Microarray expression identification of differentially expressed genes in serous epithelial ovarian cancer compared with bulk normal ovarian tissue and ovarian surface scrapings

DAN GRISARU[1,2,5], JAN HAUSPY[1,3], MONA PRASAD[2], MONIQUE ALBERT[2], K. JOAN MURPHY[1], ALLAN COVENS[4], PASCALE F. MACGREGOR[2] and BARRY ROSEN[1]

[1]Division of Gynecological Oncology, Princess Margaret Hospital, [2]Microarray Centre, [3]Division of Cellular and Molecular Biology, Ontario Cancer Institute, University Health Network; [4]Division of Gynecologic Oncology, Toronto-Sunnybrook Regional Cancer Centre, Toronto, Canada

Received March 24, 2007; Accepted April 25, 2007

Abstract. The lack of reliable early detection of ovarian cancer and the absence of specific symptoms result in diagnosis of ovarian cancer at advanced stage in the majority of the patients. Through gene expression profiling we can identify important genes that may help understand the evolution from normal ovarian tissue to ovarian cancer. The gene expression profiles of 7 normal ovaries and 26 ovaries with serous epithelial ovarian cancer (SEOC) were examined by cDNA microarrays using supervised and unsupervised analysis, with sequential significance filtering. Real-time RT-PCR was used to measure and compare the expression levels of 5 selected genes: WAP four-disulfide core domain protein HE4 (WAP, up-regulated), secreted phosphoprotein 1 (SPP1, osteopontin; up-regulated), activin A receptor, type I (ACVR1, up-regulated), tumor necrosis factor (TNF superfamily, member 2; TNF, up-regulated) and decorin (DCN, down-regulated) in 4 epithelial scrapings and in 6 bulk-extracted normal ovaries. The gene expression profile of SEOC was not dependent on the stage of the disease at diagnosis. A supervised microarray data analysis identified a subset of 329 genes showing significant differential expression between SEOC samples and bulk normal ovarian tissue and ovarian surface scrapings, including several new genes such as TNFα and activin A receptor type I. The real-time RT-PCR for the up-regulated genes did not differ significantly between normal ovarian epithelial scrapings and bulk-extracted ovaries. However, decorin showed a statistically significant difference (P=0.0073) in expression between epithelial scrapings and bulk-extracted ovaries. Previously uncharacterized genes are associated with the malignant phenotype of SEOC. Bulk normal ovarian tissue may serve as control for SEOC tissue in gene expression profiling. Gene expression profiling and sequential statistical analyses of gene subsets can identify new genes and molecular pathways affecting development of SEOC. The genes of interest can be potential targets for future research of SEOC.

Correspondence to: Dr Dan Grisaru [5]*Present address*: Gynecologic Oncology Service, Head, Tel-Aviv Sourasky Medical Center, 6 Weizman Street, Tel-Aviv 64239, Israel
E-mail: grisaro@post.tau.ac.il

Key words: serous epithelial ovarian cancer, microarrays, osteopontin, activin A receptor, tumor necrosis factor, decorin

Introduction

Epithelial cancer of the ovary is the sixth most common cancer in women and is the leading cause of death from a gynaecological malignancy (1). The lack of reliable methods of early detection and the absence of specific symptoms result in late stage at diagnosis in the majority of the patients (70%). The 5-year survival for patients with early disease [International Federation of Gynecology and Obstetrics (FIGO) stages I/II] differs from that of patients with advanced disease (FIGO stages III/IV) (2) and little progress has been made over the past decades to improve long-term survival (3).

We conducted expression profiling of 26 serous epithelial ovarian tumor samples obtained from 24 patients, and tissue from 7 normal ovary samples, using cDNA microarrays spotted with 19,200 known genes and Expressed Sequence Tags (ESTs), and used SAM (significance analysis of microarrays) (4) to reveal genes differentially expressed in SEOC as compared to normal ovaries.

Patients and methods

Ovarian tissues. Twenty-six snap-frozen ovarian cancer tissue samples from 24 patients were obtained from the University of Toronto Ovarian Tissue Bank and Database. All tissue samples in the University of Toronto Ovarian Tissue Bank are banked after consent from each patient to use tissue removed during surgery for later use in basic scientific research. This study was approved by the Research Ethics Board of the University Health Network Toronto. Samples were selected to include carcinomas of only serous histology, high grade, and to include carcinomas diagnosed at early and late patho-

Table I. Histopathological data of EOC patients and normal ovaries.

Patient sample	Type	Stage	Grade
OVD12	Ser Pap	I	2
OVD15	Ser Pap	I	3
OCA19	Ser Pap	II	3
OVD16	Ser Pap	II	3
OVD17	Ser Pap	II	3
OVS1	Ser Pap	II	2
OVS 33	Ser Pap	II	2
TGH 523	Ser Pap	III	3
OCA27A/B	Ser Pap	III	2
OVD1	Ser Pap	III	3
OVD2	Ser Pap	III	3
OVD4	Ser Pap	III	2
OVS3	Ser Pap	III	2
OVS8	Ser Pap	III	3
OVS 336	Ser Pap	III	3
OVS 92	Ser Pap	III	3
OCA17	Ser Pap	III	2
OVD3	Ser Pap	III	3
OVD5	Ser Pap	III	2
OVS6	Ser Pap	III	3
OVS 93	Ser Pap	III	3
OCA 46	Ser Pap	IIIB	3
OCA 33A	Ser Pap	IIIB	2
OCA23A/B	Ser Pap	IV	3
OVD7	Non-Ca	-	-
OVD8	Non-Ca	-	-
OVD9	Non-Ca	-	-
OVD10	Non-Ca	-	-
OVD11	Non-Ca	-	-
Ambion normal	Non-Ca	-	-
Stratagene normal	Non-Ca	-	-

Ser Pap, serous papillary; Non-Ca, normal (non-cancerous) ovarian tissue; Tumor stage is given according to the International Federation of Gynecology and Obstetrics criteria, tumor allocation is based on stage.

logical stage (Table I). Samples were included from both the primary 'A sample' and metastatic 'B sample' sites in two cases (OCA23A/B and OCA27A/B). Snap-frozen tissue samples from five normal ovaries from women undergoing oophorectomy for non-neoplastic conditions were also included. The clinical data on all patients are summarized in Table I. Following procedures well-established in our centre, all tissue samples were snap-frozen in liquid nitrogen after surgical resection and stored at -80˚C. RNA samples extracted from the ovaries of healthy donors were purchased from two sources (Stratagene, La Jolla, CA, USA; and Ambion, Austin, TX, USA). Finally, ovarian surface epithelial scrapings were obtained from four patients undergoing pelvic surgery for indications other than ovarian cancer. Because of the very low yield of epithelial scrapings, we were unable to obtain a sufficient quantity of normal ovarian surface epithelium to conduct microarray experiments.

RNA extraction. RNA was extracted from tissue samples as previously described (5). RNA quality and concentration were verified using an Agilent Bioanalyzer (Agilent BioTechnologies, Palo Alto, CA).

Microarray experiments. Ovarian tumor total-RNA (10 μg), normal ovary total-RNA or Human Universal Reference (HUR) RNA (prepared from a pool of 10 cell lines, Stratagene), were reverse transcribed with Superscript II reverse transcriptase (Invitrogen Canada, Burlington, ON) while incorporating Cy3- or Cy5-dCTP (NEN, Boston, MA). The fluorescently labeled cDNAs were co-hybridized overnight at 37˚C to 19K version 2 and 3 microarrays manufactured at the University Health Network Microarray Centre and spotted in duplicate with 19,200 cDNA fragments. A complete list of the cDNA collection used for these arrays (Human 19.2K) can be found at the University Health Network Microarray Centre web site (http://www.microarrays.ca). The microarrays were read by a confocal laser reader (ScanArray 4000, Packard BioScience, Meriden, CT) after stringent washings. We used the GenePix Pro 3.0 (Axon Instruments, Foster City, CA) image analysis software to quantify the scanned images. Low quality spots were filtered as identified by GenePix Pro 3.0 as well based on visual examination of the images.

Microarray data analysis

Initial data processing. Data warehousing, filtering and normalization were performed using the GeneTraffic software (version 2.6, Iobion, Stratagene). All hybridizations were annotated according to the MIAME (Minimum Information About a Microarray Experiment) guidelines [http://www.mged.org/Workgroups/MIAME/miame.html and in Brazma *et al* (6)]. Briefly, each of the 33 samples was assayed with two dye reversal microarray hybridizations (except one sample for which only one hybridization was successful) for a total of 65 hybridizations. Dye reversal replicate hybridizations were used to control for potential labeling differences. The UHN Human 19K microarray used included 38,400 spots corresponding to 18,114 unique spotted cDNA clones. The initial data set was filtered to exclude spots flagged in the quantification process, spots whose raw intensity was less than the local background in either of the two channels, spots which had a spot to background ratio of <1, and spots whose raw intensity was <100. LOWESS (Locally Weighted Scatter Plot Smoother) normalization by sub-array (http://stat-www.berkeley.edu/users/terry/zarray/Html and GeneTraffic 2.0 Manual, Iobion, Stratagene) was used for normalization of all microarrays. A Master 'spot ratio' table (MST) was exported from GeneTraffic with the mean normalized log2 ratio (sample vs. HUR) recorded for each unique clone ID (spot) and each unique hybridization (slide); in this table the flagged spots appear as missing values. A Master 'gene ratio' table (MGT) was generated by GeneTraffic from the Master Spot ratio table MST, by averaging all spot ratios corresponding to the same clone IDs on the arrays and

averaging all hybridizations corresponding to the same patient samples. The missing values (from flagged spots) were estimated by 10-nearest-neighbor imput.

'Tumor versus normal' study: unsupervised analysis. Unsupervised analysis comparing tumor samples to normal ovarian tissue was carried out on the gene ratio tables derived from MGT. The sub data set was median-centered and then analyzed by two-dimensional hierarchical clustering using the Cluster by Eisen Lab (http://rana.lbl.gov/EisenSoftware.htm) software. Average-linkage clustering using a Pearson correlation (uncentered) metric was applied to both the gene and sample dimension (7). The result of this unsupervised analysis are two dendrograms; one indicating the similarity between samples and the other indicating the similarity between genes. This hierarchical cluster was visualized in TreeView (http://rana.lbl.gov/EisenSoftware.htm) as a 2D 'heat map'. In the 2D view the genes and cell lines are ordered according to their position in the dendrogram hierarchy, while the color at each position indicates the level of gene expression of each sample relative to the median level of expression.

'Tumor versus normal' study: supervised analysis. Supervised analysis was carried out on the first sub-data set derived from the MST, i.e. at the spot ratio level, in order to increase the statistical power of the data set by treating each hybridization and each spot independently. In order to identify genes that were differentially expressed in EOC as compared to normal ovarian tissue, the samples were divided into two groups, normal ovaries (n=7) and tumor samples (n=26) and processed with a two-class statistical analysis of microarrays (SAM) analysis (4). SAM employs a modified t-test to identify differentially expressed genes and estimates the false discovery rate (FDR) by a permutation analysis. In the present analysis, 370 clones were identified (corresponding to 329 unique genes) as being differentially expressed, with a 4.2% FDR. These clones were then resolved using hierarchical clustering as described above, and visualized using a 2D heat map.

Quantitative real-time RT-PCR. DNaseI (1 μg) treated total-RNA from ovarian tumors and normal ovaries (or 0.1-1 μg from ovarian epithelial scrapings) was reverse transcribed in a 100-μl reaction mixture containing 1X room temperature buffer, 5.5 mM $MgCl_2$, 500 μM each dNTP, 2.5 μM random hexamers, 0.4 units/μl RNase inhibitor, and 3.125 units/μl MultiScribe Reverse Transcriptase (Applied Biosystems, Foster City, CA) in the following conditions: 25˚C for 10 min, 48˚C for 60 min, and 95˚C for 5 min. Real-time relative quantitative PCR was performed in triplicate using the ABI PRISM 7900HT Sequence Detection system (Applied Biosystems) according to the manufacturer's instructions, and data were averaged. Primers were specifically designed for each of four genes: WAP four-disulfide core domain protein HE4 (WAP), secreted phosphoprotein 1 (SPP1, osteopontin), decorin (DCN) and human peptidylprolyl isomerase A (cyclophilin A, PPIA); using Primer Express v1.5a (Applied Biosystems). The following primers were used: a) decorin: 5'-GCC AGA AAA AAT GCC CAA AAC-3' (forward primer), 5'-AGT AAC TTT TCG CAC TTT GGT GAT C-3' (reverse primer); b) osteopontin: 5'-CGT CTC AGG CCA GTT GCA-3' (forward primer), 5'-GTG ATG CCT AGG AGG CAA A-3' (reverse primer); c) WAP four-disulfide core domain 2: 5'-CAG GTG GAC AGC CAG TGT-3' (forward primer), 5'-GGA CCT CAG AAA TTG GGA GTG A-3' (reverse primer); d) human cyclophilin A: 5'-TGC TGG ACC CAA CAC AAA TG-3' (forward primer), 5'-TGC CAT CCA ACC ACT CAG TC-3' (reverse primer). Of each cDNA (0.4 μl) was amplified using SYBR® Green PCR Master Mix. TaqMan Assay-On-Demand (Applied Biosystems) was used for activin A receptor, type I (ACVR1); tumor necrosis factor (TNF superfamily, member 2; TNF) and human peptidylprolyl isomerase A (cyclophilin A, PPIA). Both SYBR Green PCR and TaqMan® assays were carried out with the following PCR conditions: 50˚C for 2 min, 95˚C for 10 min, followed by 40 cycles of 95˚C for 15 sec and 60˚C for 1 min. Human cyclophilin A was used as an endogenous control because it resulted in minimum variation throughout the samples and has been used by other investigators analyzing gene expression using real-time RT-PCR [e.g., by Trogan *et al* (8)]. Each threshold cycle (C_T), which indicates the cycle at which an increase in reporter fluorescence goes slightly over the optimal value line, was determined. The C_T value of human cyclophilin A was subtracted from each C_T value of tumor, normal tissue and ovarian epithelial scrapings for normalization and the ratio of sample to human universal reference (Stratagene) expression was calculated to compare with microarray data. All real-time products were run on the Agilent 2100 Bioanalyzer (Agilent Technologies) on the DNA 500 Assay lab chip to confirm a single product formation and amplicon size. The 1-tailed t-test was used to examine the significant difference of gene expression ratios between groups of ovary scrapings, normal and tumors. Statistical significance was defined as $p<0.05$.

Results

In this study, we performed microarray data analysis in two sequential phases. First, we used unsupervised analysis of the complete data set using two-dimensional hierarchical clustering to look at the overall gene expression profiles of SEOC samples with respect to normal ovary tissue, and second, we carried out a statistical analysis of the same complete data set using SAM [Tusher *et al* (4) and http://www-stat.stanford.edu/~tibs/SAM/] to identify those genes showing statistically significant differential expression between normal ovaries and SEOC.

SAM is a powerful statistical technique for finding genes significantly differentially expressed in a series of microarray experiments. The pre-processed (i.e., flagged, filtered and normalized) microarray data obtained for a set of samples are entered, and a response variable is chosen for each sample. The algorithm uses repeated permutations of the data to identify those genes whose expression is significantly related to SEOC. The significance cut-off can be chosen, through the selection of a false discovery rate (FDR).

To ascertain the validity of our microarray methodology, and to identify novel genes potentially implicated in SEOC, we first profiled the expression of the whole cohort of 33 samples, by unsupervised average-linkage two-dimensional hierarchical clustering, using a common reference (Human Universal Reference, Stratagene). In order to avoid possible

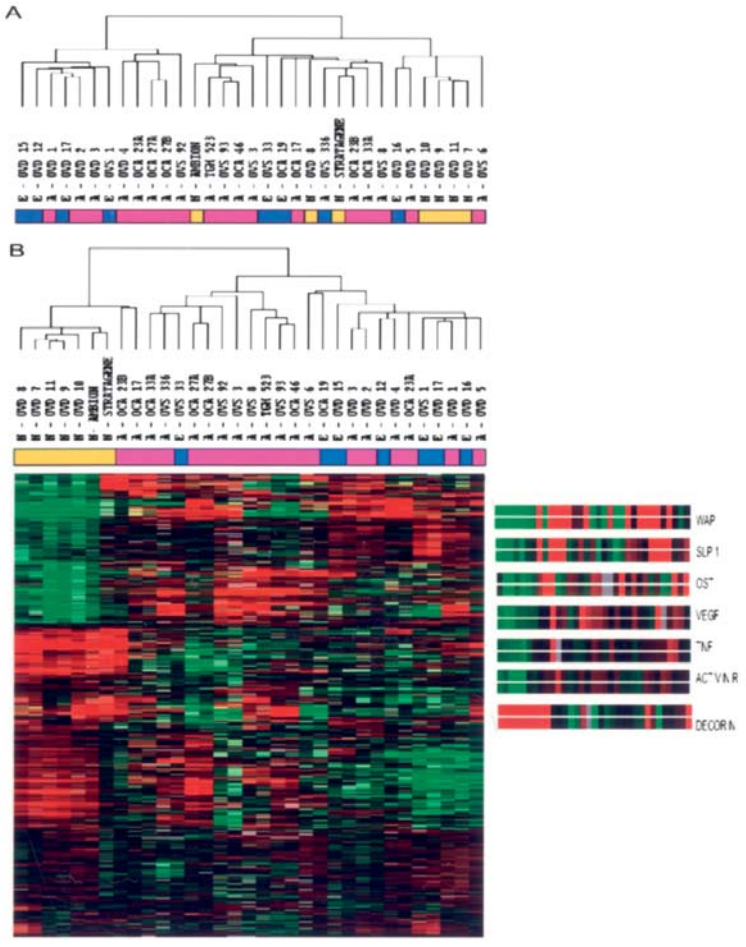

Figure 1. Differential gene expression in SEOC. (A) Unsupervised two-dimensional hierarchical clustering of all clones across all EOC samples and normal ovaries. (B) Unsupervised two-dimensional hierarchical clustering of 370 differentially expressed clones obtained by SAM analysis, for all EOC samples and normal ovaries.

data contamination or bias introduced by the use of normal ovary RNA as a reference, we used the internal reference HUR RNA described above in all microarray experiments. This approach has been commonly used in microarray studies (9-15) to allow the determination of global relative gene expression profiles in multiple samples and is an invaluable tool for tumor classification.

The dendrogram representing the overall degree of similarity in samples expression profiles is displayed in Fig. 1A and shows that the 33 ovarian samples studied here separate into 2 main groups based on their microarray expression profiles. The first cluster contains exclusively tumor samples with a larger proportion of advanced stage SEOC samples (8 advanced samples, 4 early samples). The second cluster of samples contains both tumor samples and normal ovary samples, with 4 (all tissue samples) out of 7 normal ovarian samples clustering together as a group. Universal reference samples clustered with advanced stage cancers.

To identify subsets of genes that show significant differential expression between normal and ovarian tumors, we carried out a supervised statistical analysis on the same data set separated into 2 groups, normal ovaries (group N), and SEOC samples (group T). Two-class SAM analysis identified 370 differentially expressed clones and ESTs at a false discovery rate of 4.2%. We then used hierarchical clustering to resolve the differential expression of the samples over these 370 differentially expressed clones (Fig. 1B). This clustering demonstrates, as expected, a clear separation of normal ovary samples from tumor samples. Out of these 370 clones, 234 unique genes showed a differential expression >1.3 times between normal ovaries and SEOC, with down-regulation in tumor samples observed for 144 unique genes, and up-regulation in tumor samples was observed for 90 unique genes. A partial list of these differentially expressed genes is presented in Table II. In addition to previously reported genes, such as the two WAP four-disulfide core domain proteins HE4 and SPL (16) (p=0.005; Fig. 2), osteopontin (17) (p=0.0024; Fig. 2), decorin (5) (p=0.0087; Fig. 2), and several ribosomal proteins (Table II), we also identified several genes, which to the best of our knowledge have not been previously reported in expression analysis of human ovarian cancer tissues. Of particular interest among the newly identified genes, is the up-regulation of tumor necrosis factor α (TNFα; p=0.0001; Fig. 2) concurrent with the down-regulation of decorin (p=0.0087; Fig. 2).

Five genes were selected for further validation on the basis of their gene expression differences combined with available literature of the genes in ovarian cancer (16-20). Real-time RT-PCR was used to measure and compare the expression levels of these 5 selected genes: WAP four-disulfide core

Table II. Genes statistically differentially expressed between SEOC and normal ovarian tissue.

Clone ID or accession no.	SAM score	Fold change (T/N)	Name	Symbol
		Up		
366323	3.3	5.3	WAP four-disulfide core domain 2	WFDC2
380396	3.4	2.7	CD74 antigen	CD74
200180	2.5	2.7	Secreted phosphoprotein 1 (osteopontin)	SPP1
N27733	2.5	2.4	Secretory leukocyte protease inhibitor	SLPI
380794	4.6	2.4	Oxidase (cytochrome c) assembly 1-like	OXA1L
266323	4.2	2.2	In multiple clusters Hs.194351|Hs.42502	
682763	3.2	2.2	Hypothetical protein DKFZp667E0512	DKFZp667E0512
469892	2.5	2.2	Major histocompatibility complex, class II, DQ ß 1	HLA-DQB1
357679	4.4	2.2	Cathepsin G	CTSG
338479	3.2	2.1	Cystic fibrosis transmembrane conductance regulator	CFTR
156060	3.1	2.1	In multiple clusters Hs.446471|Hs.84298	
4808400	3.1	2.0	Wilms tumor 1	WT1
5431259, 5755043	2.8	2.0	Glyceraldehyde-3-phosphate dehydrogenase	GAPD
418025	4.2	1.9	Solute carrier family 24, member 4	SLC24A4
261814	3.3	1.9	In multiple clusters Hs.42592|Hs.444649	
359898	2.8	1.9	Spermidine/spermine N1-acetyltransferase	SAT
110352, 327550	2.4	1.8	Major histocompatibility complex, class II, DP ß 1	HLA-DPB1
321230	2.2	1.8	Major histocompatibility complex, class II, DR ß 3	HLA-DRB3
470953	2.4	1.8	Major histocompatibility complex, class II, DR ß 1	HLA-DRB1
321758	2.3	1.7	Tubulin, α 1 (testis specific)	TUBA1
357442	2.6	1.7	CD47 antigen	CD47
200735	2.7	1.7	Major histocompatibility complex, class II, DP α 1	HLA-DPA1
BM683235	3.1	1.6	*Homo sapiens* cDNA FLJ37889 fis	
430083	2.5	1.6	In multiple clusters Hs.268024|Hs.37427	
125278	2.6	1.6	Claudin 4	CLDN4
488079	3.8	1.6	Voltage-gated sodium channel ß-3 subunit	HSA243396
4521436	2.3	1.6	Apoptosis inhibitor 5	API5
491066, BM982981	4.0	1.5	Activin A receptor, type I	ACVR1
5707999	2.6	1.5	Collapsin response mediator protein 1	Crmp1
505334	2.3	1.5	Regulatory factor X, 2	RFX2
23123	3.6	1.5	Hypothetical protein FLJ20047	FLJ20047
28896	2.3	1.5	Hect domain and RCC1 (CHC1)-like domain (RLD) 1	HERC1
W38749	2.5	1.5	mal,T-cell differentiation protein 2	MAL2
469329	2.4	1.5	folate receptor 2 (fetal)	FOLR2
151988	2.4	1.5	CD24 antigen	CD24
292818	3.1	1.4	Vascular endothelial growth factor	VEGF
45900	3.0	1.4	Proprotein convertase subtilisin/kexin type 7	PCSK7
342591	2.3	1.4	Kallikrein 5	KLK5
682967	2.8	1.4	Signal transducer and activator of transcription 6	STAT6
116427	2.5	1.4	LUC7-like (S. cerevisiae)	LUC7L
430216	2.8	1.4	Slingshot 1	SSH1
5221374	2.4	1.4	T cell receptor ß locus	TRB@
153328	2.4	1.4	Complement component 1, q subcomponent	C1QG
342222	2.4	1.4	*Homo sapiens*, clone IMAGE:3866125, mRNA	
341874	3.3	1.4	Major histocompatibility complex, class II, DM ß	HLA-DMB
415697	3.5	1.4	*Homo sapiens* unc93 homolog A, mRNA	
328493	2.7	1.4	Protease, serine, 1 (trypsin 1)	PRSS1
5482894	2.9	1.3	Transgelin 2	TAGLN2
682640	3.1	1.3	Tumor necrosis factor (TNF superfamily, member 2)	TNF
5732871	2.3	1.3	RAB2, member RAS oncogene family-like	RAB2L
5284947	2.4	1.3	Hypothetical protein LOC51239	LOC51239
362795	2.4	1.3	ISL1 transcription factor, LIM/homeodomain, (islet-1)	ISL1

Table II. Continued.

Clone ID or accession no.	SAM score	Fold change (T/N)	Name	Symbol
		Down		
5708746	-4.7	0.3	RIKEN cDNA 2510049I19 gene	2510049I19Rik
144221, 148425, 3108408	-4.5	0.4	Hemoglobin, ß	HBB
26052	-2.5	0.4	Microsomal glutathione S-transferase 1	MGST1
AW949701	-5.8	0.4	Neural precursor cell expressed	NEDD4
141698	-4.2	0.4	oxysterol binding protein-like 3	OSBPL3
296225, 382989, 5851583, BM720684, AL568635, BG898644	-4.7	0.5	Decorin	DCN
4730866	-3.4	0.5	Hemoglobin, γ A	HBG1
110331, 347360	-3.3	0.6	Collagen, type VI, α 1	COL6A1
258674, 308529	-3.5	0.6	Insulin-like growth factor binding protein 4	IGFBP4
418214, BQ028546	-3.8	0.6	Hypothetical protein FLJ21174	FLJ21174
324234, 485672, 502694	-3.7	0.6	Insulin-like growth factor binding protein 5	IGFBP5
5420166, 5829576	-2.8	0.6	Actin, α 2, smooth muscle, aorta	ACTA2
298556, 3546201, BM850060	-3.4	0.6	Ribosomal protein L10	RPL10
245298, 4300428	-3.8	0.6	Ribosomal protein L9	RPL9
5735844	-4.5	0.5	*Homo sapiens*, clone IMAGE:4731764, mRNA	
72664	-4.5	0.5	Ribosomal protein, large P2	RPLP2
BM314525	-3.1	0.5	Insulin-like growth factor binding protein 7	IGFBP7
5735844	-4.3	0.5	*Homo sapiens*, clone IMAGE:4731764, mRNA	
4150706	-3.6	0.5	SPARC-like 1 (mast9, hevin)	SPARCL1
5444580	-5.0	0.6	Ribosomal protein L21	RPL21
5016393	-3.7	0.6	Ribosomal protein L18	RPL18
344490	-4.1	0.6	Caldesmon 1	CALD1
256907	-4.2	0.6	Glutathione S-transferase A3	GSTA3
5635159	-3.5	0.6	Thioredoxin interacting protein	TXNIP
119255	-3.9	0.6	In multiple clusters Hs.171952|Hs.250641	
BQ128382	-3.4	0.6	Hemoglobin, α 1	HBA1
665955	-3.2	0.6	Hypothetical protein FLJ12287 similar to semaphorins	FLJ12287
590070	-4.1	0.6	Ribosomal protein L23a	RPL23A
358078	-3.1	0.6	Myristoylated alanine-rich protein kinase C substrate	MARCKS
5756295	-3.6	0.6	Ribosomal protein L13a	RPL13A
302482	-2.3	0.6	In multiple clusters Hs.119206|Hs.296014	
5923507	-3.5	0.6	Ribosomal protein L23	RPL23
AL037460	-2.5	0.6	*Homo sapiens* tissue-type brain unknown mRNA	
505564	-4.0	0.6	Platelet-derived growth factor receptor	PDGFRA
238591	-3.4	0.6	Serine (or cysteine) proteinase inhibitor	SERPINF1
358484	-2.7	0.6	Actin, α, cardiac muscle	ACTC
5827652	-3.7	0.7	Tumor protein, translationally-controlled 1	TPT1
4862761	-3.6	0.7	Dipeptidylpeptidase 7	DPP7
5798718	-3.6	0.7	Finkel-Biskis-Reilly murine sarcoma virus	FAU
153244	-3.4	0.7	In multiple clusters (only IMAGE ID)	
489907	-3.6	0.7	KIAA0084 protein	KIAA0084
270735	-3.5	0.7	Eukaryotic translation elongation factor 2	EEF2
27803	-2.4	0.7	In multiple clusters Hs.397075|Hs.44289	
502358	-3.2	0.7	Hypothetical protein DKFZp547O146	DKFZp547O146
485076	-3.0	0.7	Ribosomal protein L31	RPL31
4762963	-3.3	0.7	Ribosomal protein S13	RPS13
263884	-3.5	0.7	v-ski sarcoma viral oncogene homolog (avian)	SKI
5422775	-3.1	0.7	Ribosomal protein S5	RPS5
126403	-3.3	0.7	Glutathione S-transferase A2	GSTA2
5491196	-3.3	0.7	Ribosomal protein S23	RPS23

Table II. Continued.

Clone ID or accession no.	SAM score	Fold change (T/N)	Name	Symbol
298532	-2.8	0.7	Spectrin, ß, non-erythrocytic 1	SPTBN1
156136	-3.0	0.7	Ribosomal protein S15	RPS15
207619	-2.5	0.7	Chromosome 14 open reading frame 68	C14orf68
5588658	-2.7	0.7	Ribosomal protein L31	RPL31
5927927	-2.9	0.7	Ribosomal protein L10a	RPL10A
321259	-3.1	0.7	In multiple clusters Hs.164170|Hs.184014	
171965	-4.0	0.7	Hypothetical protein FLJ12242	FLJ12242
AL038078	-2.5	0.7	*Homo sapiens* cDNA FLJ38828 fis	
163464	-2.5	0.7	AD-015 protein	LOC55829
166910	-3.3	0.7	In multiple clusters Hs.268832|Hs.279837	
166606	-2.9	0.7	In multiple clusters Hs.279837|Hs.301961	
509663	-2.5	0.7	Zinc ribbon domain containing, 1	ZNRD1
5534440	-2.9	0.7	BET1 homolog (S. cerevisiae)	BET1
4710621	-2.6	0.7	Ribosomal protein S3	RPS3
489739	-2.9	0.7	Ribosomal protein L36a-like	RPL36AL
190641	-2.2	0.7	Biphenyl hydrolase-like	BPHL
155790	-2.2	0.7	Polymerase (DNA directed), α	POLA
469876	-3.2	0.7	Transforming growth factor, ß receptor II	TGFBR2
491155	-2.9	0.7	Ribosomal protein L27	RPL27
118081	-3.0	0.7	Hypothetical protein FLJ20265	FLJ20265
AL564444	-2.4	0.7	*Homo sapiens* PRO2743 mRNA, complete cds	
BQ028584	-2.7	0.7	Ribosomal protein S16	RPS16
279619	-3.0	0.7	In multiple clusters Hs.355554|Hs.433387	
4862574	-2.7	0.7	IMP (inosine monophosphate) dehydrogenase 2	IMPDH2
5823226	-2.3	0.7	Human HepG2 3' region cDNA, clone hmd1f06	
4862951	-2.7	0.7	Nuclear receptor subfamily 2, group F, member 2	NR2F2
470889	-2.7	0.7	Membrane protein, palmitoylated 2	MPP2
86081	-2.4	0.7	Ribosomal protein S29	RPS29
327279	-2.6	0.7	Hypothetical protein PRO1843	PRO1843
186676	-2.7	0.7	Inhibitor of DNA binding 2	ID2
270104	-2.8	0.7	pp21 homolog	LOC51186
191907,446628, 5764801	-3.4	0.7	Basic transcription factor 3	BTF3
203692, 243600	-3.4	0.7	Glutathione S-transferase A1	GSTA1
176889, 5558091	-3.0	0.7	Ribosomal protein L41	RPL41
125802, 200190, 3909011	-2.9	0.7	Nascent-polypeptide-associated complex α polypeptide	NACA
347367, 5770317	-2.7	0.7	Nerve growth factor receptor (TNFRSF16)	NGFRAP1
124692, 5208247	-2.5	0.8	Ribosomal protein L32	RPL32
241640	-2.7	0.8	Ribosomal protein L38	RPL38
5540111	-2.1	0.8	Claudin 12	CLDN12

Ratio (T/N), relative average expression ratio of SEOC versus normal ovaries. The SAM score is the T-statistic value and indicates significance.

domain protein HE4 (WAP, up-regulated), secreted phosphoprotein 1 (SPP1, osteopontin; up-regulated), activin A receptor, type I (ACVR1; up-regulated), tumor necrosis factor (TNF superfamily, member 2; TNF, up-regulated) and decorin (DCN, down-regulated) in 4 epithelial scrapings and in 6 bulk-extracted normal ovaries. The results obtained (Fig. 2) show that for the up-regulated genes the expression levels measured by real-time RT-PCR did not differ significantly between normal ovarian epithelial scrapings and bulk-extracted ovaries. Decorin however, showed a statistically significant difference ($p=0.0073$) in expression between epithelial scrapings and bulk-extracted ovaries. Decorin is a member of the small leucine-rich proteoglycan superfamily and is involved in the regulation of several important cellular functions, such as matrix assembly and cell proliferation and migration.

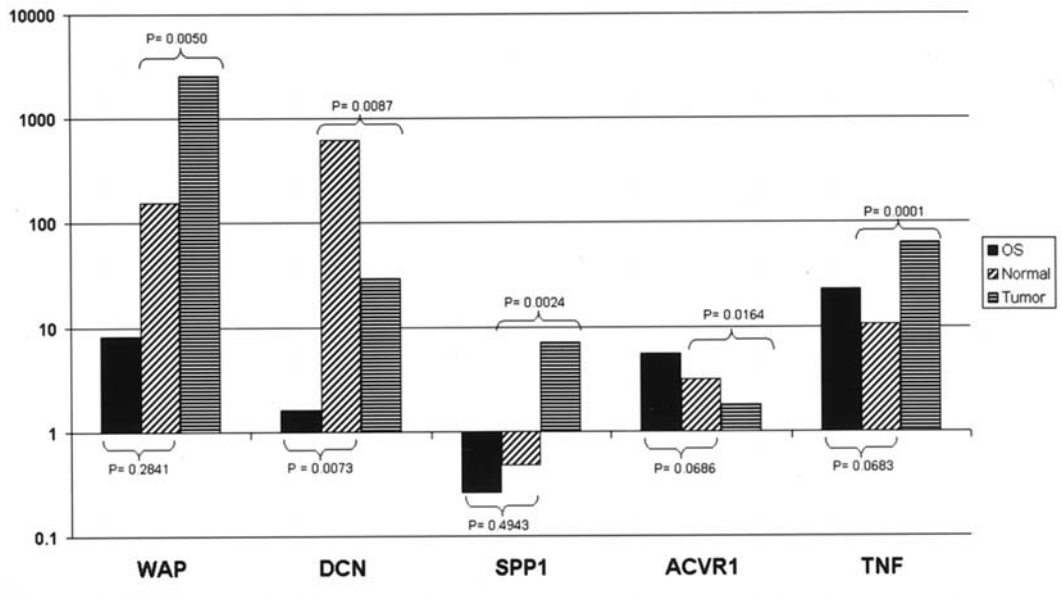

Figure 2. Real-time RT-PCR analysis. Comparison of the expression levels of the genes encoding WAP four-disulfide core domain protein HE4 (WAP), secreted phosphoprotein 1 (SPP1, osteopontin), activin A receptor, type I (ACVR1), tumor necrosis factor (TNF) and decorin (DCN) obtained by real-time RT-PCR. OS, normal ovarian surface epithelial scrapings. Normal and tumor, bulk-extracted normal ovaries. Note that only DCN showed a significant differential expression between OS and bulk-extracted normal ovaries.

Discussion

Little is known of the molecular changes that are associated with the development from normal ovarian tissue to invasive ovarian cancer. No lesion is clinically recognized that immediately precedes ovarian cancer. The late onset of symptoms (in patients with invasive cancers) makes early detection difficult. Surface epithelial tumors account for 50-55% of all ovarian tumors, and their malignant forms account for approximately 90% of all ovarian cancers in the western world. A number of well-characterized genes, differentially expressed between normal ovaries and SEOC, have been reported in the literature (5,20-30). In the present study we used gene expression profiling to reveal differential gene expression between SEOC and normal ovarian tissue (bulk and surface epithelial scrapings).

Several of the genes identified here have been previously reported in the literature by us (5) and others (17,20,26-32) and these findings provide strong experimental support for this microarray design and analytical strategy. Since we used bulk-extracted normal ovary tissue for comparison to the ovarian cancer samples, these tissues contributed considerably and possibly confounding cellular heterogeneity to the analysis. Therefore while the genes identified in the present study are relatively up- or down-regulated compared to normal ovarian tissue, their differential expression might partially reflect the differences in cell populations. Ovarian epithelial cell scrapings have been used by some investigators to generate more homologous reference RNA (22,28,33), but that approach requires access to a large number of normal ovaries. The yield of ovarian surface epithelial cells is relatively low and very often unsatisfactory for microarray use. Linear PCR techniques are being used to address this issue but they appear to be challenging and have variable reliability (34,35). Although a large part of the normal ovary consists of a stromal component, this stroma is very scant with cells. Hence most of the normal bulk-extracted ovary will resemble the expression profile of epithelial scrapings. We have chosen to perform our analysis on bulk-extracted ovaries and surface epithelial cells rather than cultured epithelial cells or immortalized cells as previous studies have shown that culturing and immortalizing ovarian epithelial cells can influence their gene expression (33). Bulk-extracted ovarian tissue contains a mixture of epithelial and stromal cells whereas the majority of cells in the brushing samples are of epithelial origin only. We argue that the differences seen in our experiment between normal ovarian tissue (regardless whether its origin is bulk-extracted or epithelial scrapings) and ovarian cancer tissue are a result of malignant features of the cells.

Of the five genes that were validated by RT-PCR only decorin showed a statistically significant difference between the epithelial scrapings and bulk-extracted ovaries. Decorin has been identified as down-regulated in ovarian cancer both in microarray studies using bulk-extracted normal ovaries (5,26) and by those using epithelial scrapings as a reference (27,28). Moreover, Nash *et al* (36) reported that decorin of ovarian tissue is made by myofibroblasts rather than by the epithelial tumor cells. The presence of stroma cell RNA in the bulk-extracted ovaries, explains our results, with expression of decorin in bulk-extracted ovaries opposed to a very low expression of decorin in epithelial scrapings. Recent studies, including by our group, have shown that decorin is down-regulated in ovarian cancer (5,26,28). Since TNFα has been previously shown *in vitro* to down-regulate decorin gene expression (37), the present study suggests that down-regulation of decorin in SEOC might be a consequence of the up-regulation of TNFα. In addition, decorin plays a role in angiogenesis (38) and it was recently shown that decorin

down-regulates the expression of VEGF in various tumor cell types (38). Our present data demonstrating up-regulation of VEGF concurrent with a down-regulation of decorin are consistent with that observation, and suggest an important role for decorin in SEOC angiogenesis and tumor progression.

An interesting finding was disclosed when we validated the microarray results on activin A receptor type I (ACVR1). While it was up-regulated in SEOC in the microarray studies, it showed down-regulation in the validation by the RT-PCR ($p=0.0164$; Fig. 2). It is noteworthy that the activin type IA receptor expression was shown to have a high variability in serous ovarian tumors (19). On the average, SEOC tend to have higher levels while normal ovaries tend to express lower levels, in a pattern similar to the microarray experiments. ACVR1 interacts with p120 of the inhibin receptor, that allows inhibin to bind with high affinity to the complex, which consequently destabilizes the complex and abrogates activin signal transduction (39). Activin, along with inhibin, is a member of the TGF-ß superfamily (40) and is produced locally in the ovary (41). It has been suggested that activin is involved in regulating cell proliferation and ovarian tumor development (18,42). Recently, Choi *et al* (43) have also shown that activin induces cell proliferation in OVCAR-3 cells but not in normal OSE cells, and these authors hypothesized that activin may be an autocrine regulator of neoplastic OSE progression.

In conclusion, our results indicate that previously uncharacterized genes are associated with the malignant phenotype of SEOC. Bulk normal ovarian tissue may serve as control for SEOC tissue in gene expression profiling. Gene expression profiling and sequential statistical analyses of gene subsets can identify new genes and molecular pathways affecting development of SEOC. In addition this approach provides additional support for emerging biomarkers with clinical applications in the future management of serous EOC.

Acknowledgements

This study was supported by the Princess Margaret Hospital Foundation and the Ontario Research and Development Challenge Fund (to M.A. and P.F.M.). We also express our sincere gratitude to the Ovarian Fashion Show Committee for their financial support (to D.G. and M.P.). The University of Toronto Ovarian Tissue Bank and Database was supported in part by the National Ovarian Cancer Association and by the St. George's Society. We would also like to thank Dr J.A. Squire (Division of Cellular and Molecular Biology, Ontario Cancer Institute, University Health Network, Department of Medical Biophysics, Department of Laboratory Medicine and Pathobiology, University of Toronto) for his help and guidance during this project.

References

1. Jemal A, Murray T, Ward E, *et al*: Cancer statistics, 2005. CA Cancer J Clin 55: 10-30, 2005.
2. Kosary CL: FIGO stage, histology, histologic grade, age and race as prognostic factors in determining survival for cancers of the female gynecological system: an analysis of 1973-87 SEER cases of cancers of the endometrium, cervix, ovary, vulva, and vagina. Semin Surg Oncol 10: 31-46, 1994.
3. Engel J, Eckel R, Schubert-Fritschle G, *et al*: Moderate progress for ovarian cancer in the last 20 years: prolongation of survival, but no improvement in the cure rate. Eur J Cancer 38: 2435-2445, 2002.
4. Tusher VG, Tibshirani R and Chu G: Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 98: 5116-5121, 2001.
5. Bayani J, Brenton JD, Macgregor PF, *et al*: Parallel analysis of sporadic primary ovarian carcinomas by spectral karyotyping, comparative genomic hybridization, and expression microarrays. Cancer Res 62: 3466-3476, 2002.
6. Brazma A, Hingamp P, Quackenbush J, *et al*: Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet 29: 365-371, 2001.
7. Eisen MB, Spellman PT, Brown PO and Botstein D: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95: 14863-14868, 1998.
8. Trogan E, Choudhury RP, Dansky HM, Rong JX, Breslow JL and Fisher EA: Laser capture microdissection analysis of gene expression in macrophages from atherosclerotic lesions of apolipoprotein E-deficient mice. Proc Natl Acad Sci USA 99: 2234-2239, 2002.
9. Garber ME, Troyanskaya OG, Schluens K, *et al*: Diversity of gene expression in adenocarcinoma of the lung. Proc Natl Acad Sci USA 98: 13784-13789, 2001.
10. Bittner M, Meltzer P, Chen Y, *et al*: Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature 406: 536-540, 2000.
11. Perou CM, Jeffrey SS, van de Rijn M, *et al*: Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc Natl Acad Sci USA 96: 9212-9217, 1999.
12. Alizadeh AA, Eisen MB, Davis RE, *et al*: Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403: 503-511, 2000.
13. Alon U, Barkai N, Notterman DA, *et al*: Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc Natl Acad Sci USA 96: 6745-6750, 1999.
14. Luo J, Duggan DJ, Chen Y, *et al*: Human prostate cancer and benign prostatic hyperplasia: molecular dissection by gene expression profiling. Cancer Res 61: 4683-4688, 2001.
15. Selaru FM, Zou T, Xu Y, *et al*: Global gene expression profiling in Barrett's esophagus and esophageal cancer: a comparative analysis using cDNA microarrays. Oncogene 21: 475-478, 2002.
16. Bingle L, Singleton V and Bingle CD: The putative ovarian tumour marker gene HE4 (WFDC2), is expressed in normal tissues and undergoes complex alternative splicing to yield multiple protein isoforms. Oncogene 21: 2768-2773, 2002.
17. Kim JH, Skates SJ, Uede T, *et al*: Osteopontin as a potential diagnostic biomarker for ovarian cancer. JAMA 287: 1671-1679, 2002.
18. Di Simone N, Crowley WF Jr, Wang QF, Sluss PM and Schneyer AL: Characterization of inhibin/activin subunit, follistatin, and activin type II receptors in human ovarian cancer cell lines: a potential role in autocrine growth regulation. Endocrinology 137: 486-494, 1996.
19. Fuller PJ, Zumpe ET, Chu S, Mamers P and Burger HG: Inhibin-activin receptor subunit gene expression in ovarian tumors. J Clin Endocrinol Metab 87: 1395-1401, 2002.
20. Hough CD, Cho KR, Zonderman AB, Schwartz DR and Morin PJ: Coordinately up-regulated genes in ovarian cancer. Cancer Res 61: 3869-3876, 2001.
21. De Cecco L, Marchionni L, Gariboldi M, *et al*: Gene expression profiling of advanced ovarian cancer: characterization of a molecular signature involving fibroblast growth factor 2. Oncogene 23: 8171-8183, 2004.
22. Huddleston HG, Wong KK, Welch WR, Berkowitz RS and Mok SC: Clinical applications of microarray technology: creatine kinase B is an up-regulated gene in epithelial ovarian cancer and shows promise as a serum marker. Gynecol Oncol 96: 77-83, 2005.
23. Lee BC, Cha K, Avraham S and Avraham HK: Microarray analysis of differentially expressed genes associated with human ovarian cancer. Int J Oncol 24: 847-851, 2004.
24. Zhang X, Feng J, Cheng Y, *et al*: Characterization of differentially expressed genes in ovarian cancer by cDNA microarrays. Int J Gynecol Cancer 15: 50-57, 2005.
25. Hough CD, Sherman-Baust CA, Pizer ES, *et al*: Large-scale serial analysis of gene expression reveals genes differentially expressed in ovarian cancer. Cancer Res 60: 6281-6287, 2000.

26. Ono K, Tanaka T, Tsunoda T, *et al*: Identification by cDNA microarray of genes involved in ovarian carcinogenesis. Cancer Res 60: 5007-5011, 2000.
27. Shridhar V, Sen A, Chien J, *et al*: Identification of under-expressed genes in early- and late-stage primary ovarian tumors by suppression subtraction hybridization. Cancer Res 62: 262-270, 2002.
28. Shridhar V, Lee J, Pandita A, *et al*: Genetic analysis of early- versus late-stage ovarian tumors. Cancer Res 61: 5895-5904, 2001.
29. Welsh JB, Zarrinkar PP, Sapinoso LM, *et al*: Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. Proc Natl Acad Sci USA 98: 1176-1181, 2001.
30. Tonin PN, Hudson TJ, Rodier F, *et al*: Microarray analysis of gene expression mirrors the biology of an ovarian cancer model. Oncogene 20: 6617-6626, 2001.
31. Jazaeri AA, Yee CJ, Sotiriou C, Brantley KR, Boyd J and Liu ET: Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers. J Natl Cancer Inst 94: 990-1000, 2002.
32. Wang K, Gan L, Jeffery E, *et al*: Monitoring gene expression profile changes in ovarian carcinomas using cDNA microarray. Gene 229: 101-108, 1999.
33. Zorn KK, Jazaeri AA, Awtrey CS, *et al*: Choice of normal ovarian control influences determination of differentially expressed genes in ovarian cancer expression profiling studies. Clin Cancer Res 9: 4811-4818, 2003.
34. Saghizadeh M, Brown DJ, Tajbakhsh J, *et al*: Evaluation of techniques using amplified nucleic acid probes for gene expression profiling. Biomol Eng 20: 97-106, 2003.
35. Rudnicki M, Eder S, Schratzberger G, *et al*: Reliability of t7-based mRNA linear amplification validated by gene expression analysis of human kidney cells using cDNA microarrays. Nephron Exp Nephrol 97: E86-E95, 2004.
36. Nash MA, Deavers MT and Freedman RS: The expression of decorin in human ovarian tumors. Clin Cancer Res 8: 1754-1760, 2002.
37. Mauviel A, Santra M, Chen YQ, Uitto J and Iozzo RV: Transcriptional regulation of decorin gene expression. Induction by quiescence and repression by tumor necrosis factor-alpha. J Biol Chem 270: 11692-11700, 1995.
38. Grant DS, Yenisey C, Rose RW, Tootell M, Santra M and Iozzo RV: Decorin suppresses tumor cell-mediated angiogenesis. Oncogene 21: 4765-4777, 2002.
39. Chapman SC and Woodruff TK: Modulation of activin signal transduction by inhibin B and inhibin-binding protein (INhBP). Mol Endocrinol 15: 668-679, 2001.
40. Woodruff TK: Regulation of cellular and system function by activin. Biochem Pharmacol 55: 953-963, 1998.
41. Eramaa M, Heikinheimo K, Tuuri T, Hilden K and Ritvos O: Inhibin/activin subunit mRNA expression in human granulosa-luteal cells. Mol Cell Endocrinol 92: R15-R20, 1993.
42. Welt CK, Lambert-Messerlian G, Zheng W, Crowley WF, Jr. and Schneyer AL: Presence of activin, inhibin, and follistatin in epithelial ovarian carcinoma. J Clin Endocrinol Metab 82: 3720-3727, 1997.
43. Choi KC, Kang SK, Tai CJ, Auersperg N and Leung PC: The regulation of apoptosis by activin and transforming growth factor-beta in early neoplastic and tumorigenic ovarian surface epithelium. J Clin Endocrinol Metab 86: 2125-2135, 2001.