Comparison of the molecular classification with FIGO stage and histological grade on endometrial cancer

B. Cai[1], *Ph.D.*; **L. Liu**[1,2], *Ph.D.*; **X. W. Xi**[1], *M.D.*; **Y. P. Zhu**[1], *M.D.*; **G. Z. Lu**[3], *M.D., Ph.D.*; **Y. X. Yang**[1], **X. P. Wan**[1], *M.D., Ph.D.*

[1]*Department of Obstetrics and Gynecology, Shanghai Jiao Tong University Affiliated First People's Hospital, Shanghai (China)*
[2]*Departments of Medicine and Medical Biophysics, University of Toronto, Toronto (Canada)*
[3]*Department of Pathology, Shanghai Jiao Tong University Affiliated First People's Hospital, Shanghai (China)*

Summary

Purpose of investigation: To classify endometrial cancers based on gene expression profiling, and to compare the prognostic value of the classification systems based on gene expression, grade, and stage. *Methods:* cDNA microarray was carried out in 32 endometrioid endometrial cancers. Differentially expressed genes were identified among tumor tissues of different grades and stages. The classification and prognosis comparison analysis was performed between histological grades, FIGO stages and gene expression profiles. *Results:* Class comparison analysis between different grade and stage endometrial cancer revealed 33 genes that are differentially expressed in tumors of different grades, ten in those of different stages, and 104 in a combined classification of grades and stages ($p < 0.001$). *Conclusion:* The cDNA microarray technique is a feasible way to generate gene expression profiles of endometrial cancer. Classification based on gene expression patterns may be more accurate than histological grade and FIGO stage classification in predicting the prognosis of tumors.

Key words: Gene expression profiling; Endometrial carcinoma; Classification; Prognosis.

Introduction

Endometrial carcinoma (EC) is the most common malignancy of the female genital tract. Each year, endometrial cancer develops in about 142,000 women worldwide, and an estimated 42,000 women die from this cancer [1]. Approximately 80% of patients with adenocarcinoma of the endometrium have endometrioid endometrial carcinoma (EEC) [1]. Abnormal uterine bleeding is the most frequent symptom of endometrial cancer, which can be detected and diagnosed in early stage, and has a 5-year survival rate of 83% [1]. It is believed to be relatively benign. In contrast, about 10%-20% early stage EEC patients experience relapse or metastasis [1, 2]. The FIGO (International Federation of Gynecology and Obstetrics) system is most commonly used for staging. The staging system we now use is the modified FIGO system (1988), which emphasizes complete surgicopathologic assessment of data, such as histological grade, myometrial invasion, as well as the extent and location of extrauterine spread, including retroperitoneal lymph node metastases [3]. Since 1988, the cure rate or the death rate of endometrial cancer has not improved significantly. Although this phenomenon can be attributed to many reasons, it indicates the need for a critical reassessment of the guidelines for managing EEC. Presently, so many clinical predictors are not adequate to predict clinical outcome accurately, and many patients receive radio- or hormonal therapy regardless of whether they need this additional treatment. In a 17-year practice, we have observed that even EEC patients with the same state of the disease (FIGO stage, myometrial invasion, histological type, and differentiation grade) can have very different responses to treatment and overall outcome.

Microarrays allow global analysis of gene expression and provide the opportunity to identify patterns that underlie biological differences among cancers [3, 4]. It has been widely used in identifying novel cancer-related genes, classifying human cancer subtypes [5-7], and prediction of the outcome [8, 9]. In the present study, we searched for a progression gene expression signature of high-risk EECs using high-density cDNA microarrays. To identify the gene expression patterns associated with high-risk EECs, that would relapse or metastasize in the future, we profiled a total of 32 EECs, and examined the differences in gene expression patterns between different stage and histological grade EECs.

Materials and Methods

Tissue specimens

Thirty-two EECs diagnosed at the First People's Hospital affiliated with Shanghai Jiaotong University from February 2002 to April 2005 were obtained from institutional tissue banks. All samples were sharp-dissected and snap-frozen in

This work was supported by the Chinese National Natural Science Foundation, Grant 30371481; the Natural Science Foundation of Shanghai, Grant 06ZR14053; and the Key Project of the Shanghai Health Bureau, Grant 2005ZD002.

Revised manuscript accepted for publication May 28, 2007

Table 1. — *Real-time PCR validation of the microarray results.*

	Microarray expression (Cy3/Cy5)						Real-time PCR expression*					
	Sam1 (EC03)	Sam2 (EC13)	Sam3 (EC12)	Sam4 (EC18)	Sam5 (EC23)	Sam6 (EC09)	Sam1 (EC03)	Sam2 (EC13)	Sam3 (EC12)	Sam4 (EC18)	Sam5 (EC23)	Sam6 (EC09)
MTA1	1.233	1.091	1.060	0.385	0.275	1.149	1.674	1.756	0.846	2.353	1.035	1.859
TGFB2	0.732	0.690	0.435	0.601	0.653	2.776	0.499	0.927	0.147	0.553	0.955	5.275
IGF1	0.175	0.037	0.109	0.055	0.122	0.031	0.145	0.038	0.165	0.064	0.056	0.007
TOB2	0.893	0.466	1.103	1.773	0.682	0.895	0.773	0.554	0.304	0.439	0.437	0.740

* the value is the fold change of target gene, which is $2^{-\Delta(\Delta Ct)}$, Where ΔCt = Ct(target)-Ct(ß-actin), $\Delta(\Delta Ct)$ = Ct(EECs)-Ct(control).

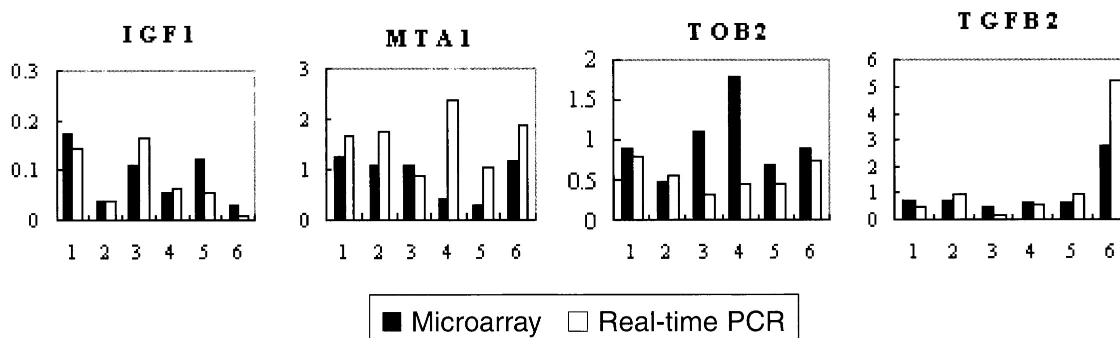

Figure 1. — Comparison of gene expression with real-time PCR and microarray. Real-time PCR expression and microarray expression comparison of four selected genes in six endometrial cancers. X and Y axes represent, respectively, the six endometrial cancers and the ratio of cancer/control. If the ratio is > 1, the gene is up-regulated; if the ratio is < 1, it is down-regulated in endometrial cancer.

liquid nitrogen within 30 minutes from resection. None of the patients had received preoperative chemotherapy or radiation. The age of the 32 patients included in this study ranged from 40 to 83 years (mean 61 years). The clinical stage of their disease was Stage I in 16, II in three, III in ten, and IV in three (Table 2). The histological grade was G_1 in 13, G_2 in nine and G_3 in ten. All patients involved in the study were continuously followed (24~61 months, median 40 months. In addition, 36 normal endometrial samples were obtained from patients (mean 50 years) undergoing hysterectomy for benign gynecological diseases (e.g., adenomyoma or uterine myoma) as a universal reference. Written informed consent was obtained from each subject, and this study was approved by the Shanghai Jiao Tong University Human Studies Committee. The frozen sections of each sample were stained with H&E to check tumor purity. Each sample was confirmed by a gynecologic pathologist, and only tumor samples containing at least 80% of tumor epithelial cells were retained for further total RNA extraction.

cDNA microarray

A set of three slides cDNA microarrays with 13,824 cDNA clones representing 11,203 genes were prepared in Shanghai Biochip by using a Generation III spotter (Amersham Biosciences, UK). All cDNA clones were sequence-verified before spotting on the microarray. Total RNA was prepared by using TRIzol (Life Technologies, USA), further purified with an RNeasy column (Qiagen, Germany), and quantified using an RNA LabChip kit (Agilent Technologies, Palo Alto, CA). For reverse-transcription labeling: approximately 30 µg of each RNA sample was reverse-transcribed into cDNA primed with oligo(dT) and labeled with either Cy3-dCTP or Cy5-dCTP (Amersham Biosciences, UK, Cy3 for endometrial cancer and Cy5 for normal endometrium) by using Superscript II reverse transcriptase (Life Technologies, USA). A detailed protocol for cDNA probe labeling, and hybridization is available on the Internet (http://nciarray.nci.nih.gov/reference/NCIReference.shtml). Data acquisition was performed using GenePix 4000B fluorescence scanner (Axon Inc, USA). The ImageQuant (Amersham Biosciences, UK) and GeneSpring (Silicon Genetics, Agilent, USA) software were applied to analyze the expression of proper genes, and normalization was based on the expression level over the whole slides for each sample.

Data analysis

First, we performed random permutation tests to distinguish genes that were expressed differently between EECs and normal endometria. The criteria for selection of discriminating genes were (1) signal/noise ratio of the gene greater than 3.0, (2) p value in a random permutation test lower than 0.001, and (3) expression level in cancer at least two-fold stronger than normal (over-expressed) or half that of normal endometria (under-expressed). Genes that fit these criteria were considered significant for discrimination (Table 1).

Then, we classified the 32 patients into different groups according to the histological grade, FIGO stage, or a combination of grade and stage (e.g., G_1, G_2, G_3 - three groups; Stage I, II, III, IV - four groups). One-way ANOVA analysis was used to identify genes with differences in expression levels among different groups. We selected a p value threshold of 0.001 for statistical significance. As a result, significant genes, which can distinguish the different groups, were selected at the 0.001 level.

Thirdly, unsupervised analysis was done using hierarchical clustering. Hierarchical clustering of the tumor samples using Pearson's correlation similarity metric and average linkage clustering was performed using CLUSTER and TREVIEW software programs obtained at http://rana.lbl.gov/EisenSoftware.htm [10].

Validation of the gene expression using real-time PCR

To verify the data obtained from microarrays, four of the differentially expressed genes: insulin-like growth factor 1 *(IGF1)*, transforming growth factor, beta 2 *(TGFB2)*, metastasis associated 1 *(MTA1)* and transducer of ERBB2, 2 *(TOB2)* were

Table 2. — *Patient information.*

Patient code	Age	Grade	Stage	Histologic type	Myometrial invasion	Pelvic lymph node metastases	Ascites	ER	PR	Operation	Outcome	Survival time (month)
EC01	42	1	Ib	endometrioid	≤ 1/3	-	-	++	+++	2002-5-14		
EC02	55	1	Ib	endometrioid	≤ 1/2	-	-	-	-	2002-6-4		
EC03	49	1	Ib	endometrioid	≤ 1/2	-	-	-	+++	2002-11-6		
EC04	40	1	Ic	endometrioid	> 1/2	-	-	-	+++	2002-10-31		
EC05	48	1	Ib	endometrioid	≤ 1/3	-	-	+	++	2004-1-28		
EC06	62	3	Ib	endometrioid	≤ 1/2	-	-	-	-	2002-10-31	2004-3-9 dead	16
EC07	66	3	Ib	endometrioid	≤ 1/3	-	-	+	+	2003-3-27		
EC08	75	3	Ib	endometrioid	≤ 1/2	-	-	+	+	2003-12-31	2004-11-25 dead	11
EC09	84	3	Ib	endometrioid	≤ 1/2	no lymphadenectomy	-	-	-	2003-12-22	2006-4-20 dead	28
EC10	79	3	IVb	endometrioid	full	no lymphadenectomy	+	-	-	2003-1-9	2003-9-12 dead	8
EC11	83	1-2	Ib	endometrioid	0	-	-	++	+++	2004-2-13		
EC12	46	2	Ib	endometrioid	≤ 1/3	-	-	-	-	2003-11-17		
EC13	54	1	IIIa	endometrioid	≤ 1/2	-	+	+	+	2003-3-19		
EC14	59	1	IVa	endometrioid	> 1/2	-	-	-	-	2003-3-25		
EC15	61	2	IIIc	endometrioid	≤ 1/2	+	-	-	-	2003-4-29		
EC16	61	1	IIb	endometrioid	≤ 1/3	-	-	++	+++	2003-10-21	2005-4-1 dead	18
EC17	51	3	IIIc	endometrioid	≤ 1/2	+	-	+	+	2004-3-18		
EC18	77	2	IIIc	endometrioid	> 1/2	+	-	++	+++	2002-4-23		
EC19	83	3	IVb	endometrioid	full	-	+	+	+	2003-9-3		
EC20	53	2	Ib	endometrioid	≤ 1/3	-	-	-	+++	2001-8-7		
EC21	75	3	IIIc	endometrioid	full	+	-	-	-	2003-11-25		
EC22	64	1-2	IIb	endometrioid	≤ 1/2	-	-	+	+++	2004-3-23		
EC23	68	2	IVb	endometrioid	≤ 1/3	no lymphadenectomy	-	+	+++	2001-8-30	2002-7-26 dead	11
EC24	59	3	IVb	endometrioid	full	+	-	+	+	2002-11-7	2004-11-12 dead	24
EC25	47	1-2	Ib	endometrioid	≤ 1/3	-	-	+	+++	2004-2-24		
EC26	53	1	IIa	endometrioid	> 1/2	no lymphadenectomy		+	+	2002-6-13	2003-10-6 dead	16
EC27	75	3	Ic	endometrioid	> 1/2	-	-	-	-	2001-7-3		
EC28	54	2	IIIa	endometrioid	full	-	-	+	++	2002-6-20		
EC29	51	1	IIIa	endometrioid	> 1/3	-	+	-	+++	2002-4-11		
EC30	47	2	IIIa	endometrioid	≤ 1/2	-	+	-	++	2003-12-25	2006-3 right groin lymph node metastasis	
EC31	72	2	Ib	endometrioid	≤ 1/3	-	-	-	-	2004-7-28		
EC32	57	1	Ib	endometrioid	≤ 1/3	-	-	+	+++	2004-6-17		

Table 3. — *Differentially expressed genes for grade.*

Parametric p-value	GB acc	UNIQID	Name
0.0000011	NM_005397	podocalyxin-like	PODXL
0.0000014	NM_005562	laminin, gamma 2	LAMC2
0.0000055	NM_002407	secretoglobin, family 2A, member 1	SCGB2A1
0.0000309	NM_000624	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	SERPINA3
0.0000518	NM_012436	sperm associated antigen 8	SPAG8
0.0000589	NM_006636	methylene tetrahydrofolate dehydrogenase (NAD+ dependent), methenyltetrahydrofolate cyclohydrolase	MTHFD2
0.0001131	NM_001454	forkhead box J1	FOXJ1
0.0001369	NM_002448	msh homeo box homolog 1 (Drosophila)	MSX1
0.0001469	NM_012144	dynein, axonemal, intermediate polypeptide 1	DNAI1
0.0001703	NM_014643	zinc finger protein 516	ZNF516
0.0002312	NM_003633	ectodermal-neural cortex (with BTB-like domain)	ENC1
0.0002324	NM_003353	urocortin	UCN
0.0002527	NM_014466	tektin 2 (testicular)	TEKT2
0.0002932	NM_001552	insulin-like growth factor binding protein 4	IGFBP4
0.0003206	AK054814	Meis1, myeloid ecotropic viral integration site 1 homolog 4 (mouse)	MEIS4
0.0003549	NM_000538	regulatory factor X-associated protein	RFXAP
0.0004422	NM_004526	MCM2 minichromosome maintenance deficient 2, mitotin (S. cerevisiae)	MCM2
0.0004437	BX647876	hypothetical protein LOC137886	LOC137886
0.0004543	BC035245	hypothetical protein LOC253782	LOC253782
0.0004627	NM_004670	3'-phosphoadenosine 5'-phosphosulfate synthase 2	PAPSS2
0.0004712	NM_016447	membrane protein, palmitoylated 6 (MAGUK p55 subfamily member 6)	MPP6
0.0004812	NM_022476	fused toes homolog (mouse)	FTS
0.000498	NM_033514	LIM and senescent cell antigen-like domains 3	LIMS3
0.0005049	NM_000186	H factor 1 (complement)	HF1
0.0005599	NM_001874	carboxypeptidase M	CPM
0.0005913	NM_024096	XTP3-transactivated protein A	XTP3TPA
0.0006824	NM_002609	platelet-derived growth factor receptor, beta polypeptide	PDGFRB
0.0007076	NM_019106	septin 3	SEPT3

continued

0.0007492	NM_000867	5-hydroxytryptamine (serotonin) receptor 2B	HTR2B
0.000765	AB058704	hypothetical protein FLJ12303	FLJ12303
0.0007657	NM_003551	non-metastatic cells 5, protein expressed in (nucleoside-diphosphate kinase)	NME5
0.0009095	AL136878	paraneoplastic antigen MA3	PNMA3
0.0009828	NM_003903	CDC16 cell division cycle 16 homolog (S. cerevisiae)	CDC16

Table 4. — *Differentially expressed genes for stage.*

Parametric p-value	GB acc	UNIQID	Name
0.0000879	NM_175061	juxtaposed with another zinc finger gene 1	JAZF1
0.0003558	NM_006855	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3	KDELR3
0.0003618	NM_002614	PDZ domain containing 1	PDZK1
0.0004044	NM_012329	monocyte to macrophage differentiation-associated	MMD
0.0004434	NM_004689	metastasis associated 1	MTA1
0.0004505	NM_014369	protein tyrosine phosphatase, non-receptor type 18 (brain-derived)	PTPN18
0.0007339	NM_000355	transcobalamin II; macrocytic anemia	TCN2
0.0007652	NM_024771	hypothetical protein FLJ13848	FLJ13848
0.0008353	NM_015079	KIAA1055 protein	KIAA1055
0.0009589	NM_002406	mannosyl (alpha-1,3-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase	MGAT1

Table 5. — *Differentially expressed genes for grade and stage.*

Parametric p-value	GB acc	UNIQID	Name
0.0000002	NM_000702	ATPase, Na+/K+ transporting, alpha 2 (+) polypeptide	ATP1A2
0.0000034	NM_014428	tight junction protein 3 (zona occludens 3)	TJP3
0.0000041	NM_005562	laminin, gamma 2	LAMC2
0.0000057	NM_005488	target of myb1 (chicken)	TOM1
0.0000105	NM_003071	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 3	SMARCA3
0.0000141	NM_005397	podocalyxin-like	PODXL
0.0000176	AB007867	plexin B1	PLXNB1
0.0000196	NM_014643	zinc finger protein 516	ZNF516
0.0000319	NM_003068	snail homolog 2 (Drosophila)	SNAI2
0.000035	NM_015889	PC2 (positive cofactor 2, multiprotein complex) glutamine/Q-rich-associated protein	PCQAP
0.0000364	NM_000624	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	SERPINA3
0.0000381	NM_014030	G protein-coupled receptor kinase-interactor 1	GIT1
0.0000381	NM_012436	sperm associated antigen 8	SPAG8
0.0000383	NM_014567	breast cancer anti-estrogen resistance 1	BCAR1
0.0000602	NM_003926	methyl-CpG binding domain protein 3	MBD3
0.0000611	NM_003012	secreted frizzled-related protein 1	SFRP1
0.0000643	NM_000102	cytochrome P450, family 17, subfamily A, polypeptide 1	CYP17A1
0.0000694	NM_014466	tektin 2 (testicular)	TEKT2
0.0000738	NM_015270	adenylate cyclase 6	ADCY6
0.0000778	BC012542	KIAA0186 gene product	KIAA0186
0.0000878	NM_013227	aggrecan 1 (chondroitin sulfate proteoglycan 1, large aggregating proteoglycan, antigen identified b	AGC1
0.0000878	NM_025241	UBX domain containing 1	UBXD1
0.0000921	NM_014176	HSPC150 protein similar to ubiquitin-conjugating enzyme	HSPC150
0.0000956	NM_002560	purinergic receptor P2X, ligand-gated ion channel, 4	P2RX4
0.000098	NM_016219	mannosidase, alpha, class 1B, member 1	MAN1B1
0.0001074	NM_017732	hypoxia-inducible factor prolyl 4-hydroxylase	PH-4
0.000129	NM_004850	Rho-associated, coiled-coil containing protein kinase 2	ROCK2
0.0001378	NM_145110	mitogen-activated protein kinase kinase 3	MAP2K3
0.0001696	NM_001999	fibrillin 2 (congenital contractural arachnodactyly)	FBN2
0.0001755	NM_032019	histone deacetylase 10	HDAC10
0.000178	NM_004126	guanine nucleotide binding protein (G protein), gamma 11	GNG11
0.0002336	NM_145342	mitogen-activated protein kinase kinase kinase 7 interacting protein 2	MAP3K7IP2
0.0002351	X52973	carboxylesterase 1 (monocyte/macrophage serine esterase 1)	CES1
0.0002379	AK074088	tumor protein p53 inducible protein 5	TP53I5
0.0002385	NM_000831	glutamate receptor, ionotropic, kainate 3	GRIK3
0.0002402	NM_003756	eukaryotic translation initiation factor 3, subunit 3 gamma, 40kDa	EIF3S3
0.0002411	XM_039877	mucin 5, subtypes A and C, tracheobronchial/gastric	MUC5AC
0.0002425	NM_003182	tachykinin, precursor 1 (substance K, substance P, neurokinin 1, neurokinin 2, neuromedin L, neuroki	TAC1
0.0002489	NM_178507	NS5ATP13TP2 protein	NS5ATP13TP2
0.0002497	NM_032186	KIAA1221 protein	KIAA1221

continued

Parametric p-value	GB acc	UNIQID	Name
0.0002605	NM_033514	LIM and senescent cell antigen-like domains 3	LIMS3
0.0002739	NM_016272	transducer of ERBB2, 2	TOB2
0.0002863	AK057681	hypothetical protein LOC338692	LOC338692
0.0002971	NM_002614	PDZ domain containing 1	PDZK1
0.0002982	NM_014815	thyroid hormone receptor associated protein 4	THRAP4
0.0003107	NM_024608	nei endonuclease VIII-like 1 (E. coli)	NEIL1
0.0003208	NM_024029	hypothetical protein MGC3262	MGC3262
0.0003267	NM_002407	secretoglobin, family 2A, member 1	SCGB2A1
0.0003277	NM_006992	B7 gene	B7
0.0003353	NM_003350	ubiquitin-conjugating enzyme E2 variant 2	UBE2V2
0.0003361	NM_005528	DnaJ (Hsp40) homolog, subfamily C, member 4	DNAJC4
0.0003501	NM_004950	dermatan sulfate proteoglycan 3	DSPG3
0.0003549	NM_003682	MAP-kinase activating death domain	MADD
0.0003609	NM_003153	signal transducer and activator of transcription 6, interleukin-4 induced	STAT6
0.0003682	NM_014409	TAF5-like RNA polymerase II, p300/CBP-associated factor (PCAF)-associated factor, 65kDa	TAF5L
0.0003717	NM_001874	carboxypeptidase M	CPM
0.000374	NM_002096	general transcription factor IIF, polypeptide 1, 74kDa	GTF2F1
0.0003831	NM_004187	Jumonji, AT rich interactive domain 1C (RBP2-like)	JARID1C
0.0003838	NM_032309	coiled-coil-helix-coiled-coil-helix domain containing 5	CHCHD5
0.0003873	NM_005003	NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1, 8kDa	NDUFAB1
0.0004454	NM_003651	cold shock domain protein A	CSDA
0.0004485	NM_000216	Kallmann syndrome 1 sequence	KAL1
0.0004524	NM_012413	glutaminyl-peptide cyclotransferase (glutaminyl cyclase)	QPCT
0.0004532	NM_018004	hypothetical protein FLJ10134	FLJ10134
0.0004707	NM_002448	msh homeo box homolog 1 (Drosophila)	MSX1
0.0004949	NM_006200	proprotein convertase subtilisin/kexin type 5	PCSK5
0.0005134	NM_030762	basic helix-loop-helix domain containing, class B, 3	BHLHB3
0.0005262	NM_014732	KIAA0513 gene product	KIAA0513
0.0005297	NM_005496	SMC4 structural maintenance of chromosomes 4-like 1 (yeast)	SMC4L1
0.0005453	NM_014372	ring finger protein 11	RNF11
0.000547	NM_003946	nucleolar protein 3 (apoptosis repressor with CARD domain)	NOL3
0.0005471	NM_023034	Wolf-Hirschhorn syndrome candidate 1-like 1	WHSC1L1
0.0005591	NM_033209	Thy-1 cell surface antigen	THY1
0.0005687	NM_014369	protein tyrosine phosphatase, non-receptor type 18 (brain-derived)	PTPN18
0.0005711	NM_002690	polymerase (DNA directed), beta	POLB
0.0006034	NM_005770	small EDRK-rich factor 2	SERF2
0.0006743	NM_004613	transglutaminase 2 (C polypeptide, protein-glutamine-gamma-glutamyltransferase)	TGM2
0.0006856	NM_014630	KIAA0211 gene product	KIAA0211
0.000692	NM_004425	extracellular matrix protein 1	ECM1
0.0007	NM_014759	phytanoyl-CoA hydroxylase interacting protein	PHYHIP
0.0007138	NM_002085	glutathione peroxidase 4 (phospholipid hydroperoxidase)	GPX4
0.000719	NM_014766	secernin 1	SCRN1
0.0007439	NM_015391	DKFZP566D193 protein	DKFZP566D193
0.0007547	BC035245	hypothetical protein LOC253782	LOC253782
0.0007745	NM_018841	guanine nucleotide binding protein (G protein), gamma 12	GNG12
0.0008328	NM_007263	coatomer protein complex, subunit epsilon	COPE
0.0008377	NM_003542	histone 1, H4f	HIST1H4F
0.0008423	NM_175852	taxilin	DKFZp451J0118
0.0008444	NM_006855	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3	KDELR3
0.0008462	NM_022761	chromosome 11 open reading frame 1	C11orf1
0.0008785	NM_004470	FK506 binding protein 2, 13kDa	FKBP2
0.0008808	NM_023037	hypothetical protein CG003	13CDNA73
0.0008875	NM_020935	ubiquitin specific protease 37	USP37
0.0008876	NM_006807	chromobox homolog 1 (HP1 beta homolog Drosophila)	CBX1
0.0009004	NM_003130	sorcin	SRI
0.0009025	NM_003070	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2
0.0009048	NM_002277	keratin, hair, acidic, 1	KRTHA1
0.0009097	NM_001454	forkhead box J1	FOXJ1
0.0009163	NM_000954	prostaglandin D2 synthase 21kDa (brain)	PTGDS
0.0009618	NM_002730	protein kinase, cAMP-dependent, catalytic, alpha	PRKACA
0.0009646	NM_007085	follistatin-like 1	FSTL1
0.0009675	NM_005079	tumor protein D52	TPD52
0.0009804	NM_002775	protease, serine, 11 (IGF binding)	PRSS11
0.0009806	XM_291141	KIAA0303 protein	KIAA0303

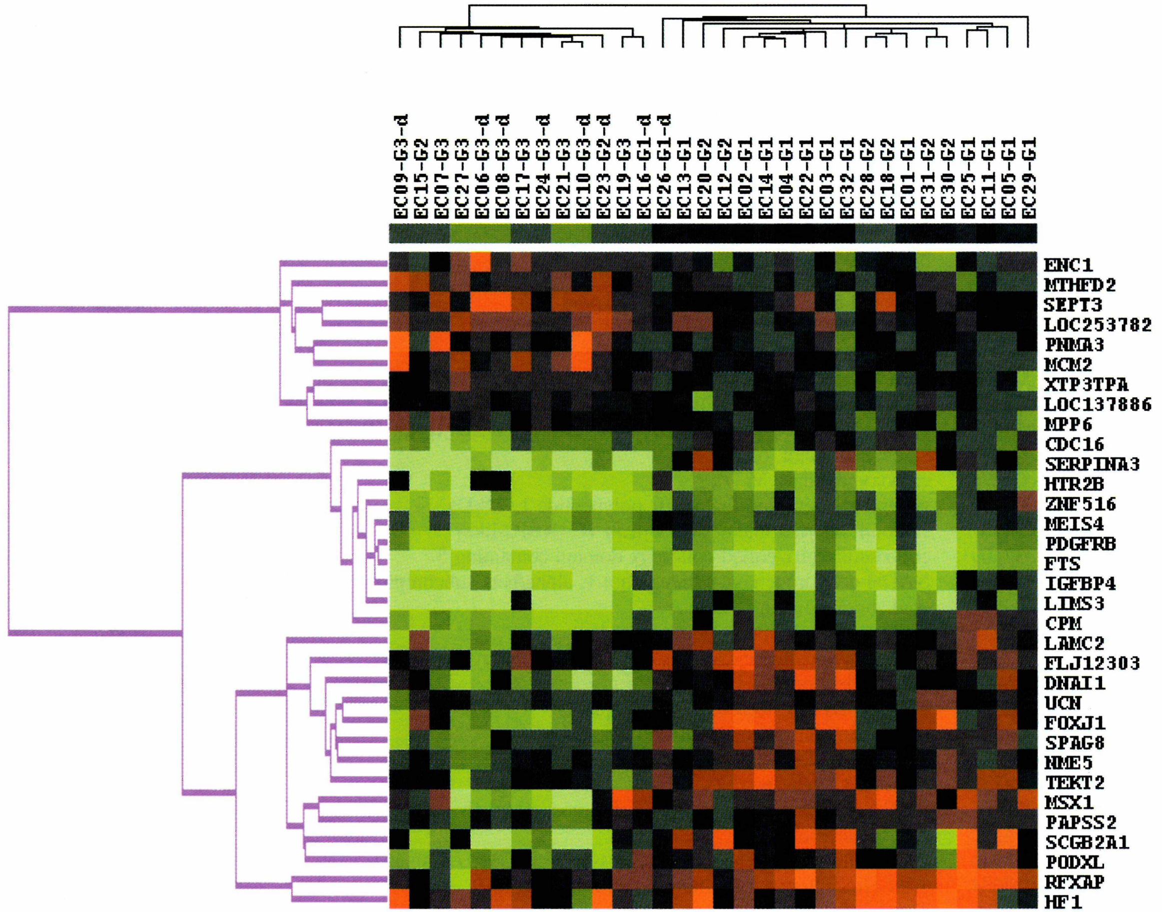

Figure 2a. — Dendrograms of unsupervised hierarchical clustering with the 33 DE genes for grade. On top are individual samples; on the side are 33 DE genes. Genes in red were increased, genes in green decreased, and genes in black invariant compared with the mean expression level. Thirty-two tumor samples were separated into two groups: the left poorly-differentiated (G_3) EECs and the right well/moderately-differentiated ($G_{1/2}$) EECs. Seven of the eight decreased EEC patients fell into the left group.

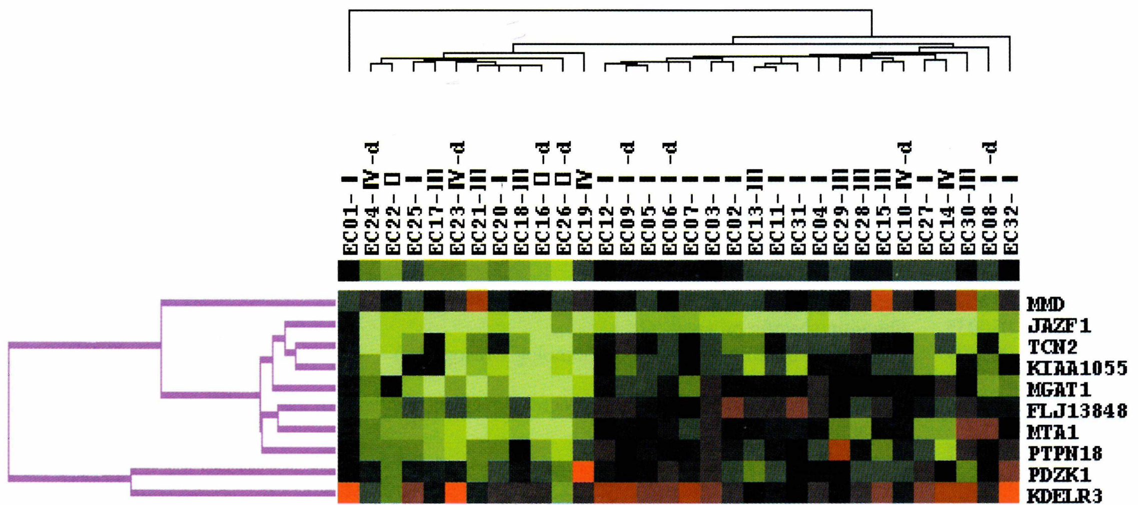

Figure 2b. — Classification with the ten genes identified from the comparison of different stage EECs. Thirty EECs were separated into two groups, and two EECs (EC01 and EC32) did not belong to either group. Most of the Stage I EECs (12/16) were separated into the right group, and 13 advanced EECs (Stage III/IV) and samples from the eight deceased patient EECs were distributed into both groups.

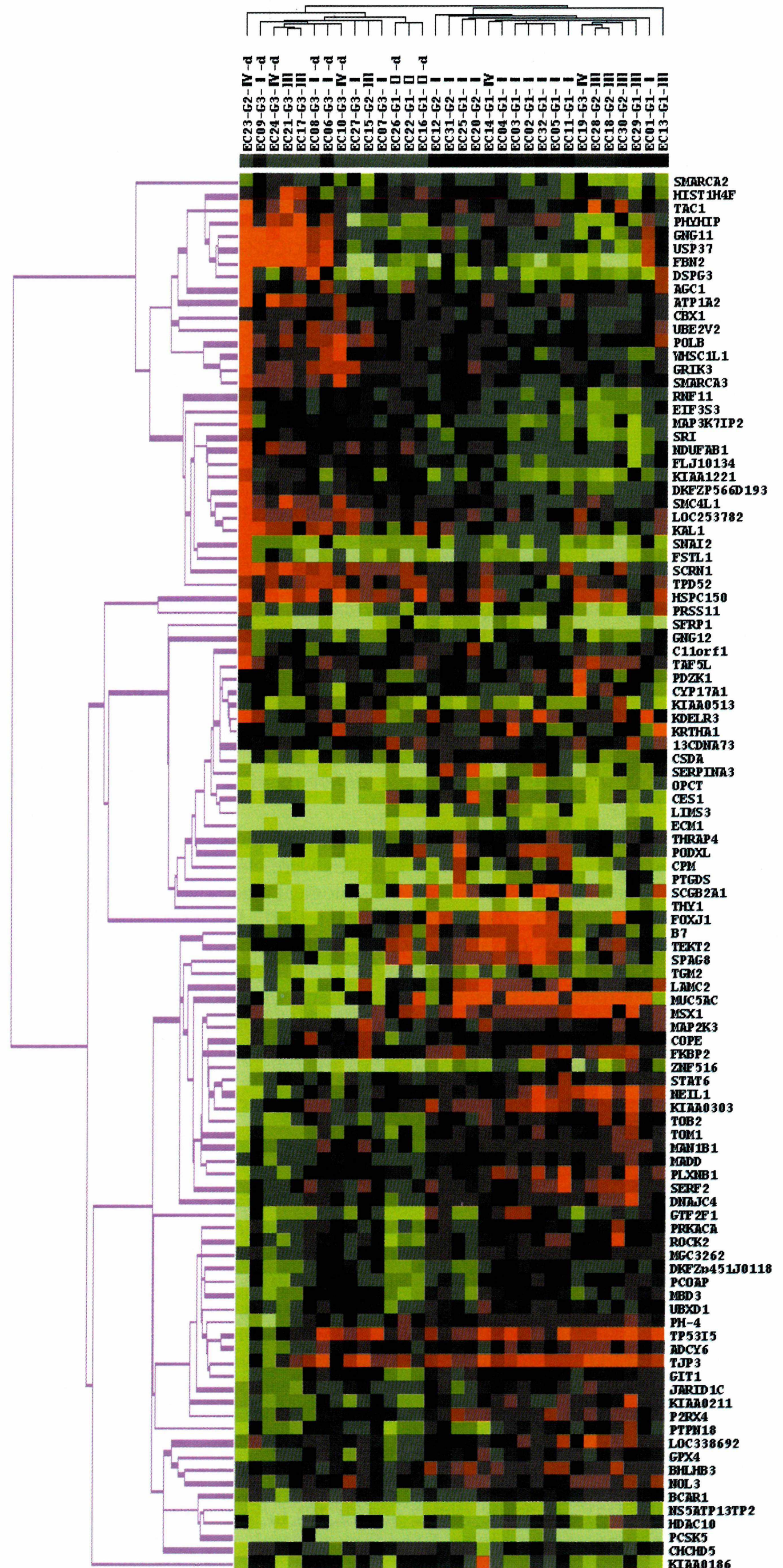

Figure 2c. — Classification with the 104 genes identified from the comparison of different grade and stage EECs. Thirty-two tumor samples were separated into two groups. Within the right group, 11 of the 18 tumors were early stage, seven were advanced tumors (Stage III/IV), and they were all alive with intense follow-up. Within the left group, 11 of the 14 tumors were poorly-differentiated or advanced tumors. Very interestingly, eight dead patients were all classified in the left group.

chosen for further validation in six EECs (EC03, EC09, EC12, EC13, EC18, EC23) of different grade and stage and normal endometria samples with real-time PCR. The first strand cDNA was prepared with SuperScript III Reverse Transcription Reagents (invitrogen, USA). The primers were designed with PrimerExpress software (Applied Biosystems, USA) and synthesized locally: MTA1 (F-GTCGTCTCTGCGCATCTTGTT, R-TGCAGATAGAAAGCCTGCCTG), TGFB2 (F-ACACT-GTCCCTGCTGCACTTT, R-TGGAGGTGCCATCAATAC-CTG, IGF1F-AAGCCTGTCCACCCTTGAGAA, R-ACATTG-GCATAGCTGGCCAA), TOB2 (F-AGCCCCTGATGCCTA-TGAGAA, R-AACTCTAAAATCCCACCCTGCC). The ß-actin gene (primers: F-GCAGAAGGAAATCACAGCCCT, R-GCTGATCCACATCTGCTGGAA) was used as an internal control for expression level. Real-time PCR was carried out in an ABI Prism 7000 Real-time PCR System using SYBR Green I (Applied Biosystems, USA) for detection. Amplification was carried out in a total volume of 25 µl containing 1xSYBR Green I, 100 nM each primer, 10 nM calibration dye (Bio-Rad) and 2 µl of 1:2 diluted cDNA. The thermal profile consisted of a 10 min of *Taq* polymerase activation at 95°C, followed by 45 cycles of PCR at 95°C for 15 sec and 60°C for 1 min. At the end of the PCR, the temperature was increased 0.5°C every 10 sec from 60°C to 95°C, and the fluorescence was measured every 10 sec to construct the melting curve. A non-template control was run with every assay, and all determinations were performed at least in duplicates to achieve reproducibility. The relative quantitation was calculated for each sample using the comparative C_T method.

Results

Differentially expressed (DE) genes in different grade and stage EECs

We then explored the gene expression profiles of EECs with different grades and stages. The 32 EECs were classified into different groups according the histological grade and stage (e.g., G_1, G_2, G_3 - three groups; Stage I, II, III, IV - four groups). Analysis of variance was used to identify genes with differences in expression levels among different groups. We found 33 differentially expressed genes in different grade EECs, and ten differentially expressed genes in different stage EECs at $p = 0.001$ (Tables 3 and 4). When we classified 32 EECs with the combination of grade and stage, 104 differentially expressed genes were identified at the 0.001 level (Table 5).

Validation of the gene expression

To verify the data obtained from microarrays, four of the differentially expressed genes *(MTA1, TGFB2, IGF1* and *TOB2)* due to their function in the tumor were selected for real-time PCR in six EECs. The growth factors, *IGF1* and *TGFB2* were universally down-regulated in endometrial cancer compared with normal endometria. *MTA1* was differentially expressed in the different stages of endometrial cancer, and *TOB2* was differentially expressed in a combination of stage and grade analysis. The expression levels of *IGF1* and *TGFB2* were correlated in real-time PCR for each sample. There were slight discrepancies with the expression level of *MTA1* (EC18 and EC23) and *TOB2* (EC12 and EC18) in real-time PCR and microarray (Table 1 and Figure 1).

Classification of EECs with gene expression profiles

We further investigated whether clinical EECs could be classified into groups on the basis of their gene-expression profiles, which were compared with the traditional classification. For this purpose, unsupervised analysis was done using hierarchical clustering with the 33, 10 and 104 differentially expressed genes. The 32 EECs were separated into two groups according to the expression of 33 differentially expressed genes (Figure 2a). The result of hierarchical clustering analysis agreed with classification of the histological grade (conformity rate was 90%).

The EECs were also separated into two groups according to the expression of ten differentially expressed genes, and two EECs did not fit into any group (Figure 2b), but the conformity rate of hierarchical clustering analysis and FIGO stage was low (only 60%). Furthermore, hierarchical clustering was also done with the 104 differentially expressed genes, which were identified with the combination of grade and stage. Thirty-two EECs were also separated into two groups. It was interesting that the six dead EECs all fell into the left high-risk EEC group when the clinical information was added into each case of EEC (Figure 2c).

Discussion

EEC is the type of endometrial cancer from which the majority of patients suffer. Its prognosis and treatment regimens depend on clinical and histological characterization of the tumors. Combined with differentiation and stage, EECs can be divided into low (G_1/IA IB), medium (G_2/IB IC G_1/IC) and high (G_3/IB IC) risk [11], but this system is not adequate to predict the clinical outcome accurately and results in empirical treatment for EECs. Gene expression profiling using cDNA microarray is a powerful approach to survey the biological characteristics of tumor cells. cDNA microarrays have been used to investigate the differential gene expression profile of endometrial cancer [12-16], including different histological types [17, 18], prediction of lymph node metastasis [19], and stratification of risk for recurrence [20]. In this study, we have compared the gene expression profile with two known prognosis factors (FIGO stage and histological grade), and identified 104 gene expression signatures, which may be used to assess risk and prognosis.

We explored the differential gene expression profiles of EECs with different grades and stages. Thirty-three and ten genes were found, respectively, to be differentially expressed in EECs of different grades and stages (Table 3, 4). Similarly, 104 differentially expressed genes were identified in EECs when the classification was based on a combination of grade and stage (Table 5). Among these differentially expressed genes, genes of cell cycle, adhesion molecule, growth factor, transcriptional factor, protein kinase and metabolized enzyme were up-regulated, while the tumor suppressor, extracelluar matrix protein and enzyme, gap junction protein, and growth differentiation factor were down-regulated. Many genes, including LAMC2 [21], MTA1 [22], TPD52 [23], SERPINA3 [24],

MSX1 [25], TOB2 [26], ECM1 [27], have previously been shown to be involved in carcinogenesis are present on our lists of differentially expressed genes, providing validity to our microarray analysis. To further verify the quality of our microarray data, a subset of four genes (*IGF1, TGFB2, MTA1* and *TOB2*) was examined in six samples using quantitative real-time PCR.

Since FIGO stage, histological type and grade cannot adequately predict the outcome of EEC patients, with the 33, ten and 104 differentially expressed genes, unsupervised analysis was done using hierarchical clustering to further understand the underlying biology of the heterogeneity of EECs. The classification of the 32 EECs on the basis of their gene-expression profiles was compared with the traditional classification.

In Figure 2a, the 32 EECs were separated into two groups according to genes differentially expressed in various grades. All of the ten poorly-differentiated (G_3) EECs were classified into the left high-risk group, and the samples from seven of the eight deceased patients also fell into this group. However, most of the well and medium-differentiated ($G_{1/2}$) EECs were separated into the right group. Sample 26 was well-differentiated EEC, but the patient had rectum cancer and lung metastasis before being diagnosed with EEC, and died 16 months after surgery. Surprisingly, one G_1 sample (EC16) was clustered close to the G3 EECs and fell into the left group. In Figure 2b, EECs were separated into two groups according to genes differentially expressed in various FIGO stages, and two EECs (EC01 and EC32) with Stage I did not belong to either group. Most of the Stage I EECs (12/16) were separated into the right group, and 13 advanced EECs (Stage III/IV) and samples from the eight deceased EECs were distributed into both groups. Our observations are different from those of Smid-Koopman *et al.* [15], who reported that classification of endometrial tumors on the basis of their gene expression profiles showed similarity with the FIGO staging system. However, the total number of samples and the number of the Stage III tumors in that study were limited, as there were only 12 endometrial cancer samples, including ten EECs (one IA, seven IB, one IC and one IIIA) and two papillary serous adenocarcinomas (one IB and one IIIC). In our study, there were variances in the classification of 32 EECs on the basis of the gene expression profile and classification with the FIGO staging system. Here, we show that the conformity rate of the classification with gene expression profile was higher in grade than FIGO stage (90% and 60%, respectively). Not all patients had surgical staging which may be one of the factors (patients EC08 and EC09 were 75 and 84 years old, both had severe hypertension and diabetes mellitus. Patient EC08 received a pelvic lymph node biopsy and patient EC09 only underwent total abdominal hysterectomy and bilateral salpingo-oophorectomy) which also suggests that FIGO stage may not adequately reflect the real characteristics of EECs. It was necessary to combine the histological grade and FIGO stage and to identify the differentially expressed genes.

With the selected 104 genes, 32 EECs were also separated into two groups (Figure 2c). Six advanced tumors and nine poorly-differentiated EECs were separated into the left group. It was interesting that the samples from the eight deceased patients (EC24, EC06, EC10, EC08, EC09, EC23, EC26, EC16) all fell into the left high-risk EEC group when the clinical information was added into each tumor sample (eight patients died from tumor recurrence or metastasis 24, 16, 8, 11, 28, 11, 16 and 18 months after surgery). It was 32, 28, 61, 39, 40, and 28 months after surgery for patients EC21, EC17, EC27, EC15, EC07, and EC22 up to July 2006. There was no evidence of recurrence or metastasis for the six EECs in the follow-up. Long-time follow-up may need to determine whether these individuals in fact had high-risk EECs. The right group comprised 11 Stage I EECs and seven advanced EECs (Stage III/IV). As of July 2006, it had been 40, 34 and 49, 51, 31, 51, 40 months, respectively, after surgery for the two Stage IV EECs (EC14 and EC19) and five Stage III EECs (EC28, EC18, EC30, EC29 and EC13). Except that sample EC30 had right groin lymph node metastasis 31 months postoperation, there was no evidence of recurrence or metastasis for the other six patients in the subsequent follow-up. The fact that the seven advanced EECs were classified into the same group as the 11 early ones may indicate that histological grade and FIGO stage are both insufficient in accurately predicting the outcome of endometrial cancer. The classification on the basis of the gene expression profile with the 104 differentially expressed genes may be of benefit for the identification of high-risk EECs.

In conclusion, the current study has established the clinical feasibility of the cDNA microarray technique to generate gene expression profiles of EEC samples. The molecular classification of EECs has the potential to serve as a classification system that is more accurate in predicting the biological behavior of the tumors than histological grade and FIGO stage classification. However, the number of tissue samples and the time of clinical follow-up were limited in the current study. A larger number of EEC samples, combined with long-term clinical follow-up, are needed to confirm current conclusion. The roles of the differentially expressed genes in EEC should also be explored in the near future. Expression analysis of key genes that predict the prognosis of endometrial cancer will allow the customization of treatments based on individual needs, thereby improving the long-time survival rate for endometrial cancer patients.

Acknowledgment

We thank D. Hogg and J. Loo of the University of Toronto for reviewing and revising the manuscript.

This work was supported by the Chinese National Natural Science Foundation, Grant 30371481; the Natural Science Foundation of Shanghai, Grant 06ZR14053; and the Key Project of the Shanghai Health Bureau, Grant 2005ZD002.

References

[1] Amant F., Moerman P., Neven P., Timmerman D., Van Limbergen E., Vergote I.: "Endometrial cancer". *Lancet,* 2005, *366*, 491.

[2] American College of Obstetricians and Gynecologists: "ACOG practice bulletin, clinical management guidelines for obstetrician-gynecologists, number 65, August 2005: management of endometrial cancer". *Obstet. Gynecol.*, 2005, *106,* 413.

[3] Bucca G., Carruba G., Saetta A., Muti P., Castagnetta L., Smith C.P.: "Gene expression profiling of human cancers". *Ann. N Y Acad Sci.,* 2004, *1028,* 28.

[4] Quackenbush J.: "Microarray analysis and tumor classification". *N. Engl. J. Med.*, 2006, *354*, 2463.

[5] Sorlie T., Perou C.M., Tibshirani R., Aas T., Geisler S., Johnsen H. *et al.*: "Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications". *Proc. Natl. Acad. Sci. USA,* 2001, *98,* 10869.

[6] Alizadeh A.A., Eisen M.B., Davis R.E., Ma C., Lossos I.S,. Rosenwald A. *et al.*: "Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling". *Nature,* 2000, *403,* 503.

[7] Dave S.S., Fu K., Wright G.W., Lam L.T., Kluin P., Boerma E.J. *et al.*: "Molecular diagnosis of Burkitt's lymphoma". *N. Engl. J. Med.,* 2006, *354*, 2431.

[8] van't Veer L.J., Dai H., van de Vijver M.J., He Y.D., Hart A.A., Mao M. *et al.*: "Gene expression profiling predicts clinical outcome of breast cancer". *Nature*, 2002, *415*, 530.

[9] Ouellet V., Provencher D.M., Maugard C.M., Le Page C., Ren F., Lussier C. *et al.*: "Discrimination between serous low malignant potential and invasive epithelial ovarian tumors using molecular profiling". *Oncogene,* 2005, *24,* 4672.

[10] Eisen M.B., Spellman P.T., Brown P.O., Botstein D.: "Cluster analysis and display of genome-wide expression patterns". *Proc. Natl. Acad. Sci USA,* 1998, 95, 14863.

[11] Kitchener H.: "Management of endometrial cancer". *Eur. J. Surg. Oncol.*, 2006, 32, 838.

[12] Mutter G.L., Baak J.P., Fitzgerald J.T. *et al.*: "Global expression changes of constitutive and hormonally regulated genes during endometrial neoplastic transformation". *Gynecol. Oncol.*, 2001, *83*, 177.

[13] Gunter Kieback D., Fischer D.C.: "Gene expression profile in endometrioid endometrial carcinoma". *Gynecol. Oncol.,* 2001, *83,* 175.

[14] Planaguma J., Diaz-Fuertes M., Gil-Moreno A., Abal M., Monge M., Garcia A. *et al.*: "A differential gene expression profile reveals overexpression of RUNX1/AML1 in invasive endometrioid carcinoma". *Cancer Res.,* 2004, *64,* 8846.

[15] Smid-Koopman E., Blok L.J., Helmerhorst T.J., Chadha-Ajwani S., Burger C.W., Brinkmann A.O. *et al.*: "Gene expression profiling in human endometrial cancer tissue samples: utility and diagnostic value". *Gynecol. Oncol.,* 2004, *93,* 292.

[16] Ferguson S.E., Olshen A.B., Viale A., Awtrey C.S., Barakat R.R., Boyd J.: "Gene expression profiling of tamoxifen-associated uterine cancers: evidence for two molecular classes of endometrial carcinoma". *Gynecol. Oncol.*, 2004, *92,* 719.

[17] Risinger J.I., Maxwell G.L., Chandramouli G.V., Jazaeri A., Aprelikova O., Patterson T. *et al.*: "Microarray analysis reveals distinct gene expression profiles among different histologic types of endometrial cancer". *Cancer Res.,* 2003, *63,* 6.

[18] Maxwell G.L., Chandramouli G.V., Dainty L., Litzi T.J., Berchuck A., Barrett J.C. *et al.*: "Microarray analysis of endometrial carcinomas and mixed mullerian tumors reveals distinct gene expression profiles associated with different histologic types of uterine cancer". *Clin. Cancer Res.,* 2005, *11*, 4056.

[19] Bidus M.A., Risinger J.I., Chandramouli G.V., Dainty L.A., Litzi T.J., Berchuck A. *et al.*: "Prediction of lymph node metastasis in patients with endometrioid endometrial cancer using expression microarray". *Clin. Cancer Res.,* 2006, *12,* 83.

[20] Ferguson S.E., Olshen A.B., Viale A., Barakat R.R., Boyd J.: "Stratification of intermediate-risk endometrial cancer patients into groups at high risk or low risk for recurrence based on tumor gene expression profiles". *Clin. Cancer Res.,* 2005, *11*, 2252.

[21] Takahashi S., Hasebe T., Oda T., Sasaki S., Kinoshita T., Konishi M. *et al.*: "Cytoplasmic expression of laminin gamma2 chain correlates with postoperative hepatic metastasis and poor prognosis in patients with pancreatic ductal adenocarcinoma". *Cancer,* 2002, *94,* 1894.

[22] Hofer M.D., Kuefer R., Varambally S., Li H., Ma J., Shapiro G.I. *et al.*: "The role of metastasis-associated protein 1 in prostate cancer progression". *Cancer Res.*, 2004, *64,* 825.

[23] Byrne J.A., Balleine R.L., Schoenberg Fejzo M., Mercieca J., Chiew Y.E., Livnat Y. *et al.*: "Tumor protein D52 (TPD52) is over-expressed and a gene amplification target in ovarian cancer". *Int. J. Cancer,* 2005, *117*, 1049.

[24] Yamamura J., Miyoshi Y., Tamaki Y., Taguchi T., Iwao K., Monden M. *et al.*: "mRNA expression level of estrogen-inducible gene, alpha 1-antichymotrypsin, is a predictor of early tumor recurrence in patients with invasive breast cancers". *Cancer Sci.,* 2004, *95,* 887.

[25] Park K., Kim K., Rho S.B., Choi K., Kim D., Oh S.H. *et al.*: "Homeobox Msx1 interacts with p53 tumor suppressor and inhibits tumor growth by inducing apoptosis". *Cancer Res.,* 2005, *65*, 749.

[26] Kawamura-Tsuzuku J., Suzuki T., Yoshida Y., Yamamoto T.: "Nuclear localization of Tob is important for regulation of its antiproliferative activity". *Oncogene,* 2004, *23,* 6630.

[27] Wang L., Yu J., Ni J., Xu X.M., Wang J., Ning H. *et al.*: "Extracellular matrix protein 1 (ECM1) is over-expressed in malignant epithelial tumors". *Cancer Lett.,* 2003, *200*, 57.

Address reprint requests to:
X.P. WAN, M.D.
Department of Obstetrics and Gynecology
Shanghai Jiao Tong Univ.
Affiliated First People's Hospital
#85, Wujin Road
Shanghai 200080 (China)