S1. Gene Expression Omnibus (GEO) data selection:

Our results were validated using seven publicly available datasets discussed in Bentink, et al. [1] These datasets were queried from GEO (Gene Expression Omnibus), using the GEOquery extension package.[2] Using the annotation data queried by the GEOquery package, all probe sets linked to the gene symbols in our 31 gene panel were mapped. The validation analyses were conducted using five out of seven studies: GSE17260, GSE26712, GSE19161, GSE14764, and GSE18520. GSE9899 was not used because survival outcomes could not be queried. GSE13876 was not used because the number of unique patients in the queried dataset was not consistent with other publications that used this dataset.[1] For each dataset used, patients with serous carcinoma (stage IIIa, b, c, IV, and tumor grade 2 or 3 only) were selected. There was one limitation when analyzing GSE19161. A summary table in “Integrated Analysis of Multiple Microarray Datasets Identifies a Reproducible Survival Predictor in Ovarian Cancer” shows that GSE19161 contains one patient with endometroid carcinoma.[3] However, no information regarding the patients’ histology was found in the clinical data. Thus, this patient was not excluded from our analysis. A probe set associated with a gene was considered to be “significant” if a) the unadjusted p-value was less than 0.05, and b) its estimated effect size was concordant in the direction with that observed for the linked gene. All the genes that fulfilled the significance criteria are reported in Table 3. The median follow-up times for GSE17260, GSE 26712, GSE19161, GSE14764, GSE18520 were 52, 38, 37, 53, and 25 months, respectfully. No significant probe set was found in the GSE17260

data set. No significant probe set was found in the GSE19161 data set. Probe sets that were significantly differentially expressed are presented in Table 3 and are related to GSE 26712 and GSE 14764.

References Cited:

1. Bentink S, Haibe-Kains B, Risch T, Fan JB, Hirsch MS, Holton K, Rubio R, April C, Chen J, Wickham-Garcia E, Liu J, Culhane A, Drapkin R, Quackenbush J, Matulonis UA: Angiogenic mRNA and microRNA gene expression signature predicts a novel subtype of serous ovarian cancer. PLoS One 7:e30269, 2012

2. Davis S, Meltzer PS: GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics 23:1846-7, 2007

3. Konstantinopoulos PA, Cannistra SA, Fountzilas H, Culhane A, Pillay K, Rueda B, Cramer D, Seiden M, Birrer M, Coukos G, Zhang L, Quackenbush J, Spentzos D: Integrated analysis of multiple microarray datasets identifies a reproducible survival predictor in ovarian cancer. PLoS One 6:e18202, 2011