UCLA
UCLA Previously Published Works

Title
Integrated multiomic predictors for ovarian cancer survival

Permalink
https://escholarship.org/uc/item/5bm2z5f9

Journal
Carcinogenesis, 39(7)

ISSN
0143-3334

Authors
Fu, Alan
Chang, Helena R
Zhang, Zuo-Feng

Publication Date
2018-07-03

DOI
10.1093/carcin/bgy055

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

OXFORD

Carcinogenesis, 2018, Vol. 39, No. 7, 860–868

doi:10.1093/carcin/bgy055
Advance Access publication April 20, 2018
Original Article

ORIGINAL ARTICLE

Integrated multiomic predictors for ovarian cancer survival

Alan Fu[1,*], Helena R. Chang[2] and Zuo-Feng Zhang[1]

[1]Department of Epidemiology, UCLA Fielding School of Public Health, CHS 73-271, 650 Charles E. Young Dr. South, Los Angeles, CA 90095, USA and [2]Department of Surgery, Revlon/UCLA Breast Center, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095, USA

*To whom correspondence should be addressed. Tel: +1 203 343 1066; Fax: +1 203 343 1066; Email: alan.nuo.fu@gmail.com

Abstract

Increasingly affordable high-throughput molecular profiling technologies have made feasible the measurement of omics-wide interindividual variations for the purposes of predicting cancer prognosis. While multiple types of genetic, epigenetic and expression changes have been implicated in ovarian cancer, existing prognostic biomarker strategies are constrained to analyzing a single class of molecular variations. The extra predictive power afforded by the integration of multiple omics types remains largely unexplored. In this study, we performed integrative analysis on tumor-based exome-, transcriptome- and methylome-wide molecular profiles from The Cancer Genome Atlas (TCGA) for variations in cancer-relevant genes to construct robust, cross-validated multiomic predictors for ovarian cancer survival. These integrated polygenic survival scores (PSSs) were able to predict 5-year overall (OS) and progression-free survival in the Caucasian subsample with high accuracy (AUROC = 0.87 and 0.81, respectively). These findings suggest that the PSSs are able to predict long-term OS in TCGA patients with accuracy beyond that of previously proposed protein-based biomarker strategies. Our findings reveal the promise of an integrated omics-based approach in enhancing existing prognostic strategies. Future investigations should be aimed toward prospective external validation, strategies for standardizing application and the integration of germline variants.

Introduction

Of the gynecologic cancers, ovarian cancer remains the most lethal, owing, in part, to the advanced clinical stage at presentation (1). The heterogeneity in disease outcomes, for which there is a strong biological component, prompts the need for methods that can reliably predict prognosis. Studies in the past have identified a number of biomarker strategies with prognostic value. The first and best known of these is the serum-based glycoprotein CA125 (1,2), which, despite providing limited sensitivity and specificity, is generally accepted to have clinical utility in predicting patient survival and response to chemotherapy (3,4). Since the 1981 identification of CA125, a number of other prognostic biomarkers have been identified. However, individually, none of these biomarkers have been able to surpass CA125 in performance, and the pool of clinically useful markers remains pitifully small (1,5).

As rapidly maturing high-throughput technologies become affordable for clinical use, a natural course of action is to look for strategies to leverage the prognostic power of combinations of molecular markers. While the prognostic value of these markers may be weak individually, their integration in the form of marker panels may yet improve upon existing single marker-based strategies. A number of studies have been proposed promising serum- and tumor-based biomarker panels for ovarian cancer survival (2,6–10) (Table 1). However, these investigations were restricted to single-type analyses of protein-, mRNA- and methylation-level signatures, and the prognostic potential of integrating of multiple levels of molecular variation data remains largely unexplored.

In the current study, we construct and evaluate a polygenic scoring system, designed to maximize the likelihood of robust predictions, for ovarian cancer prognosis based on the integrative analysis of tumor-based exome-, transcriptome- and methylome-wide molecular profiles from 488 high-grade ovarian serous cystadenocarcinoma patients in The Cancer Genome

Received: May 2, 2017; **Revised:** March 18, 2018; **Accepted:** April 13, 2018

© The Author(s) 2018. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com.

Abbreviations

AIC	Akaike information criterion
AUC/AUROC	area under the receiver operating characteristic
FDR	false discovery rate
OS	overall survival
PFS	progression-free survival
PSSs	polygenic survival scores
ROC	receiver operating characteristic
TCGA	The Cancer Genome Atlas

Atlas (TCGA). The prognostic performance of our multi-omics polygenic survival scores (PSSs) suggests superior predictive accuracy compared with previously proposed biomarker-based strategies, motivating further investigative efforts toward external validation and clinical standardization.

Materials and methods

A workflow of the methods we employed can be found in Figure 1. All analyses were performed using SAS (v9.3) (SAS Institute) unless stated otherwise.

Ovarian cancer patient characteristics in the TCGA database

Our analyses were based on tumor molecular variation, demographic and clinical data of 488 high-grade ovarian serous cystadenocarcinoma patients who were included for analysis in the original TCGA publication (13). This data is hosted by TCGA in an online repository (https://tcga-data.nci.nih.gov/docs/publications/ov_2011/). Select demographic and clinical characteristics of the patient sample are given in Table 2. Of the 488 patients, 380 (77.9%) were documented as being of Caucasian descent and 82 (16.8%) were documented as being of other races, while 26 (5.3%) patients had no data on race. At the time of diagnosis, the average age for the overall patient sample was 60.2 and 60.8 years for the Caucasian subsample. A total of 268 (54.9%) patients in the overall sample and 221 (58.2%) patients in the Caucasian subsample had died of any cause by the end of follow-up. Three hundred and forty-nine (71.5%) patients in the overall sample and 272 (71.6%) patients of the Caucasian subsample were identified as having progressed or recurring disease at the end of follow-up with median follow-up times of 30.26 and 30.05 months, respectively. Of the overall sample, 244 (50%) patients experienced disease progression or recurrence before death, while 199 (52.4%) of the Caucasian patients experienced both progression or recurrence and death. No statistical differences between Caucasians and the overall patient population were found for any patient characteristics.

TCGA molecular variation data

In addition to clinical and demographic information, the TCGA repository hosts data on tumor-specific molecular variations, including somatic mutations and variations in DNA methylation, mRNA expression and miRNA expression. The details of the biospecimen collection and processing and molecular variation measurement processes can be found in the Supplementary Information, available at *Carcinogenesis* Online, of the original TCGA publication. In brief, biospecimens were collected from patients with newly diagnosed ovarian serous cystadenocarcinoma who were undergoing surgical resection and had received no prior chemotherapy or radiotherapy. Each frozen tumor specimen was paired with a companion 'normal' tissue specimen, which could be adjacent normal tissue, peripheral lymphocytes or previously extracted germline DNA. Each case of ovarian serous adenocarcinoma was histologically confirmed by a board-certified pathologist. The AllPrep DNA/RNA mini kit (Qiagen) was used to isolate DNA and RNA fractions from tissue. Exome sequencing was performed for tumor and matched normal tissues by The Genome Center at Washington University, the Broad Institute and the Human Genome Center at Baylor College of Medicine. Somatic mutation calls were performed using an automated pipeline and subsequently validated and annotated, as described in the original TCGA publication. Only functional mutations, defined as missense mutations most likely to generate functional changes that enhance tumor proliferation, as identified by the Cancer-specific High-throughput Annotation of Somatic Mutations (CHASM) method, were retained for our analyses. CpG-site methylation variations were accessed using the Illumina Infinium HumanMethylation27 platform, as described in the original publication. For the purposes of this study, only CpG sites within the promoter region, 5′ UTR, and first exon were evaluated. Tumor mRNA expression levels were measured using the Agilent 244K Whole Genome Expression, Affymetrix HT-HG-U133A and Affymetrix HuEx arrays. Normalized cross-platform expression measurements were combined to create a unified standardized expression set for 11864 genes. miRNA expression in tumor tissues was measured using the Agilent 8 × 15K Human array and subsequently quantile normalized, and log2 transformed using the Subio Platform (v1.18) (Subio).

Selection of candidate genes

In total, we analyzed 451 genes and the miRNAs known to target them (Supplementary Table I, available at *Carcinogenesis* Online). These genes were selected from two sources: the original TCGA ovarian cancer report (13) and Vogelstein *et al.* (14). The TCGA gene set is composed of genes with functional somatic mutations, copy number variations or mRNA expression that formed a part of a transcriptional signature associated with overall survival (OS). The Vogelstein gene set consists of validated driver oncogenes and tumor suppressors. miRNAs targeting the TCGA and Vogelstein gene sets were identified from experimentally validated miRNA–gene interactions cataloged in the starBase (v2.0) noncoding RNA interactions browser (http://starbase.sysu.edu.cn). Specifically, miRNA–gene interactions are considered validated if they are confirmed by two or more independent chromatin-linked immunoprecipitation experiments. In all, TCGA molecular variation data was available for 451 genes selected based on the above-mentioned criteria.

Endpoints

The definitions for OS and progression-free survival (PFS) we used are identical to those used in the original TCGA publication (13). In brief, OS was defined for each patient as the interval from the date of initial surgical resection to the date of death or last known contact with the patient. Progression-free time was defined as the interval from the date of initial surgical resection to the date of progression or recurrence or the date of the last known contact. Patients who had died but whose progression or recurrence statuses are unknown were excluded from our analyses of PFS.

Table 1. Previously proposed prognostic biomarker strategies for ovarian cancer

Strategy	Study	Sample size	Endpoint(s)	Accuracy	Validation
Longitudinal serum CA125	Mano *et al.* (11)	92	1-, 3-, 5-Y OS	0.67, 0.75, 0.73	None
Longitudinal serum CA125	Chiang *et al.* (12)	218	3-Y OS	0.85	Cross
Panel: tumor baseline CA125, KLKs, Spondin-2	Oikonomopoulo *et al.* (8)	98	1-Y OS	0.65	Cross
Panel: tumor protein signature	Riester *et al.* (10)	1525	5-Y OS	0.62	External
Panel: tumor baseline KLKs	Oikonomopoulo *et al.* (8)	98	1-Y PFS	0.75	Cross
Panel: tumor KLKs + clinical, including chemo response	Zheng *et al.* (9)	259	1-Y, 5-Y PFS	0.80, 0.88	Cross

KLK, kallikrein.

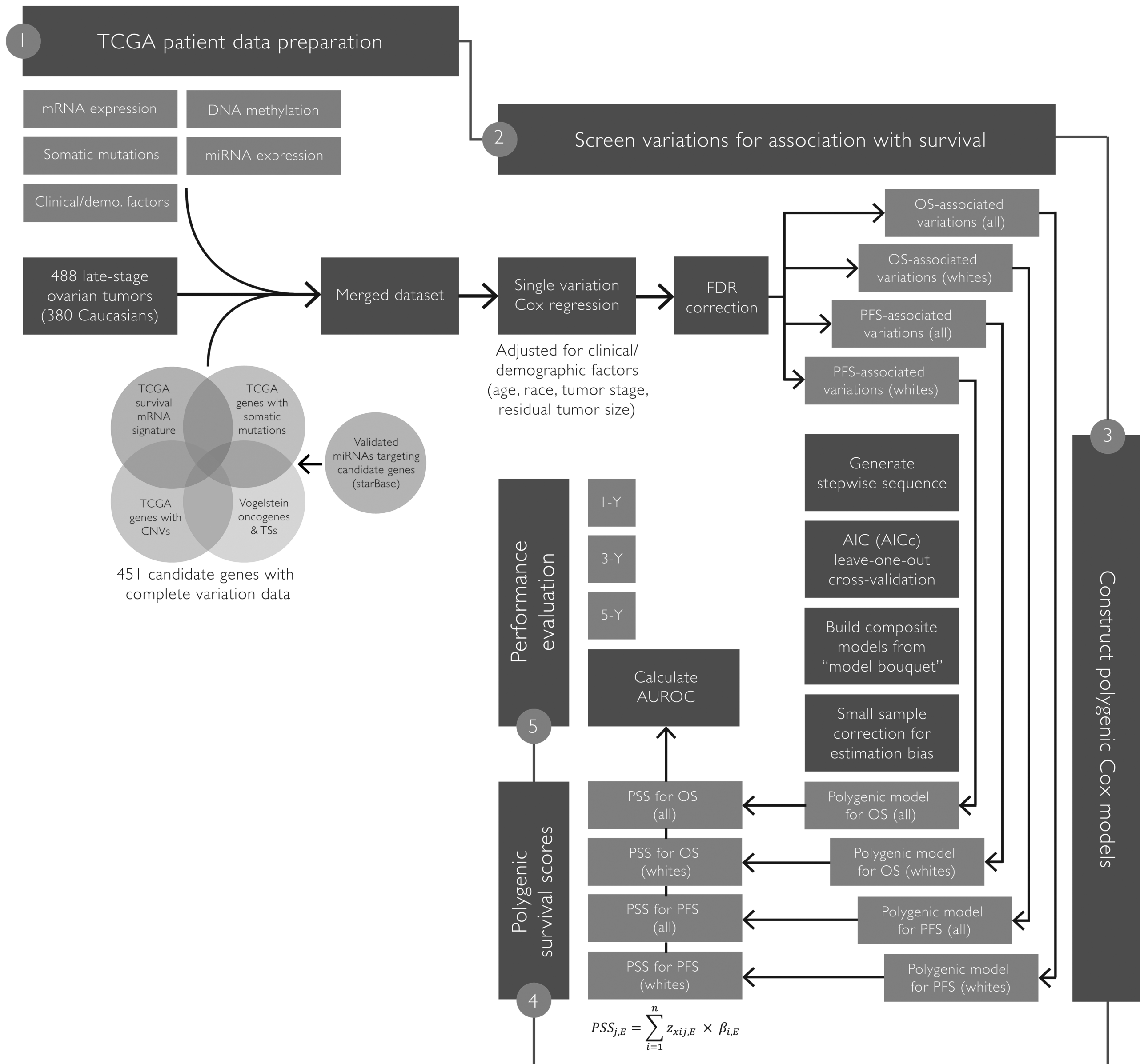

Figure 1. A workflow of the methods we used to build and evaluate the PSS.

Screening individual variations for survival association

Somatic mutation, mRNA expression and promoter DNA methylation data for the 451 candidate genes and miRNA expression data for the miRNAs targeting them were individually evaluated for association with patient survival. This was done by individually fitting each variation into a minimally adjusted Cox regression model with covariates for age at diagnosis (continuous), tumor stage, residual tumor size following primary surgical resection and self-reported race. Adjusted hazard ratios were calculated as a measure of association between each of the variations and OS or PFS for both the total patient population and the Caucasian subsample. As the frequencies of certain somatic mutations are low, rather than regressing individual somatic mutations, an ordinal gene-level variable was created that reflects the total number of functional mutations each patient has for a particular gene. This method enhanced our statistical power and reduced the total number of variations to be analyzed. mRNA and miRNA expression were treated as continuous variables, and hazard ratios were calculated based on one unit increments of normalized microarray signal intensities. Similarly, methylation was individually analyzed at each promoter, 5′ UTR, and first exon CpG site as a continuous variable with a range from 0 (completely unmethylated) to 1 (completely methylated). To reduce the number of potential false positive results stemming from multiple testing as well as the number of variations retained for subsequent analyses, the Benjamini–Hochberg method for controlling the false discovery rate (FDR) (15) was used to generate an FDR-adjusted *P*-value (q-value) for each variation. Only variations exhibiting a q-value < 0.10 were considered to be associated with survival and retained for further analyses.

Polygenic Cox model construction

We used the model building procedure described by Shtatland *et al.* (16). to select predictors from the pool of survival-associated variations to construct endpoint-specific polygenic Cox models. In brief, this procedure consists of four distinct steps designed to reduce estimation bias and generate models capable of robust predictions, performed using the SAS (v9.3) statistical analysis software.

First, we applied endpoint-specific stepwise Cox regression to FDR-qualified variations from the single variation analyses, as well as age at diagnosis, tumor stage and residual tumor size with critical *P*-values of

Table 2. Select demographic and clinical characteristics of TCGA ovarian cancer patients

Characteristic	Total population (*N* = 488)	Caucasians (*N* = 380)	*P*-value*
Age at diagnosis, mean ± SD	60.22 ± 11.43	60.80 ± 11.31	1.00
Tumor stage, *N* (%)			
IIA	3 (0.61)	2 (0.53)	0.93
IIB	4 (0.82)	3 (0.79)	
IIC	17 (3.48)	7 (1.84)	
IIIA	7 (1.43)	6 (1.58)	
IIIB	21 (4.30)	14 (3.68)	
IIIC	353 (72.34)	280 (73.68)	
IV	79 (16.19)	65 (17.11)	
Unknown/missing	4 (0.82)	3 (0.79)	
Residual tumor size, *N* (%)			
No macroscopic disease	90 (18.44)	62 (16.32)	0.61
1–10 mm	223 (45.70)	192 (50.53)	
11–20 mm	30 (6.15)	20 (5.26)	
>20 mm	89 (18.24)	70 (18.42)	
Unknown/missing	56 (811.48)	36 (9.47)	
Progression, *N* (%)			
Yes	349 (71.52)	272 (71.58)	0.46
No	137 (28.07)	108 (28.42)	
Unknown/missing	2 (0.41)	0 (0.00)	
Vital status, *N* (%)			
Alive	215 (44.06)	156 (41.05)	0.61
Deceased	268 (54.92)	221 (58.16)	
Unknown/missing	5 (1.02)	3 (0.79)	
Progression-associated survival, *N* (%)			
Progression + deceased	244 (50.00)	199 (52.37)	0.86
No progression + alive	112 (22.95)	85 (22.37)	
No progression + deseased	24 (4.92)	22 (5.79)	
Progression + alive	103 (21.11)	71 (18.68)	
Unknown/missing	5 (1.11)	3 (0.79)	
Follow-up months, median (IQR)	30.26 (12.10–47.57)	30.05 (12.35–47.47)	1.00
Chemotherapy, *N*[a]			0.85
Intraperitoneal (IP)	59	50	
Intravenous (IV)	411	331	
Oral	29	27	
Other	2	2	
Unknown/missing	104	72	

[a]Some patients were administered chemotherapy through multiple routes.
*P-values derived from *t*-test for continuous variables and chi-square test for categorical variables.

SLENTRY and SLSTAY close to 1 (0.99 and 0.995, respectively) to build a complete stepwise sequence starting with the null model with no predictors and ending with the full model consisting of all potential predictors. The models in this sequence were ordered in a way that maximized the increment in likelihood at every step. Using this method, in the case of *k* potential predictors, we would have *k* potential candidate models rather than the 2^k models in a typical stepwise regression procedure:

$$\begin{array}{ll} 0 & \text{Null} \\ 1 & \beta_1 x_1 \\ 2 & \beta_1 x_1 + \beta_2 x_2 \\ \ldots & \ldots \\ k & \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \end{array}$$

We also obviate the need to choose an arbitrary critical *P*-value cutoff. The result is a procedure that generates a list of candidate models for model selection (Step 2) without torturing the data, an important step toward building robust models when working with a large number of potential predictors, as in our case.

The second step involved calculating the Akaike Information Criterion (AIC) of each of the *k* models in the full stepwise sequence. AIC is asymptotically equivalent to leave-one-out cross-validation (16,17), but when the sample size *n* is small, AIC can lose its asymptotic optimal properties and become severely biased. In such cases, a corrected AIC, the AIC_c is applied:

$$\text{AIC}_c = \text{AIC} + \frac{2k(k+1)}{(n-k-1)}$$

Where *n* is the sample size and *k* is the number of potential predictors. The use of AIC_c also protects against overfitting, as a penalty is imposed on models with higher *k*. The AIC_c was applied to our models for OS, where the ratio *n*/*k* was smaller than the recommended threshold of 40 (18).

In the third step, we identified a set of the top AIC-optimized models for each endpoint. Visually, these 'model bouquets' represent the cluster of models with the lowest AICs (or AIC_cs) for each endpoint. As a further measure to increase model stability, we built our final endpoint-specific polygenic Cox models by retaining all predictors from the endpoint-specific model bouquets.

In the fourth and final step, a shrinkage factor was imposed to reduce the magnitude of the regression coefficients for the polygenic models for OS as a further measure against estimation bias:

$$l_{\text{AIC}} = \frac{\text{AIC}(M) - \text{AIC}(0) - k}{\text{AIC}(M) - \text{AIC}(0)}$$

Where AIC(*M*) is the AIC of the full model, and AIC(0) is the AIC of the null model.

Polygenic survival score

A system was created to score patients based on the molecular and clinical/demographic predictors in the polygenic Cox models. The PSS represents a continuous predictor of survival; the higher a patient's PSS, the poorer the predicted survival. Values for the predictors (tumor variations + age, tumor stage, residual tumor size) in each polygenic Cox model were first converted to z-scores and regressed together in endpoint-specific z-score-standardized polygenic Cox models. Afterward, PSSs were calculated by summing the products of patients' z-score-standardized predictor values × their respective coefficients from the endpoint-specific polygenic Cox models:

$$\mathrm{PSS}_{j,E} = \sum_{i=1}^{n} z_{xij,E} \times \beta_{i,E}$$

Where $\mathrm{PSS}_{j,E}$ is the PSS for the *j*th patient for endpoint *E*, z_{xi} is the z-score of the ith predictor, and β_i is the regression coefficient for z_{xi}. In all, four types of PSSs were calculated: one for OS in the total patient sample, one for OS in the Caucasian subsample, one for PFS in the total patient sample and one for PFS in the Caucasian subsample.

Performance evaluation

The ability of the PSSs to predict survival was evaluated by ROC analysis using the procedure described by Lu *et al*. (19). Additional scores were constructed from baseline clinical characteristics (age at diagnosis, tumor stage, and/or residual tumor size) only or from baseline clinical characteristics + single molecular variations from the polygenic Cox models. Model stability and prediction optimism were evaluated using the bootstrap method with 10000 resamples of the training sets.

Results

Associations between individual variations and survival

Each candidate marker was individually fitted into a minimally adjusted Cox proportional hazards model adjusted for age at diagnosis, tumor stage and residual tumor size following primary resection. A graphical overview of the molecular variations individually associated with OS or PFS following FDR adjustment is found in Supplementary Figure 1, available at *Carcinogenesis* Online.

Somatic mutations

Our analysis of somatic mutations was restricted to genes in the TCGA dataset with validated mutations predicted to be functional *in silico*. Mutations in *ABCA3* and *NIPBL* were found to be associated with poor OS in both the total TCGA patient population and the Caucasian subsample (Supplementary Figure 2A, available at *Carcinogenesis* Online). Only mutations in the tumor suppressor *CREBBP* were linked to PFS. *CREBBP* mutations were associated with decreased survival in both the total patient sample and the Caucasian subsample.

mRNA expression

Our analysis incorporated genes that formed a transcriptional profile predictive of OS in the original TCGA publication (13). Of these, 33 were found to be individually associated with OS in the total patient population, and 40 were associated with OS in the Caucasian subsample (Supplementary Figures 2B and D, available at *Carcinogenesis* Online). Included among these genes are a number of canonical oncogenes and tumor suppressors, including *BRCA2*, *EGFR* and *RB1*. No genes were identified to be associated with PFS in the total patient population. Only the expression of *GALNT6* was found to be associated with PFS in Caucasians.

miRNA expression

Of the miRNAs that target one or more of the protein-coding genes in our analysis, miR-198 and miR-422a, both putative tumor suppressors, were associated with decreased OS in the total sample. miR-198 was also associated with decreased OS in the Caucasian subsample. miR-519e-5p was associated with decreased OS in both the total sample and the Caucasian subsample, while miR-518c-5p was associated with poor OS only in the Caucasian subsample (Supplementary Figure 2C, available at *Carcinogenesis* Online). Expressions of miR-302-5p and the putative tumor suppressor miR-449a were associated with PFS in both the total sample and the Caucasian subsample (Supplementary Figure 2C, available at *Carcinogenesis* Online).

DNA methylation

No FDR-adjusted associations with either OS or PFS were identified for variations in methylation.

Polygenic scores for ovarian cancer survival

A model building procedure, designed to maximize the likelihood of robust predictions, was applied to select subsets of variations from the total pool of survival-associated variations. These subsets were used in combination with age, tumor stage and residual tumor size information to calculate endpoint-specific PSS for each patient. The molecular components of the PSSs and the functional role of the genes harboring them are given in Supplementary Table II, available at *Carcinogenesis* Online. In addition, a descriptive summary of the PSSs with regard to their correlation with individual-level survival, as well as the relative impact of the individual PSS components on survival, is given in Figure 2.

We observed a decrease in median OS and PFS duration with each successive increase in PSS tertile (log-rank $P < 0.0001$) (Figure 3A and Supplementary Figure 3, available at *Carcinogenesis* Online). For OS in the total patient population, we observed an 18.9-month decrease in median survival from the lowest to the middle tertile (63.8 versus 44.9 months) and an additional 18-month decrease going from the middle to highest tertile (44.9 versus 26.9 months). This pattern was replicated in the Caucasian subsample. We observed the similar tertile-dependent decline for PFS, although the intertertile decrease is less marked (Figure 3A and Supplementary Figure 3, available at *Carcinogenesis* Online).

We performed ROC analysis to evaluate the prognostic performance of the PSSs. The accuracy of the PSSs to predict survival beyond a certain timepoint generally increased with respect to the duration of time following resection. We observed good predictive accuracy for 5-year OS in the total patient sample (AUC = 0.80) and the Caucasian subsample (AUC = 0.87) and for 5-year PFS in the Caucasian subsample (AUC = 0.81) (Figure 3B and C). In contrast, hazard scores constructed from baseline clinical characteristics (age, tumor stage and/or residual tumor size) alone offered poor predictions ($0.60 \le \mathrm{AUC} < 0.70$) at the 1-year mark and no predictive value at the 5-year mark (Figure 3B). The accuracy of the PSS was poor for PFS in the total sample for all timepoints ($0.6 \le \mathrm{AUC} < 0.7$) and failed to outperform miR-302c at the 5-year mark (Supplementary Figures 4 and 5, available at *Carcinogenesis* Online).

Prediction optimism was quantified for each model using bootstrapped resamples of its respective training set. The optimism for predictive accuracy, as measured by Somer's D, was minimal (Supplementary Table IV, available at *Carcinogenesis* Online).

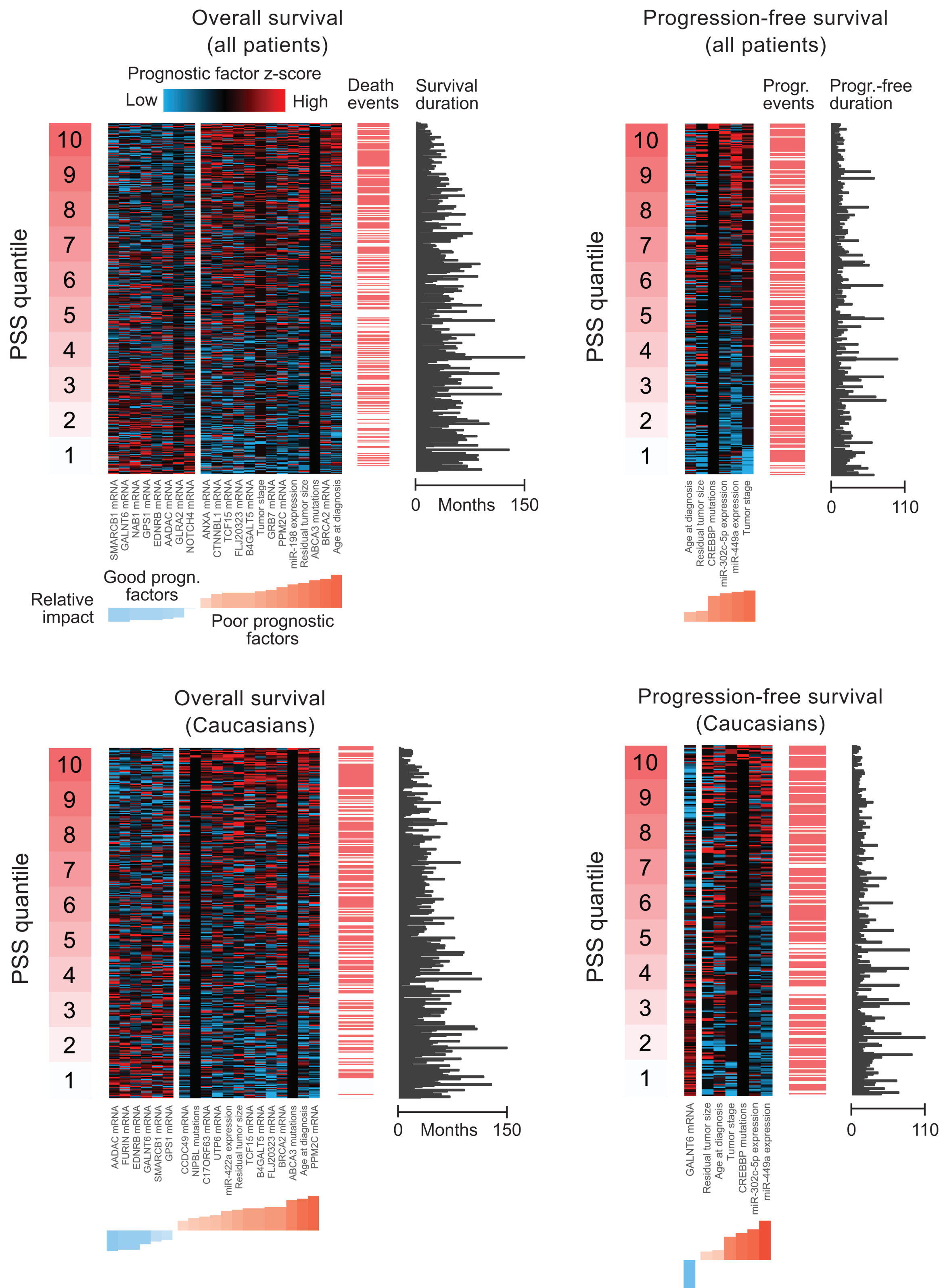

Figure 2. Correlation between PSS and z-scores of the individual PSS components and patient-level survival. As expected, endpoint-specific PSSs were positively correlated with the z-scores of protective component prognostic factors and negatively correlated with poor component prognostic factors. Similarly, endpoint-specific PSSs were positively correlated with death and progression events and negatively correlated with survival and progression-free duration. The relative impact of the component prognostic factors were calculated from the coefficients of z-score-standardized predictors in the endpoint-specific polygenic Cox models.

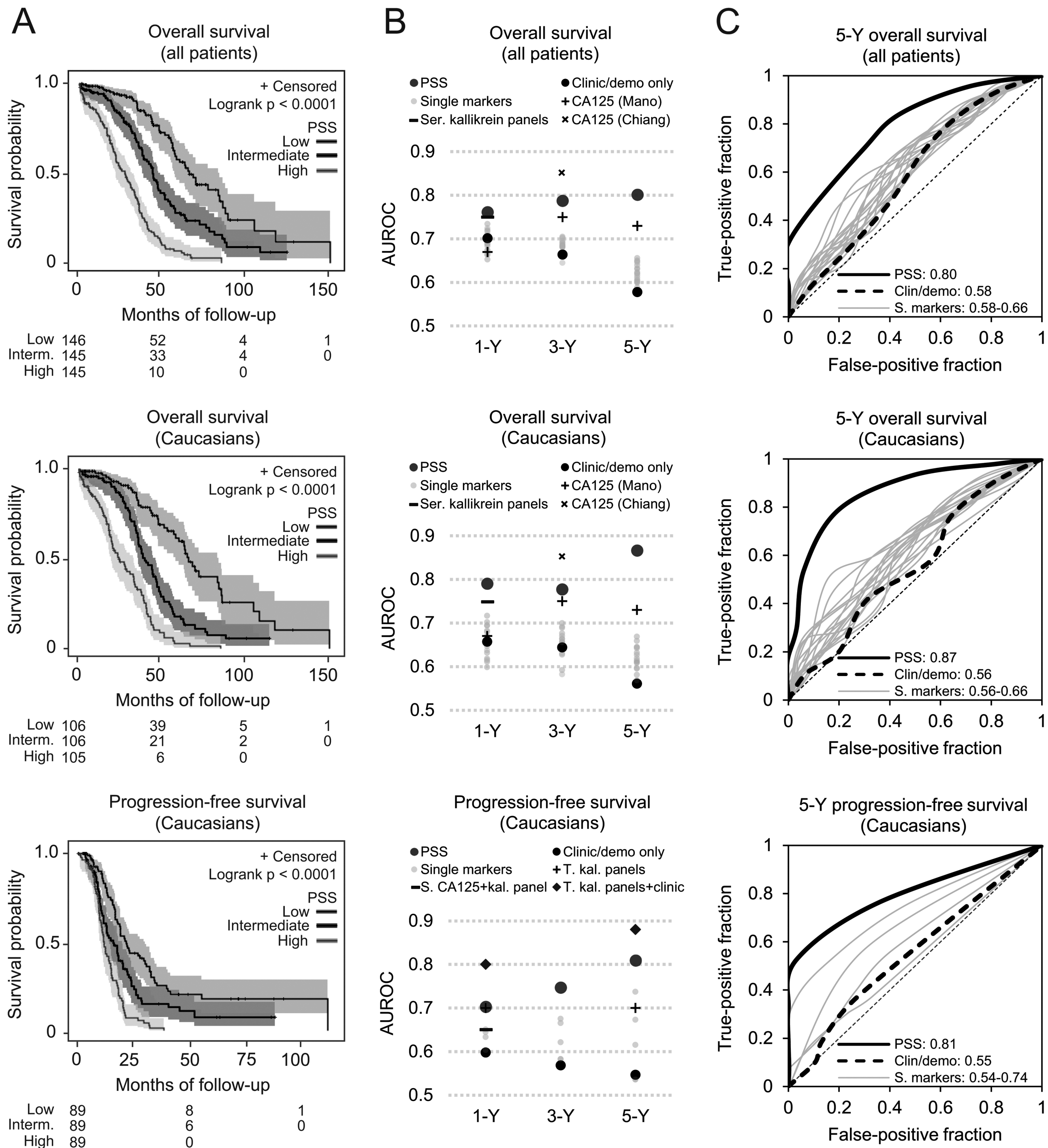

Figure 3. (**A**) Kaplan–Meier curves stratified by PSS tertiles. A decrease in median OS and PFS duration was observed with each successive increase in PSS tertile. (**B**, **C**) Prognostic performance of our PSSs. We observed good predictive accuracy for 5-year OS in the total (AUC = 0.80) and Caucasian (AUC = 0.87) samples and for 5-year PFS in the Caucasian subsample (AUC = 0.81). Age, tumor stage, and residual tumor size following primary resection ('Clinic/demo only'/'Clin/demo'; red) provided no predictive advantage for 5-year OS and PFS. The AUCs of single component markers in combination with age, tumor, stage and residual tumor size ('Single markers'/'S. markers') are given in gray. The performance of our PSSs appears to be superior to existing CA125- and kallikrein-based biomarker strategies (black markers) for long-term survival but fall short of a protein panel based primarily on tumor kallikreins combined with select clinical characteristics ('T. kal. panels + clinic'; black diamonds), which included responsiveness to chemotherapy, for PFS.

Discussion

CA125 serum levels and decay kinetics remain the highest performing individual prognostic biomarkers for ovarian cancer (1,2). An analysis of 92 ovarian cancer patients diagnosed at Coimbra University Hospital in Portugal found successive postdiagnosis measurements of serum CA125 to predict 1-, 3- and 5-year OS with AUROCs of 0.67, 0.75 and 0.73, respectively (11) (Figure 3B and Supplementary Figure 4, available at *Carcinogenesis* Online). A more recently proposed longitudinal model combining serum CA125 levels with tumor stage and

residual tumor size was able to predict 3-year OS with a cross-validated AUROC of 0.85 in 218 patients (12). Efforts have also been made to incorporate serum and tumor protein signatures into prognostic marker panels. Protein panels of serum CA125, kallikreins and/or Spondin-2 were shown to predict 1-year OS and PFS with cross-validated AUROCs of 0.75 and 0.65, respectively, based on 98 ovarian cancer patients (8). A recent meta-analysis of tumor protein signatures across 13 studies and 1525 patients identified an expression signature capable of predicting 5-year OS with a cross-validated AUROC of 0.62 (10). Additionally, a protein panel based primarily on tumor kallikreins combined with select clinical characteristics, which included responsiveness to chemotherapy, yielded cross-validated AUROCs of 0.80 and 0.88 for 1- and 5-year PFS in a sample of 259 patients (Figure 3B and Supplementary Figure 4, available at *Carcinogenesis* Online) (9).

The cross-validated AUROCs of our PSSs for long-term OS appear to exceed those reported for existing protein-based strategies. It is conceivable that our tumor-based PSSs could be applied alongside or integrated with existing strategies to help refine prognostic accuracy, a scenario that is made more feasible by the ready availability of tumor tissues, as tumor debulking remains the main first-line treatment for ovarian cancer. While meticulous external validation is necessary before more concrete inferences on clinical utility can be made, our study offers an important first look into the prognostic value of a multi-omics approach in ovarian cancer prognosis.

Aside from CA125 (MUC16), preoperative HE4 (WFDC2) protein levels in the serum have also been shown to have independent prognostic value (20–23). However, we found no associations between MUC16 and WFDC2 molecular signatures, including mRNA expression, and OS or PFS, and incorporation of these signatures into our integrated models failed to improve their discriminative performance (Supplementary Table III, available at *Carcinogenesis* Online). Given that the prognostic value of CA125 and HE4 has been validated in independent samples, our findings reveal an interesting if unsurprising disjunction between the protein and pretranslational signatures of MUC16 and MFDC2.

Considering no other study has performed the multilayered molecular profiling to the depth achieved by TCGA, true external validation of the PSSs is not possible. This limitation is not a trivial one, as it is well known that predictor and model selection for regression can be unstable and sensitive to small changes in data, particularly if the number of predictors is large relative to the sample. The underlying cause, data idiosyncrasy, also results in what Ioannidis (24) refers to as type B biomarker failures—failures of external validation. To offset this disadvantage, we took several precautionary measures to reduce estimate bias and increase the likelihood of robust predictions, including shunning the use of standard stepwise predictor selection in favor of building composite models from 'bouquets' of near AIC-optimal candidate models and a two-step small sample correction for overfitting and bias. Bootstrapping the training sets revealed minimal optimism bias, suggesting predictive stability given subtle changes in the data structure.

An additional roadblock to clinical translation stems from the difficulty in developing a standardized protocol for clinical implementation, particularly when working with gene expression profiles (24). Our PSSs utilized gene transcriptional profiles from a combination of Affymetrix and Agilent microarrays. While Agilent two-channel arrays can be individually normalized via the LOWESS method, expression signatures from Affymetrix single-channel arrays are normalized across arrays, making the analysis of new Affymetrix patient arrays using standard normalization methods problematic. To enable the prospective analysis of individual single-channel arrays, frozen robust multiarray analysis, which leverages a 'frozen' reference signal distribution from microarray samples in the GEO and ArrayExpress repositories, was proposed as an alternative to true across-sample normalization (25).

The discordance of gene expression measurements across different technologies represents another potential hindrance to implementation, with the major concern being whether valid comparisons could be made between prospective expression measurements and the original measurements used to train the PSSs when the platforms are different. The likelihood of disagreement, however, is reduced when we take into consideration the concordance measures that were used in generating the unified TCGA expression set and findings, which suggest generally good concordance among commonly used array (26,27) and RNA-seq platforms (28).

To our knowledge, our analysis represents the first investigation of the prognostic potential of an integrative multi-omics approach in ovarian cancer. Multiple measures were taken to maximize the predictive robustness of our integrated PSSs, which appear to exceed previously proposed biomarker strategies in prognostic performance for long-term OS. Our findings warrant further investigative pursuit with regard to external validation, the standardization of application, particularly with respect to the measurement of transcriptional profiles, and the predictive benefit of integrating germline variations.

Funding

This research was supported by the US National Institutes of Health (Grant Numbers T32 CA09142) and Alper Research funds for Environmental Genomics.

Acknowledgements

The authors gratefully acknowledge members of the TCGA Research Network for their role in the generation of data used in this study and for their promotion of open access data sharing within the scientific community. The authors would also like to thank Dr Shen-Chih Chang and Ms Aileen Baecker of the Department of Epidemiology at the UCLA Fielding School of Public Health for their contributions.
Conflict of Interest Statement: None declared.

Supplementary material

Supplementary data can be found at *Carcinogenesis* online.

References

1. Huang, J. *et al*. (2010) Prognostic biomarkers in ovarian cancer. Cancer Biomark., 8, 231–251.
2. Kulasingam, V. *et al*. (2010) Integrating high-throughput technologies in the quest for effective biomarkers for ovarian cancer. Nat. Rev. Cancer, 10, 371–378.
3. Jacobs, I. *et al*. (1989) The CA 125 tumour-associated antigen: a review of the literature. Hum. Reprod., 4, 1–12.
4. Rustin, G.J. *et al*. (2004) Use of CA-125 in clinical trial evaluation of new therapeutic drugs for ovarian cancer. Clin. Cancer Res., 10, 3919–3926.
5. Høgdall, E. (2008) Cancer antigen 125 and prognosis. Curr. Opin. Obstet. Gynecol., 20, 4–8.
6. Spentzos, D. *et al*. (2004) Gene expression signature with independent prognostic significance in epithelial ovarian cancer. J. Clin. Oncol., 22, 4700–4710.

7. Su, H.Y. *et al*. (2009) An epigenetic marker panel for screening and prognostic prediction of ovarian cancer. Int. J. Cancer, 124, 387–393.
8. Oikonomopoulou, K. *et al*. (2008) Prediction of ovarian cancer prognosis and response to chemotherapy by a serum-based multiparametric biomarker panel. Br. J. Cancer, 99, 1103–1113.
9. Zheng, Y. *et al*. (2007) A multiparametric panel for ovarian cancer diagnosis, prognosis, and response to chemotherapy. Clin. Cancer Res., 13, 6984–6992.
10. Riester, M. *et al*. (2014) Risk prediction for late-stage ovarian cancer by meta-analysis of 1525 patient samples. J. Natl. Cancer Inst., 106. doi:10.1093/jnci/dju048; https://academic.oup.com/jnci/article/106/5/dju048/2606959.
11. Mano, A. *et al*. (2005) CA-125 AUC as a new prognostic factor for patients with ovarian cancer. Gynecol. Oncol., 97, 529–534.
12. Chiang, A.J. *et al*. (2014) A longitudinal analysis with CA-125 to predict overall survival in patients with ovarian cancer. J. Gynecol. Oncol., 25, 51–57.
13. Cancer Genome Atlas Research Network. (2011) Integrated genomic analyses of ovarian carcinoma. Nature, 474, 609–615.
14. Vogelstein, B. *et al*. (2013) Cancer genome landscapes. Science, 339, 1546–1558.
15. Benjamini, Y. *et al*. (2001) Controlling the false discovery rate in behavior genetics research. Behav. Brain Res., 125, 279–284.
16. Ernest, S. *et al*. (2004) A new strategy of model building in proc logistic with automatic variable selection, validation, shrinkage and model averaging. SUGI 29 Proceedings.
17. Fang, Y. (2011) Asymptotic equivalence between cross-validations and Akaike information criteria in mixed-effects models. J. Data Sci., 9, 15–21.
18. Wagenmakers, E.J. (2003) Model selection and multimodel inference: a practical information-theoretic approach. J. Math Stat. Psychol., 47, 580–586.
19. Li Lu, C.L. (2006) Using the time dependent ROC curve to build better survival Model in SAS. NESUG 2006.
20. Steffensen, K.D. *et al*. (2011) Prognostic impact of prechemotherapy serum levels of HER2, CA125, and HE4 in ovarian cancer patients. Int. J. Gynecol. Cancer, 21, 1040–1047.
21. Kalapotharakos, G. *et al*. (2012) High preoperative blood levels of HE4 predicts poor prognosis in patients with ovarian cancer. J. Ovarian Res., 5, 20.
22. Bandiera, E. *et al*. (2011) Serum human epididymis protein 4 and risk for ovarian malignancy algorithm as new diagnostic and prognostic tools for epithelial ovarian cancer management. Cancer Epidemiol. Biomarkers Prev., 20, 2496–2506.
23. Kong, S.Y. *et al*. (2012) Serum HE4 level is an independent prognostic factor in epithelial ovarian cancer. Ann. Surg. Oncol., 19, 1707–1712.
24. Ioannidis, J.P. (2013) Biomarker failures. Clin. Chem., 59, 202–204.
25. McCall, M.N. *et al*. (2010) Frozen robust multiarray analysis (fRMA). Biostatistics, 11, 242–253.
26. Petersen, D. *et al*. (2005) Three microarray platforms: an analysis of their concordance in profiling gene expression. BMC Genomics, 6, 63.
27. Yauk, C.L. *et al*. (2004) Comprehensive comparison of six microarray technologies. Nucleic Acids Res., 32, e124.
28. Su, Z.Q. *et al*. (2014) A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. Nat. Biotechnol., 32, 903–914.