Supplementary Information:

SUPPLEMENTARY EXPERIMENTAL PROCEDURES AND SUPPLEMENTAL REFERENCES

Development and Evaluation of a BAD-Pathway Gene Signature

Using genomic data from the panel of OVCA cells previously subjected to serial cisplatin treatment, principal components analysis (PCA) methodology was used to derive a *BAD-pathway gene expression signature* with a corresponding "*pathway score*" to represent an overall gene expression level for the BAD-pathway genes (or subsets thereof for data sets generated by U133A or U95A). The generation of the signature used data from cell lines only; no patient data were used. That is, no data from the Duke/MCC samples were used in the initial development/generation of the 47-gene signature; the Duke/MCC ovarian data was a completely independent evaluation set. Specifically, using genomic and EC_{50} data from OVCA cell lines, Pearson correlation was used to identify genes associated with cisplatin resistance (EC_{50}). Expression was calculated using the robust multi-array average algorithm (1) implemented in Bioconductor (http://www.bioconductor.org) extensions to the R-statistical programming environment as described previously (2). Probe sets with expression ranges <2-fold (maximum/minimum) and control probes (i.e., AFFX_* probe sets) were excluded from the analysis. For each cell line, Pearson correlation coefficients were calculated for expression data and cisplatin EC_{50}. Genes/probe sets demonstrating expression/EC_{50} correlations ($|R| > 0.85$) were subjected to biological pathway analysis using GeneGo/MetaCore™ software, and maps/pathways were identified using the GeneGo/MetaCore™ statistical test for significance ($P < 0.001$). In this way, the BAD apoptosis pathway ($P < 0.001$) was found to be associated with cell line cisplatin resistance. To build a pathway-specific PCA score, initially GeneGo/MetaCore™-defined objects (genes) within the BAD pathway (associated with OVCA cell line CIS EC_{50}) were identified. Next, for each BAD pathway object identified in this way, **all** probesets were selected and used for generation of the PCA score. For the BAD pathway, these 98 probesets represented 47 genes. Using all 98 probesets, we performed principal component analysis to reduce data dimension into a small set of uncorrelated principal components. This set of principal components was generated based on its ability to account for variation. *We used the first principal component (1st PCA), as it accounts for the largest variability in the data, as a pathway score to represent the overall expression level for the BAD pathway*. That is, pathway score= $\sum w_i x_i$, a weighted average expression among the BAD

pathway genes, where x_i represents gene *i* expression level, w_i is the corresponding weight (loading coefficient) with $\sum w_i^2 = 1$, and the w_i values maximize the variance of $\sum w_i x_i$. This approach has been used to derive a malignancy pathway gene signature in a breast cancer study (3). The BAD-pathway gene expression signature score developed in OVCA cell lines was evaluated in an independent set of 142 OVCA samples from MCC and Duke University Medical Center (OVCA 142 dataset). For the clinical-genomic OVCA dataset, log-rank test with Kaplan-Meier survival curves was used to test any association between the BAD-pathway score ("high versus "low" based upon a median value cut-off) and overall survival for patients with OVCA. NO data from the Duke/MCC samples was used to identify the 47-gene signature; the Duke/MCC ovarian dataset was a completely independent evaluation set.

Primary OVCA Patient Samples

Two clinical sample/datasets were used: **1)** genome-wide expression data from 142 patients treated at Duke and Moffitt Cancer Centers (including 114, previously reported, Dressman et al. 2007 (4), and 28 new samples), and **2)** 147 OVCA samples obtained from the University of Minnesota (UMN, n = 49) and Moffitt Cancer Center (n = 98) and analyzed by immunoflourescence for pBAD protein levels. As such, this study included analysis of data/specimens from 289 (142 + 148) OVCA patients treated at Moffitt, Duke, and UMN.

Inclusion criteria for all 289 patients (including those treated at Moffitt, Duke, UMN) included:

- a pathologically confirmed diagnosis of serous epithelial ovarian cancer,
- age >18 years,
- surgically confirmed advanced stage (III/IV) disease,
- primary surgical cytoreductive surgery prior to chemotherapy,
- primary chemotherapy with a platinum-based regimen (+/- taxane or cyclophosphamide).

Exclusion criteria for all 289 patients (including those treated at Moffitt, Duke, UMN) included:

- Non-epithelial cancer,
- borderline tumors,
- non-serous tumors,
- early stage (I/II) disease,
- absence of pathologic documentation of diagnosis,
- recurrent disease,
- receipt of neoadjuvant chemotherapy,
- unknown clinical response to primary therapy.

1) **The 142 patients treated at Duke and Moffitt Cancer Centers** for whom genomic data were analyzed in the current study had a mean age of 56 years and included 101 patients who demonstrated a CR to primary therapy and 41 who demonstrated an IR. Cytoreductive surgery was optimal for 73 patients and suboptimal for 68. The number of patients with grade 1 disease was 6, grade 2 was 61, grade 3

was 73, with grade unknown for two patients. Race data for this group included: Caucasian, 117; African-American, 18; Asian, 4; Hispanic, 1; and unknown, 2.

2) **The 147 OVCA samples obtained from UMN and Moffitt** and analyzed by immunoflourescence included 86 patients who demonstrated a CR to primary therapy and 61 who demonstrated an IR. Cytoreductive surgery was optimal for 109 patients, suboptimal for 36, and unknown for 2. The mean survival for patients who demonstrated a CR to primary therapy was 54 months, and for patients who demonstrated an IR, 28 months.

Defining Clinical response: Using medical record review, we evaluated overall survival and characterized all 289 OVCA samples as CR or incomplete responder (IR) to primary platinum-based therapy using criteria described previously (4). Clinical response to primary therapy (surgery plus platinum-based chemotherapy) was therefore established for all 289 patients using standard WHO criteria for patients with measurable disease (5). CA-125 was used to classify responses only in the absence of a measurable lesion (e.g. patients subject to optimal cytoreductive surgery); CA-125 response criteria were based on established guidelines (6, 7). A complete-response (CR) was defined as a complete disappearance of all measurable and assessable disease or, in the absence of measurable lesions, a normalization of the CA-125 level after adjuvant therapy. Patients were considered to have an incomplete-response (IR) if they demonstrated only a partial response, had stable disease, or demonstrated progressive disease during primary therapy. A partial response was considered a 50% or greater reduction in the product obtained from measurement of each bi-dimensional lesion for at least 4 weeks or a decrease in the CA-125 level by at least 50% for at least 4 weeks. Disease progression was defined as a 50% or greater increase in the product from any lesion documented within 8 weeks of initiation of therapy, the appearance of any new lesion within 8 weeks of initiation of therapy, or any increase in the CA-125 from baseline at initiation of therapy. Stable disease was defined as disease not meeting any of the above criteria. All tissues, acquired with Institutional Review Board approval, were processed as previously reported (4, 8). Microarray gene expression data (Affymetrix HG-U133A) were analyzed for 142 patients (114 samples previously reported (4) and 28 Moffitt Cancer Center (MCC) samples; GEO accession number GSE23554).

SUPPLEMENTAL REFERENCES

1. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U*, et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4: 249-64.

2. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003;19: 185-93.

3. Chen DT, Nasir A, Culhane A, Venkataramu C, Fulp W, Rubio R*, et al.* Proliferative genes dominate malignancy-risk gene signature in histologically-normal breast tissue. Breast Cancer Res Treat;119: 335-46.

4. Dressman HK, Berchuck A, Chan G, Zhai J, Bild A, Sayer R*, et al.* An integrated genomic-based approach to individualized treatment of patients with advanced-stage ovarian cancer. J Clin Oncol 2007;25: 517-25.

5. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. Cancer 1981;47: 207-14.

6. Rustin GJ, Nelstrop AE, Bentzen SM, Piccart MJ, Bertelsen K. Use of tumour markers in monitoring the course of ovarian cancer. Ann Oncol 1999;10 Suppl 1: 21-7.

7. Rustin GJ, Nelstrop AE, McClean P, Brady MF, McGuire WP, Hoskins WJ*, et al.* Defining response of ovarian carcinoma to initial chemotherapy according to serum CA 125. J Clin Oncol 1996;14: 1545-51.

8. Boren T, Xiong Y, Hakam A, Wenham R, Apte S, Chan G*, et al.* MicroRNAs and their target messenger RNAs associated with ovarian cancer response to chemotherapy. Gynecol Oncol 2009;113: 249-55.