
mRNA splicing	GO:0006371	1	2
Transcription	GO:0006350	0	1
Regulation of transcription, DNA-dependent	GO:0006355	5	7
RNA elongation	GO:0006354	0	1
Protein metabolism	GO:0019538		
Protein modification	GO:0006464	2	2
Protein folding	GO:0006457	0	1
Protein complex assembly	GO:0006461	0	2
Response to stress	GO:0006950	0	1
Response to heat	GO:0006951	1	0
Transport	GO:0006810	10	9
Nucleotide-sugar transport	GO:0015780	1	0
Hydrogen transport	GO:0006818	1	0
Protein transport	GO:0015031	1	1
Intracellular protein transport	GO:0006886	1	1
Death	GO:0016265		
Cell death	GO:0008219		
Apoptosis	GO:0006915	0	1
Development	GO:0007275	1	0
Morphogenesis	GO:0009653	0	2
Epidermal differentiation	GO:0008544	0	1
Organogenesis	GO:0009887	2	2
Neurogenesis	GO:0007399	1	0
Skeletal development	GO:0001501	1	1

Table 3. Summarized molecular function ontology of up- and down-regulated transcripts

Molecular function ontology	GO code	Up	Down
Apoptosis regulator	GO:0016329	0	1
Apoptosis inhibitor	GO:0008189	0	1
Binding	GO:0005488	0	2
Calcium ion binding	GO:0005509	0	1
Calcium ion storage	GO:0005514	0	1
Glycosaminoglycan binding	GO:0005539		
Zinc binding	GO:0008270	0	3
Magnesium binding	GO:0000287	0	2
Nucleic acid binding	GO:0003676	2	0
DNA binding	GO:0003677	6	12
Damaged DNA binding	GO:0003684	0	1
Single-stranded DNA binding	GO:0003697	1	0
AT DNA binding	GO:0003680	1	0
Exonuclease	GO:0004527	0	1
RNA binding	GO:0003723	10	5
mRNA binding	GO:0003729	1	2
rRNA binding	GO:0019843	1	0
Translation elongation factor	GO:0003746	1	1
Translation initiation factor	GO:0003743	1	0
Nucleotide binding	GO:0000166		
ATP binding	GO:0005524	0	4
GTP binding	GO:0005525	1	0
Protein binding	GO:0005515	2	4
Transcription cofactor	GO:0003712	0	1
Chaperone	GO:0003754	1	2
Heat shock protein	GO:0003773	1	0
Hsp70/Hsc70 protein regulator	GO:0030192	0	2
Enzyme	GO:0003824	2	0
Hydrolase	GO:0016787	4	3
Isomerase	GO:0016853	1	0
Ligase	GO:0016874	2	2
Lyase	GO:0016829	1	1
Oxidoreductase	GO:0016491	0	5
Ubiquitin conjugating enzyme	GO:0004840	5	3
Transferase	GO:0016740	3	4
Protein phosphatase	GO:0004721	0	1
Enzyme regulator	GO:0030234		
Serine protease inhibitor	GO:0004867	0	1
Signal transducer	GO:0004871	0	1
Receptor	GO:0004872	0	3
Ligand-dependent nuclear receptor	GO:0004879	0	1
Benzodiazepine receptor	GO:0008503	0	1
Cell adhesion receptor	GO:0004895	0	1
MHC class 1 receptor	GO:0030106	0	1
Receptor signaling protein	GO:0005057	1	0
Structural molecule	GO:0005198	0	2
Structural constituent of cytoskeleton	GO:0005200	0	1
Structural constituent of ribosome	GO:0003735	5	2
Transcription regulator	GO:0030528		
Transcriptional repressor	GO:0016564	0	1
Translation regulator	GO:0045182		
Translation factor, nucleic acid binding	GO:0008135	2	0
Translation initiation factor	GO:0003743	1	0
Transporter	GO:0005215	2	0
Auxiliary transport protein	GO:0015457	2	1
Carrier	GO:0005386	0	1
Channel/pore class transporter	GO:0015267	2	1
Electron transporter	GO:0005489	1	1
Ion transporter	GO:0015075	4	1
Protein transporter	GO:0008565	1	1

Table 4. Summarized cellular component ontology of up- and down-regulated transcripts

Cellular component ontology	GO code	Up	Down
Cell fraction	GO:0000267		
Membrane fraction	GO:0005624	3	2
Soluble fraction	GO:0005625	0	2
Intracellular	GO:0005622	4	3
Chromosome	GO:0005694	2	0
Centromere	GO:0005698	1	0
Chromatin	GO:0005717	1	0
Kinetochore	GO:0005699	1	0
Cytoplasm	GO:0005737	5	4
Cytoskeleton	GO:0005856	4	3
Intermediate filament	GO:0005882	0	1
Spindle	GO:0005819	1	0
26S proteasome	GO:0005837	1	0
Endoplasmic reticulum	GO:0005783	2	3
Endoplamic reticulum lumen	GO:0005788	1	0
Endoplasmic reticulum membrane	GO:0005789	1	1
Golgi apparatus	GO:0005794	2	1
Mitochondrion	GO:0005739	3	8
Mitochondrial outer membrane	GO:0005741	3	4
Ribosome	GO:0005840	5	2
Lysosome	GO:0005764	11	0
Nucleus	GO:0005634	0	13
Nuclear membrane	GO:0005635	0	1
Nuclear outer membrane	GO:0005640	1	0
Nuclear pore	GO:0005643	1	1
Nucleolus	GO:0005730	1	0
Small nucleolar ribonucleoprotein complex	GO:0005732	1	2
Nucleoplasm	GO:0005654	1	1
DNA-directed RNA polymerase III complex	GO:0005666	0	1
Alpha DNA polymerase:primase complex	GO:0005658	0	1
Spliceosome complex	GO:0005681	1	1
Integral to plasma membrane	GO:0005887	0	4
Integrin complex	GO:0008305	0	1
Sodium/potassium-exchanging ATPase complex	GO:0005890	1	0
Respiratory chain complex I	GO:0045271	0	2
Membrane	GO:0016020	2	3
Inner membrane	GO:0019866	0	2
Mitochondrial inner membrane	GO:0005743	1	2
Mitochondrial electron transport chain complex	GO:0005746	0	1
Integral membrane protein	GO:0016021	5	12
Outer membrane	GO:0019867	1	0
Plasma membrane	GO:0005886		
Cell-cell adherens junction	GO:0005913	0	1
Tight junction	GO:0005923	0	1
Extracellular	GO:0005576	0	1

DISCUSSION

Most of clones and transcripts with less than a 2-fold change were not used in this assay, indicating that the transcriptional regulation of many transcripts was actively maintained at certain levels, regardless of the HPV-16-derived cellular process. Several of the up-regulated transcripts, such as 16S ribosomal RNA, casein kinase 2 subunit beta and NADH dehydrogenase subunit 4, were previously reported to be over-expressed in squamous cell cervical carcinomas (8). Also, a number of genes, known as being differentially up- or down-regulated compared to normal, such as XBP1, CENPF, RPS25, HEXB, ITGA6, CNTNAP2, MCM5, MCM3 and IKBKAP, were identified. Reliable gene expression data shows the

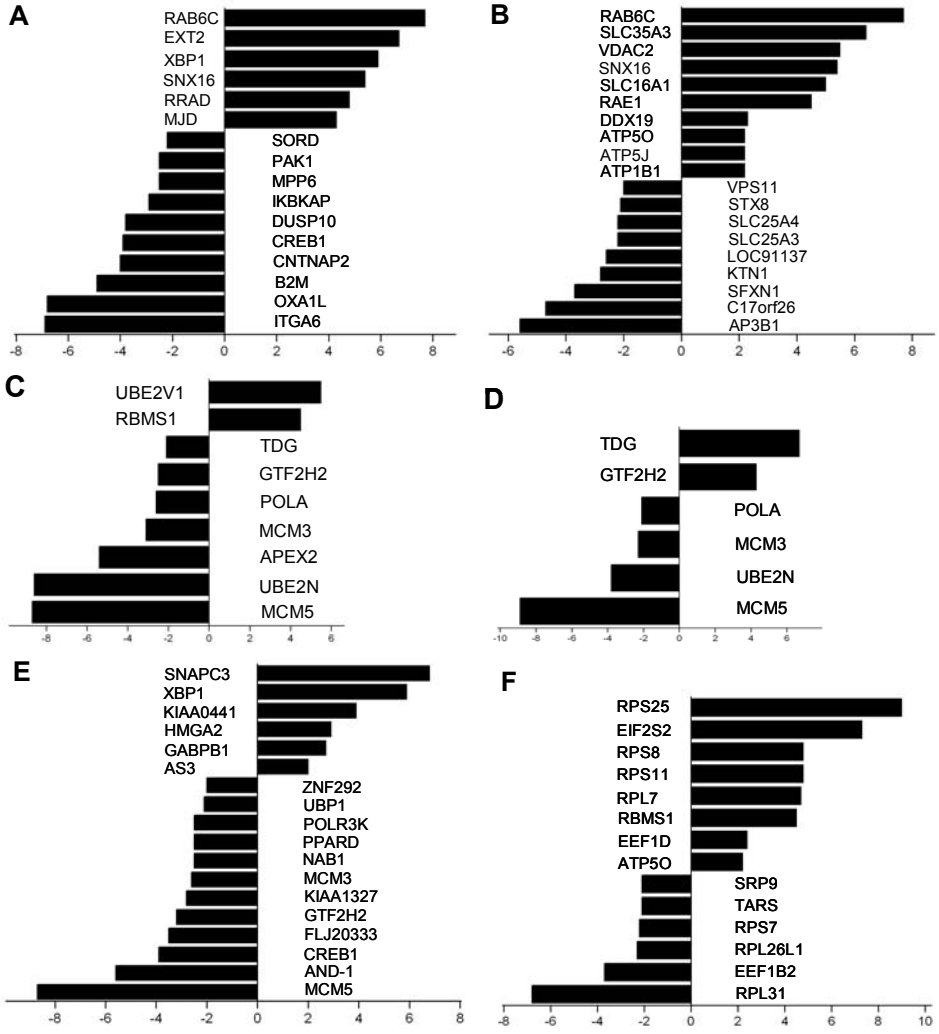

Fig. 3. Relative expression levels for genes with function-dependent transcript representation; (A) cell communication, (B) transport, (C) DNA metabolism, (D) development, (E) DNA binding, (F) protein biosynthesis.

consistency of the above-regulated transcripts to those of the previously reported results, and validates the profiling method for the study of the cervical cancer-specific pathway.

As usual, the gene expression profiles have been classified into reasonable groups using statistical data mining methods, such as the hierarchical clustering and K-Means clustering. These are, however, based on the statistical relationships between the genes, resulting in a biological description limit, as well as an analytical reproducibility problem. Thus, in order to obtain the biological interconnectivity between the gene expression patterns, the differentially regulated genes should be categorized using the Gene Ontology. Correlations are sought between the genes in cellular process groups and their expression patterns. The changes observed in this analysis provide important insights into the more specific cancer-related gene activity and HPV-16-derived functional change patterns. Closer examination of the 157 genes resulted in a number of reciprocally dependent cellular processes, revealing that the expression patterns are not randomly distributed with respect to their function.

On the other hand, as shown in Table 1, the potentially significant genes with unknown functions are shown in bold type. The reported genes can be noted as new putative cervical cancer-related genes, as there were no matches between the genes and any other known functional activities in the GO database search. Of these genes, for instance, the down-regulation of LASS2 in hepatocellular carcinoma cell lines was only reported to lead to the progression of cancer cell growth. Also, MPP6, a member of the p55-like membrane-associated guanylate kinase (MAGUK) subfamily, which functions in tumor suppression and receptor clustering by forming multi-protein complexes, contains distinct sets of transmembrane, cytoskeletal and cytoplasmic signaling proteins (9). The MPP6 was down-regulated in this study. Future studies will be required to clarify these genes regulatory mechanisms and their role in cervical carcinogenesis.

With cell communication, the down-regulated cell adhesion activity was likely to be affected by the HPV 16-associated cellular pathway, indicating a decrease in the adhesive properties of cervical carcinogenesis. For instance, the alpha-6 integrin (ITGA6) and contactin-associated protein 2 (CNTNAP2), two other cell-to-matrix attachment-related components, were highly down-regulated. It was also reported that mice deficient in alpha-6 integrin, through a targeted disruption, died with

severe blistering of the skin and other epithelia (10). The phenotype was reminiscent of human epidermolysis bullosa. Transcripts in the immune response (1/1: XBP1/IKBKAP) contained only two genes, but a previous study reported that inflammatory and chemotactic-related transcripts were not detected in cervical cancer cells compared to normal cells (11). Especially, the up-regulated XBP1, involved in B-cell biology, was reported to be able to differentiate myeloma from non myeloma cell lines (12). Recent studies have shown that two mutations in the IKBKAP gene are responsible for a developmental disorder of the sensory and autonomic nervous system (13). One of the highly up-regulated genes in normal uterus tissue (14), CREB1, was down-regulated in the cell communication function.

With cell growth and maintenance, many cell cycle checkpoints are deregulated in oncogenesis, which results from interference with the cell cycle regulation of the oncoproteins by destruction of the p53 and pRb tumor suppressors. Their inactivation leads to alteration in the cellular gene expression. It was reported that CENPF, which is involved in DNA replication and mitosis, was induced by E6 and/or E7 retroviruses, after immortalization, thus confirming the consistency of our study. Also, the MCM3 protein is down-regulated in cells that have not ceased to proliferate. As shown in our study, the transcription level of VDAC2 in the malignant tumor cell line was significantly higher than that in normal cells (15). There are several notable gene expression patterns in response to stress function. Nearly all the stress response genes, including those involved in the response to pathogens and bacteria, were not expressed, regardless of HPV 16. Only H11, the eighth known human small Hsp, described as a serine-threonine protein kinase (16), was up-regulated, whereas the DUSP10 in the stress response was down-regulated.

With the metabolism, HEXB has not been fully described in the context of cervical cancer. Recently, it was represented, in the GenBank (accession number: AF378118.1), to be up-regulated in human cervical cancer, as a proto-oncogene. Its role in cervical carcinogenesis is being investigated further. For example, the question is to test if the HEXB gene maps to the cervical cancer susceptibility locus, which would support a potential link in the malfunctioned metabolism related to disease development, and could be regarded as being responsible for an increased susceptibility to the HPV 16 in cervical cancer. A highly over-expressed clone, NADH dehydrogenase subunit 4, as a mitochondrial enzyme, is involved in the metabolic pathway. It has been suggested that metabolic transformation is likely to occur during the early malignant stage of carcinogenesis (17). Also, the level of NADH dehydrogenase subunits expression in pre-malignant colon adenomas was up regulated, indicating that the tumorigenicity of cells was associated with the overexpression of the NADH dehydrogenase subunit 4 (18). All these evidence suggest the NADH dehydrogenase subunit 4 serves an important role in cervical carcinogenesis.

With the development, an overexpressed clone, the *Mus musculus* casein kinase 2 beta subunit, is highly homologous to that of the human casein kinase 2 subunit beta, an enzyme that is involved in the transcription, signaling and proliferation, and in various stages of the development. More significantly, it has been reported that the phosphorylation of HPV E7, by casein kinase 2, enhances the disruption of the G1/S transition by HPV E7, suggesting that the human casein kinase 2 subunit beta plays an important role in human cervical neoplasia (19).

With apoptosis, a gene involved with the Hsp70/Hsc70 protein regulator, BAG3, which binds with high affinity to the ATPase domain of the Hsc70, inhibiting its chaperone activity in a Hip-repressible manner, allows opportunities for the specification and diversification of the Hsp70/Hsc70 chaperone functions (20). CANX (calnexin) functions as a chaperone, regulating the transit of proteins through the ER. Comparisons of the sequences from cDNA clones of human, mouse and rat CANXs have demonstrated a high level of conservation in the sequence identity, suggesting that CANX performs important cellular functions (21). The gene expression profiles showed that SiHa cells do not induce significant levels of apoptosis, as the rate of protein synthesis (nearly up-regulated) was promoted at the polypeptide chain initiation level.

Transcription is a complex category that can lead to global alterations in the whole network of gene expressions. It has been reported that transcriptional modification was accomplished by SNAPC3 proteins, which acted at a specific promotor, as well as by regional activation (CENPF, HMGA2) of the chromosomal domain. Some of these transcriptional regulators are expected to be highly relevant in the cervical cancer pathway, and their regulation may affect different cellular functions. The HMGA2 gene has been considered, for several reasons, to be an especially good candidate for involvement in lipomas. Firstly, its location, 12q15, is involved in a variety of malignant and benign solid tumors, such as lipomas, pleomorphic salivary adenomas and uterine leiomyomas. Secondly, the chromosomal abnormalities of the 12q13-q15 region, frequently found in human benign tumors, are the critical events in the genesis of the tumors (22). Conversely, the transcriptional repressor, NAB1, was down-regulated, and whose aberrant expression may serve to activate other downstream pathways that further contribute to the development of a tumor. One of the highly expressed clones is homologous to the human 16S ribosomal RNA required for the translation of mitochondrial subunits.

Several genes, coding for cellular structure proteins, changed their expression in the pathway to that of the HPV-16-dependent transcription. Transcripts in this category can be subdivided into two groups: cytoskeletal and nuclear-related genes. The cytoskeleton integrity is known to play an important role in cell cycle progression, cell death and cell differentiation. An abnormal cytoskeleton is often observed in cancer cells. In this study, however, the gene expression profiles involved in the cytoskeleton were balanced, whereas a clone, similar to the human molecule, showing 98% homology on beta-tubulin, was highly down-regulated in the SiHa cells, which was consistent with a previous study (23).

Genes involved in the proteasome and ubiquitin degradation pathways form another important group. For protein degradation arrest, the repression of the ubiquitin-conjugating enzyme, UBE2E1, which is involved in the selective protein destruction by a specific protease complex, and the 26S proteasome, PSMB4, may be required in a dominant negative form for these proteins to prevent the recognition into

degradation (24). However, the transcripts in this category were up-regulated 2-fold, whereas, the degradation of ubiquitin fusion proteins were down-regulated. On the other hand, the protein biosynthesis function in the metabolism was relatively up-regulated. The over-expression of RPS25 especially, can result in significant effects on protein synthesis and SRP9, which leads to an inhibition of polypeptide elongation, known as translation arrest, were down-regulated, indicating that they are essential for overcoming the translation arrest.

Taken together, the expression behaviors of genes in protein degradation suggests that SiHa cells may increase the turnover of many proteins, which could be due to either the replacement of damaged molecules, or to the need to significantly change the proteome of the cell, probably to obtain new molecules that are involved in the cervical carcinogenesis or the replacement of damaged proteins.

CONCLUSIONS

Our results have shown that GO analysis can describe the cellular processes that occur in HPV 16-positive cervical cancer cells, and overcomes the complexity of gene expression profiles, and can be an alternative to the hierarchical or K-Means clustering. Further systematic approaches, including cervical cancer lesions at various stages and genome-wide analyses, using the GO, can certainly elucidate new connections between gene expression profiles and the cellular pathways of cervical carcinogenesis, in order to identify valuable prognostic candidate genes.

REFERENCES

1. Münger K, Howley PM. Human papillomavirus immortalization and transformation functions. Vir Res 2002;89:213-228.
2. Lee SJ, Namkoong SE, Lee WC, Sul JW, Jee SH, You YK, Lee JE, Park JS. Polymorphisms of p53, p21 and IRF-1 and cervical cancer susceptibility in Korean women. Cancer Res Treat 2002;34:357-364.
3. Ponnazhagan S, Curiel DT, Shaw DR, Alvarez RD, Gene PS. Adeno-associated virus for cancer gene therapy. Cancer Research 2001;61:6313-6321
4. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. Nat Genet 2000;25:25-29.
5. Boukamp P, Petrussevska RT, Breitkreutz D, Hornung J, Markham A, Fusenig NE. Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. J Cell Biol 1988;106:761-771.
6. Liang P, Pardee AB. Differential display of eukaryotic messenger RNA by means of polymerase chain reaction. Science 1992;257:967-971.
7. Pletcher SD, Macdonald SJ, Marguerie R, Certa U, Stearns SC, Goldstein DB, Partridge L. Genome-wide transcript profiles in aging and calorically restricted Drosophila melanogaster. Curr Biol 2002;12:712-723.
8. Cheng Q, Lau WM, Tay SK, Chew SH, Ho TH, Hui KM. Identification and characterization of genes involved in the carcinogenesis of human squamous cell cervical carcinoma. Int J Cancer 2002;98:419-426.
9. Tseng TC, Marfatia SM, Bryant PJ, Pack S, Zhuang A, O'Brien JE, Lin L, Hanada T, Chishti AH. VAM-1: a new member of the MAGUK family binds to human Veli-1 through a conserved domain. Biochim Biophys Acta 2001;1518:249-259.
10. Georges-Labouesse E, Messaddeq N, Yehia G, Cadalbert L, Dierich A, Le Meur M. Absence of integrin alpha-6 leads to epidermolysis bullosa and neonatal death in mice. Nat Genet 1996;13:370-373.
11. Hazelbag S, Fleuren GJ, Baelde JJ, Schuuring E, Kenter GG, Gorter A. Cytokine profile of cervical cancer cells. Gyne Oncol 2001;83:235-243.
12. Claudio JO, Masih-Khan K, Tang H, Goncalves J, Voralia M, Li ZH, Nadeem V, Cukerman E, Francisco-Pabalan O, Liew CC, Woodgett JR, Stewart AK. A molecular compendium of genes expressed in multiple myeloma. Blood 2002;100:2175-2186.
13. Coli R, Anderson SL, Volpi SA, Rubin BY. Genomic organization and chromosomal localization of the mouse IKBKAP gene. Gene 2001;279:81-89.
14. Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A, Patapoutian A, Hampton GM, Schultz PG, Hogenesch JB. Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci USA 2002;99:4465-4470
15. Shinohara Y, Ishida T, Hino M, Yamazaki N, Baba Y, Terada H. Characterization of porin isoforms expressed in tumor cells. Eur J Biochem 2000;267:6067-6073.
16. Kappe G, Verschuure P, Philipsen RL, Staalduinen AA, Van de Boogaart P, Boelens WC, De Jong WW. Characterization of two novel human small heat shock proteins: protein kinase-related HspB8 and testis-specific HspB9. Biochim Biophys Acta 2001;1520:1-6.
17. Costello LC, Franklin RB. Bioenergetic theory of prostate malignancy. Prostate 1994;25:162-166.
18. Mills KI, Woodgate LJ, Gilkes AF, Walsh V, Sweeney MC, Brown G, Burnett AK. Inhibition of mitochondrial function in HL60 cells is associated with an increased apoptosis and expression of CD14. Biochem Biophys Res Commun 1999; 263:294-300.
19. Massimi P, Pim D, Storey A, Banks L. HPV-16 E7 and adenovirus E1a complex formation with TATA box binding protein is enhanced by casein kinase 2 phosphorylation. Oncogene 1996;12:2325-2330.
20. Takayama S, Xie Z, Reed JC: An evolutionarily conserved family of Hsp70/Hsc70 molecular chaperone regulators. J Biol Chem 1999;274:781-786.
21. Tjoelker LW, Seyfried CE, Eddy RL, Byers MG, Shows TB, Calderon J, Schreiber RB, Gray PW. Human, mouse, and rat calnexin cDNA cloning: identification of potential calcium binding motifs and gene localization to human chromosome 5. Biochemistry 1994;33:3229-3236.
22. Schoenmakers EPPM, Wanschura S, Mols R, Bullerdiek J, Van den Berghe H, Van de Ven WJM. Recurrent rearrangements in the high mobility group protein gene, HMGI-C, in benign mesenchymal tumours. Nat Genet 1995;10:436-444.
23. Alazawi W, Pett M, Arch B, Scott L, Freeman T, Stanley MA, Coleman N. Changes in cervical keratinocyte gene expression associated with integration of human papillomavirus 16. Cancer Res 2002;62:6959-6965.
24. Nuber U, Schwarz S, Kaiser P, Schneider R, Scheffner M. Cloning of human ubiquitin-conjugating enzymes UbcH6 and UbcH7 (E2-F1) and characterization of their interaction with E6-AP and RSP5. J Biol Chem 1996;271:2795-2800.