Gene Expression Profiles Predict Early Relapse in Ovarian Cancer after Platinum-Paclitaxel Chemotherapy


1Mayo Clinic Cancer Center, Rochester, Minnesota; 2University of Texas M.D. Anderson Cancer Center, Houston, Texas; and 3Millennium Pharmaceuticals, Cambridge, Massachusetts

ABSTRACT

Purpose: Women with advanced epithelial ovarian cancer are routinely treated with platinum-paclitaxel chemotherapy following cytoreductive surgery, yet only ~20% achieve long-term disease-free survival. We hypothesized that differences in gene expression before treatment could distinguish patients with short versus long time to recurrence after administration of platinum-paclitaxel combination chemotherapy.

Experimental Design: To test this hypothesis, gene expression profiling of 79 primary surgically resected tumors from women with advanced-stage, high-grade epithelial ovarian cancer was done using cDNA microarrays containing 30,721 genes. Supervised learning algorithms were applied in an effort to develop a binary classifier that could discriminate women at risk for early relapse (<21 months) versus late (>21 months) relapse after initial chemotherapy.

Results: A 14-gene predictive model was developed using a set of training samples (n = 51) and subsequently tested using an independent set of test samples (n = 28). This model correctly predicted the outcome of 24 of the 28 test samples (86% accuracy) with 95% positive predictive value for early relapse.

Conclusions: Predictive markers for early recurrence can be identified for platinum-paclitaxel combination chemotherapy in primary ovarian carcinoma. The proposed 14-gene model requires further validation.

INTRODUCTION

Of the cancers unique to women, epithelial ovarian cancer has the highest mortality rate, with 16,090 estimated deaths in the United States for the year 2004 (1). The initial treatment of advanced ovarian cancer includes aggressive cytoreductive surgery followed by combination chemotherapy. Randomized clinical trials provide clear evidence that platinum-based combination chemotherapy after surgery offers the highest clinical benefit based on response rate, time to recurrence (TTR), and overall survival (2–4). Currently, the combination of carboplatin and paclitaxel is the preferred standard regimen in North America (5). Unfortunately, most women with advanced ovarian cancer ultimately relapse and die of disease progression despite an initial response to first-line platinum-paclitaxel and only a minority (~20%) are cured with the standard combination. If those women who will derive little benefit from standard platinum-paclitaxel could be identified before therapy (i.e., those with early recurrence), then early intervention with alternative approaches, including novel agents targeted to molecular pathways that are aberrantly activated in ovarian cancer (6), could be used in them.

Numerous prognostic factors have been studied in advanced ovarian cancer, but none are predictive of outcome after platinum-paclitaxel chemotherapy (7–10). Recent studies in several malignancies suggest that unique (RNA) gene expression patterns are often associated with prognosis (11). Emerging data in non-Hodgkin’s lymphoma, for example, suggest that gene expression profiling can predict disease-free survival and overall survival in patients treated with anthracycline-based combination chemotherapy (12, 13). The application of whole-genome technologies, such as gene expression profiling, allows one to generate a multivariate predictor based on the measurement of 30,000 (or more) genes in each sample. Through statistical testing and selection of an analytic gene set for the entire sample group, hierarchical clustering can be applied to find previously unrecognized relationships for the purpose of class comparison (14). This has implications for molecular classification (diagnosis) as well as better prognostic classification (good versus poor). However, the generation of a multigene classifier could also be tested as a predictive model for various outcomes, such as response rate, TTR, and overall survival, after drug administration (15).

To test the hypothesis that gene expression profiling of primary surgically resected ovarian cancers (obtained before chemotherapy) could identify a molecular signature that would distinguish women at risk for early relapse after platinum-paclitaxel chemotherapy, we analyzed 79 ovarian cancer specimens using cDNA microarrays corresponding to 30,721 transcripts. Our results identify a 14-gene predictive model that has 86% accuracy when tested on an independent series of
patients treated in a similar manner. Interestingly, of the 19 patients that had early relapse in our test set, 18 patients were correctly identified based on the gene expression profile obtained from the primary tumor.

MATERIALS AND METHODS

Patients and Tumor Specimens. Tumors removed at primary surgery were snap frozen in the surgical pathology suite of the Mayo Clinic Cancer Center or University of Texas M.D. Anderson Cancer Center and stored at −80°C. Only tumor blocks containing at least 70% malignant tissue were included in this study. Tumors were staged according to International Federation of Gynecology and Obstetrics standards (16). Optimal cytoreductive surgery was defined as residual disease of ≤1 cm in greatest diameter. No patient had received chemotherapy before surgery. All patients were treated with paclitaxel plus either carboplatin or cisplatin as initial therapy after surgery. Because most patients were optimally debulked and had no measurable disease at the initiation of chemotherapy, TTR was selected as the primary end point for the classifier. TTR was the time from diagnosis to recurrence and initiation of second-line therapy. The collection of tumor samples and clinical data was done under institutional review board–approved protocols.

For the present study, we studied two sample sets. A training set of 51 samples was used for marker identification. The median TTR in these patients was 21 months. Because 21 months is the median TTR for optimally debulked, advanced-stage ovarian patients treated with carboplatinum-paclitaxel in the recent Gynecologic Oncology Group phase III trial (5), we used 21 months as the cutoff between early and late relapse. A second test set contained 28 samples. For women who had a positive second look laparotomy after completion of their chemotherapy (n = 11 of 79), the date of their second look was not considered their date of recurrence. These women were included in the early relapse group, but TTR was not available for them.

As noted, the patient samples for the training data were selected to carry out a two-group analysis with a 21-month cutoff. As a test of sensitivity of this selection, we carried out analyses using different cut points of TTR, including patients who had very early relapse (<12 months). The results of those analyses showed no gain in prediction accuracy. We also considered TTR as a continuous variable. However, no single marker or set of markers was significantly associated with TTR after correction for multiple comparison artifacts.

cDNA Arrays and Transcriptional Profiling. RNA was isolated from primary ovarian tumors using the RNeasy kit (Qiagen, Valencia, CA) and quality controlled as described (17). Probes for profiling were generated by reverse transcribing RNA with SuperScript II (Invitrogen, Carlsbad, CA) into cDNA in the presence of [α-33P]dCTP (100 mCi/mL, Amersham, Little Chalfont, United Kingdom) from 2 μg total RNA. The cDNA probes were hybridized (without further amplification) to high-density nylon cDNA arrays as described (18). The arrays used in this study contained 30,721 human sequence clones corresponding to 21,594 unique UniGene clusters. Array hybridizations were done in duplicate and were highly reproducible as reflected by the average coefficient of variation (SD/mean) of 0.2.

Expression results were normalized to the median of all genes in each array. After normalization, expression data were transformed using the expression log (expression + 1) and used in all subsequent analyses.

Data Analysis. The strategy used for data analysis was similar to the one described in our earlier report on building a predictive model for neoadjuvant chemotherapy in breast cancer (15). A multiple random sampling methodology was used to determine the best model (i.e., the model with the lowest prediction error rate based on the 51 training samples). A simpler approach was considered initially that uses all the data for model selection, which is then tested on an independent validation data set. However, simulation studies (data not shown) showed superior performance of the multiple random selection scheme for model selection in data sets with relatively low signal-to-noise levels such as reported in this study. For the random sampling method, a 70% training fraction was sampled, blocking on outcome, debulking status, histology, and tumor stage, from the 51 training samples (see Fig. 1). Signal-to-noise ratio (SNR) [absolute (μlate − μearly)/(σlate + σearly)] where μ and σ represent the mean and SD of expression for each class, respectively] was calculated for each potential marker. The top SNR markers, up to 500, were then used in a support vector machine (SVM) using a second-order polynomial for classification (19). This process was repeated 100 times resulting in 100 possible models. Each of the 100 models was then refit using all 51 training samples and studied against the 28 test samples. The model with the fewest prediction errors was selected as the best model.

To determine whether our results could be attributable to chance alone, we did a permutation test in which the entire random sampling procedure outlined above was repeated but permuting the outcome labeling of the training set 10 times. This generated 1,000 (100 models per permutations) random models. Average prediction rates and point-wise 99% confidence intervals for the permuted results were compared with results using the true outcome labels. Kaplan-Meier curves, stratified by predicted class from the selected model, were constructed based on TTR or follow-up time. Statistically significant differences between predicted class strata were assessed using the log-rank test. To assess general patterns in the data, average linkage clustering was done with CLUSTER and TREEVIEW software (20). Expression data were log (expression + 1) transformed, features were median centered and a Pearson correlation was used.

RESULTS

Patient and Sample Characteristics. The present study examined a total of 79 tumor samples obtained from women with advanced stage, high-grade epithelial ovarian cancer before the initiation of systemic chemotherapy. The characteristics of all 79 ovarian cancer patients are listed in Table 1. Each of these patients underwent debulking surgery followed by chemotherapy with paclitaxel in combination with either carboplatin or cisplatin on standard schedules and dosages as first-line treatment. Samples from these patients were divided into a training set (n = 51), which was used for marker identification and model building, and a test set (n = 28), which was used to assess the top independent models.
The median TTR was 21 months for the 51 patients in the training set and 14 months for the 28 test patients. The median follow-up for the 79 patients was 31 months. We chose 21 months as the time point to distinguish early versus late recurrence because this time point represents the median TTR from the literature for patients with optimally debulked advanced-stage disease treated with platinum-paclitaxel (5, 21). The training set included two distinct subsets of early and late relapse patients. The median TTR was 35.5 months in the late relapse group and 8.7 months in the early relapse group. Twenty-four of 25 (96%) women in the late relapse group were alive at last follow-up compared with 9 of 26 (35%) still alive in the early relapse group.

**Gene Expression Profiling.** Individual features (“genes”) whose expression correlated with a given sample belonging to one class or the other were ranked using the SNR metric. We then developed a SVM algorithm using the SNR features and built a binary model to predict recurrence (early versus late recurrence defined by TTR <21 versus >21 months). With the SVM algorithm, we identified a top model that contained 14 markers. Each gene with annotation is listed in Table 2 and box plots of gene expression by early versus late recurrence are displayed in Fig. 2. In 4-fold cross-validation studies, the model correctly classified 40 of the 51 training samples (78% accuracy; \( P < 0.001 \); see Fig. 3B).

The test set contained 28 samples, of which 7 are associated with late relapse and 21 are associated with early relapse. In this study, the 14-gene SVM model correctly categorized 6 of the 7 (86%) samples of the late relapse group and 18 of the 21 (86%) samples of the early relapse group for an overall accuracy of 86% \( (P < 0.05; \text{Table 3}) \). The model predicted that 19 (of the 28 test cases) would have early recurrence and 18 had disease recurrence <21 months for a positive predictive value of 95%. The model predicted late relapse in nine cases and six had disease recurrence beyond 21 months for a negative predictive value of 67%. A Kaplan-Meier plot of the TTR for the predicted late and early relapse patients in the test set is shown in Fig. 3C \( (P < 0.05) \). In addition, we did a permutation analysis on the data using 1,000 random models. Our two best SVM-derived models lay outside the 99% confidence interval for the best permuted model, indicating a very small likelihood that a model such as ours could be discovered by chance.

A clustering analysis was conducted on the gene expression data to look for systematic differences due to known or unknown sources of variability. Results (not shown) showed no formation of major cluster groups that would

---

**Table 1**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Training</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. subjects</td>
<td>51</td>
<td>28</td>
</tr>
<tr>
<td>Median age at diagnosis</td>
<td>62 (39-81)</td>
<td>62 (39-76)</td>
</tr>
<tr>
<td>Median TTR (mo)*</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td>No. alive at time of analysis (%)</td>
<td>33 (65)</td>
<td>16 (57)</td>
</tr>
<tr>
<td>Median length of follow-up (mo)†</td>
<td>30.8</td>
<td>29.6</td>
</tr>
<tr>
<td>Stage, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>1 (2)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>III</td>
<td>42 (82)</td>
<td>20 (71)</td>
</tr>
<tr>
<td>IV</td>
<td>8 (16)</td>
<td>8 (29)</td>
</tr>
<tr>
<td>Histologic characteristics, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serous</td>
<td>36 (71)</td>
<td>24 (86)</td>
</tr>
<tr>
<td>Endometrioid</td>
<td>7 (14)</td>
<td>3 (11)</td>
</tr>
<tr>
<td>Mixed serous-endometrioid</td>
<td>8 (16)</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Primary surgery, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal debulking</td>
<td>37 (73)</td>
<td>25 (89)</td>
</tr>
<tr>
<td>Suboptimal</td>
<td>14 (27)</td>
<td>3 (11)</td>
</tr>
<tr>
<td>Median pre-CA125</td>
<td>943.5</td>
<td>858</td>
</tr>
<tr>
<td>Median post-CA125</td>
<td>13.6</td>
<td>13</td>
</tr>
</tbody>
</table>

*Excludes positive second look laparotomy patients \((n = 11 \text{ of 79})\).
†Follow-up refers to time from date of initial surgery to death or to most recent follow-up visit.
suggest systematic effects due to technical processes. Patient characteristics (outcome, debulking status, etc.) also showed no formation of major clusters.

**Subset Analyses.** We applied the top 14-gene model to the following sample subsets: stage III disease patients, optimally debulked patients, and, finally, patients with serous histology tumors. There was no significant improvement in the performance of the model in these patient subgroups (data not shown). Moreover, we generated a new model using only the 37 optimally debulked training samples, which was then tested in the 25 optimally debulked test samples using cross-validation. The best performing SVM-derived model had an overall accuracy (84%) similar to the top model constructed using all the 51 training set patient samples.

**DISCUSSION**

At present, platinum-paclitaxel chemotherapy is given to the large majority of women with advanced-stage ovarian cancer after cytoreductive surgery. With this approach, ~20% remain free of recurrence at 5 years (5). Using gene expression profiling, we sought to develop a predictive model that would identify women with ovarian carcinoma destined for early relapse if treated with platinum-paclitaxel chemotherapy. If such individuals could be identified reliably at diagnosis, then investigational agents, or alternative strategies with existing chemotherapy regimens, could be offered as initial therapy.

The 14-gene model developed in the present study accurately predicted early or late recurrence in 86% of patients in the test set ($P < 0.05$). Significantly, the positive predictive value of the proposed model is 95% (of the 19 patients who tested positive, 18 relapsed at ≤21 months). To our knowledge, this study represents the first multigene predictive model for the platinum-paclitaxel regimen in ovarian cancer.

The present study differs from previous reports in which gene expression has been profiled in ovarian cancer. We have previously used gene expression profiling with cDNA microarrays in an attempt to elucidate early events in ovarian carcinogenesis and to distinguish early-stage from late-stage disease (22). Interestingly, none of the 14 genes in the current model were listed in our earlier report. Several other groups have employed gene expression profiling to analyze epithelial ovarian carcinoma. Collectively, these reports describe differences in gene expression between normal ovarian epithelium and epithelial ovarian cancers and among stages, grades, histologic subtypes, and BRCA1/2 status of ovarian cancers (23–32). Two groups have also previously correlated gene expression profiles with patient outcomes in ovarian cancer. Lancaster et al. recently analyzed 31 advanced-stage serous cancers using oligonucleotide arrays with unsupervised learning approaches (33). Fourteen of the patients survived >7 years and 17 died within 2 years after initial diagnosis. When hierarchical clustering was done using 2,655 genes that were differentially expressed among the 31 tumor samples, a cluster of 43 genes seemed to distinguish the long-term (>84 months) from the short-term (<24 months) survivors. It is important to note, however, that there were significant differences between debulking status (78% optimally debulked in the good outcome group versus 29% optimally debulked in the poor outcome group) that could have affected the results of this earlier study. Moreover, formal class prediction by supervised learning approaches was not reported. In the second report, methylation microarray analysis was done on 18 stage III to IV epithelial ovarian cancers in an attempt to identify markers for prognosis (34). Notably, five of the samples were obtained at relapse, and progression-free survival data were not available for all the patients. Nevertheless, hierarchical clustering revealed that the group with high levels of tumor methylation had an inferior outcome.

In the present study, there were high rates of optimal debulking surgery in both outcome groups (Table 1), all samples were obtained before chemotherapy, and outcome data were available for all patients. Using this sample set, supervised learning approaches identified a set of 14 differentially expressed genes that distinguish patients based on TTR (Table 2).

Interestingly, the list of 14 differentially expressed genes does not include polypeptides such as ATP binding cassette transporters, metallothionein, DNA repair components, or apoptotic regulators (e.g., Bcl-2 family members), which have been implicated in drug resistance based on studies in model systems. In this regard, the present results parallel those of Holleman et al., who identified a gene expression profile in pediatric acute lymphoblastic leukemia that correlated with drug resistance and treatment outcome.
sensitivity and treatment outcome but also contained a paucity of transcripts commonly implicated in resistance (35). Nonetheless, several markers contained within the 14-gene model identified in our study are plausible determinants of treatment outcome in ovarian carcinoma. Protein kinase C\(\text{D}\) has been implicated previously as a poor prognostic factor in ovarian carcinoma (36, 37) and its overexpression in the present study is observed in tumors that recur early after initial therapy (Table 2). Protein kinase C family members are known to affect diverse signaling pathways that could affect not only cell cycle progression and differentiation (38) but also response to immune system death signals (39). Another of the top markers, inhibitor of DNA binding 4, negatively regulates the expression of \(\text{BRCA1}\) (40, 41). Interestingly, two major studies have shown longer TTR (42) and overall survival (43) in BRCA-linked cancers possibly because of enhanced sensitivity of the tumor cells to DNA damaging agents. In our study, elevated expression of inhibitor of DNA binding 4, which would be expected to down-regulate \(\text{BRCA1}\), also correlated with longer TTR. In addition, FLJ20241, showing higher expression in early recurrence in our study, has been identified as a nuclear factor-\(\kappa\)B-activating protein in a large-scale experimental study (44). Interestingly, Mabuchi et al. (45) recently linked nuclear factor-\(\kappa\)B inhibition to an increase in cisplatin efficacy in \textit{in vitro} and \textit{in vivo} ovarian cancer models. Together, our data and these findings suggest that higher FLJ20241 expression would lead to nuclear factor-\(\kappa\)B activation and consequently decreased cisplatin efficacy, resulting in early recurrence of ovarian cancer after therapy. The exact role of the remaining genes in the biology of ovarian cancer requires further investigation.

In summary, application of gene expression profiling with supervised learning algorithms to a set of ovarian cancers obtained before chemotherapy allowed development of a 14-gene profile that identified patients at risk for early recurrence after combination chemotherapy. The current model predicted outcome in an independent set of patients (\(n = 28\)) with an
overall accuracy of 86%. These observations need to be validated independently with other sample sets obtained from women with ovarian cancer given platinum-paclitaxel as initial therapy. A second validation attempt could be made using primary ovarian tumors from women treated with other platinum-based regimens to determine the specificity of the 14-gene model. If future studies validate this model, stratification based on pretreatment gene expression would be feasible in patients with ovarian cancer, thereby allowing women who are destined for early relapse to receive alternative primary treatment strategies.

ACKNOWLEDGMENTS

We thank Dr. Jeffrey Brown for his invaluable assistance and guidance in this work and Jill Dowdy, Jennifer Burkhardt, and Karin Goodman for their assistance with data collection.

REFERENCES


Gene Expression Profiles Predict Early Relapse in Ovarian Cancer after Platinum-Paclitaxel Chemotherapy

Lynn C. Hartmann, Karen H. Lu, Gerald P. Linette, et al.


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/11/6/2149

Cited articles
This article cites 44 articles, 21 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/11/6/2149.full.html#ref-list-1

Citing articles
This article has been cited by 15 HighWire-hosted articles. Access the articles at:
/content/11/6/2149.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.