

AD \_\_\_\_\_

Award Number: DAMD17-02-1-0183

TITLE: Molecular Biology and Prevention of Endometrial Cancer

PRINCIPAL INVESTIGATOR: George L. Maxwell, M.D.

CONTRACTING ORGANIZATION: Henry M. Jackson Foundation for the  
Advancement of Military Medicine  
Rockville, Maryland 20852-1428

REPORT DATE: July 2003

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20040618 183

**REPORT DOCUMENTATION PAGE**Form Approved  
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

<b>1. AGENCY USE ONLY</b> (Leave blank)		<b>2. REPORT DATE</b> July 2003	<b>3. REPORT TYPE AND DATES COVERED</b> Annual (17 Jun 2002 - 16 Jun 2003)	
<b>4. TITLE AND SUBTITLE</b> Molecular Biology and Prevention of Endometrial Cancer			<b>5. FUNDING NUMBERS</b> DAMD17-02-1-0183	
<b>6. AUTHOR(S)</b> George L. Maxwell, M.D.				
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Henry M. Jackson Foundation for the Advancement of Military Medicine Rockville, Maryland 20852-1428  <i>E-Mail:</i> George.Maxwell@na.amedd.army.mil			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012			<b>10. SPONSORING / MONITORING AGENCY REPORT NUMBER</b>	
<b>11. SUPPLEMENTARY NOTES</b> Original contains color plates: All DTIC reproductions will be in black and white.				
<b>12a. DISTRIBUTION / AVAILABILITY STATEMENT</b> Approved for Public Release; Distribution Unlimited				<b>12b. DISTRIBUTION CODE</b>
<b>13. ABSTRACT (Maximum 200 Words)</b> Objective: To increase our understanding of the molecular aberrations associated with endometrial carcinogenesis and the biologic mechanisms underlying the protective effect of oral contraceptive therapy. Methods: 1) Both cDNA and oligonucleotide microarray analysis was performed on a panel of endometrial cancers. 2) A subset of adenocarcinoma cases from the International DES Registry was analyzed for MSI and the poor quality of the DNA from many of these cases prohibited an adequate analysis of the entire set. The cases are currently being evaluated for PTEN expression. 3) We analyzed the Cancer and Steroid Hormone (CASH) database to determine if higher progestin potency confers a more protective effect against endometrial cancer. 4) Finally, we are in the process of analyzing the endometrium from cymologous macaques that were exposed to long term progestins as part of the three year randomized controlled trial. Results: 1) Different histological types of endometrial cancer have unique genomic expression patterns. 2) The rate of MSI in cancers that develop after a perinatal exposure to DES is lower than previously reported. 3) The potency of the progestin in most oral contraceptives appears adequate to provide a protective effect against endometrial cancer.				
<b>14. SUBJECT TERMS</b> Endometrial cancer, hormone, prevention			<b>15. NUMBER OF PAGES</b> 21	
			<b>16. PRICE CODE</b>	
<b>17. SECURITY CLASSIFICATION OF REPORT</b> Unclassified	<b>18. SECURITY CLASSIFICATION OF THIS PAGE</b> Unclassified	<b>19. SECURITY CLASSIFICATION OF ABSTRACT</b> Unclassified	<b>20. LIMITATION OF ABSTRACT</b> Unlimited	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)  
Prescribed by ANSI Std. Z39-18  
298-102

**TABLE OF CONTENTS**

<b>Standard Form 298</b>	<b>2</b>
<b>Table of contents</b>	<b>3</b>
<b>Introduction</b>	<b>4</b>
<b>Body</b>	<b>4-6</b>
<b>Key Research Accomplishments</b>	<b>7</b>
<b>Reportable Outcomes</b>	<b>7</b>
<b>Conclusions</b>	<b>8</b>
<b>References</b>	<b>8</b>
<b>Appendices</b>	<b>9-16</b>

## INTRODUCTION

Endometrial cancer is the most common type of gynecologic cancer in the United States and was estimated by the American Cancer Society to have been newly diagnosed among 36,000 American women in the year 2000 and lead to approximately 6500 cancer related deaths (1). Approximately 25% of all endometrial cancers occur in premenopausal women (2). Major advances in our understanding and treatment of endometrial cancer have occurred over the past decade, yet the frequency of this cancer in the general population has not been altered appreciably. Despite the known protective effect of oral contraceptives, little has been learned regarding the underlying mechanism. We believe that an understanding of the molecular profiles of endometrial cancers and the molecular events underlying the protective effect of oral contraceptives against endometrial cancer could facilitate the development of effective chemopreventives and significantly decrease the incidence of endometrial cancer in women.

## BODY

**Aim 1: To characterize and compare the molecular profiles of Type I endometrioid endometrial cancers, which often develop in an estrogen milieu, to that of Type II endometrial cancers. In addition, we will use microarray to examine the molecular changes in the endometrium associated with progestin exposure in order to gain insight into the biologic mechanism underlying the chemopreventive effect of the oral contraceptive pill (OCP).**

- A pilot study using a cDNA microarray platform was performed on 35 cases of endometrial cancer and 7 normal endometrium samples (appendix 1). Following the notification of this award an amended request for additional funds was submitted in an effort to obtain funds to facilitate a more global genomic expression screen using the oligonucleotide 2-chip Affymetrix system. Excess CDRMP funds were not identified to cover this additional expense. However, NCI intramural funding was made available to supplement our microarray project and the set of 80 cancers (including the 35 cases in the pilot study) was analyzed using the Affymetrix chips. In addition, 40 extra endometrial tumors were added to the analysis using available supplementary funds for a total analysis of 120 tumors. The data using the Affymetrix oligonucleotide methodology has confirmed our pilot study results: different histologic types of endometrial cancer have distinct molecular expression patterns. Differential gene expression was obtained from over 18,000 genes that were detected in at least 50% of the cases. Multi-dimensional scaling revealed grossly separated clusters for each of the histologic groups (Appendix 2). The number of genes that were differentially expressed at least 3-fold from normal epithelium ( $p < 0.001$ ) was 716 in endometrioid, 562 in papillary serous, and 697 in MMT. In a comparison of endometrioid and MMT cases, 63 genes were differentially expressed by at least 5-fold at  $p < 0.001$ . Genes previously noted to be elevated in endometrioid cancers were differentially expressed in this comparison (TFF2 and MUC1). In a comparison of 12 pairs of endometrioid and papillary serous cases matched for stage and grade, 65 genes were differentially expressed by at least 2-fold at  $p < 0.005$  (Appendix 3). Several genes encoding IGF binding proteins were down regulated in endometrioid cancers. Class prediction for both MMT versus endometrioid and papillary serous versus endometrioid comparisons was  $>90\%$ . Following a cross-reference of the two supervised comparisons, 14 genes (several known genes and some with unknown function) were found to be associated with an endometrioid expression pattern and present on both gene lists. This data has been accepted for oral plenary presentation at the annual clinical meeting for the Society of Gynecologic Oncologists.

- A subset of the endometrioid endometrial cancer cases was analyzed for microsatellite instability (MSI). The other types (papillary serous, clear cell, etc) were not analyzed since MSI is uncommon in these histologic types. Both unsupervised (appendix 4) and supervised analysis (appendix 5) has been completed on this microarray data and the manuscript is currently being completed. The abstract has been accepted for poster presentation at the annual clinical meeting for the Society of Gynecologic Oncologists.
- Since the rate of PTEN alteration was found to be approximately 95%, an analysis of the array data based on PTEN status was not felt to have an adequate number cases with wild type PTEN to enable an accurate analysis. Perhaps a future microarray study aimed at an analysis of PTEN in very early stage endometrial cancers and precancerous endometrial conditions would facilitate an expression profile associated with alterations in PTEN.

**Aim 2: To analyze vaginal and cervical adenocarcinomas, that have arisen in women exposed to DES in-utero for methylation and mutation of PTEN and MLH1 in order to determine if estrogen induces genetic alterations in these tumors characteristic of Type I endometrioid carcinomas.**

Although a pilot study aimed at an analysis of MSI in 7 cases from the International DES Registry was successful with repetitive attempts at DNA amplification (Appendix 6), the analysis of the entire set was not successful presumably secondary to the quality of the DNA which reflects the old age of the specimens and the various methods that were used in their preservation. Less than 50% of the samples amplified at any one of the markers making the data inadequate for designation of MSI status. The experience of our lab and others in the use of MLH1 immunostaining indicates that this approach is an option but not an optimal one given the difficulties that we have historically had with the antibodies available for MLH1 immunostaining.

Our current intent is to evaluate the cases for PTEN immunostaining with the intent to evaluate for methylation of the PTEN promoter in cases with altered PTEN expression. The methodology for PTEN staining using human cancer has recently been optimized in preparation for analysis of the samples that are being evaluated as part of Aim 4.

**Aim 3: Using data from the Centers for Disease Control Cancer and Steroid Hormone Study, we will determine if the protective effect of OCP's against endometrial cancer are impacted by the progestin or estrogen potency of OCP formulations.**

Subjects included 417 endometrial cancer cases and 2,452 controls, between 20 and 54 years of age, identified from the Centers for Disease Control Cancer and Steroid Hormone (CASH) study. OCP's were classified into four categories according to the individual potencies of each hormonal constituent (high or low progestin or estrogen potency). Logistic regression was used to calculate the odds ratio (OR) and 95% confidence interval (CI) for associations between endometrial cancer risk and combination OCP formulations with high versus low estrogen and progestin potencies. With non-users as the referent group, all combinations of OCPs were significantly protective against endometrial cancer. Overall, OCP's containing high potency progestin were not significantly more protective against endometrial cancer as compared to OCP's with low progestin potency (OR 0.3, 95% CI 0.2-0.6; and OR 0.5, 95% CI 0.3-0.6 respectively) (Appendix 7). In patients with a body mass index (BMI) at or above the median, there was a trend toward greater protection afforded by use of high potency progestin OCP's (BMI at the median or above: OR=0.6; 95% CI=0.3-1.2; BMI below the median: OR=1.3; 95% CI=0.5-3.3). However the trend was not statistically significant and the study was not powered to demonstrate this association relative to BMI. Patients taking OCP's for a duration >18 months

had a greater protective effect than those patients who used OCP formulations for a shorter duration (Appendix 8). The protective effect persisted for at least 16 years after cessation of OCP use. Both the duration of OCP use and the interval of protection were not influenced significantly by the potency of the progestin component. This manuscript was accepted for oral plenary presentation at the annual clinical meeting for the Society of Gynecologic Oncologists.

**Aim 4: To test the hypothesis that the oral contraceptives and hormone replacement therapy progestins provide a chemoprotective effect against endometrial cancer through induction of apoptosis, PTEN, and TGF-beta in the endometrium.**

Epidemiological studies have demonstrated that OCP use lowers the risk of subsequent endometrial and ovarian cancer. Although the biologic mechanism(s) underlying the protective effect of OCP's on the risk of both of these cancers have not been well defined, there is evidence to suggest that biologic effects related to the progestin component may underlie the cancer preventive effects of the OCP. Recent studies have reported the progestin-mediated activation of apoptosis in endometrial cancer cell lines and endometrial hyperplasias. The finding that progestin activates the apoptosis pathway in endometrial cells raises the possibility that this may be a major mechanism underlying the therapeutic effect of progestins against endometrial hyperplasia. Similarly, our group has found that progestins markedly activate both apoptosis and TGF-beta expression in the ovarian epithelium leading to the hypothesis that progestins may act as chemopreventives for ovarian cancer (see preliminary data below). It is interesting that tumors arising from the ovary and endometrium share common epidemiological risk factors, and that both the endometrium and ovarian surface epithelium share a common embryological precursor. It is thus plausible that progestins activate similar molecular pathways relevant to cancer prevention in both of these organ sites. Recent evidence suggests that expression *PTEN* appears to be upregulated in the secretory phase of the menstrual cycle. It is plausible that the chemopreventive effects of OCP's are mediated through overexpressed *PTEN* with resultant suppression of cell cycle progression and activation of apoptosis in endometrial cells.

- The long term effects of progestins on apoptosis as well as the expression of *PTEN* and TGF-b in the endometrium were to be evaluated using uterine specimens from 210 cynomolgus macaques (80 premenopausal and 130 postmenopausal) previously part of a three-year randomized trial designed to evaluate the effects of the combination oral contraceptive pill and hormone replacement therapy on reproductive organs. The assay conditions for the analysis of these samples has been optimized and the macaque slides are currently being interpreted.
- The short-term effects of progestins on apoptosis as well as the expression of *PTEN* and TGF- $\beta$  in the endometrium will be evaluated using uterine specimens collected from patients enrolled in a double-blinded prospective randomized trial. The initiation of this trial has been delayed secondary to multiple revisions requested by multiple IRB committees. However, the protocol recently has been completed and the study will soon be underway at both Walter Reed and Evanston Northwestern. The tissue collected as part of this trial will also be used in part 2 of Aim 1, which will involve a microarray analysis of microdissected endometrium samples.

## KEY RESEARCH ACCOMPLISHMENTS

- Identified 191 genes that are differentially expressed in endometrial cancer vs. normal endometria.
- Determined that expression differences in only 24 transcripts could distinguish between the two most common types of endometrial cancers.
- Demonstrated that the use of oral contraceptives (OC) reduces the risk of endometrial cancer regardless of progestin potency.
- Demonstrated that the chemopreventive effects of OCPs persist for at least 16 years after cessation of therapy.

## REPORTABLE OUTCOMES

### Research

#### *Manuscript*

- Risinger JI, Maxwell GL, Chandramouli GVR, et al. 2003. Microarray analysis reveals distinct gene expression profiles among different histologic types of endometrial cancer. *Cancer Research* 63:6-11.

#### *Abstracts*

- Risinger JI. 2004. Gene expression profiles detect separate modulators of carcinogenesis in microsatellite unstable versus microsatellite stable endometrial cancers. Society for Gynecologic Oncologists. Abstract "In press". *Gynecologic Oncology*
- Maxwell GL. 2004. Progestin and estrogen potency of combination oral contraceptives and endometrial cancer risk. Society of Gynecologic Oncology. Abstract "In press". *Gynecologic Oncology*
- Maxwell GL. 2004. Microarray analysis of different histologic types of uterine cancer reveals unique expression profiles. Abstract "In press". *Gynecologic Oncology*

#### *Presentations*

- Maxwell GL. Microarray analysis reveals distinct gene expression profiles among different histologic types of endometrial cancer. Society of Gynecologic Oncologists Annual Meeting. New Orleans, Louisiana.
- Maxwell GL. Microarray analysis reveals distinct gene expression profiles among different histologic types of endometrial cancer. American College of Obstetricians and Gynecologists, Armed Forces District Annual Meeting.

## CONCLUSIONS

Different histological types of cancer have genomic expression patterns that reflect unique pathways of carcinogenesis. Likewise, cancers characterized by microsatellite instability result in the expression of genes most likely to be affected by alterations in mismatch repair. As we improve our understanding of the alterations that accompany endometrial carcinogenesis, it is likely that future chemopreventives may be developed for several types of endometrial cancer, each of which develops by specific pathways.

## REFERENCES

1. American Cancer Society: Cancer Facts and Figures 2000:16-17.
2. Gallup DG, Stock RJ: Adenocarcinoma of the endometrium in women 40 years of age or younger. *Obstet Gynecol* 1984;64:417

## Microarray Analysis Reveals Distinct Gene Expression Profiles among Different Histologic Types of Endometrial Cancer

John I. Risinger, G. Larry Maxwell, G. V. R. Chandramouli, Amir Jazaeri, Olga Aprelikova, Tricia Patterson, Andrew Berchuck, and J. Carl Barrett<sup>1</sup>

Laboratory of Biosystems and Cancer, National Cancer Institute, Bethesda, Maryland 20892 [J. I. R., G. L. M., G. V. R. C., A. J., O. A., T. P., J. C. B.]; Walter Reed Army Medical Center, Washington, DC 20013 [G. L. M.]; Department of Obstetrics and Gynecology/Division of Gynecologic Oncology, Duke University, Durham, North Carolina 27710 [A. B.]

### Abstract

Previous studies of oncogene and tumor suppressor gene alterations have suggested that differences exist in the molecular pathogenesis of the various histological types of endometrial cancer. To elucidate further the molecular events involved in endometrial carcinogenesis, we examined global expression patterns of 16 nonendometrioid cancers (13 serous papillary and 3 clear cell), 19 endometrioid cancers, and 7 age-matched normal endometria using cDNA microarrays. Unsupervised analysis of gene expression identified 191 genes that exhibited >2-fold differences ( $P < 0.001$ ) between the histological groups. Many genes were similarly dysregulated in both nonendometrioid and endometrioid cancers relative to normal endometria. Gene expression differences in only 24 transcripts could distinguish serous from endometrioid cancers, the two most common subgroups. These data provide the basis for investigation of previously unrecognized novel pathways involved in the development of endometrial cancers.

### Introduction

Two subtypes of endometrial carcinoma have been described based on both clinical and histopathologic variables (1). Type I endometrial cancers account for the majority of cases, and these cancers are usually well differentiated and E<sup>2</sup> in histology. These cancers are associated frequently with a history of unopposed estrogen exposure or other hyperestrogenic risk factors, such as obesity. Patients, with type I endometrial cancer, typically have early stage disease and a favorable prognosis with appropriate therapy. In contrast, Type II endometrial cancers are often poorly differentiated, non-E, and are not associated with hyperestrogenic factors. These cancers are more likely to be metastatic at presentation and often recur despite aggressive clinical interventions.

Molecular genetic evidence indicates that endometrial carcinoma likely develops as the result of a multistep process of oncogene activation and tumor suppressor gene inactivation (2). Our group and others have described some of these changes and demonstrated that these molecular alterations appear to be specific for Type I (E) and Type II (non-E) cancers. Type I cancers are characterized by mutation of *PTEN* (3-5), *KRAS2* (6, 7), and *CTNNB1* (8, 9) defects in DNA mismatch repair (Ref. 10; as evidenced by the microsatellite instability phenotype) and a near diploid karyotype (11). Type II cancers often contain mutations of *TP53* (12-14) and *Her-2/neu* (15) and are usually nondiploid (11).

Received 9/30/02; accepted 11/14/02.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>1</sup>To whom requests for reprints should be addressed, at NIH, 9000 Rockville Pike, Building 37, Room 5032, Mail Stop Code 4262, Laboratory of Biosystems and Cancer, National Cancer Institute, Bethesda, Maryland 20892.

<sup>2</sup>The abbreviations used are: E, endometrioid; NCI, National Cancer Institute; PS, papillary serous; PEG3, polyethylene glycol 3; STAT12, signal transducers and activators of transcription 12; TFF, intestinal trefoil factor; FOLR, folate binding protein; IGF, insulin-like growth factor; AGR2, anterior gradient 2; CC, clear cell.

The molecular pathogenesis of endometrial cancer is incompletely understood. Although alterations in several genes noted above have been described, none are present in the majority of cases. In addition, some endometrial cancers lack evidence of alterations in any of these genes. In a recent survey of our databases, we found that 44 of 87 endometrial cancers (50%) were lacking mutations in *PTEN*, *TP53*, *CTNNB1*, or the microsatellite instability phenotype, although about half of these cancers was advanced stage. This suggests the existence of unrecognized pathways that can lead to the development of endometrial cancer. To elucidate further the molecular pathogenesis of endometrial cancers, we have used cDNA microarrays to examine patterns of gene expression in E and non-E cancers and normal endometrium.

### Materials and Methods

#### Specimens and Nucleic Acid Isolation

Flash frozen endometrial cancers were obtained from patients undergoing hysterectomy at Duke University Medical Center. None of the patients had received preoperative chemotherapy or radiation. In addition, samples of normal endometrium (N) were obtained from patients undergoing hysterectomy for benign gynecological diseases. Tissues were obtained at Duke with Institutional Review Board-approved informed consent, and this study was approved by the NCI Institutional Review Board. Endometrial cancers were examined by a gynecologic pathologist to confirm the histologies as being PS, E, or CC. Tissue samples were subjected to RNA isolation using TRIzol and an additional purification using the RNeasy Kit (Qiagen, Valencia, CA) following the manufacturer's recommendation. We examined 19 histologically normal endometria that were age matched to the cancers and able to isolate a sufficient quantity of dissected glandular epithelium in only seven that were used in this study. After isolation of RNA, the integrity of each RNA sample was verified by denaturing gel electrophoresis.

#### Microarray Processing

Total RNA was amplified linearly with a modification of the Eberwine method. Briefly, total RNA was reverse transcribed by using a 63 nucleotide synthetic primer containing the T7 RNA polymerase binding site, 5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG(T)<sub>24</sub>-3'. Second-strand cDNA synthesis (producing double-stranded cDNA) was performed with RNase H, *Escherichia coli* DNA polymerase I, and *E. coli* DNA ligase (Invitrogen, Carlsbad, CA). After cDNA was made blunt ended with T4 DNA polymerase (Invitrogen), it was purified by extraction with a mixture of phenol, chloroform, and isoamyl alcohol and by precipitation in the presence of ammonium acetate and ethanol. The double-stranded cDNA was then transcribed with T7 RNA polymerase (T7 Megascript kit; Ambion, Austin, TX), yielding linearly amplified antisense RNA, which was purified with RNeasy mini columns (Qiagen). The cDNA microarray chips contained 9984 total features representing the human GEM2 set of clones (InCyte) and were manufactured at the NCI microarray facility. Four micrograms of amplified RNA were reverse transcribed and directly labeled with cyanine 5-conjugated dUTP (endometrial RNA) or cyanine 3-conjugated dUTP (Stratagene Universal Reference RNA). Hybridization was performed in 5 × saline sodium citrate

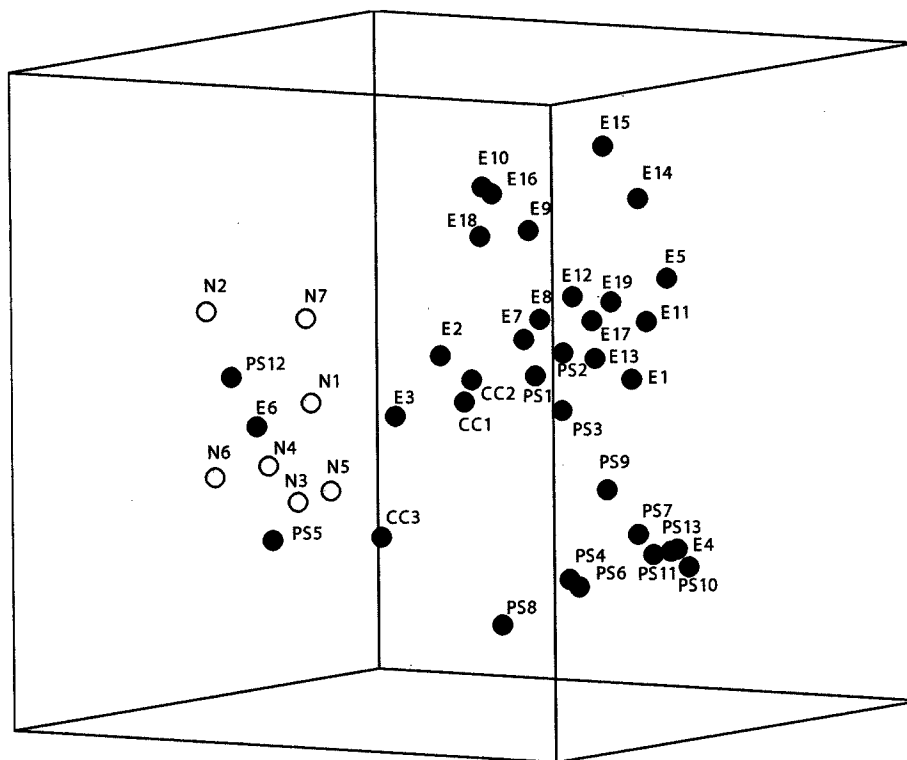


Fig. 1. Multidimensional scaling model based on the overall gene expression in PS, E, CC, and normal endometrium (N).

containing 25% formamide for 14–16 h at 42°C. Slides were washed, dried, and scanned. A detailed protocol for RNA amplification, cDNA probe labeling, and hybridization is available on the Internet.<sup>3</sup> Genepix software (Axon Instruments, Inc., Union City, CA) was used to analyze the raw data, which were then uploaded to a relational database maintained by the Center for Information Technology at the NIH (Bethesda, MD).

#### Data Analysis

The microarrays were scanned on an Axon Instruments Genepix 4000A scanner (Axon Instruments, Foster City, CA) at wavelengths 635 and 532 nm for Cy5 and Cy3 dyes, respectively, to obtain images of 10- $\mu$ m resolution. The quantification of spot intensities, qualities, and local background was performed automatically by Genepix software using variable spot diameter in the range 70–180  $\mu$ m and a manual supervision for any inaccuracies in the automatic spot detection. Local background correction was applied to spot intensities before the calculation of expression ratios. The spots having a minimum signal level of 250 counts and well defined with  $\geq 70\%$  of the pixels above SD were used for expression analysis. The expressions were normalized by median centering the logarithmic expression ratios of Cy3 and Cy5 signals within each array. The logarithmic ratio *versus* logarithmic signal scatters and Cy5 signal *versus* Cy3 signal scatters did not indicate any abnormalities in distributions. Statistical analysis was performed using Biometric Research Branch Array Tools software (NCI) using logarithmic values of expression ratios to the base 2.

A set of 9431 cDNA clones with good spot quality from a total of 9984 spotted on microarray was used to study differential gene expression. The significance of expression differences between tissues was computed by F tests at a two-tailed  $P < 0.001$ . All of the genes found significant by F test were ranked by the magnitude of average expression differences, which gave 412 transcripts above a 1.5-fold ratio threshold and 191 genes above a 2-fold threshold level. The SE values computed for 99% of average expressions of the different histological groups were  $< 1.25$ -fold, indicating this value as a general error level of averaged expressions. The random chance of obtaining significant differential expression at  $P < 0.001$  and above ratio 1.5 threshold level was computed as follows. The percentage of average expression differ-

ences of all of the 9431 genes on arrays  $> 1.5$  threshold was estimated as 8%. Considering this as high limit of random possibility of expression change  $> 1.5$  fold, at the most, only 1 of 9431 clones on the array could be identified as significant by random chance at  $P < 0.001$ . At a 2-fold ratio threshold level, the random chance is computed as 0.2 of 9431 genes.

**Hierarchical Clustering.** The set of genes differentially expressed  $> 2$ -fold was clustered by the similarity of their expression profiles. Hierarchical clustering was performed on logarithmic values of expressions using  $1-\rho$  as distance metric, where  $\rho$  is the correlation coefficient of any two gene expressions (16). The expression data were shown relative to average normal endometria expression. The cluster is color coded using red for up-regulation from normal endometria and green for down-regulation.

**Multidimensional Scaling.** The similarities of gene expression profiles of samples were studied by a multidimensional scaling procedure (Young and Hamer details available on the Internet).<sup>4</sup> Each sample was considered as an  $n$ -dimensional vector where each gene represents a dimension, and  $n$  is the number of genes. The dissimilarities of samples were computed using  $1-\rho$  as distance metric, where  $\rho$  is the correlation coefficient of the two samples. Each sample is represented as a coordinate in a three-dimensional space as shown in Fig. 1. Samples with similar gene expressions on microarray are placed at closer proximity compared with the dissimilar ones.

**Classification.** Binary class comparison and prediction was performed on PS and E pairs. The genes distinguishing PS and E were selected by F test ( $P < 0.001$ ). The class prediction was performed by computing compound covariate of gene expressions (17). The compound covariate is defined as  $\sum_i t_i \times (x_{ij} - m_i)$ , where  $x_{ij}$  is the expression of gene  $i$  in sample  $j$ ,  $m_i$  is the midpoint of two classes for gene  $i$ ,  $t_i$  is the t score, and the summation is overall genes selected for classification. The prediction was performed by leaving one sample out at a time for cross-validation and using all other samples for classification. The cross-validated misclassification error rate was estimated by 10,000 random permutations of class labels as  $< 0.01\%$ . The class prediction using 232 differentially expressed genes between PS and E could successfully predict 94% of the samples.

**Data availability.** The complete dataset can be accessed on the Internet.<sup>5</sup>

<sup>3</sup> Internet address: <http://nciarrray.nci.nih.gov/reference/index.shtml>.

<sup>4</sup> Internet address: <http://forrest.psych.unc.edu/teaching/p208a/mds/mds.html>.

<sup>5</sup> Internet address: <http://home.ccr.cancer.gov/risingerdata1102>.

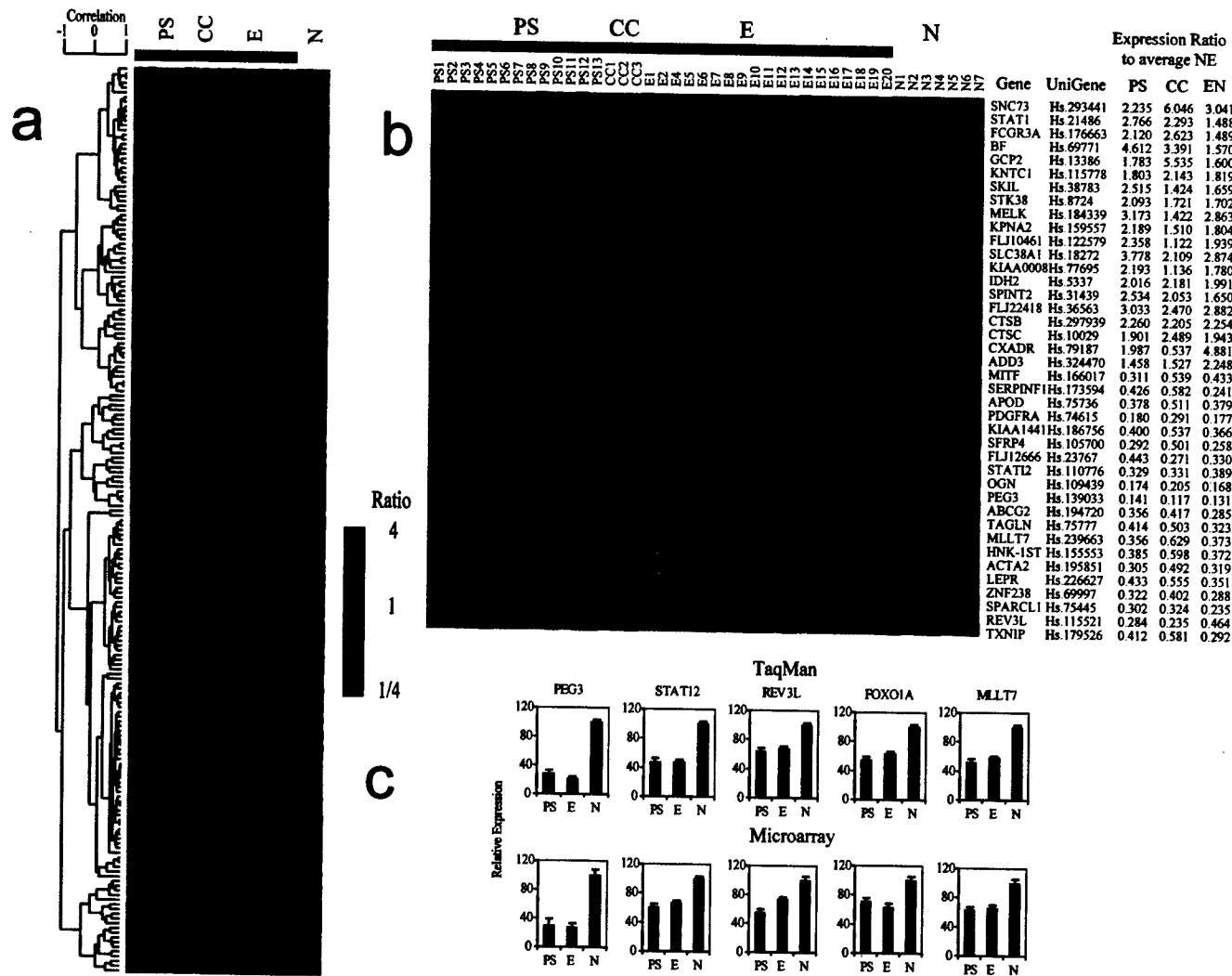


Fig. 2. Molecular profiles of all 42 specimens: 13 PS, 19 E, 3 CC, and 7 normal endometrium (N). a, hierarchical clustering of 191 genes with differential expression between each of the four groups ( $P < 0.001$ ) using a threshold of 2-fold. The cluster is color coded using red for up-regulation from normal endometrium, green for down-regulation, and black for median expression. Agglomerative clustering of genes was illustrated with dendrograms. b, hierarchical clustering of the 20 most up-regulated and down-regulated genes across in the comparison of the four groups. The symbol for each gene is followed by the UniGene clone number of the corresponding DNA spotted on the array. Expression ratios comparing each of the two groups to normal endometrium as well as to each other are listed. c, quantitative PCR (TaqMan) and microarray expression analysis of five selected genes differentially expressed between PS and E cancers with normal endometria.

**Quantitative Real-time PCR**

The relative expression of genes shown in Fig. 2 was determined as follows. The concentrations of genes *PEG3*, *STAT12*, *REV3L*, *FOXO1A*, *MLL7*, and *glyceraldehyde-3-phosphate dehydrogenase* for all samples were determined using the standard curve method for normalization. Sequences for primers and probes are available on request. The gene expressions were then compared with average threshold PCR amplification cycle time of normal endometrial samples. Fig. 2c shows relative gene expressions (on logarithmic scale to base 2) compared with normal endometria. The SE values are shown as error bars. Gene expression assays (assay on demand) for the analysis of samples that distinguish PS and E cases were purchased from Applied Biosystems (Foster City, CA). The concentrations of genes *TFF3*, *IGF-II*, *dual specificity phosphatase 6*, *AGR2*, *ubiquitin COOH-terminal esterase-like 1*, and *FOLR1* were determined using multiplex PCR (same tube) method, where  $\beta$ -actin was used as a reference. The relative expressions were compared with average normal endometrial expression as above.

**Results and Discussion**

The endometrium is one of the most dynamic tissues present in placental mammals and poses interesting questions regarding experimental approaches with gene expression array technologies, because

levels of ovarian steroids would be predicted to drastically affect gene expression. Previous global expression profile analysis has examined the normal endometria of normal cycling women and in one case compared these data to a limited set of E endometrial cancers (18–20). We chose to examine a set of cancers and normal tissues matched for age and typical of the median age of onset for endometrial cancers, which is generally after the period of normal cycling. We analyzed the RNA from 42 endometrial samples and hybridized them with a universal reference RNA against a cDNA gene chip with 9984 features, most of which are known genes. RNA preparation and processing of arrays were stringently controlled for quality and reproducibility. After array processing and scanning, 9431 features provided gene expression data sufficient for analysis (expression was present in either the endometrial or universal RNA).

Initially, to seek similarities of expression profiles, we classified the 35 endometrial cancers and 7 normal endometria using unsupervised multidimensional scaling. Multidimensional scaling using all of the 9431 genes on arrays revealed distinctively separated clusters for PS, CC, E, and normal endometria, indicating consistent global expression changes for each type of sample. This analysis indicated that gene

expression patterns between samples were somewhat variable, as expected, but differences were sufficient to cluster the majority of cases into their respective subgroups (Fig. 1). These data are in accordance with the differing morphological characteristics of these samples yet still show the heterogeneity predicted of any cancer. Somewhat surprisingly, two PS cancers clustered close to the normal endometrial samples. Additionally, the distinct clustering of the three CC cancers indicates that their underlying biology may be distinct from the majority of serous papillary and E lesions.

Next, we compared the relative gene expressions based on the individual histology. The significance of expression differences of genes between samples was computed using F tests. The significant differential gene expression from normal endometrial tissue (N) to PS, E, or CC histology was separately determined with a two-tailed F test ( $P < 0.001$ ). At this statistical significance level, 293 genes in PS, 281 genes in E, and 67 genes in CC were found to differ significantly from normal endometria. F tests comparing PS and E, PS and CC, and N and CC revealed 232, 16, and 113 differentially expressed genes at the same significance level. These data indicate the least significant differences in the expression profiles were between PS and CC, whereas the greatest differences were between PS and normal endometria.

To further understand the underlying biology of these groupings, the 191 genes that statistically differed between one or more of the four subgroups and exhibited at least a 2-fold gene expression difference were analyzed using hierarchical clustering (Fig. 2a). We grouped these and depicted the 20 most highly up-regulated and 20 most highly down-regulated genes (Fig. 2b). Many genes shown previously to be involved in carcinogenesis are present on these lists, providing some validity to our array analysis, including the *platelet-derived growth factor-A* and *AXL* genes, which were included in a previous array study (20). To further verify the quality of our array data, a subset of five genes (*PEG3*, *STAT12*, *REV3L*, *FOXO1A*, and *MLLT7*) was examined in all 42 tissue specimens using quantitative real-time PCR (TaqMan; Fig. 2c). These genes were chosen based on their universal down-regulation in both E and non-E cancers and on their potentially interesting roles in carcinogenesis.

Defects in DNA mismatch repair characterize only ~25% of E endometrial cancers (10), suggesting that other unrecognized DNA repair pathways are impaired in endometrial cancers. To examine this concept, we scanned our array data for genes involved in DNA repair processes. Microarray analysis indicated that most PS cancers under-expressed the DNA polymerase  $\zeta$  catalytic subunit *REVL*, as did a significant portion of the E lesions. Loss of this activity could result in a hypermutagenic phenotype as the result of defective translesion repair synthesis, a major route of DNA resynthesis after DNA repair. Real-time PCR analysis confirmed the down-regulation of this gene in endometrial cancers (Fig. 2c).

Endometrial cancers likely evolve in part as a result of many epigenetic defects, including loss of imprinting and promoter hypermethylation, e.g., most endometrial cancers with the microsatellite instability phenotype do not contain mutation of a mismatch repair gene but instead contain hypermethylated *MLH1* promoter alleles that silence transcription from this locus (21, 22). Evidence for other hypermethylated loci in endometrial cancers involves the estrogen and progesterone receptors and the adenomatous polyposis coli tumor suppressor gene (23–25). Our array and a previous array study that included only E cancers (20) indicate many down-regulated genes, some of which may be silenced by similar epigenetic mechanisms. We chose to examine the *PEG3* gene as one of these down-regulated transcripts in which to verify our array data. *PEG3*, a kruppel type zinc-finger transcription factor, is transcribed from a paternally imprinted locus on chromosome 19. *PEG3* is down-regulated in gliomas by a hypermethylation mechanism (26), and reintroduction of *PEG3* suppresses tumor formation in glioma cell lines (27). Our microar-

ray and real-time PCR analysis show a dramatic reduction in *PEG3* mRNA almost universally in endometrial cancers, perhaps indicating identification of a novel unrecognized pathway in endometrial carcinogenesis. Several genes like *PEG3* are candidates for epigenetic down-regulation by promoter hypermethylation in endometrial cancers.

Several genes that may impact on a central signaling pathway involving the mammalian counterparts of the C-elegans insulin/IGF-I cell survival or longevity pathway were identified on our array. *PTEN*, an essential lipid phosphatase that has activity in many signal transduction paths is a key component of this pathway and is mutated in many E endometrial cancers (3–5). *PTEN* negatively regulates signals through its lipid phosphatase activity (28). We hypothesize that many risk factors for endometrial cancer act by elevating growth factors like estrogens and IGF-I, which serve to constitutively activate antiapoptosis pathways. Normally, negative regulators of these pathways like *PTEN* suppress carcinogenesis, and endometrial cancers can only arise clonally from rare cells that escape the tight regulation of these pathways either through gene mutations or epigenetic changes. Mutation of *PTEN* has been demonstrated in histologically normal and hyperplastic endometrium, suggesting that alteration of this tumor suppressor gene may be an early event in carcinogenesis of E endometrial cancer (29, 30).

Interestingly, array data indicated the down-regulation of two forkhead transcription factors, *MLLT7* and *FOXO1A*, also known as *AFX* and *FKHR*. These transcripts represent two of the three mammalian homologues to the C-Elegans *DAF-16* forkhead transcription factor involved in insulin/IGF-I signaling and cell longevity. *PTEN*, a negative upstream regulator of this pathway, affects downstream signaling presumably through augmentation of cell survival and antiapoptosis programs. This activity is mediated in part through the phosphorylation of the v-AKT homologues, which are augmented in the absence of functional *PTEN*. Phosphorylated AKT acts on *FKHRs* by directing their ubiquitination in the cytoplasm and subsequently allowing the transcription of antiapoptotic genes. Interestingly, our array data indicate that these two forkhead transcripts are down-regulated in endometrial cancers, suggesting an additional mechanism by transcriptional repression. This observation would result in the diminution of the activity of *DAF-16* homologues. The exact downstream changes that might result from this diminution and contribute to endometrial carcinogenesis are obscure. Independent real-time PCR analysis confirms the down-regulation of these genes in our cancer set and the need to further explore this observation (Fig. 2c).

Loss of tumor suppressor activities is a fundamental property of most cancer types. Our gene array analysis indicated a down-regulation of the *STAT12* or *SOCS2* transcript in endometrial cancers. *SOCS2* possesses functional characteristics of a tumor suppressor gene. Specifically, *SOCS2*-deficient mice exhibit a giant phenotype as the result of nonabrogated growth (31). This phenotype is identical to mice given exogenous IGF-I or growth hormone. Subsequent study based in part on this observation indicated that *SOCS2* also directly interacts with the IGF receptor (32). This observation also potentially links this gene to the insulin survival pathway. Additionally, we also noted dysregulation of the *YWHAZ* (14–3-3 zeta) transcript in our array data. Limited evidence also implicates this gene to the insulin survival pathway by its association with *IRS1* (33). Dereglulation of growth factor pathways caused by inappropriate epithelial-stromal interactions have been proposed a potential mediator of endometrial carcinogenesis (34, 35). *SOCS2* may also function in this capacity. Down-regulation of *SOCS2*, Forkhead transcription factors, and mutant *PTEN* indicate significant convergence on the IGF-I cell survival pathway in this tumor type.

Determining the gene expression differences between E and serous cancers represents a potential step in determining the differences between these two most common histologies of endometrial cancer. A relatively few transcripts could distinguish between these two types of

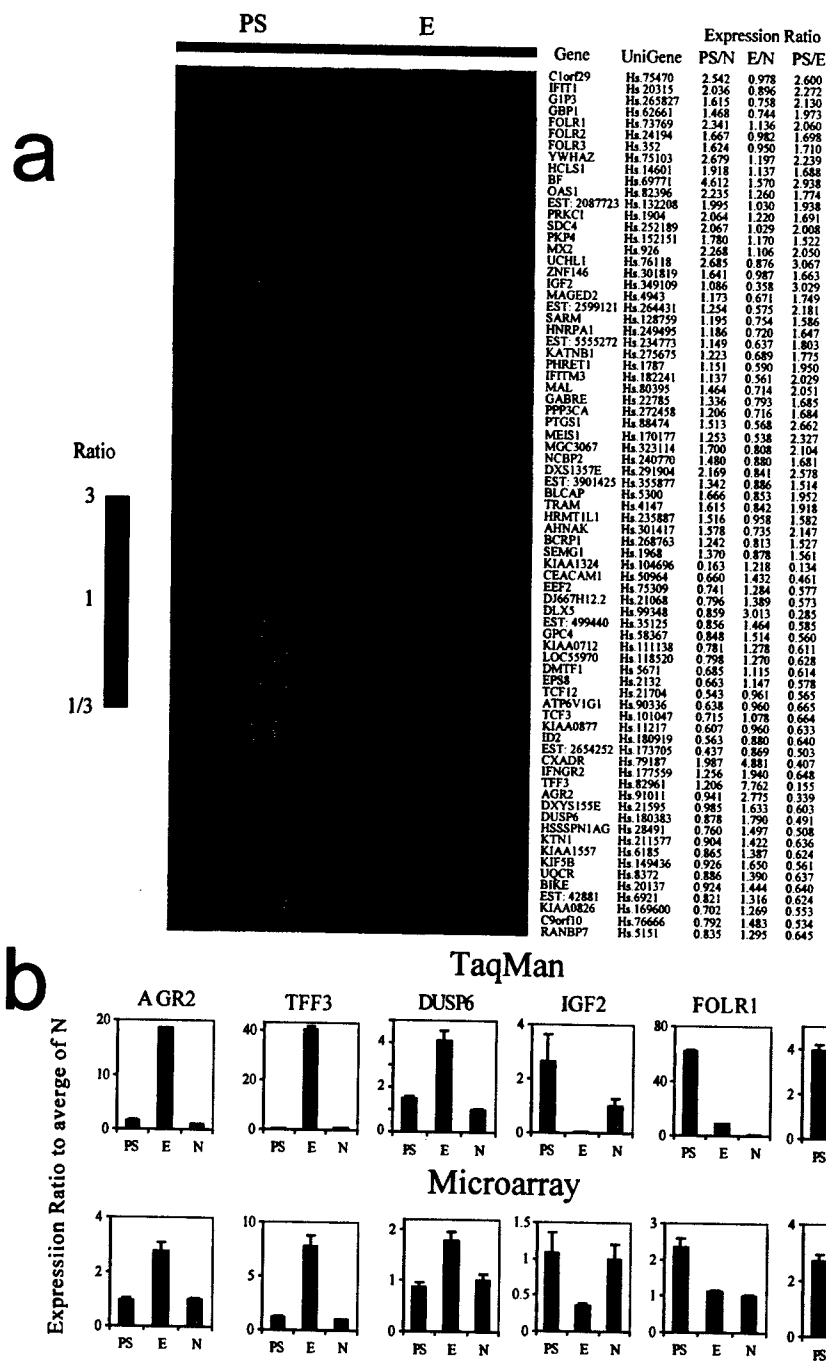


Fig. 3. Gene expression differences that distinguish PS versus E endometrial cancer. a, hierarchical clustering of 75 nonredundant genes with statistically differential expression between the PS and E cancer groups (modified F test,  $P < 0.001$ ) using a threshold of 1.5-fold. The cluster is color coded using red for up-regulation from normal endometrium (N), green for down-regulation, and black for median expression. Agglomerative clustering of genes was illustrated with dendrograms. b, quantitative PCR (TaqMan) and microarray expression analysis of six selected genes differentially expressed between PS and E cancers also showing their relative expression in normal endometrium (N).

endometrial cancer (at  $\geq 2$ -fold difference in 24 genes and 1.5-fold difference in 75 genes). The entire list of 75 genes is depicted in Fig. 3, including the relative expression of all cancers in each histological group. We validated the expression of six of the transcripts, including the *FOLR*, *dual specificity phosphatase 6*, *IGF-II*, *TFF3*, *AGR2*, and *ubiquitin COOH-terminal esterase-like 1*. All six transcripts were analyzed by real-time PCR and confirmed our microarray analysis (Fig. 3b). Class prediction models could reproducibly define 30 of 32 cancers into either serous or E histology (data not shown). Interestingly, the two cancers that were not accurately classified in this model were the same two serous cancers that grouped with normal atrophic endometrium in multidimensional scaling analysis. These two cancers may represent a grouping of serous cancers that could be distinguished with a larger sample set or may represent anomalies because of some other undetermined factors. Despite these

two samples, the successful analysis on such a limited number of samples suggests that class prediction models could be applied to future endometrial arrays with the idea of predicting clinical outcomes.

Quantitative PCR analyses are more reflective of the absolute level of gene expression differences than are cDNA arrays. This analysis validated our array data in all cases but indicated substantial gene expression differences in three of the transcripts that we chose to validate. In particular, transcript for the intestinal trefoil protein, *TFF3*, was dramatically up-regulated in the cancers with E histology as was the *AGR2* developmental gene (Fig. 3c). The trefoil peptides are implicated in other cancer types and down-regulation of *TFF1* and *TFF2* is often accompanied by up-regulation of *TFF3*, which mediates various cell adhesion and other signaling pathways (36–39). The observed ~40-fold up-regulation of this gene deserves more study in E cancers.

Overexpression of the folate receptor and its fetal homologues was noted in the PS cancers. Real-time quantitative PCR assays revealed that overexpression of FOLR in PS cancers was even more striking than suggested by the microarray, with 9 of 13 PS tumors showing a significantly elevated level of expression compared with normal endometrium. These data are represented as a group in Fig. 3b, where these cancers express this gene at levels > 60-fold, as compared with normal endometria. Vaccines that target FOLR have been developed for use in ovarian carcinoma (40–42) and could potentially be used in the adjuvant treatment of PS endometrial cancers that overexpress this gene.

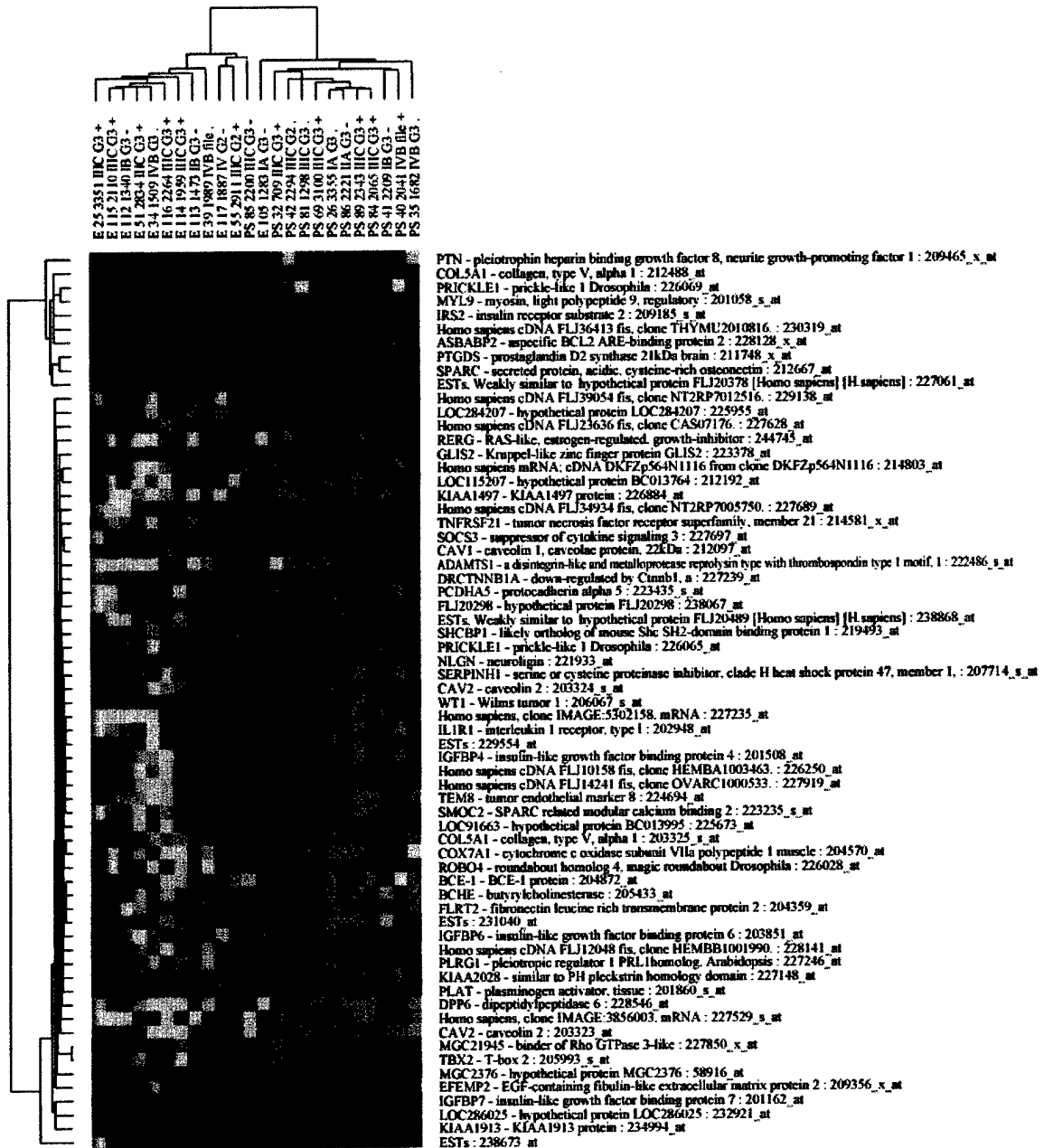
In summary, these array data will be useful for investigating pathways to be targeted by small molecules or for future gene reactivation strategies yet to be realized. As expected, many genes were dysregulated in various histological groups, but our data also indicate many genes similarly dysregulated in the various histologies. These data identify several additional pathways important in the development of endometrial cancer and suggest multiple avenues of investigation.

## References

- Deligdisch, L., and Holinka, C. F. Endometrial carcinoma: two diseases? *Cancer Detect. Prev.*, 10: 237–246, 1987.
- Berchuck, A., and Boyd, J. Molecular basis of endometrial cancer. *Cancer*, 76: 2034–2040, 1995.
- Kong, D., Suzuki, A., Zou, T. T., Sakurada, A., Kemp, L. W., Wakatsuki, S., Yokoyama, T., Yamakawa, H., Furukawa, T., Sato, M., Ohuchi, N., Sato, S., Yin, J., Wang, S., Abraham, J. M., Souza, R. F., Smolinski, K. N., Meltzer, S. J., and Horii, A. PTEN1 is frequently mutated in primary endometrial carcinomas. *Nat. Genet.*, 17: 143–144, 1997.
- Risinger, J. I., Hayes, A. K., Berchuck, A., and Barrett, J. C. PTEN/MMAC1 mutations in endometrial cancers. *Cancer Res.*, 57: 4736–4738, 1997.
- Tashiro, H., Blazes, M. S., Wu, R., Cho, K. R., Bose, S., Wang, S. I., Li, J., Parsons, R., and Ellenson, L. H. Mutations in PTEN are frequent in endometrial carcinoma but rare in other common gynecological malignancies. *Cancer Res.*, 57: 3935–3940, 1997.
- Enomoto, T., Inoue, M., Perantoni, A. O., Buzard, G. S., Miki, H., Tanizawa, O., and Rice, J. M. K-ras activation in premalignant and malignant epithelial lesions of the human uterus. *Cancer Res.*, 51: 5308–5314, 1991.
- Ignar-Trowbridge, D., Risinger, J. I., Dent, G. A., Kohler, M., Berchuck, A., McLachlan, J. A., and Boyd, J. Mutations of the Ki-ras oncogene in endometrial carcinoma. *Am. J. Obstet. Gynecol.*, 167: 227–232, 1992.
- Fukuchi, T., Sakamoto, M., Tsuda, H., Maruyama, K., Nozawa, S., and Hirohashi, S. Beta-catenin mutation in carcinoma of the uterine endometrium. *Cancer Res.*, 58: 3526–3528, 1998.
- Kobayashi, K., Sagae, S., Nishioka, Y., Tokino, T., and Kudo, R. Mutations of the beta-catenin gene in endometrial carcinomas. *Jpn. J. Cancer Res.*, 90: 55–59, 1999.
- Risinger, J. I., Berchuck, A., Kohler, M. F., Watson, P., Lynch, H. T., and Boyd, J. Genetic instability of microsatellites in endometrial carcinoma. *Cancer Res.*, 53: 5100–5103, 1993.
- Lukes, A. S., Kohler, M. F., Pieper, C. F., Kerns, B. J., Bentley, R., Rodriguez, G. C., Soper, J. T., Clarke-Pearson, D. L., Bast, R. C., Jr., and Berchuck, A. Multivariable analysis of DNA ploidy, p53, and HER-2/neu as prognostic factors in endometrial cancer. *Cancer*, 73: 2380–2385, 1994.
- Okamoto, A., Sameshima, Y., Yamada, Y., Teshima, S., Terashima, Y., Terada, M., and Yokota, J. Allelic loss on chromosome 17p and p53 mutations in human endometrial carcinoma of the uterus. *Cancer Res.*, 51: 5632–5635, 1991.
- Kohler, M. F., Berchuck, A., Davidoff, A. M., Humphrey, P. A., Dodge, R. K., Iglehart, J. D., Soper, J. T., Clarke-Pearson, D. L., Bast, R. C., Jr., and Marks, J. R. Overexpression and mutation of p53 in endometrial carcinoma. *Cancer Res.*, 52: 1622–1627, 1992.
- Kohler, M. F., Carney, P., Dodge, R., Soper, J. T., Clarke-Pearson, D. L., Marks, J. R., and Berchuck, A. p53 overexpression in advanced-stage endometrial adenocarcinoma. *Am. J. Obstet. Gynecol.*, 175: 1246–1252, 1996.
- Berchuck, A., Rodriguez, G., Kinney, R. B., Soper, J. T., Dodge, R. K., Clarke-Pearson, D. L., and Bast, R. C., Jr. Overexpression of HER-2/neu in endometrial cancer is associated with advanced stage disease. *Am. J. Obstet. Gynecol.*, 164: 15–21, 1991.
- Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA*, 95: 14863–14868, 1998.
- Radmacher, M. D., McShane, L. M., and Simon, R. A paradigm for class prediction using gene expression profiles. *J. Comput. Biol.*, 9: 505–511, 2002.
- Carson, D. D., Lagow, E., Thathiah, A., Al-Shami, R., Farach-Carson, M. C., Vernon, M., Yuan, L., Fritz, M. A., and Lessey, B. Changes in gene expression during the early to mid-luteal (receptive phase) transition in human endometrium detected by high-density microarray screening. *Mol. Hum. Reprod.*, 8: 871–879, 2002.
- Kao, L. C., Tulac, S., Lobo, S., Imani, B., Yang, J. P., Germeyer, A., Osteen, K., Taylor, R. N., Lessey, B. A., and Giudice, L. C. Global gene profiling in human endometrium during the window of implantation. *Endocrinology*, 143: 2119–2138, 2002.
- Mutter, G. L., Baak, J. P., Fitzgerald, J. T., Gray, R., Neubergh, D., Kust, G. A., Gentleman, R., Gullans, S. R., Wei, L. J., and Wilcox, M. Global expression changes of constitutive and hormonally regulated genes during endometrial neoplastic transformation. *Gynecol. Oncol.*, 83: 177–185, 2001.
- Katabuchi, H., van Rees, B., Lambers, A. R., Ronnett, B. M., Blazes, M. S., Leach, F. S., Cho, K. R., and Hedrick, L. Mutations in DNA mismatch repair genes are not responsible for microsatellite instability in most sporadic endometrial carcinomas. *Cancer Res.*, 55: 5556–5560, 1995.
- Simpkins, S. B., Bocker, T., Swisher, E. M., Mutch, D. G., Gersell, D. J., Kovatic, A. J., Palazzo, J. P., Fishel, R., and Goodfellow, P. J. MLH1 promoter methylation and gene silencing is the primary cause of microsatellite instability in sporadic endometrial cancers. *Hum. Mol. Genet.*, 8: 661–666, 1999.
- Sasaki, M., Kotcherquina, L., Dharia, A., Fujimoto, S., and Dahiya, R. Cytosine-phosphoguanine methylation of estrogen receptors in endometrial cancer. *Cancer Res.*, 61: 3262–3266, 2001.
- Zysman, M., Saka, A., Millar, A., Knight, J., Chapman, W., and Bapat, B. Methylation of adenomatous polyposis coli in endometrial cancer occurs more frequently in tumors with microsatellite instability phenotype. *Cancer Res.*, 62: 3663–3666, 2002.
- Sasaki, M., Dharia, A., Oh, B. R., Tanaka, Y., Fujimoto, S., and Dahiya, R. Progesterone receptor B gene inactivation and CpG hypermethylation in human uterine endometrial cancer. *Cancer Res.*, 61: 97–102, 2001.
- Maegawa, S., Yoshioka, H., Itaba, N., Kubota, N., Nishihara, S., Shirayoshi, Y., Nanba, E., and Oshimura, M. Epigenetic silencing of PEG3 gene expression in human glioma cell lines. *Mol. Carcinog.*, 31: 1–9, 2001.
- Kohda, T., Asai, A., Kuroiwa, Y., Kobayashi, S., Aisaka, K., Nagashima, G., Yoshida, M. C., Kondo, Y., Kagiya, N., Kirino, T., Kaneko-Ishino, T., and Ishino, F. Tumor suppressor activity of human imprinted gene PEG3 in a glioma cell line. *Genes Cells*, 6: 237–247, 2001.
- Myers, M. P., Pass, I., Batty, I. H., Van der Kaay, J., Stolarow, J. P., Hemmings, B. A., Wigler, M. H., Downes, C. P., and Tonks, N. K. The lipid phosphatase activity of PTEN is critical for its tumor suppressor function. *Proc. Natl. Acad. Sci. USA*, 95: 13513–13518, 1998.
- Maxwell, G. L., Risinger, J. I., Gumbs, C., Shaw, H., Bentley, R. C., Barrett, J. C., Berchuck, A., and Futreal, P. A. Mutation of the PTEN tumor suppressor gene in endometrial hyperplasias. *Cancer Res.*, 58: 2500–2503, 1998.
- Mutter, G. L., Ince, T. A., Baak, J. P., Kust, G. A., Zhou, X. P., and Eng, C. Molecular identification of latent precancers in histologically normal endometrium. *Cancer Res.*, 61: 4311–4314, 2001.
- Horvat, S., and Medrano, J. F. Lack of Socs2 expression causes the high-growth phenotype in mice. *Genomics*, 72: 209–212, 2001.
- Dey, B. R., Spence, S. L., Nissley, P., and Furlanetto, R. W. Interaction of human suppressor of cytokine signaling (SOCS)-2 with the insulin-like growth factor-I receptor. *J. Biol. Chem.*, 273: 24095–24101, 1998.
- Ogihara, T., Isobe, T., Ichimura, T., Taoka, M., Funaki, M., Sakoda, H., Onishi, Y., Inukai, K., Anai, M., Fukushima, Y., Kikuchi, M., Yazaki, Y., Oka, Y., and Asano, T. 14–3–3 protein binds to insulin receptor substrate-1, one of the binding sites of which is in the phosphotyrosine binding domain. *J. Biol. Chem.*, 272: 25267–25274, 1997.
- Arnold, J. T., Kaufman, D. G., Seppala, M., and Lessey, B. A. Endometrial stromal cells regulate epithelial cell growth in vitro: a new co-culture model. *Hum. Reprod.*, 16: 836–845, 2001.
- Arnold, J. T., Lessey, B. A., Seppala, M., and Kaufman, D. G. Effect of normal endometrial stroma on growth and differentiation in Ishikawa endometrial adenocarcinoma cells. *Cancer Res.*, 62: 79–88, 2002.
- Kirikoshi, H., and Katoh, M. Expression of TFF1, TFF2 and TFF3 in gastric cancer. *Int. J. Oncol.*, 21: 655–659, 2002.
- Leung, W. K., Yu, J., Chan, F. K., To, K. F., Chan, M. W., Ebert, M. P., Ng, E. K., Chung, S. C., Malfetheriner, P., and Sung, J. J. Expression of trefoil peptides (TFF1, TFF2, and TFF3) in gastric carcinomas, intestinal metaplasia, and non-neoplastic gastric tissues. *J. Pathol.*, 197: 582–588, 2002.
- Taupin, D., Pedersen, J., Familari, M., Cook, G., Yeomans, N., and Giraud, A. S. Augmented intestinal trefoil factor (TFF3) and loss of pS2 (TFF1) expression precedes metaplastic differentiation of gastric epithelium. *Lab. Invest.*, 81: 397–408, 2001.
- Efstathiou, J. A., Noda, M., Rowan, A., Dixon, C., Chinery, R., Jawhari, A., Hattori, T., Wright, N. A., Bodmer, W. F., and Pignatelli, M. Intestinal trefoil factor controls the expression of the adenomatous polyposis coli-catenin and the E-cadherin-catenin complexes in human colon carcinoma cells. *Proc. Natl. Acad. Sci. USA*, 95: 3122–3127, 1998.
- Kim, D. K., Lee, T. V., Castilleja, A., Anderson, B. W., Peoples, G. E., Kudelka, A. P., Murray, J. L., Sittisomwong, T., Wharton, J. T., Kim, J. W., and Ioannides, C. G. Folate binding protein peptide 191–199 presented on dendritic cells can stimulate CTL from ovarian and breast cancer patients. *Anticancer Res.*, 19: 2907–2916, 1999.
- Peoples, G. E., Anderson, B. W., Lee, T. V., Murray, J. L., Kudelka, A. P., Wharton, J. T., and Ioannides, C. G. Vaccine implications of folate binding protein, a novel cytotoxic T lymphocyte-recognized antigen system in epithelial cancers. *Clin. Cancer Res.*, 5: 4214–4223, 1999.
- Peoples, G. E., Anderson, B. W., Fisk, B., Kudelka, A. P., Wharton, J. T., and Ioannides, C. G. Ovarian cancer-associated lymphocyte recognition of folate binding protein peptides. *Ann. Surg. Oncol.*, 5: 743–750, 1998.

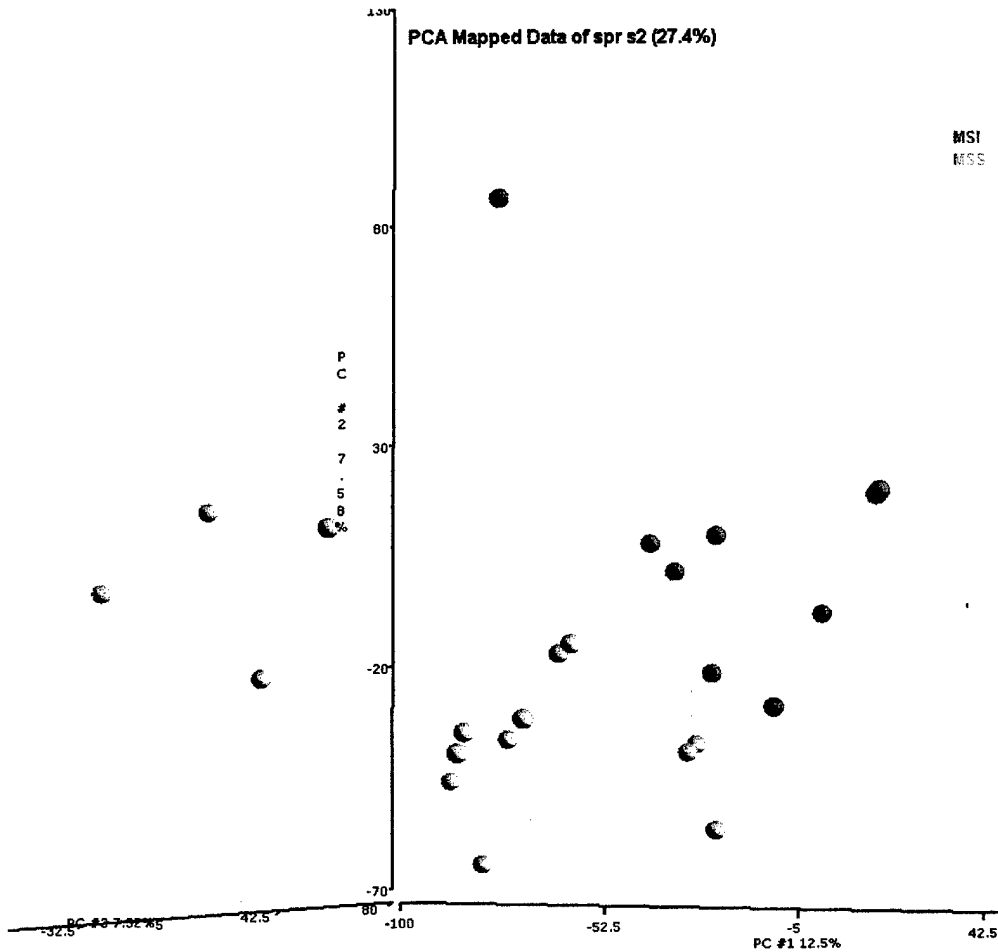


## APPENDIX 3



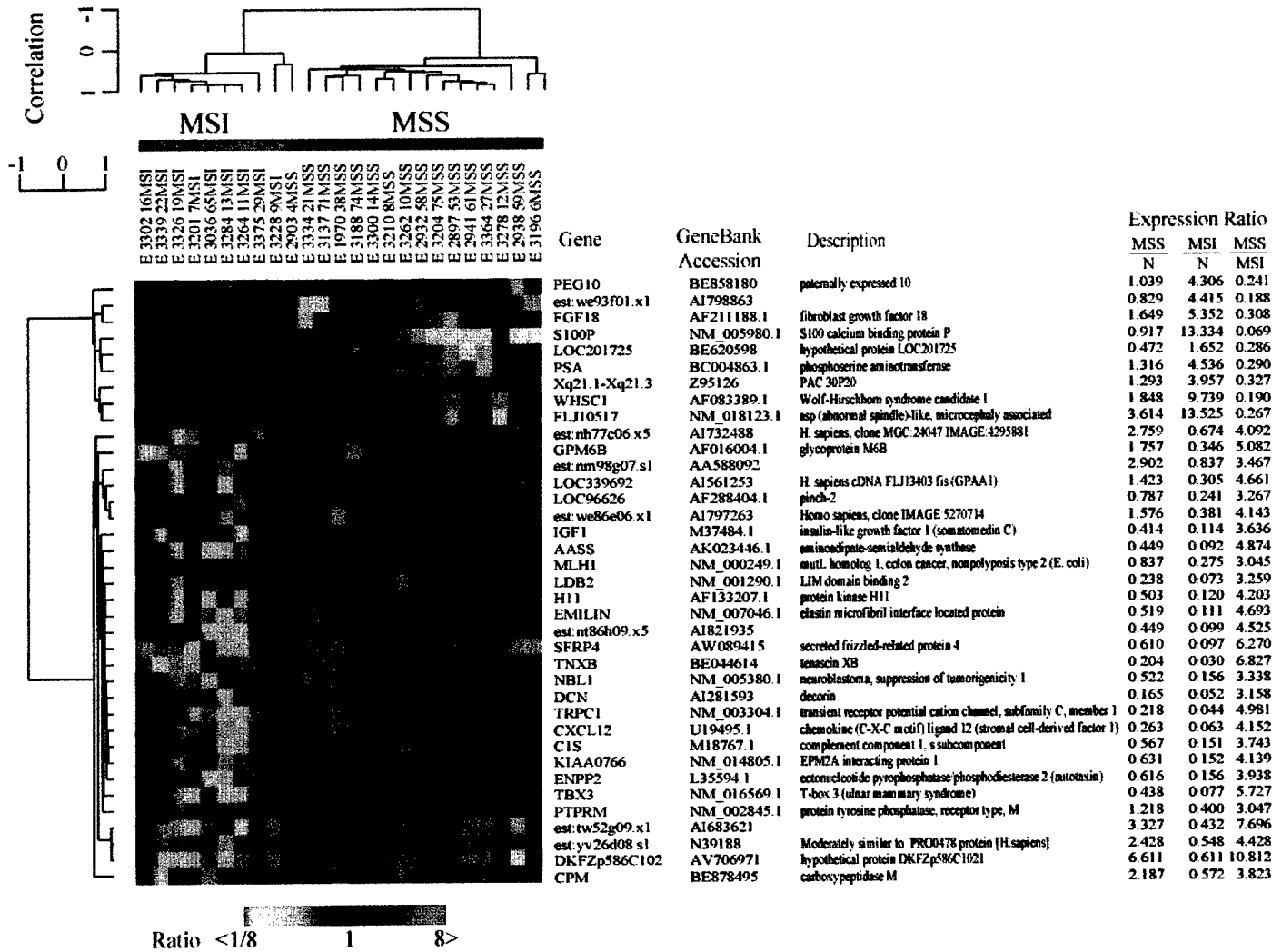
Gene expression differences that distinguish papillary serous (PS) versus endometrioid (E) endometrial cancer. Hierarchical clustering of 65 nonredundant genes with statistically significant differential expression between the papillary serous and endometrioid cancer groups (modified F test,  $P < 0.005$ ) using a threshold of 2-fold. The cluster is color-coded using red for up-regulation from normal endometrium (N), green for down-regulation and black for median expression. Agglomerative clustering of genes was illustrated with dendrograms.

## APPENDIX 4



The similarities of gene expression profiles of samples were studied by a principle component analysis (PCA) procedure. Each sample was considered as an  $n$ -dimensional vector where each gene represents a dimension and  $n$  is number of genes. Each sample is represented as a coordinate in a three dimensional space as shown. Samples with similar gene expressions on microarray are placed at closer proximity compared to the dissimilar ones. PCA based on the overall gene expression in microsatellite instability (MSI) (red) and microsatellite stable (MSS) (green) endometrioid endometrial cancers.

# APPENDIX 5



Hierarchical clustering of 37 nonredundant genes with statistically significant differential expression (modified F test,  $P < 0.005$ ) between the microsatellite instability (MSI) and microsatellite stable (MSS) groups using a threshold of 3-fold. The cluster is color-coded using red for up-regulation from normal endometrium (N), green for down-regulation and black for median expression. Agglomerative clustering of genes was illustrated with dendrograms.

## APPENDIX 6

	D2S123	D5S346	D17S250	BAT25	BAT26	D17S1330	D7S1794
1.	X	X	X	0	X	X	X
2.	0	0	0	0	0	0	0
3.	0	0	0	0	0	0	0
4.	0	0	0	0	0	0	X
5.	0	0	0	0	0	0	X
6.	0	0	0	0	0	0	0
7.	0	0	X	0	X	0	0

X: Microsatellite instability

0: Microsatellite stable

## APPENDIX 7

Age adjusted odds ratios and 95% CI for endometrial cancer according to OC hormone potency, with non-users as referent group

<u>Progestin/Estrogen</u>	<u>Cases, n (%)</u>	<u>Controls, n (%)</u>	<u>OR</u>	<u>95% CI</u>
Non-user	347 (83)	1430 (58)	1.0	Referent
High/high	16 (4)	309 (13)	0.4	(0.2-0.6)
High/low*	0 (0)	11 (0)	-	-
Low/high	37 (9)	419 (17)	0.5	(0.3-0.7)
Low/low	17 (4)	283 (12)	0.4	(0.3-0.7)

<u>Progestin</u>	<u>OR</u>	<u>95% CI</u>
Non-user	1.0	Referent
Low	0.3	(0.2-0.6)
High	0.5	(0.3-0.6)

<u>Estrogen</u>	<u>OR</u>	<u>95% CI</u>
Non-user	1.0	Referent
Low	0.4	(0.3-0.6)
High	0.4	(0.2-0.7)

Note

\* The low frequency of high progestin/low estrogen formulation use (<1% of OC users) prevented an independent statistical assessment of this group

## APPENDIX 8

Age adjusted odds ratios for endometrial cancer according to OC hormone potency and duration of OC with nonuser as the referent

prog/estrog	<u>3-18 mos. use</u>			<u>19-59 mos. use</u>			<u>60+ mos. use</u>		
	[case/con]	OR	95% CI	[case/con]	OR	95% CI	[case, cont]	OR	95% CI
<b>Progestin</b>									
High	[9,71]	0.8	(0.4 - 1.7)	[2,122]	0.1	(0.0 - 0.5)	[5,126]	0.2	(0.1 - 0.6)
Low	[23,212]	0.6	(0.4 - 1.0)	[13,228]	0.4	(0.2 - 0.7)	[17,255]	0.4	(0.2 - 0.6)
<b>Estrogen</b>									
High	[24,180]	0.8	(0.5 - 1.2)	[11,262]	0.3	(0.1 - 0.5)	[18,279]	0.4	(0.2 - 0.6)
Low	[8,103]	0.5	(0.2 - 1.0)	[4,88]	0.5	(0.2 - 1.3)	[4,102]	0.2	(0.1 - 0.7)