High-Resolution Mapping of Genomic Imbalance and Identification of Gene Expression Profiles Associated with Differential Chemotherapy Response in Serous Epithelial Ovarian Cancer[1]*

Marcus Bernardini[*,†], Chung-Hae Lee*[*], Ben Beheshti*[*], Mona Prasad*[‡], Monique Albert*[‡], Paula Marrano*[*], Heather Begley*[£], Patricia Shaw*[£,†,¶], Al Covens*[§], Joan Murphy*[¶], Barry Rosen*[¶], Salomon Minkin*[*,#], Jeremy A. Squire*[*,†] and Pascale F. Macgregor*[‡,2]

*Ontario Cancer Institute, Princess Margaret Hospital, University Health Network, Toronto, Ontario, Canada; †Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada; ‡Microarray Centre, Clinical Genomics Centre, University Health Network, Toronto, Ontario, Canada; £Department of Pathology, University Health Network, Toronto, Ontario, Canada; §Division of Gynecological Oncology, Sunnybrook and Women's College Hospital, Toronto, Ontario, Canada; ¶Division of Gynecological Oncology, Department of Obstetrics and Gynecology, University of Toronto, Toronto, Ontario, Canada; #Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

Abstract

Array comparative genomic hybridization (aCGH) and microarray expression profiling were used to subclassify DNA and RNA alterations associated with differential response to chemotherapy in ovarian cancer. Two to 4 Mb interval arrays were used to map genomic imbalances in 26 sporadic serous ovarian tumors. Cytobands 1p36, 1q42-44, 6p22.1-p21.2, 7q32.1-q34 9q33.3-q34.3, 11p15.2, 13q12.2-q13.1, 13q21.31, 17q11.2, 17q24.2-q25.3, 18q12.2, and 21q21.2-q21.3 were found to be statistically associated with chemotherapy response, and novel regions of loss at 15q11.2-q15.1 and 17q21.32-q21.33 were identified. Gene expression profiles were obtained from a subset of these tumors and identified a group of genes whose differential expression was significantly associated with drug resistance. Within this group, five genes (*GAPD*, *HMGB2*, *HSC70*, *GRP58*, and *HMGB1*), previously shown to form a nuclear complex associated with resistance to DNA conformation–altering chemotherapeutic drugs in *in vitro* systems, may represent a novel class of genes associated with *in vivo* drug response in ovarian cancer patients. Although RNA expression change indicated only weak DNA copy number dependence, these data illustrate the value of molecular profiling at both the RNA and DNA levels to identify small genomic regions and gene subsets that could be associated with differential chemotherapy response in ovarian cancer.

Neoplasia (2005) **7,** 603–613

Keywords: cisplatin, taxol, gene amplification, gene deletion, microarray data mining.

Introduction

Ovarian cancer is the second most frequently diagnosed gynecologic malignancy, and causes more deaths than any other cancer of the reproductive system. The lack of reliable methods of early detection and the absence of specific symptoms result in late-stage diagnosis in 70% of patients. Although initial response rates to conventional chemotherapy among advanced stage patients are high, resistance to chemotherapy remains the primary factor accounting for the low 5-year survival in this patient population [1].

Ovarian cancer chemotherapy most commonly involves a first-line combination of platinum-based compounds plus paclitaxel following cytoreductive surgery. Response to chemotherapy varies among patients, and initial treatment response is often the most important consideration in choosing second-line therapies. The role of CA 125 serum levels as a surrogate

Abbreviations: aCGH, array comparative genomic hybridization; FFPE, formalin-fixed paraffin-embedded; *GAPD*, glyceraldehyde phosphate dehydrogenase; *GRP58*, glucose-regulatory protein 58; *HMGB1*, high-mobility group beta 1; *HMGB2*, high-mobility group beta 2; *HSC70*, heat shock cognate protein 70; PAM, prediction analysis of microarrays; SAM, significance analysis of microarrays; SEOC, serous epithelial ovarian cancer

Address all correspondence to: Jeremy A. Squire, PhD, Ontario Cancer Institute, Princess Margaret Hospital, Room 9-721, 610 University Avenue, Toronto, Ontario, Canada M5G 2M9. E-mail: jeremy.squire@utoronto.ca

[1]This study was supported by a grant award from the Ontario Cancer Research Network (funded through the Government of Ontario), Genome Canada, and the Ovarian Fashion Show Committee (Princess Margaret Hospital). M.B. was supported by the Kristi Pia Memorial Fellowship. The Toronto Ovarian Tissue Bank and Database was funded, in part, by the National Ovarian Cancer Association and the St. George's Society.

[2]Present address: Canadian Breast Cancer Foundation – Ontario Chapter, 790 Bay Street, Suite 1000, Toronto, Ontario, M5G 1N8, Canada.

*This article refers to supplementary material, which is designated by "W" (ie, Table W1, Figure W1) and is available online at www.bcdecker.com.

Received 6 December 2004; Revised 21 March 2005; Accepted 21 March 2005.

DOI 10.1593/neo.04760

marker to assess chemotherapy response is well established (reviewed in Ref. [2]). Both the rate of decline as well as the absolute value of CA 125, determined after the first courses of chemotherapy, are generally considered predictors of the final clinical response [3].

Most investigations of drug resistance in ovarian cancer have used anticancer drugs *in vitro* to select for subclones of cell lines with resistance to the selected agent [4–9]. A disadvantage of these approaches is that the cultured cells used are often genomically unstable and may have acquired *in vitro* genetic and epigenetic alterations that are not representative of *in vivo* conditions. In addition, such models primarily address *acquired* drug resistance, and do not provide direct insights into the expression and genomic alterations associated with *intrinsic* drug resistance.

In recent years, cytogenetic study of solid tumors has been directed toward the identification of recurrent chromosomal rearrangements and patterns of copy number imbalance that may pinpoint genomic regions involved in cancer initiation, progression, drug resistance, and patients' outcome [10,11]. Molecular cytogenetic methods such as spectral karyotyping and comparative genomic hybridization (CGH) have provided useful insights concerning genomic alterations in ovarian cancer [12,13]. However, because metaphase CGH has a resolving power of 10 to 20 Mb [14], it has not been possible to determine genomic imbalance patterns within cytobands. Recently, genomic and cDNA arrays (reviewed in Ref. [15]) have provided more detailed maps of genomic copy number alterations in tumors and, in due course, will provide comprehensive maps of genomic imbalance in a variety of tumors [16–18]. Moreover, high-resolution maps of copy number imbalance are now being integrated with expression profile data to identify clinically relevant subsets of genes based on concomitant alterations at the DNA and RNA levels [19–23]. Microarray expression profiling has been utilized in a number of recent studies in ovarian cancer (reviewed in Ref. [24]). However, no study to date has performed parallel microarray expression and array comparative genomic hybridization (aCGH) analyses to address genomic imbalance and concurrent expression alterations associated with intrinsic drug resistance in ovarian cancer.

Materials and Methods

This study was designed in three phases (Figure 1). In the first phase, a 2- to 4-Mb genomic interval aCGH map of genomic imbalance in 26 serous epithelial ovarian cancer (SEOC) tumors was generated. In the second phase, statistical analysis of aCGH data sets was used to identify cytobands in which imbalance was associated with drug resistance. In the third phase, gene expression profiles were obtained from a subset of tumors, patterns of gene expression associated with drug response were identified, and a concordance analysis of the relationship between genomic imbalance and expression levels was performed. Finally, expression microarray prediction analysis was carried out to identify a subset of classifier genes that could predict chemotherapy response in ovarian cancer patients.

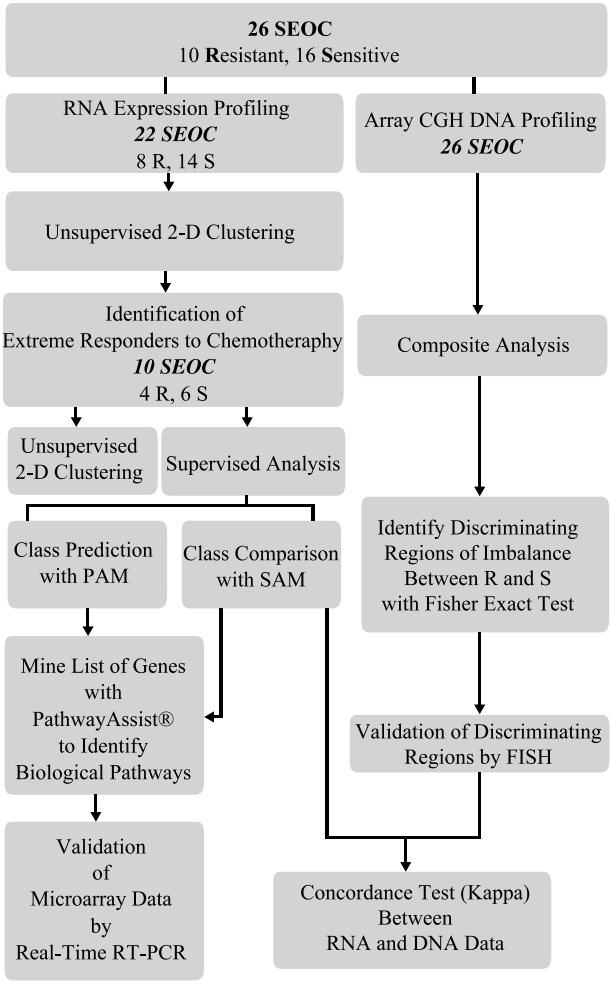


Figure 1. *Flowchart of the experimental design.*

SEOC Tumor Samples

Snap-frozen tumor tissue samples from 26 sporadic SEOC tumors naïve to chemotherapy were selected from the Toronto Ovarian Tissue Bank and Database. No patient included in this study had a family history of either breast or ovarian cancer. All samples were acquired according to the institutional guidelines of the Research Ethics Board. The tumor specimens selected for this study contained at least 75% tumor content as assessed by the surface area of histology slides corresponding to the snap-frozen tissues (the available clinical data are summarized in Table 1). Patients received standard SEOC chemotherapy (carboplatin + taxol). To be classified as sensitive, CA 125 values from patient tumor samples had to meet two criteria. First, the CA 125 values had to fall to below the normal reference ($\sim$35 U/ml) within three cycles of chemotherapy, regardless of the initial baseline. Second, the values had to remain below the normal reference of a period of at least 6 months from the initiation of chemotherapy. Using these criteria within our group of samples, 16 met the criteria for sensitivity and 10 were thus classified as resistant. Due to the accepted variability of CA 125 values, especially in those classified as resistant, a subset of samples was used for a more detailed class comparison. In this group of six sensitive and four resistant samples, the resistant tumors displayed

CA 125 levels that failed to decline below 50% of their original postsurgical values, whereas the selected subset of sensitive samples comparatively displayed the highest rate of decline from initial baseline [3].

Tissue Arrays

A tissue array comprising 1-mm-diameter bores through tumor-rich areas of formalin-fixed paraffin-embedded (FFPE) tumors was designed following published methods [25] and constructed using a standard arraying device (Beecher Instruments, Sun Prairie, WI). Duplicate tissue cores from each donor block were included in the tissue microarray, and sections (5 μm) were cut from the recipient tissue array block for hematoxylin and eosin staining and interphase fluorescence *in situ* hybridization (FISH) analysis.

FISH

Interphase FISH was performed on unstained 5-μm sections from the FFPE tissue array using a commercially available Spectrum Green *RB1* locus probe mapping to cytoband 13q14 (Vysis, Downers Grove, IL) according to the manufacturer's instructions. Slides were imaged using the Vysis Quips SmartCapture (Vysis) imaging system. The scoring criteria used for the interpretation of FISH results on the FFPE sections have been previously described [19]. Chromosomal gains were assigned when more than 10% of the nuclei exhibited more than two signals.

aCGH

Genomic DNA was obtained from all tumor samples using standard phenol chloroform extraction methods. The normal human reference DNA was comprised of an equimolar mixture of DNA derived from multiple male donors (Promega, Madison, WI). The genomic array slides were obtained from Spectral Genomics (Houston, TX) and comprised 1300 large insert clones (BACs/PACs) spaced ~2 to 4 Mb apart. Supplier-recommended protocol was used. In brief, 2 μg each of genomic tumor and normal DNA was directly labeled with Cy3-dCTP or Cy5-dCTP (Amersham, Baie D'Urfe, Canada) using random priming. Following hybridization, the microarrays were washed using 50% formamide/2× SSC (20 minutes), 0.1% Igepal/2× SSC (20 minutes), and 0.2× SSC (10 minutes), all prewarmed to 50°C. A final wash with deionized distilled water was carried out. Air-dried microarray slides were scanned with an Axon GenePix 4000A confocal scanner, and fluorescence intensities were quantified with the GenePix Pro 3.0 software (Axon Instruments, Union City, CA). Hybridizations were carried out in duplicate with fluor reversals to ensure that labeling differences did not affect imbalance assignments. Repetitive spots showing >20% variation in their signal ratio were removed prior to value averaging. Details concerning software, normalization, and imbalance assignments have been described previously [16,18] and are available at http://www.utoronto.ca/cancyto/. The analysis software provides data in two formats: 1) the raw normalized data for each feature on the array, and 2) the feature intensity data represented as significant gain (two baseline standard deviations) and loss per individual experiment. This second output was the primary analysis format used in the aCGH portion of the study. To compensate for possible interexperimental variability, data were normalized to show areas of significant gain or loss in relation to each given experiment. Individual array features were assigned a positive value for significant gain and a negative value for significant loss based on 2 SD from a baseline determined for each individual experiment. The baseline threshold for each experiment is determined by the software using the largest chromosomal region of contiguous clones having the smallest deviation in their intensity ratios. Spots with fluorescence intensity ratios of greater than ±2 SD threshold are assigned as copy number imbalance and given a score of +1 for gain and −1 for loss.

For group comparisons, the differences in $\log_2$ ratios as well as the Fisher exact test were used to determine whether there was any significant gain or loss of genomic content within particular cytobands between resistant and sensitive tumors. The Fisher exact test utilized three categories (gain, loss, and no change), with the null hypothesis that the relative proportions of each of the three imbalance categories would be expected to be the same in both groups. The statistical package S-Plus was used for these group comparisons. We reported uncorrected *P* values and used the permutation-based stepdown method to correct the *P* values for multiple comparisons [26].

Expression Microarrays

RNA was extracted using Trizol (Invitrogen Canada, Burlington, Ontario, Canada). RNA quality and concentration were verified using an Agilent Bioanalyzer (Agilent

Table 1. Patient Sample Information.

Sample Number	Stage	Grade	Surgery	Age	Classification
OVCA 3	IIIB	3	SOPT	N/A	R
OVCA 8	IIIC	2	SOPT	N/A	S
OVCA 33*	IIIB	2	SOPT	N/A	S
OVCA 38	III	2	SOPT	66	R
OVCA 46*	IIIB	3	N/A	N/A	S
OVCA 93	IIIC	3	SOPT	67	S
OVCA 123*	III	3	OPT	59	R
OVCA 130*	III	3	OPT	46	R
OVCA 161*	IIC	3	SOPT	78	S
OVCA 162*	III	3	SOPT	44	S
OVCA 180	IIC	3	OPT	88	S
OVCA 209*	III	1	OPT	46	S
OVCA 237	III	3	SOPT	57	R
OVCA 239	III	3	OPT	63	S
OVCA 249*	III	3	N/A	52	R
OVCA 261	IV	3	SOPT	50	R
OVCA 263	III	3	OPT	44	S
OVCA 304	III	1	SOPT	57	R
OVCA 329	III	3	SOPT	59	S
OVCA 354	III	2	SOPT	65	S
OVCA 363	III	3	SOPT	55	S
OVCA 365	III	3	SOPT	68	R
OVCA 371*	III	3	N/A	N/A	S
OVCA 384	III	3	SOPT	N/A	S
OVCA 390	III	1	N/A	37	S
OVCA 498*	IIIC	3	OPT	52	S

N/A: not available; NC: no change; OPT: optimally debulked; SOPT: suboptimally debulked.
*Extreme responders to chemotherapy.

BioTechnologies, Palo Alto, CA). High-quality RNA was obtained from 22 of 26 tumors. Standard optimized protocols and a full description of the cDNA arrays used in this study can be found at the University Health Network (UHN) Microarray Centre (http://www.microarrays.ca). Ten micrograms of ovarian tumor total RNA or Human Universal Reference (HUR) RNA (Stratagene, La Jolla, CA) was reverse-transcribed with Superscript II reverse transcriptase (Invitrogen Canada) while incorporating Cy3-dCTP or Cy5-dCTP (NEN, Boston, MA). The fluorescently labeled cDNA were cohybridized overnight at 37°C to human 19K UHN microarrays comprising 19,200 sequence-verified cDNA fragments spotted in duplicate. Each of the 22 samples was assayed with dye reversal microarray hybridizations (to control for potential labeling differences) for a total of 44 hybridizations. Microarrays were scanned by a confocal laser reader (ScanArray 4000; Packard BioScience, Meriden, CT) after stringent washes. Quantification was carried out using GenePix Pro 3.0 (Axon Instruments). Low-quality spots were filtered using GenePix Pro 3.0 and by visual examination of the images. Paired dye reversal correlations were examined by unsupervised cluster analysis. Samples displaying correlation less than .65 were repeated (data not shown).

Expression Microarray Data Analysis

Data warehousing, filtering, and normalization were performed using the GeneTraffic software (version 2.7, Iobion; Stratagene). Hybridizations were annotated according to the Minimum Information About a Microarray Experiment (MIAME) guidelines (http://www.mged.org/Workgroups/MIAME/miame.html and Ref. [27]). 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 one of the two channels, spots that had an intensity-to-background ratio of less than 2, and spots whose raw intensity was less than 500. Locally Weighted Scatter Plot Smoother (LOWESS) normalization by subarray (for background, see http://www.stat.berkeley.edu/users/terry/zarray/Html/normspie.html and GeneTraffic 2.7 Manual, Iobion; Stratagene) was used for normalization between the arrays. Expression array data are available at http://www.utoronto.ca/cancyto/OVCA2004NEO/.

Unsupervised two-dimensional hierarchical clustering was carried out as described in Ref. [28], using Cluster 2.01 available at http://rana.lbl.gov/EisenSoftware.htm, on gene expression values that were present in at least 80% of the tumors (10,806 genes). Gene expression ratios were median-centered across all samples and arrays before agglomerative average linkage clustering using uncentered Pearson correlation. All observations for a given item were weighted equally. Clustering results were visualized using the Treeview software available at http://rana.lbl.gov/EisenSoftware.htm. Significance analysis of microarrays (SAM) (Ref. [29] and http://www-stat.stanford.edu/~tibs/PAM/) and prediction analysis of microarrays (PAM) (Ref. [30] and http://www-stat.stanford.edu/~tibs/PAM/) were performed using the software available and published methods.

Validation by Real-Time Reverse Transcription Polymerase Chain Reaction (RT-PCR)

Two micrograms of total RNA from six sensitive and four resistant ovarian tumors classified as extreme responders according to the patients' CA 125 profiles was reverse-transcribed in a 100-μl reaction mixture comprising 5.5 mM $MgC1_2$, 500 μM of each dNTP, 2.5 μM random hexamers, 0.4 U/μl RNase inhibitor, and 3.125 U/μl MultiScribe Reverse Transcriptase (Applied Biosystems, Foster City, CA) under the following conditions: 25°C for 10 minutes, 48°C for 30 minutes, and 95°C for 5 minutes. 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. A subset of genes was chosen for validation using commercially available Assays-on-Demand probe primer sets with provided master mix (Applied Biosystems). The following PCR conditions were used: 50°C for 2 minutes, 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Human cyclophilin A was used as an endogenous control because it resulted in minimum variation throughout the samples and has been previously used to validate cancer microarray expression data by real-time RT-PCR [31]. The initial copy numbers of RNA targets can be quantified using real-time PCR analysis based on threshold cycle (C_t) determinations. C_t is defined as the cycle at which a statistically significant increase in fluorescence (above background signal contributed by the fluorescence-labeled oligonucleotides within the PCR reaction) is detected. The threshold cycle is inversely proportional to the log of the initial copy number. The C_t value of human cyclophilin A was subtracted from each C_t value of OVCA or HUR sample for normalization and the ratio of OVCA tumor:HUR RNA expression was calculated so that real-time RT-PCR and microarray data could be compared.

Results

In this study, aCGH analysis improved the resolution of such regions including 9q21.11-q33.1 and 11p15.1-pter, as well as identified novel regions of loss at 15q11.2-q15.1 and 17q21.32-q21.33 that have not been reported in ovarian cancer. When imbalance profiles of resistant tumors were compared to sensitive tumors, 13 regions of the genome were strongly associated with differentiating responses. Parallel expression analysis by cDNA microarrays revealed a nuclear complex comprised of GAPD, HMGB2, HSC70, GRP58, and HMGB1 whose RNA levels were lower in the resistant tumors in comparison to the sensitive group.

Overall aCGH Analysis of 26 SEOC Tumors

The patterns of genomic imbalances of DNA overrepresentation and underrepresentation at 2- to 4-Mb intervals in 26 SEOC tumors were identified by aCGH (Figure 2). All imbalance data from individual aCGH profiles of each tumor are published as supporting information at http://www.utoronto.ca/cancyto/OVCA2004NEO/. Losses at 1p, 4q, 6q, 8p, 9q, 13q, 16q, 17p, and 18q were present. Gains at

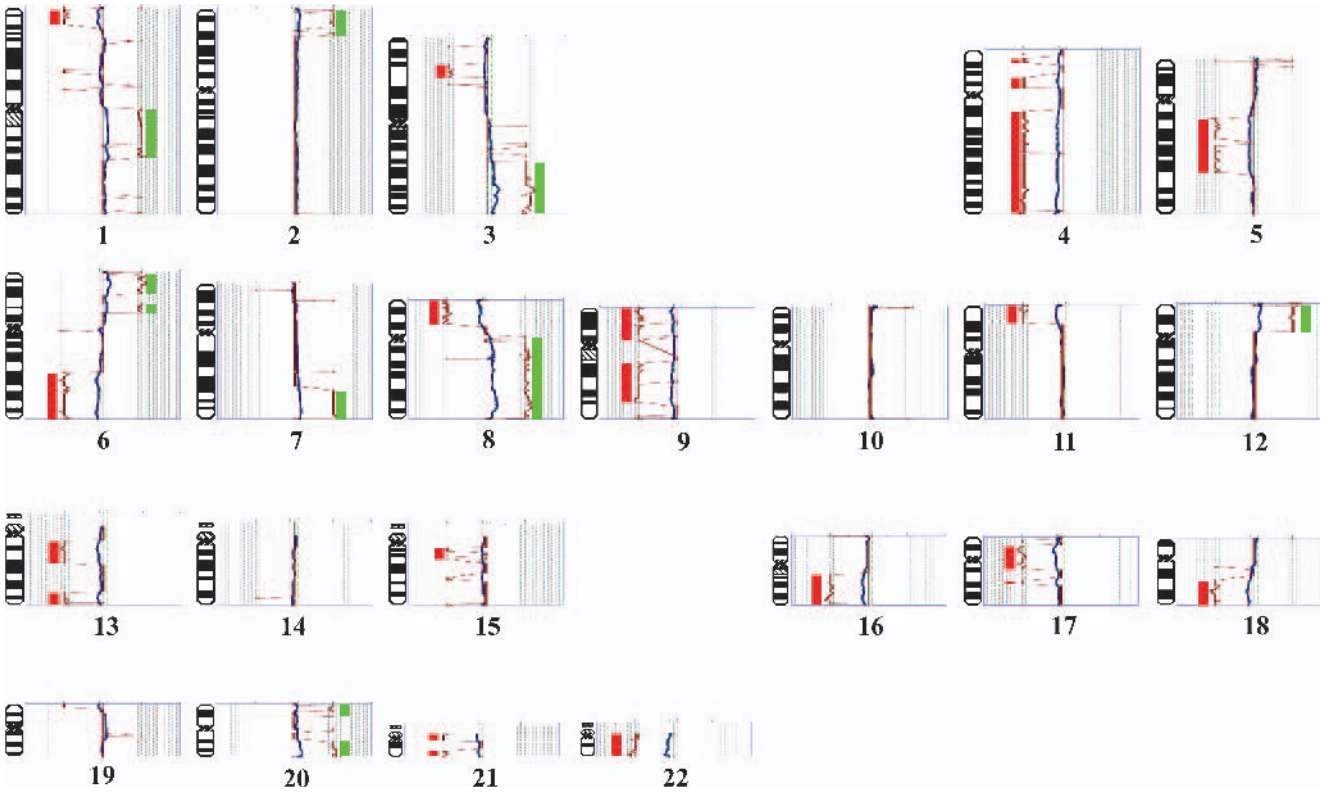


Figure 2. *Summary of all aCGH findings using 26 SEOC samples. Overall gains and losses as determined by mean values for individual features are shown to the right of each chromosome ideogram as green and red bars, respectively. In this analysis, closely linked BACs that consistently exhibited fluorescence intensities that deviated by 2 SD or more were used to generate the average imbalance profile (traced in red).*

1q, 3q, 8q, 12p, and 20q were also detected, but no focal high copy number gene amplification was evident within this study group. To validate the imbalances detected at 13q14 by aCGH, interphase FISH analysis was performed using a 13q14-specific probe (*RB1* gene) (data not shown). In 17 of 23 samples studied by interphase FISH, imbalances were in agreement with aCGH. For three samples, alterations in ploidy levels or cellular heterogeneity within tissue sections were identified. The remaining three samples could not be scored as a result of poor signal intensity.

DNA Copy Number Changes Associated with Differential Treatment Response

The samples were divided into sensitive and resistant groups, as described earlier, for a more detailed analysis of the patterns of genomic imbalances associated with differential responses to chemotherapy (Figure 3). Based on the number of BAC clones subject to imbalance, in comparison to the total number of clones analyzed, a group analysis of the overall percentage of the genome altered indicated that the resistant group had an increased level of genomic imbalance (8.3% gain, 2.4% loss) compared to the sensitive group (4.4% gain, 1.1% loss) (data not shown). Moreover, consistent with this elevated percentage of genomic change, twice as many imbalances were identified in the resistant group (55 losses/gains) compared to the sensitive group (28 losses/gains). The Fisher exact test was used to compare the resistant and sensitive groups in three categories (gain, loss, and no change) to determine which contiguous genomic regions were statistically concordant with differential treatment response (Table 2, Figure 3). Three particular regions of imbalance are identified as 13q12.2-13q13, 1p36.33, and 17q11.2.

Microarray Expression Profiling

Unsupervised two-dimensional hierarchical clustering (Figure W1) was performed using expression data derived from 22 of 26 SEOC tumor cohorts classified as sensitive or resistant based on their CA 125 patterns. The clustering pattern observed did not clearly separate tumors based on response to chemotherapy and, consistent with these findings, supervised analysis using SAM only identified a limited number of genes differentially expressed in this group comparison (data not shown). A subset of 10 tumor samples was then selected from patients exhibiting the most extreme differences in CA 125 response. Unsupervised two-dimensional hierarchical clustering (Figure 4) clearly stratified this subset into a resistant and a sensitive group. A cluster of 1301 genes (highlighted in yellow) largely overlapped with those of a similar-sized gene cluster apparent in the hierarchical clustering performed on the complete sample cohort (highlighted in yellow in Figure W1).

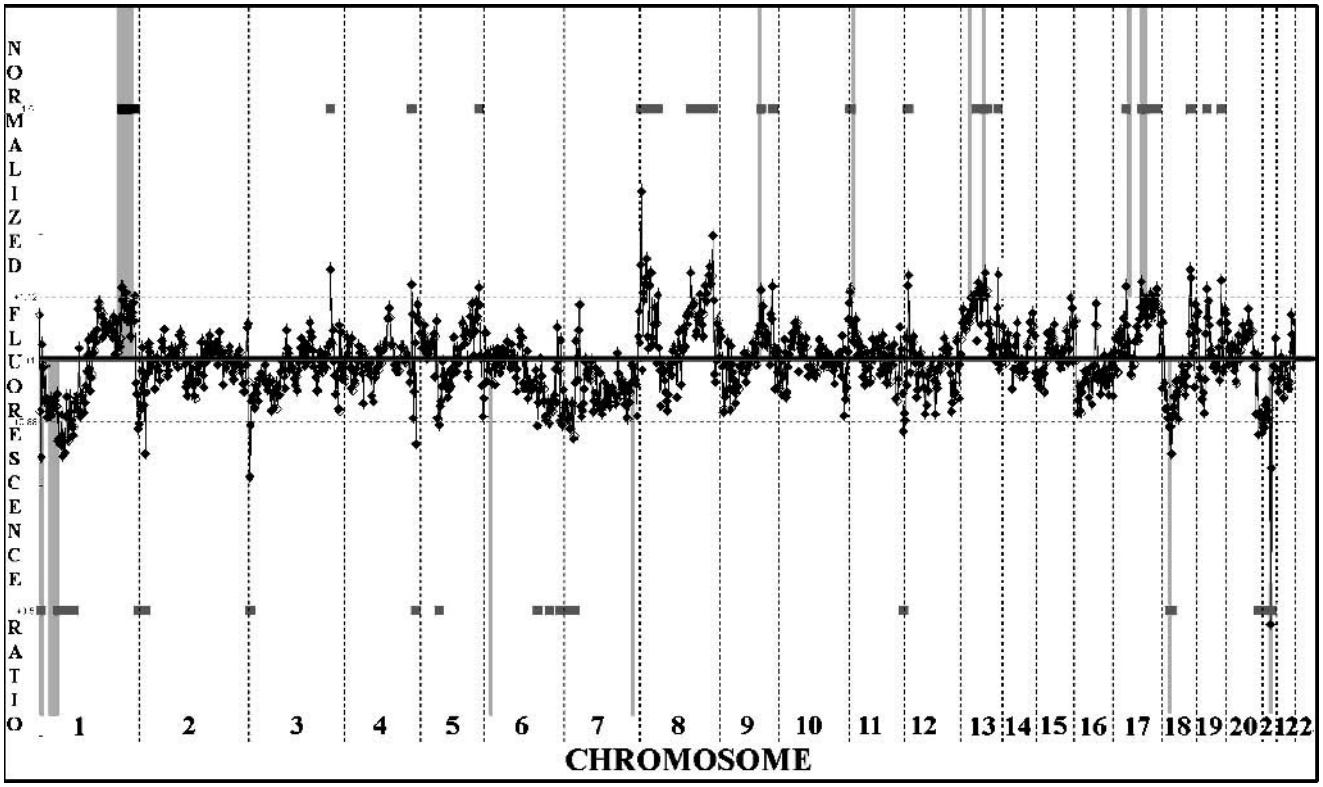


Figure 3. *Whole genome plot of the relative difference in normalized log_2 average ratios between the 10 resistant and 16 sensitive samples. Positive (above baseline) and negative (below baseline) deflections of the profile indicate the mean overrepresentation and underrepresentation of the region in resistant versus sensitive samples, respectively. The vertical pale gray bars correspond to the contiguous array features differentially identified using the Fisher exact test (see Table 2), with bars above the profile indicating overrepresentation and bars below indicating underrepresentation. Horizontal dark gray bars above (gain) and below (loss) the profile highlight the features with fluorescence intensity ratios greater than threshold.*

Identification of Differentially Expressed Genes Associated with Treatment Response

The large discriminating gene cluster identified in Figure 4 contains a majority of the statistically significant expression changers identified by SAM analysis, and this large cluster contains child nodes with a preponderance of genes involved in: 1) nucleus/DNA binding; 2) nucleus/metal ion binding; 3) cell cycle/cyclin–dependent protein kinase; 4)

Table 2. Fisher Exact Test: Differential Cytoband Regions Between Resistant and Sensitive Groups.

Cytoband	Mb Size (Mb Range)	*P* Range*	Gene of Interest within Cytoband (Cytoband)†	Relative Copy Nb Change (R/S)‡	Copy Nb-Resistant *Versus* Normal§	Copy Nb-Sensitive *Versus* Normal§
1p36.33	0.515 (1.164–1.679)	0.049–0.056	TP73 (1p36.3)	U	–	NC
1p36.13	1.448 (19.705–18.257)	0.0009–0.04		U	–	NC
1q42-44	20.518 (222.771–243.289)	0.001–0.06	LGALS8 (1q43)	O	+	NC
6p22.1-p21.2	11.491(28.342–39.833)	0.019–0.09	BAK1 (6p21.31)	U	NC	+
7q32.1-q34	11.873 (126.656–138.529)	0.012–0.029	BRAF (7q34)	U	NC	+
9q33.3-q34.3	13.285 (122.613–135.898)	0.006–0.09		O	NC	–
11p15.2	1.559 (14.206–15.765)	0.043–0.043	RRAS2 (11p15.2)	O	NC	–
13q12.2-q13.1	5.22 (25.719–30.939)	0.004–0.0062	BRCA2 (13q12.3)	O	NC	–
13q21.31	0.385 (61.538–61.923)	0.025–0.046		O	NC	–
17q11.2	1.421 (30.848–32.269)	0.017–0.059	NF1 (17q11.2)	O	NC	–
17q24.2-q25.3	13.974 (64.589–78.563)	0.021–0.15	BIRC5 (17q25.3)	O	NC	–
18q12.2	1.143 (31.407–32.550)	0.011–0.043		U	–	NC
21q21.2-q21.3	3.982 (23.885–27.867)	0.0085–0.080		U	–	NC

(+) Gain in copy number when compared to normal; (–) loss in copy number when compared to normal.

*Uncorrected *P* values from the Fisher exact test for the features representing the region.

†BAK1: BCL2-like 7 protein cell death inhibitor 1 apoptosis regulator BAK Bcl-2 homologous antagonist/killer; BIRC5: baculoviral IAP repeat-containing protein 5, survivin; BRAF: v-raf murine sarcoma viral oncogene homolog B1; BRCA2: breast cancer 2, early onset; LGALS8: lectin, galactoside-binding, soluble, 8 (galectin 8); NF1: neurofibromin 1 (neurofibromatosis, von Recklinghausen disease, and Watson disease); RRAS2: related RAS viral (r-ras) oncogene homolog 2; TP73: tumor protein p73.

‡U: underrepresented copies in resistant compared to the sensitive group; O: overrepresented copies in the resistant compared to the sensitive group.

§NC: no change in copy number when compared to normal.

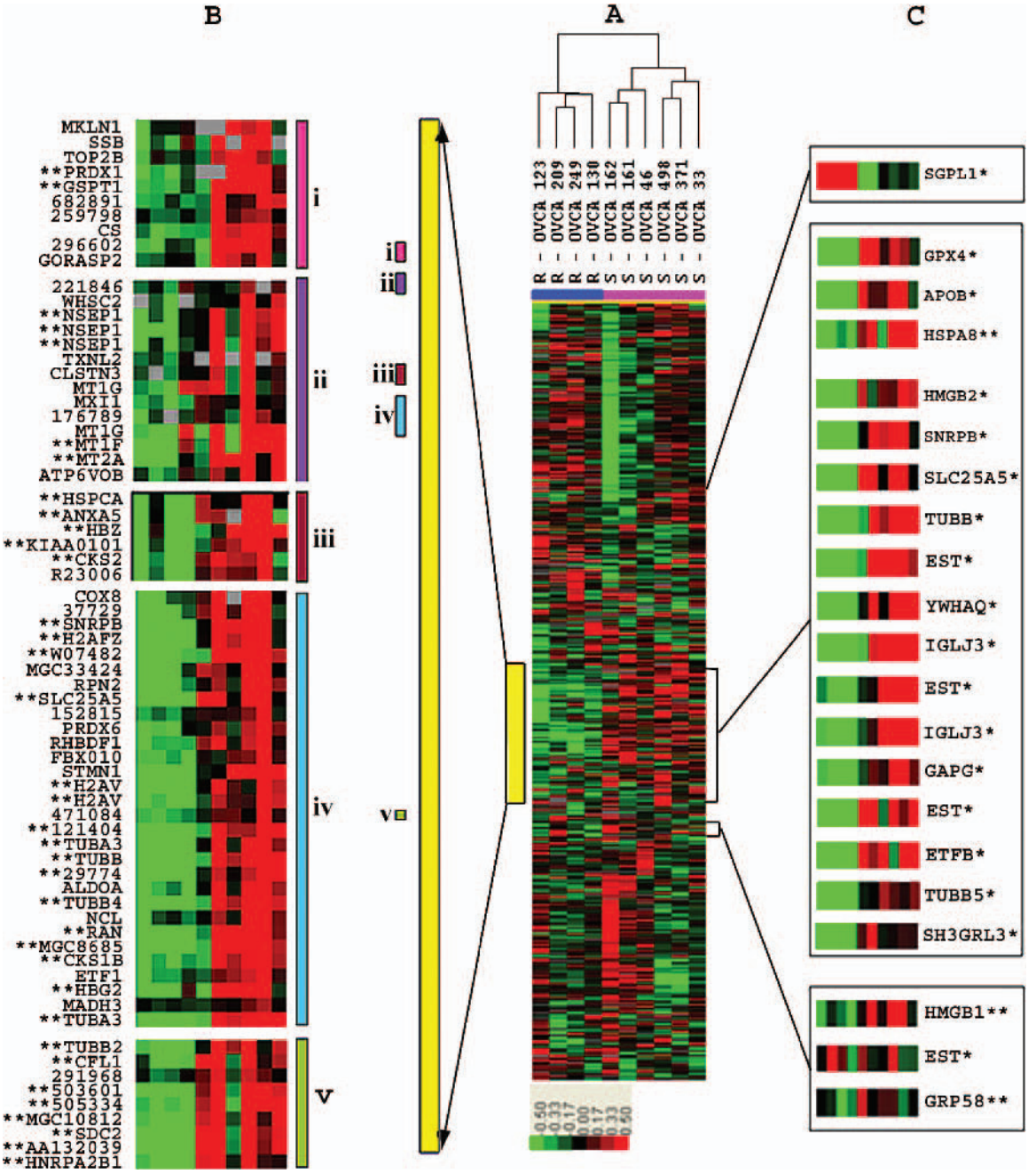


Figure 4. *Analysis of the 10 extreme responders using unsupervised hierarchical clustering. (A) The relative expression patterns of genes that are color-coded in red (up), green (down), black (no change), or grey (data missing) clearly stratifies the sensitive (S; pink) and resistant (R; blue) samples into two major nodes. A major gene cluster that includes most genes identified by SAM analysis is highlighted with a vertical yellow bar. (B) A magnified view of subclusters that include large numbers of genes belonging to functional categories of 1) nucleus/DNA binding; 2) nucleus/metal ion binding; 3) cell cycle/cyclin–dependent protein kinase; 4) microtubule/cytoskeleton; or 5) nucleus/actin cytoskeleton engineering. (C) Relative position and expression patterns of genes identified by PAM analysis. Genes identified by PAM analysis are indicated by an asterisk (*). Genes identified only by SAM analysis are indicated by two asterisks (**).*

microtubule/cytoskeleton; and 5) nucleus/actin cytoskeleton engineering.

Class comparison by statistical supervised analysis using SAM identified 173 clones (corresponding to 123 unique identified genes), which were statistically differentially expressed between the two classes with a fold change (FC) difference of at least 2 and a false discovery rate (FDR) <1%. The complete list of differentially expressed genes identified is available in supplementary materials (Table W1).

Microarray Prediction Analysis

We employed the "nearest shrunken centroid" methodology [30] that employs leave-one-out cross validation (LOOCV) to identify the most relevant classifier genes capable of predicting chemotherapy resistance in SEOC patients. When applied to the present expression data set from the 10 extreme cases, the PAM algorithm identified a set of 22 genes and ESTs that could predict with 100% accuracy the class of the test sample during LOOCV on this patient

sample set. The complete list of these clones is presented in Table 3.

Data Mining

The possible functional roles of 15 of 22 clones were examined. Seven clones could not be further analyzed because ambiguity in their DNA sequence prevented their proper annotation. By applying a data mining software (PathwayAssist, Iobion Informatics; Stratagene) to the list of 22 clones identified by PAM, a nuclear complex comprised of two predictive genes identified by PAM (*GAPD* and *HMGB2*) and three additional genes identified by SAM (*HSC70*, *GRP58*, and *HMGB1*) was revealed. This nuclear complex was previously reported as being involved in resistance to DNA conformation–altering chemotherapeutic drugs [32].

To validate the expression findings derived from microarray analysis, the levels of expression of *GAPD*, *HMGB2*, *HSC70*, *GRP58*, and *HMGB1* were measured by real-time RT-PCR. The results obtained are shown in Figure W2 and both real-time RT-PCR and microarray data agree in the direction of expression (i.e., up or down). All five genes displayed lower expression levels in resistant samples than in sensitive samples.

Correlation between Overall Pattern Gene Expression and DNA Copy Number

The level of agreement between expression and copy number changes was tested with the simple κ coefficient (Table W2). Overall comparison between expression level and aCGH copy numbers, including the no changers, showed 91% agreement. If expressed genes identified as differential in the two groups were solely considered, 8.2% of the genes agreed with copy number differences. When corrected for chance agreement for the two methods by Cohen's κ, a slight agreement between the two data sets ($\kappa\text{^} = 0.0230$; $0 < \kappa\text{^} < 1$) was indicated but was not statistically significant ($\kappa\text{^*} = 0.1316$; $\kappa\text{^*} < 1.96$).

Discussion

In keeping with classic cytogenetic studies and metaphase-based CGH findings, the imbalance profiles identified in this study are complex and characterized by low-level gains and losses that affect all chromosomes except chromosome 10. The consensus pattern of imbalance was, in general, consistent with findings previously reported in ovarian cancer by metaphase CGH (reviewed in Ref. [33] and available from CGH databases: http://www.ncbi.nlm.nih.gov/; http://www.progenetix.net; http://amba.charite.de/~ksch/cghdatabase/index.htm; and http://www.helsinki.fi/cmg/cgh_data.htm). aCGH permitted identification of cytoband imbalances (Table W3) within the larger genomic intervals identified in published metaphase CGH studies [34,35] and allelic imbalance findings [36,37]. For example, losses at distal 1p have been consistently reported, but aCGH localizes the minimal region of consistent loss to 1p36.11-pter in 18/26 tumors. Similarly, the recurrent gain at 6p was localized to two intervals at 6p21.1-p21.31 and 6p22.1-pter. Moreover, aCGH identified small focal genomic imbalances not previously detected by metaphase CGH, such as losses at 15q11.2-q15.1

Table 3. List of Genes Identified by PAM Analysis That Discriminate between Resistant and Sensitive Groups.

Clone ID or Accession Number*	Fold Change†	Name	Symbol	Cytoband
Up‡				
BM976649	2.01892	Sphingosine-1-phosphate lyase 1	SGPL1	10q21
Down§				
162892	10.39	Highly similar to Ig lambda chain C regions	N/A	N/A
182721	6.24	Immunoglobulin lambda joining 3	IGLJ3	22q11.1-q11.2
135961	4.99	Solute carrier family 25	SLC25A5	Xq24-q26
5922013	4.84	Tubulin, beta polypeptide	TUBB	6p25
300017	4.18	Apolipoprotein B	APOB	2p24-p23
4876644	4.08	Tubulin, beta 5	TUBB5	19p13.3
261822	3.79	*Homo sapiens*, clone IMAGE: 5728597	N/A	22q11.23
AL536766	3.72	Tubulin, beta polypeptide paralog	MGC8685	6p25
5806720	3.31	Glyceraldehyde-3-phosphate dehydrogenase	GAPD	12p13
5740157	3.26	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein	YWHAQ	2p25.2-p25.1
5556148	3.23	Small nuclear ribonucleoprotein polypeptides B and B1	SNRPB	20p13
504180	2.75	*H. sapiens*, clone IMAGE: 5742072, mRNA	N/A	N/A
267145	2.74	High-mobility group box 2	HMGB2	4q31
471144	2.64	Electron transfer flavoprotein, beta polypeptide	ETFB	19q13.3
270849	2.48	*H. sapiens*–transcribed sequences	N/A	N/A
AL558551	2.48	Glutathione peroxidase 4 (phospholipid hydroperoxidase)	GPX4	19p13.3
5932330	2.2	SH3 domain binding glutamic acid-rich protein like 3	SH3BGRL3	1p35-p34.3

N/A: not available.
*Four of 22 PAM clones are not listed above because they were redundant or not sequence-verified.
†Relative fold change between the two classes (resistant/sensitive).
‡List of genes whose transcript levels are greater in the resistant group compared to the sensitive group.
§List of genes whose transcript levels are lower in the resistant group compared to the sensitive group.

and 17q21.32-q21.33. Imbalances such as gains at 1q11-q25.3, 3q22.3-qter, 8q11-qter, and 20q12-qter have been previously detected in a wide variety of epithelial tumors [33] including ovarian cancer, but improved resolution has been obtained in other areas including losses at 9q21.11-q33.1 and 11p15.1-pter.

Resistant tumors were found to have twice as much genomic imbalance as sensitive tumors. These data suggest that the resistant group of SEOC tumors would have a greater capacity to adapt to the selective pressures of chemotherapy by virtue of their elevated frequency of genomic rearrangement in comparison to the sensitive group.

The interval 13q12.2-13q13.1, which comprises ~400 kb of DNA, contains the *BRCA2* gene (13q12.3), a tumor-suppressor gene known to be mutated in a high percentage of hereditary ovarian cancers [38]. This region was subject to loss in 72% of the 16 sensitive tumors. It is possible that acquired loss of *BRCA2* and cognate cellular repair functions could enhance susceptibility to chemotherapy. In this context, Kudoh et al. [34] investigated ovarian tumors resistant or sensitive to chemotherapy using metaphase CGH, and found that the region 13q12-14 was more often gained in the resistant ovarian cancer in comparison to sensitive tumors.

The ~500-kb region, 1p36.33, was found to be underrepresented in resistant tumors relative to sensitive tumors (Table 2). This small region of chromosome 1 contains the *TP73* gene, and it has recently been reported that deregulation and overexpression of specific p73 isoforms are associated with reduced overall survival in ovarian cancer [39]. By analogy with p53, it is plausible that genomic imbalance (loss or gain) of 1p36.33 may alter the spectrum of *TP73* isoforms and influence treatment response. Similarly, the *NF1* gene situated in 17q11.2 has also been studied previously in ovarian cancer cell lines [40]. The authors of this study reported that overexpression and imbalance of type II and type I isoforms of *NF1* were associated with differentiation arrest and, potentially, treatment response. The observation of overrepresentation of 17q11.2 in resistant tumors may implicate a role for *NF1* in chemotherapy resistance.

Unsupervised two-dimensional hierarchical clustering (Figure 4) analyses of expression microarray data using a subset of 10 tumor samples exhibiting the most extreme differences in CA 125 response revealed a large cluster of 1301 genes (highlighted in yellow) that was particularly important in the segregation of these two groups. Interestingly, the genes in this cluster largely overlap with those of a similar-sized gene cluster apparent in the hierarchical clustering performed on the complete sample cohort (highlighted in yellow in Figure W1). However, it is likely that, due to the relative heterogeneity of the 22 samples, the influence of this cluster was insufficient in completely stratifying the two groups.

Class comparison by SAM identified 173 clones, which included 123 unique identified genes (Table W1). Importantly, there was a significant difference ($P < .05$) in the proportion of genes involved in DNA binding, regulation of transcription, cell cycle and growth, and metal ion binding between the 173 differentially expressed genes and the gene population on the H19K microarray (data not shown). An in-depth analysis of the 123 genes differentially expressed between sensitive samples and resistant samples is beyond the scope of this study. We have focused instead on the subset of these genes that can *predict* chemotherapy response as described below.

We employed the "nearest shrunken centroid" methodology [30] to identify the most relevant classifier genes capable of predicting chemotherapy resistance in SEOC patients. This approach was used by Tibshirani et al. [30] on microarray data obtained by Khan et al. [41] on blue cell tumors of childhood and by Golub et al. [42] on leukemia. Tibshirani et al. demonstrated that this method was superior to both a neural network method and to the approach taken by Golub et al. [42]. The complete methodology of PAM is described in details in Ref. [30] and uses a LOOCV, a strategy particularly useful when identifying classifiers genes in smaller sample sizes [43]. When applied to the present expression data set from the 10 extreme cases, the PAM algorithm identified a set of 22 genes and ESTs that could predict with 100% accuracy the class of the test sample during LOOCV. The complete list of these clones is presented in Table 3.

Of the 22 clones identified by PAM analysis, possible functional roles for 15 of the clones were examined (seven had ambiguous DNA sequence, which prevented their proper annotation). β-Tubulin subtypes accounted for 3 of 15 discriminating identified genes differentially expressed between the sensitive and resistance groups. This is of interest because taxanes are thought to function by stabilizing microtubules—a process that eventually leads to apoptosis. Previous studies have indicated that resistance to taxanes may be a consequence of altering the relative amount of the various subtypes of β-tubulin, thereby decreasing the efficiency of microtubule stabilization by taxol (reviewed in Ref. [44]). In these studies, altered levels of the different β-tubulin isotypes were observed as a consequence of acquiring resistance to chemotherapy. Because the SEOC tumors in the present study were naïve to chemotherapy, the results presented here suggest that, before chemotherapy, some patients may already express varying levels of β-tubulin isotypes, which could result in differential response to taxol.

With PathwayAssist, a software that identifies links between the user's genes of interest based on mining PubMed's abstracts and public biologic databases, a nuclear complex comprised of two predictive genes identified by PAM (*GAPD* and *HMGB2*) and three additional genes (*HSC70*, *GRP58*, and *HMGB1*) was revealed. This nuclear complex was previously reported as being involved in resistance to DNA conformation–altering chemotherapeutic drugs [32]. In contrast, Sugimura et al. [45] reported increased *in vitro* expression of *HSC70* in a human ovarian adenocarcinoma cell line rendered resistant to paclitaxel. Concordant with our results, Vargas-Roig et al. [46] reported decreased *HSC70* levels in breast cancer patient tumors resistant to DNA-targeting drugs. In addition, it was shown that the potency of platinum-related drugs could be increased by inducing *HMGB1* transcription [47].

Although *HSC70*, *GRP58*, and *HMGB1* were not identified by PAM as *predictive* of drug resistance, *HSC70* was identified by SAM as significantly differentially expressed between resistant and sensitive samples (Table W1). *GRP58* and *HMGB1* were also found to be differentially expressed between the two groups, but with FCs of 1.7 and 1.3, respectively, and with a slightly higher FDR (5–10%). The present finding, that all five genes are significantly downregulated in patients resistant to chemotherapy, strongly suggests their involvement in ovarian cancer resistance to cisplatin-related drugs. Although overall comparison between expression level and aCGH copy numbers, including the no changers, showed 91% agreement, when expressed genes identified as differential in the two groups were solely considered, 8.2% of the genes agreed with copy number differences. This level of agreement was not statistically significant ($\kappa\text{^*} = 0.1316$; $\kappa\text{^*} < 1.96$) by Cohen's κ method. Because the general distribution of the clones on the H19K arrays was not significantly different from the distribution of genes in the genome (based on Build 34b, version 2), lack of concordance cannot be due to poor representation of the genome in the H19K clone set. Previous studies that examined concordance between expression changes and genomic imbalance have been conflicting [23,48,49], eluding to varying degrees of epigenetic regulation in different cancers. The class comparison presented in this report was applied to a homogeneous group of advanced stage SEOC tissues that were prospectively identified as resistant or sensitive to chemotherapy. Although a subset within this study group identified a set of genes predictive for extreme nonresponsiveness, this same subset did not provide any further information at the DNA level. It is possible that partial chemotherapy response may be determined based on copy number differences at the DNA level. However, extreme nonresponsiveness may be mediated by different processes more dependent on RNA expression.

In conclusion, these data illustrate the value of molecular profiling at both the RNA and DNA levels to identify small genomic regions, and gene subsets that could be associated with differential chemotherapy response in ovarian cancer.

Acknowledgements

We are grateful to Amit Oza for helpful discussions and consultation regarding the interpretation of CA 125 responses. We thank all patients for their participation in the study.

References

[1] McGuire WP III and Markman M (2003). Primary ovarian cancer chemotherapy: current standards of care. *Br J Cancer* **89** (Suppl 3), S3–S8.

[2] Verheijen RH, von Mensdorff-Pouilly S, van Kamp GJ, and Kenemans P (1999). CA 125: fundamental and clinical aspects. *Semin Cancer Biol* **9,** 117–124.

[3] Rustin GJ (2003). Use of CA-125 to assess response to new agents in ovarian cancer trials. *J Clin Oncol* **21,** 187–193.

[4] Yamamoto K, Kikuchi Y, Kudoh K, Hirata J, Kita T, and Nagata I (2000). Treatment with paclitaxel alone rather than combination with paclitaxel and cisplatin may be selective for cisplatin-resistant ovarian carcinoma. *Jpn J Clin Oncol* **30,** 446–449.

[5] Yamamoto K, Kikuchi Y, Kudoh K, and Nagata I (2000). Modulation of cisplatin sensitivity by taxol in cisplatin-sensitive and -resistant human ovarian carcinoma cell lines. *J Cancer Res Clin Oncol* **126,** 168–172.

[6] Maubant S, Cruet-Hennequart S, Poulain L, Carreiras F, Sichel F, Luis J, Staedel C, and Gauduchon P (2002). Altered adhesion properties and alpha$_v$ integrin expression in a cisplatin-resistant human ovarian carcinoma cell line. *Int J Cancer* **97,** 186–194.

[7] Rogers PM, Beale PJ, Al-Moundhri M, Boxall F, Patterson L, Valenti M, Raynaud F, Hobbs S, Johnston S, and Kelland LR (2002). Overexpression of BclXL in a human ovarian carcinoma cell line: paradoxic effects on chemosensitivity *in vitro versus in vivo*. *Int J Cancer* **97,** 858–863.

[8] Kumar A, Soprano DR, and Parekh HK (2001). Cross-resistance to the synthetic retinoid CD437 in a paclitaxel-resistant human ovarian carcinoma cell line is independent of the overexpression of retinoic acid receptor-gamma. *Cancer Res* **61,** 7552–7555.

[9] Yasui K, Mihara S, Zhao C, Okamoto H, Saito-Ohara F, Tomida A, Funato T, Yokomizo A, Naito S, Imoto I, et al. (2004). Alteration in copy numbers of genes as a mechanism for acquired drug resistance. *Cancer Res* **64,** 1403–1410.

[10] Taetle R, Aickin M, Yang JM, Panda L, Emerson J, Roe D, Adair L, Thompson F, Liu Y, Wisner L, et al. (1999). Chromosome abnormalities in ovarian adenocarcinoma: I. Nonrandom chromosome abnormalities from 244 cases. *Genes Chromosomes Cancer* **25,** 290–300.

[11] Mertens F, Johansson B, Hoglund M, and Mitelman F (1997). Chromosomal imbalance maps of malignant solid tumors: a cytogenetic survey of 3185 neoplasms. *Cancer Res* **57,** 2765–2780.

[12] Bayani J, Brenton JD, Macgregor PF, Beheshti B, Albert M, Nallainathan D, Karaskova J, Rosen B, Murphy J, Laframboise S, et al. (2002). Parallel analysis of sporadic primary ovarian carcinomas by spectral karyotyping, comparative genomic hybridization, and expression microarrays. *Cancer Res* **62,** 3466–3476.

[13] Bernardini M, Weberpals J, and Squire JA (2004). The use of cytogenetics in understanding ovarian cancer. *Biomed Pharmacother* **58,** 17–23.

[14] Parente F, Gaudray P, Carle GF, and Turc-Carel C (1997). Experimental assessment of the detection limit of genomic amplification by comparative genomic hybridization CGH. *Cytogenet Cell Genet* **78,** 65–68.

[15] Mantripragada KK, Buckley PG, de Stahl TD, and Dumanski JP (2004). Genomic microarrays in the spotlight. *Trends Genet* **20,** 87–94.

[16] Beheshti B, Braude I, Marrano P, Thorner P, Zielenska M, and Squire JA (2003). Chromosomal localization of DNA amplifications in neuroblastoma tumors using cDNA microarray comparative genomic hybridization. *Neoplasia* **5,** 53–62.

[17] Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, and Brown PO (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* **23,** 41–46.

[18] Squire JA, Pei J, Marrano P, Beheshti B, Bayani J, Lim G, Moldovan L, and Zielenska M (2003). High-resolution mapping of amplifications and deletions in pediatric osteosarcoma by use of CGH analysis of cDNA microarrays. *Genes Chromosomes Cancer* **38,** 215–225.

[19] Squire JA, Bayani J, Luk C, Unwin L, Tokunaga J, MacMillan C, Irish J, Brown D, Gullane P, and Kamel-Reid S (2002). Molecular cytogenetic analysis of head and neck squamous cell carcinoma: by comparative genomic hybridization, spectral karyotyping, and expression array analysis. *Head Neck* **24,** 874–887.

[20] Heiskanen M, Kononen J, Barlund M, Torhorst J, Sauter G, Kallioniemi A, and Kallioniemi O (2001). CGH, cDNA and tissue microarray analyses implicate FGFR2 amplification in a small subset of breast tumors. *Anal Cell Pathol* **22,** 229–234.

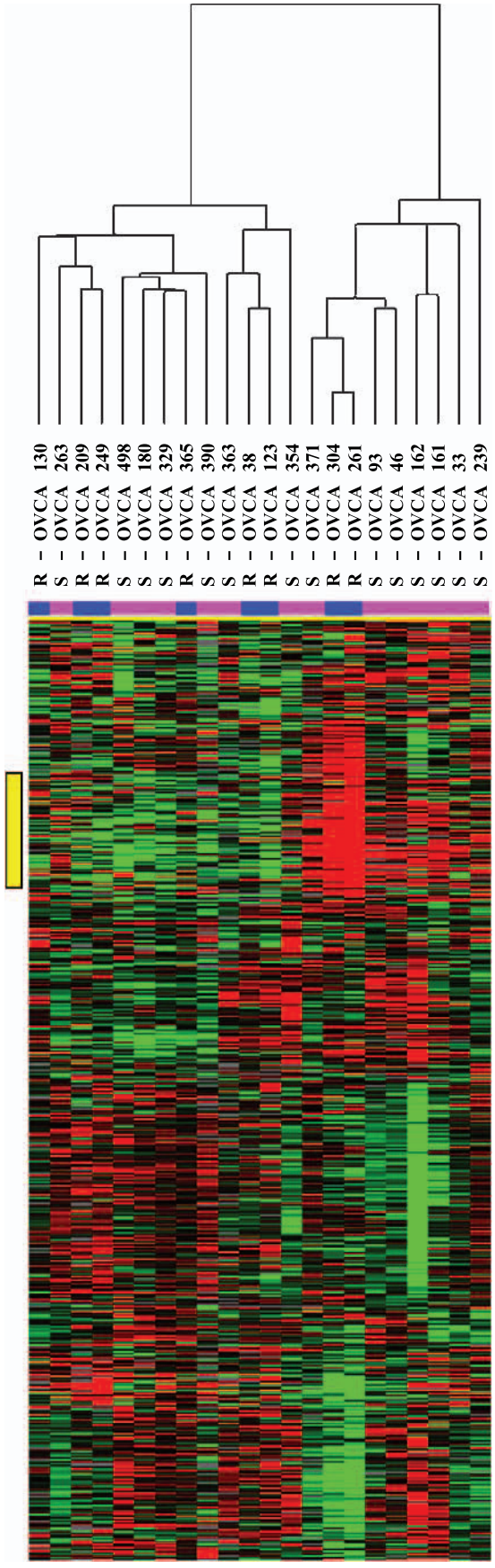
[21] Fritz B, Schubert F, Wrobel G, Schwaenen C, Wessendorf S, Nessling M, Korz C, Rieker RJ, Montgomery K, Kucherlapati R, et al. (2002). Microarray-based copy number and expression profiling in dedifferentiated and pleomorphic liposarcoma. *Cancer Res* **62,** 2993–2998.

[22] Clark J, Edwards S, John M, Flohr P, Gordon T, Maillard K, Giddings I, Brown C, Bagherzadeh A, Campbell C, et al. (2002). Identification of amplified and expressed genes in breast cancer by comparative hybridization onto microarrays of randomly selected cDNA clones. *Genes Chromosomes Cancer* **34,** 104–114.

[23] Pollack JR, Sorlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE, Tibshirani R, Botstein D, Borresen-Dale AL, and Brown PO (2002). Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci USA* **99,** 12963–12968.

[24] Le Page C, Provencher D, Maugard CM, Ouellet V, Mes-Masson AM (2004). Signature of a silent killer: expression profiling in epithelial ovarian cancer. *Expert Rev Mol Diagn* **4,** 157–167.

[25] Simon R, Mirlacher M, and Sauter G (2004). Tissue microarrays. *Biotechniques* **36,** 98–105.
[26] Westfall PH and Young SS (1993). Resampling-Based Multiple Testing: Examples and Methods for *P*-Value Adjustment. Wiley-Interscience, New York.
[27] Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, et al. (2001). Minimum Information About a Microarray Experiment (MIAME)—toward standards for microarray data. *Nat Genet* **29,** 365–371.
[28] Eisen MB, Spellman PT, Brown PO, and Botstein D (1998). Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* **95,** 14863–14868.
[29] Tusher VG, Tibshirani R, and Chu G (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* **98,** 5116–5121.
[30] Tibshirani R, Hastie T, Narasimhan B, and Chu G (2002). Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci USA* **99,** 6567–6572.
[31] Trogan E, Choudhury RP, Dansky HM, Rong JX, Breslow JL, and Fisher EA (2002). 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.
[32] Krynetski EY, Krynetskaia NF, Bianchi ME, and Evans WE (2003). A nuclear protein complex containing high mobility group proteins B1 and B2, heat shock cognate protein 70, ERp60, and glyceraldehyde-3-phosphate dehydrogenase is involved in the cytotoxic response to DNA modified by incorporation of anticancer nucleoside analogues. *Cancer Res* **63,** 100–106.
[33] Struski S, Doco-Fenzy M, and Cornillet-Lefebvre P (2002). Compilation of published comparative genomic hybridization studies. *Cancer Genet Cytogenet* **135,** 63–90.
[34] Kudoh K, Takano M, Koshikawa T, Hirai M, Yoshida S, Mano Y, Yamamoto K, Ishii K, Kita T, Kikuchi Y, et al. (1999). Gains of 1q21-q22 and 13q12-q14 are potential indicators for resistance to cisplatin-based chemotherapy in ovarian cancer patients. *Clin Cancer Res* **5,** 2526–2531.
[35] Staebler A, Heselmeyer-Haddad K, Bell K, Riopel M, Perlman E, Ried T, and Kurman RJ (2002). Micropapillary serous carcinoma of the ovary has distinct patterns of chromosomal imbalances by comparative genomic hybridization compared with atypical proliferative serous tumors and serous carcinomas. *Hum Pathol* **33,** 47–59.
[36] Launonen V, Mannermaa A, Stenback F, Kosma VM, Puistola U, Huusko P, Anttila M, Bloigu R, Saarikoski S, Kauppila A, et al. (2000). Loss of heterozygosity at chromosomes 3, 6, 8, 11, 16, and 17 in ovarian cancer: correlation to clinicopathological variables. *Cancer Genet Cytogenet* **122,** 49–54.
[37] Dent J, Hall GD, Wilkinson N, Perren TJ, Richmond I, Markham AF, Bell SM, and Bell SM (2003). Cytogenetic alterations in ovarian clear cell carcinoma detected by comparative genomic hybridisation. *Br J Cancer* **88,** 1578–1583.
[38] Rhei E, Bogomolniy F, Federici MG, Maresco DL, Offit K, Robson ME, Saigo PE, and Boyd J (1998). Molecular genetic characterization of *BRCA1*- and *BRCA2*-linked hereditary ovarian cancers. *Cancer Res* **58,** 3193–3196.
[39] Concin N, Becker K, Slade N, Erster S, Mueller-Holzner E, Ulmer H, Daxenbichler G, Zeimet A, Zeillinger R, Marth C, et al. (2004). Transdominant DeltaTAp73 isoforms are frequently up-regulated in ovarian cancer. Evidence for their role as epigenetic p53 inhibitors *in vivo*. *Cancer Res* **64,** 2449–2460.
[40] Iyengar TD, Ng S, Lau CC, Welch WR, Bell DA, Berkowitz RS, and Mok SC (1999). Differential expression of NF1 type I and type II isoforms in sporadic borderline and invasive epithelial ovarian tumors. *Oncogene* **18,** 257–262.
[41] Khan J, Simon R, Bittner M, Chen Y, Leighton SB, Pohida T, Smith PD, Jiang Y, Gooden GC, Trent JM, et al. (1998). Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. *Cancer Res* **58,** 5009–5013.
[42] Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, et al. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286,** 531–537.
[43] Simon R (2003). Diagnostic and prognostic prediction using gene expression profiles in high-dimensional microarray data. *Br J Cancer* **89,** 1599–1604.
[44] Orr GA, Verdier-Pinard P, McDaid H, and Horwitz SB (2003). Mechanisms of Taxol resistance related to microtubules. *Oncogene* **22,** 7280–7295.
[45] Sugimura M, Sagae S, Ishioka S, Nishioka Y, Tsukada K, and Kudo R (2004). Mechanisms of paclitaxel-induced apoptosis in an ovarian cancer cell line and its paclitaxel-resistant clone. *Oncology* **66,** 53–61.
[46] Vargas-Roig LM, Gago FE, Tello O, Aznar JC, and Ciocca DR (1998). Heat shock protein expression and drug resistance in breast cancer patients treated with induction chemotherapy. *Int J Cancer* **79,** 468–475.
[47] He Q, Liang CH, and Lippard SJ (2000). Steroid hormones induce HMG1 overexpression and sensitize breast cancer cells to cisplatin and carboplatin. *Proc Natl Acad Sci USA* **97,** 5768–5772.
[48] Phillips JL, Hayward SW, Wang Y, Vasselli J, Pavlovich C, Padilla-Nash H, Pezullo JR, Ghadimi BM, Grossfeld GD, Rivera A, et al. (2001). The consequences of chromosomal aneuploidy on gene expression profiles in a cell line model for prostate carcinogenesis. *Cancer Res* **61,** 8143–8149.
[49] Platzer P, Upender MB, Wilson K, Willis J, Lutterbaugh J, Nosrati A, Willson JK, Mack D, Ried T, and Markowitz S (2002). Silence of chromosomal amplifications in colon cancer. *Cancer Res* **62,** 1134–1138.



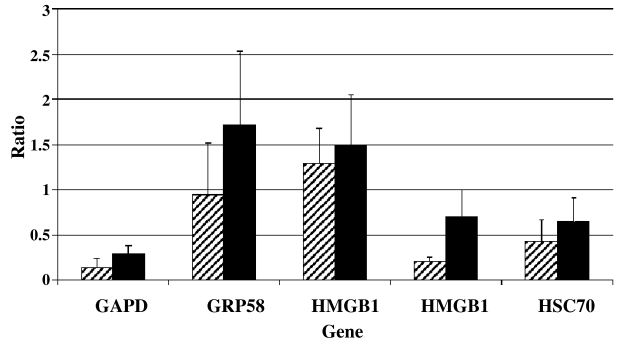


Figure W2. *Validation by real-time RT-PCR of members of the nuclear complex. Arithmetic means representing relative expressions of each gene in the resistant and sensitive groups are depicted with hashed (▨) and solid (■) bars, respectively. Expression levels were calibrated against Human Universal Reference. The error bars represent standard deviation.*

Figure W1. *Analysis of 22 samples using unsupervised hierarchical clustering. Sensitive (S; pink) and resistant (R; blue) samples are interspersed throughout the dendogram. Relative expression levels of genes are depicted in green (down), red (up), black (no change), or grey (data missing). The yellow vertical bar highlights a major gene cluster within which most genes identified as differentially expressed between the sensitive and resistant groups by SAM analysis were localized.*

Table W1. Genes Identified by SAM* Analysis that are Differentially Expressed between the Resistant and Sensitive Groups.

Clone ID	Fold Change†	Name	Gene Symbol	Cytoband
UP‡				
296275	15.43	H19, imprinted maternally expressed untranslated mRNA	LOC283120	11p15.5
109357	4.73	placenta-specific 4	PLAC4	21q22.3
53087	3.65	hypothetical protein LOC283922	LOC283922	16q22.3
158123	3.32	ATPase, H+ transporting, lysosomal 56/58kDa	ATP6V1B1	2p13.3
470393	3.29	matrix metalloproteinase 7	MMP7	11q22.2
5301677	2.85	retinol binding protein 1, cellular	RBP1	3q23
148914	2.76	Homo sapiens mRNA; cDNA DKFZp564B076	N/A	6p21.1
29988	2.42	pre-B-cell leukemia transcription factor 1	PBX1	2q23.1
149421	2.09	dual specificity phosphatase 1	DUSP1	5q35.1
2262605	2.07	Homo sapiens EMX2OS mRNA, complete sequence	N/A	10q26.11
501985	2.02	sphingosine-1-phosphate lyase 1	SGPL1	10q22.1
DOWN§				
162892	10.39	Homo sapiens cDNA FLJ26905 fis	N/A	22q11.22
146669	6.73	immunoglobulin lambda joining 3	IGLJ3	22q11.22
308924	6.47	hemoglobin, epsilon 1	HBE1	11p15.4
293298	5.62	Homo sapiens IGL mRNA for immunoglobulin lambda light chain	N/A	22q11.22
4701351	4.94	ferritin, light polypeptide	FTL	19q13.33
5922013	4.84	tubulin, beta polypeptide	TUBB	6p25.2
33203	4.77	tubulin, alpha 3	TUBA3	12q13.12
5784795	4.57	metallothionein 2A	MT2A	16q12.2
163221	4.56	Homo sapiens clone P2-114 anti-oxidized LDL immunoglobulin light chain	N/A	22q11.22
110243	4.47	Homo sapiens cDNA FLJ26905 fis, clone RCT01427, highly similar to	N/A	22q11.22
300017	4.18	apolipoprotein B	APOB	2p24.1
5923231	4.10	RAN, member RAS oncogene family	RAN	12q24.33
4876644	4.08	tubulin, beta, 5	TUBB5	19p13.3
261822	3.79	Homo sapiens, clone IMAGE:5728597	N/A	22q11.23
503601	3.74	glyceraldehyde-3-phosphate dehydrogenase	WT1	11p13
184151	3.72	tubulin, beta, 4	TUBB4	16q24.3
33204	3.72	tubulin, beta polypeptide paralog	MGC8685	6p25.2
303100	3.68	fibronectin 1	FN1	2q35
5762339	3.65	membrane targeting (tandem) C2 domain containing 1	MTAC2D1	14q32.12
198960	3.42	hemoglobin, gamma G	HBG2	11p15.4
470628	3.39	Homo sapiens cDNA clone IMAGE:5288757	N/A	1p32.2
5926230	3.36	peroxiredoxin 1	PRDX1	1p34.1
229610	3.35	Homo sapiens transcribed sequence	N/A	11p15.4
5749861	3.34	immunoglobulin lambda locus	IGL@	22q11.22
5422775	3.32	ribosomal protein S5	RPS5	19q13.43
208059	3.27	heat shock 70kDa protein 8	HSC70	11q24.1
5740157	3.26	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation pro	YWHAQ	2p25.1
5421424	3.19	enolase 1	ENO1	1p36.23
5454961	2.88	Homo sapiens transcribed sequence	N/A	15q25.1
120559	2.87	metallothionein 1F	MT1F	16q12.2
5406583	2.82	H2A histone family, member Z	H2AFZ	4q23
5806641	2.81	histone H2A.F/Z variant	H2AV	15q26.1
212856	2.79	hemoglobin, zeta	HBZ	16p13.3
5925346	2.77	chromosome 7 open reading frame 17	LOC51142	7p11.2
504180	2.75	Homo sapiens, clone IMAGE:5742072	N/A	6p22.3
267145	2.74	high-mobility group box 2	HMGB2	4q34.1
154975	2.69	Homo sapiens transcribed sequence	N/A	22q11.22
471144	2.64	electron-transfer-flavoprotein, beta polypeptide	ETFB	19q13.41
307027	2.59	heterogeneous nuclear ribonucleoprotein A2/B1	HNRPA2B1	7p15.2
27920	2.57	mitogen-activated protein kinase kinase kinase 7 interacting protein 1	MAP3K7IP1	22q13.1
297215	2.55	annexin A5	ANXA5	4q27
145972	2.54	KIAA0101 gene product	KIAA0101	15q22.31
5111837	2.52	cofilin 1	CFL1	11q13.1
5934101	2.50	pyruvate kinase, muscle	PKM2	15q23
4734377	2.49	ubiquitin B	UBB	17p11.2
270551	2.48	CD44 antigen	CD44	11p13
155222	2.48	glutathione peroxidase 4	GPX4	19p13.3
3546201	2.45	ribosomal protein L10	RPL10	Xq28
5494320	2.45	heat shock 90kDa protein 1, alpha	HSPCA	14q32.32
488225	2.44	ribosomal protein S19	RPS19	19q13.2
139590	2.43	RAN binding protein 1	RANBP1	22q11.21
29774	2.43	eukaryotic translation initiation factor 2	EIF2S2	2q31.1
341798	2.34	ubiquinol-cytochrome c reductase hinge protein	UQCRH	1p36.13
5809010	2.31	putative translation initiation factor	SUI1	17q21.2
5533075	2.30	small nuclear ribonucleoprotein polypeptides B and B1	SNRPB	20p13
490368	2.29	ribosomal protein S16	RPS16	19q13.2
489898	2.27	nuclease sensitive element binding protein 1	NSEP1	1p34.2
300041	2.26	non-metastatic cells 1	NME1	17q21.33

(*continued on next page*)

Table W1. (*continued*).

Clone ID	Fold Change†	Name	Gene Symbol	Cytoband
343295	2.25	lectin, galactoside-binding, soluble, 1 (galectin 1)	LGALS1	22q13.1
138280	2.23	DKFZp564J157 protein	DKFZP564J157	12q13.13
3916469	2.23	NHP2 non-histone chromosome protein 2-like 1	NHP2L1	22q13.2
5540049	2.22	Homo sapiens transcribed sequence	HCP15	7p15.3
5475005	2.22	Homo sapiens similar to ribosomal protein L18a	RPL18A	19p13.11
4126974	2.22	malate dehydrogenase 2, NAD	MDH2	7q11.23
230534	2.22	hemoglobin, alpha 2	HBA1	16p13.3
5785893	2.21	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit c	ATP5G3	2q31.1
5431259	2.21	glyceraldehyde-3-phosphate dehydrogenase	GAPD	12p13.31
278772	2.21	stearoyl-CoA desaturase	SCD	10q24.31
5770525	2.21	activating transcription factor 4	ATF4	22q13.1
5932330	2.20	SH3 domain binding glutamic acid-rich protein like 3	SH3BGRL3	1p36.11
5704441	2.20	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	DDX5	17q24.1
5528790	2.19	excision repair cross-complementing rodent repair deficiency, comp	ERCC1	19q13.32
4762963	2.19	ribosomal protein S13	RPS13	11p15.1
505381	2.19	syndecan 2	SDC2	8q22.1
4702445	2.19	ferritin, heavy polypeptide 1	FTH1	11q12.3
5806598	2.18	keratin 8	KRT8	12q13.13
5927927	2.17	ribosomal protein L10a	RPL10A	6p21.31
200573	2.16	CDC28 protein kinase regulatory subunit 1B	CKS1B	1q22
346784	2.15	alpha-2-HS-glycoprotein	AHSG	3q27.3
209207	2.15	superoxide dismutase 1	SOD1	21q22.11
295004	2.14	hepatoma-derived growth factor	HDGF	1q23.1
137317	2.13	chromosome 6 open reading frame 49	C6orf49	6p21.1
186682	2.13	trefoil factor 1	TFF1	21q22.3
5788140	2.13	ras homolog gene family, member A	ARHA	3p21.31
5420166	2.12	actin, alpha 2, smooth muscle, aorta	ACTA2	10q23.31
4772211	2.12	calmodulin 2	CALM2	2p21
4347418	2.10	high-mobility group nucleosome binding domain 1	HMGN1	21q22.2
5455241	2.09	solute carrier family 25	SLC25A5	Xq24
471551	2.09	Homo sapiens transcribed sequence	N/A	8q13.2
505115	2.09	chaperonin containing TCP1, subunit 8 (theta)	CCT8	21q21.3
5927445	2.08	peroxiredoxin 6	PRDX6	1q25.1
200190	2.07	nascent-polypeptide-associated complex alpha polypeptide	CA	12q13.3
151473	2.06	G1 to S phase transition 1	GSPT1	16p13.13
221318	2.05	CDC28 protein kinase regulatory subunit 2	CKS2	9q22.2
5926362	2.04	solute carrier family 25 (mitochondrial carrier; phosphate carrier	SLC25A3	12q23.1
4304010	2.04	stathmin 1/oncoprotein 18	STMN1	1p36.11
5270796	2.03	syndecan binding protein	SDCBP	8q12.1
504916	2.03	hypothetical protein MGC10812	MGC10812	19p13.11
4279563	2.02	tissue factor pathway inhibitor 2	TFPI2	7q21.3
172079	2.01	triosephosphate isomerase 1	TPI1	12p13.31
503300	2.01	hypothetical protein LOC51315	LOC51315	1p36.33
200586	2.00	metallothionein 1B	MT1B	16q12.2
23904	2.00	ELOVL family member 5, elongation of long chain fatty acids	ELOVL5	6p12.1

N/A Not available.
*SAM analysis with FDR 0.07-0.8%, fold change 2.
†Relative fold change between the two classes (resistant/sensitive).
‡List of genes whose transcript levels are greater in the resistant compared to the sensitive group.
§List of genes whose transcript levels are less in the resistant compared to the sensitive group.

Table W2. Test of Agreement between Expression Changers and Discriminate Genomic Regions.

Expression Microarray	BAC CGH*			
	Cytoband Regions Gained	Cytoband Regions with No Change	Cytoband Regions Lost	
Number of Genes with Greater Transcript Levels†	2‖**	10	0	12
Number of Genes with No Change	320	6194‖	175	6689
Number of Genes with Lower Transcript Levels†	7	95	8‖**	110
TOTAL	329	6299	183	6811

Cohen's Kappa = 0.0230‡
Test of H_o: Kappa = 0: $\kappa\hat{}^* = 0.1316$§
*Cytoband regions gained, lost, or with no change in the resistant compared to the sensitive group that were identified by the Fisher Exact test ($P < 0.05$).
†Genes whose transcript levels were determined to be significantly greater or lower in the resistant compared to the sensitive group by SAM analysis (FDR 0.7-0.08%, FC 2).
‡Slight agreement between the expression and BAC CGH data ($0 < \kappa\hat{}^* < 1$).
§Agreement does not exceed chance ($\kappa\hat{}^* < 1.96$).
‖ Overall agreement between aCGH and expression data was 91%.
**8.2% of SAM changers agree with changes in copy number.

Table W3. Overall Losses and Gains.

Loss	Gain
1p36.11-pter; 3p22.1-p21.31; 4p15.33-p16.1; 4p12-p14; 4q13.2-qter; 5q13.1-q23.3; 6q22.1-qter; 8p21.1-pter; 9p; 9q21.11-q33.1; 11p15.1-pter; 13q12.2-q21.1; 13q32.2-qter; 15q11.2-q15.1; 16q; 17p13.2-q11.2; 17q21.32-q21.33; 18q21.1-qter; 22	1q11-q25.3; 2p22.3-pter; 3q22.3-qter; 6p22.1-pter; 6p21.1-p21.31; 7q31.33-qter; 8q; 12p; 20p12.1-pter; 20q12-qter