

Development and validation of a DNA methylation signature to predict the progression-free survival in patients with advanced-stage ovarian cancer

Suzhen Fan

Zhengzhou University

Xiang Gao

Zhengzhou University

Qiaohong Qin

Zhengzhou University

Hongyu Li

Zhengzhou University

Zhongfu Yuan

Zhengzhou University

Shujun Zhao (✉ zhaoshujun0710@163.com)

Zhengzhou University Third Hospital and Henan Province Women and Children's Hospital

<https://orcid.org/0000-0003-4595-5954>

Research

Keywords: DNA methylation, ovarian cancer, progression-free survival, prognostic biomarker

Posted Date: September 8th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-71465/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Ovarian cancer (OC) is a disease characterized by late-stage presentation and poor prognosis. Our aim was to identify a DNA methylation signature for predicting progression-free survival (PFS) of patients with advanced-stage OC.

Methods: A bioinformatics analysis was performed to identify methylation sites that are relevant to PFS and develop a DNA methylation signature. A total of 501 patients with advanced-stage OC who were identified from the Cancer Genome Atlas (TCGA) were enrolled as a training cohort, and 108 patients with advanced-stage OC from GSE65820 were used as a validation cohort.

Results: A DNA methylation signature was constructed on the basis of five methylation sites. We found that patients with OC in the training and validation cohorts could be stratified, based on the DNA methylation signature, into high- and low-risk groups with distinct prognosis. Different cancer-related pathways were enriched between these two groups. Finally, a nomogram that integrated the methylation signature risk score and clinical stage was established, and the nomogram performed well.

Conclusions: The DNA methylation signature provides a promising prognostic biomarker for patients with advanced-stage OC and may help to optimize clinical management.

Background

Ovarian cancer (OC) is the most common cancer of the female reproductive system [1]. Due to a lack of typical symptoms, patients with OC are usually diagnosed at a late stage [2]. The survival of patients with advanced-stage OC (stages III or IV) is poor: the ten-year survival rate of patients with OC is 15% in the advanced stage, and 55% in the early stage [3]. About 20% of patients with advanced-stage OC survive for more than ten years, following standard treatment [3], and the mortality of patients with advanced-stage OC after ten years is close to that of the general population. Identifying patients with advanced-stage OC who survive beyond 10 years through patient stratification is beneficial to clinical management.

Surgery and subsequent platinum or taxane chemotherapy are standard therapies for patients with advanced-stage OC [4, 5]. It has been reported that 34% of patients with advanced-stage OC who were treated with surgery and chemotherapy underwent progression and recurrence within 12 months [6]. Therefore, an effort to determine effective biomarkers to identify high-risk progression and recurrence is necessary.

DNA methylation is a common epigenetic modification that involves the addition of a methyl group to a DNA molecule. Aberrant DNA methylation can silence tumor suppressor genes or activate oncogenes and affect cellular processes, such as cell cycle, DNA repair, and cell apoptosis [7]. Aberrant DNA methylation of a single gene has been extensively identified in OC. For instance, hypermethylation of tumor suppressors BRCA1 and RASSF1A was significantly higher in OC tissues than in normal tissues [8]. Hypermethylation of BRCA1 and RASSF1A causes downregulation of genes involved in the cell cycle,

thus promoting OC [8]. Numerous studies have reported that DNA methylation can be a promising biomarker in cancer. DNA methylation of NKD1, RUNX3, and ZNF671 were reported to be predictors of progression-free survival (PFS) in patients with OC [9–11]. Furthermore, a previous study reported that a DNA methylation signature can predict overall survival in patients with OC [12]. However, the study did not explore the correlation between DNA methylation status and PFS in OC.

Therefore, the DNA methylation profiles of tumor samples from patients with advanced-stage OC were obtained from The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO). A prognostic signature on the basis of DNA methylation sites was constructed and verified using bioinformatics methods and was used to predict the 3-year, 5-year, and 10-year PFS of patients with advanced-stage OC.

Results

Constructing and evaluating the five-CpG site signature

The clinical characteristics of patients with advanced-stage OC in the training and validation cohorts are summarized in Table S1. In the training cohort, after filtering low-quality probes, 501 samples with 18,914 CpG sites, were analyzed by univariate Cox regression analysis, and the results indicated that 1340 CpG sites were associated with PFS ($P < 0.05$). After LASSO regression analysis, five CpG sites were selected as candidate CpG sites (Figure S1A and B). Finally, these five methylation sites (cg02218324, cg04586563, cg07172280, cg14378057, and cg26803305) were used to construct a multivariate Cox regression model (Figure S1C). The formula was calculated as follows: DNA methylation signature risk score = $-0.7337 \times \beta\text{-value of cg02218324} + 1.6231 \times \beta\text{-value of cg04586563} - 5.1071 \times \beta\text{-value of cg07172280} - 2.2472 \times \beta\text{-value of cg14378057} + 0.8637 \times \beta\text{-value of cg26803305}$. The genes corresponding to five methylation sites (cg02218324, cg04586563, cg07172280, cg14378057, and cg26803305) were *RSPH6A* (radial spoke head 6 homolog A), *DIP2C* (disco interacting protein 2 homolog C), *DPP7* (dipeptidyl peptidase 7), *BICRAL* (BRD4 interacting chromatin remodeling complex associated protein like), and *SLC14A1* (solute carrier family 14 member 1). The information on the five CpG sites is displayed in Table 1. The associations between the methylation levels of the five CpG sites and the expression level of the corresponding genes were measured by Pearson correlation analysis. We found that the higher methylation levels of cg02218324 and cg26803305 were correlated with lower expression levels of *RSPH6A* and *SLC14A1* ($R < -0.10$, $P < 0.05$, Figure 1). Patients with advanced-stage OC were grouped into high- and low-risk groups. The Kaplan–Meier curves demonstrated that low-risk patients had longer 5-year and 10-year PFS than high-risk patients ($P < 0.0001$; Figure 2A). The 3-year area under the curve (AUC), 5-year AUC, and 10-year AUC were 0.722, 0.77, and 0.854, respectively (Figure 2B). Patients were ranked according to their risk scores (Figure 2C), and the dotplot represents their progression status (Figure 2D). A heatmap of methylation profiles of the five CpG sites is shown in Figure 2E. Moreover, methylation status of the five CpG sites in the high- and low-risk groups in the training cohort was assessed. As a result, the high-risk group presented relatively higher methylation levels of cg26803305, while relatively higher methylation levels of cg02218324, cg07172280, and cg14378057 were identified in the low-risk group (Figure S2, all $P < 0.01$). The GSE65820 validation dataset was used

to evaluate the robustness and effectiveness of the five- methylation site signature. The 5-year PFS of low-risk patients was better than that of high-risk patients (Figure 3A, $P = 0.029$). The AUCs of the signature at 3, 5, and 10 years were 0.623, 0.67, and 0.688, respectively (Figure 3B). The distribution of methylation signature risk score, progression status, and heatmap of methylation profiles of the five CpG sites in the validation dataset are displayed in Figure 3C-E.

Nomogram development

Several clinicopathological factors, including age, clinical stage, histological grade, and methylation signature risk score were included in univariate (Figure 4A) and multivariate Cox regression analyses (Figure 4B). As a result, the calculated risk score provides an independent prognostic marker for 5-year PFS in the training cohort ($P = 0.001$, Figure 4B). To predict PFS in patients with advanced-stage OC quantitatively and more precisely, a nomogram was developed (Figure 4C) that integrated the DNA methylation signature risk score and clinical stage, which correlated with PFS in the univariate analysis (Figure 4A). The calibration curves for the 5- and 10-year PFS rates demonstrated good performance of the nomogram in the training cohort (Figure 4D).

DNA methylation signature performance among patient subgroups

In addition, to evaluate the effectiveness of this signature among patients in different groups, a subgroup analysis was conducted. Patients were classified into two groups according to age: < 60 years and ≥ 60 years. High-risk patients had significantly reduced 5-year PFS for the < 60 group ($P < 0.001$, Figure 5A) and ≥ 60 group, compared with low-risk patients ($P < 0.0001$, Figure 5B). In addition, in the high-risk group, patients with Grade 1/2 ($P = 0.026$, Figure 5C) and Grade 3/4 ($P < 0.0001$, Figure 5D) OC also exhibited significantly decreased 5-year PFS rates. The methylation signature performance for patients with stage III and IV OC was also assessed. Similarly, high-risk patients exhibited a reduced survival rate compared to low-risk patients ($P < 0.01$, Figure 5E and F).

Validation of the five CpG sites in blood samples from OC patients

To investigate whether the methylation levels of these five sites that we observed in tumor tissue might be present in blood, DNA methylation data in whole blood samples from patients with OC and healthy females were collected from GSE19711. As a result, we found that the healthy group presented relatively higher methylation levels of cg02218324, cg04586563, cg07172280, while relatively higher methylation levels of cg26803305 were identified in patients with OC (Figure 6).

Signaling pathways implicated in the high- and low-risk groups

Differences in pathway activities between high- and low-risk patients with OC were scored using gene set variation analysis (GSVA). Consequently, hedgehog signaling, epithelial-mesenchymal transition (EMT), angiogenesis, KRAS signaling, hypoxia, TNF- α signaling via NF- κ B, Notch signaling, IL-6/JAK/STAT3 signaling, and TGF β signaling were significantly activated in the high-risk group (Figure 7). Unfolded

protein response, E2F targets, mTORC1 signaling, DNA repair, G2/M checkpoint, MYC targets, and interferon-alpha response were significantly activated in the low-risk group (Figure 7).

Discussion

OC is a complex disease with different histological and molecular patterns, and these features cause inconsistent results when applying a single class of drugs. Risk stratification and characterization of the potential mechanisms related to high- and low-risk groups are crucial for our understanding advanced-stage OC. A thorough identification of the molecular mechanisms related to each group will facilitate the appropriate use of targeted therapies, immunotherapies, epigenetic therapies, and combination therapies. Thus, more emphasis should be placed on genomics, epigenomics, and the analysis of patient tumor samples before treatment regimens. A few DNA methylation-related prognostic signatures have been reported in cancer, including melanoma [13], breast cancer [14], and lung adenocarcinoma [15]. These studies suggest that DNA methylation-related prognostic signatures can provide promising cancer biomarkers.

In this study, a prognostic signature was compiled that focused on five CpG sites and positively predicted PFS for patients with advanced-stage OC. The genes corresponding to the five CpG sites (cg02218324, cg04586563, cg07172280, cg14378057, and cg26803305) were five protein coding genes: RSPH6A, DIP2C, DPP7, BICRAL, and SLC14A1. Multiple studies have reported that hypermethylation causes lower gene expression, while hypomethylation causes higher expression. Thus, we explored the association of methylation level and gene expression at these five CpG sites. Inverse correlations were found only between the methylation levels of cg02218324 and cg26803305 and the expression levels of the corresponding genes. These findings indicated that methylation levels do not necessarily cause expected gene expression changes, and other factors affect gene expression.

DNA methylation signatures derived directly from blood will facilitate the prediction of OC patient survival. We found that these five CpG sites could be detected in the blood samples from patients with OC. Although we did not explore the correlation between methylation levels of these five CpG sites in the peripheral blood and the prognosis of patients with OC due to a lack of survival information, these results suggest that the five CpG sites provide potential prognostic biomarkers in both tumor tissue and body fluids.

A previous study reported that loss of DIP2C causes DNA methylation and gene expression changes and EMT in human colorectal cancer cells [16]. Furthermore, SLC14A1 has been reported to inhibit colony formation in lung cancer cell lines and have tumor-suppressor activity [17]. However, to date, these five genes and five CpG sites have not been reported in OC. Further functional studies will be necessary to clarify the biological and pathological implications of methylation in these genes and its relationship to OC progression.

The methylation signature investigated in this study was proven to function as an independent prognostic marker in the multivariate Cox model. Moreover, although patients with OC were regrouped

according to different clinicopathological variables, the subgroup analyses demonstrated that the signature still provided independent predictive ability. Furthermore, high- and low-risk patients with OC who were stratified by the methylation signature presented significant differences in cancer-related pathways. TGF β signaling, hedgehog signaling, notch signaling, and EMT were significantly activated in the high-risk group. Furthermore, MYC targets, E2F targets, G2/M checkpoint, and DNA repair were significantly activated in the low-risk group. EMT is a key step in OC metastasis [18], and TGF- β signaling has been reported to promote EMT and metastasis in advanced OC [19]. Furthermore, TGF- β signaling causes global changes in DNA methylation during the EMT in OC [20]. Notch and TGF β signaling can form a positive regulatory loop and cooperatively regulate EMT in OC cells [21]. In addition, hedgehog signaling has been reported to promote EMT in OC, through the PI3K/AKT pathway [22]. Furthermore, MYC expression has been reported to influence patient responsiveness to platinum chemotherapy and the prognosis of patients with OC [23]. E2F targets and G2M checkpoint pathways are important cell cycle signaling pathways that are crucial to OC progression [24, 25]. These signaling pathways may play important roles in OC tumorigenesis and progression, and the GSVA analysis results were in accordance with those of previous studies. We reasoned that aberrant DNA methylation can cause gene expression changes in these signaling pathways and affect cellular processes, such as cell cycle, DNA repair, and EMT, to promote tumorigenesis. In summary, our DNA methylation signature has potential functional relevance in predicting altered activities in these cancer-related pathways.

Taken together, our results suggest that this methylation signature provides a promising biomarker for patients with advanced-stage OC. In addition, the signature may improve the clinical management of patients with OC. This study has a few limitations. First, the signature was developed based on public databases and retrospective cohort studies. To apply the five-CpG sites signature in clinical practice, we will need to improve its predictive power and confirm the reliability of the signature in a prospective study. Thus, a future prospective study of a cohort of patients with OC is necessary to verify our findings, and further functional research will be necessary to explore the roles of the five CpG sites.

The DNA methylation signature that was developed in this study provides a promising prognostic marker for patients with advanced-stage OC and may help optimize their clinical management.

Methods

Patients with OC from TCGA and GEO

TCGA DNA methylation data of samples from patients with stage III–IV were obtained from UCSC Xena (<https://xena.ucsc.edu/>). The GSE65820 dataset, including DNA methylation profiles and clinical information, was downloaded from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). DNA methylation profiles in whole blood samples of patients with OC and healthy females were collected from GSE19711. The platforms of the three datasets were based on the Infinium HumanMethylation27 BeadChip (Illumina Inc., CA, USA) or the Illumina Human Methylation 450 BeadChip (Illumina Inc., CA, USA) and the genomic annotation was on the basis of GRCh38. DNA methylation status were calculated

as $M/(M + U + 100)$ and represented as β -values ranging from 0 to 1 [26]. The gene expression profiles and clinical information of patients with stage III–IV OC were obtained from TCGA data portal (<https://portal.gdc.cancer.gov/>). Patients from the TCGA cohort, which included PFS time, complete clinical information (age, clinical stage, and histological grade), and DNA methylation profiles, were screened to explore the association between DNA methylation site β -values and PFS in advanced-stage OC. In addition, 108 patients with advanced-stage OC from the GSE65820 dataset were analyzed as an external validation cohort.

Data preprocessing

The ChAMP R package [27] was used to remove low-quality probes according to the following criteria: (1) CpG site methylation undetected in any sample, (2) single-nucleotide polymorphisms at the assayed CpG dinucleotide [28], and (3) location in sex chromosomes [29].

Construction and validation of DNA methylation signature

First, a univariate Cox regression analysis was performed in the training cohort to identify prognostic methylation sites ($P < 0.05$). Then, by using the “glmnet” R package, the LASSO method was applied to decrease the number of prognostic methylation sites [30]. Finally, a multivariate Cox regression analysis was performed to establish an optimal predictive model. The risk score for every OC patient in the training and external validation cohorts was calculated. Patients with advanced-stage OC were grouped into high- and low-risk groups, with the median risk score as the cutoff. Using the “survival” R package, a survival analysis was performed to compare the differences in PFS between the two groups [31]. By using the “timeROC” R package, we calculated the AUC, to measure the prognostic capability of the signature [32].

Nomogram construction

To predict the PFS of patients with OC using a quantitative tool, we developed a nomogram. Univariate and multivariate Cox analyses were performed, based on DNA methylation signature risk score and other variables.

GSVA

Pathway analyses were performed on the hallmark gene set, which were acquired from the molecular signature database [33]. To explore enriched signaling pathways in high- and low-risk groups, we applied GSVA in the “GSVA” R package using standard settings [34].

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets supporting the conclusions of this article are available in the TCGA and GEO databases. TCGA DNA methylation data of samples from patients with stage III–IV were obtained from UCSC Xena (<https://xena.ucsc.edu/>). The GSE65820 dataset, including DNA methylation profiles and clinical information, was downloaded from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>).

Competing interests

The authors declare that they have no competing interests.

Funding

None

Authors' contributions

Suzhen Fan collected and analyzed the data and wrote the manuscript. Xiang Gao analyzed the data and reviewed the manuscript. Qiaohong Qin participated in analyzing the data. Hongyu Li and Zhongfu Yuan participated in preparation of the figures and tables. Shujun Zhao designed the study and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

Abbreviations

OC ovarian cancer

PFS progression-free survival

TCGA The Cancer Genome Atlas

GEO Gene Expression Omnibus

AUC area under the curve

GSVA gene set variation analysis

EMT epithelial-mesenchymal transition

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394–424.
2. Cannistra SA. Cancer of the ovary. *N Engl J Med.* 2004;351(24):2519–29.
3. Narod S. Can advanced-stage ovarian cancer be cured? *Nat Rev Clin Oncol.* 2016;13(4):255–61.
4. Bookman MA, Brady MF, McGuire WP, Harper PG, Alberts DS, Friedlander M, et al. Evaluation of new platinum-based treatment regimens in advanced-stage ovarian cancer: a Phase III Trial of the Gynecologic Cancer Intergroup. *J Clin Oncol.* 2009;27(9):1419–25.
5. Katsumata N, Yasuda M, Isonishi S, Takahashi F, Michimae H, Kimura E, et al. Long-term results of dose-dense paclitaxel and carboplatin versus conventional paclitaxel and carboplatin for treatment of advanced epithelial ovarian, fallopian tube, or primary peritoneal cancer (JGOG 3016): a randomised, controlled, open-label trial. *Lancet Oncol.* 2013;14(10):1020–6.
6. Yoshihara K, Tajima A, Yahata T, Kodama S, Fujiwara H, Suzuki M, et al. Gene expression profile for predicting survival in advanced-stage serous ovarian cancer across two independent datasets. *PLoS One.* 2010;5(3):e9615.
7. Singh A, Gupta S, Sachan M. Epigenetic Biomarkers in the Management of Ovarian Cancer: Current Prospectives. *Front cell Dev Biol.* 2019;7:182.
8. Ibanez de Caceres I, Battagli C, Esteller M, Herman JG, Dulaimi E, Edelson MI, et al. Tumor cell-specific BRCA1 and RASSF1A hypermethylation in serum, plasma, and peritoneal fluid from ovarian cancer patients. *Cancer Res.* 2004;64(18):6476–81.
9. Mase S, Shinjo K, Totani H, Katsushima K, Arakawa A, Takahashi S, et al. ZNF671 DNA methylation as a molecular predictor for the early recurrence of serous ovarian cancer. *Cancer Sci.* 2019;110(3):1105–16.
10. Dai W, Zeller C, Masrour N, Siddiqui N, Paul J, Brown R. Promoter CpG island methylation of genes in key cancer pathways associates with clinical outcome in high-grade serous ovarian cancer. *Clin cancer Res.* 2013;19(20):5788–97.
11. Häfner N, Steinbach D, Jansen L, Diebolder H, Dürst M, Runnebaum IB. RUNX3 and CAMK2N1 hypermethylation as prognostic marker for epithelial ovarian cancer. *Int J cancer.* 2016;138(1):217–28.
12. Guo W, Zhu L, Yu M, Zhu R, Chen Q, Wang Q. A five-DNA methylation signature act as a novel prognostic biomarker in patients with ovarian serous cystadenocarcinoma. *Clin Epigenetics.* 2018;10(1):142.
13. Guo W, Zhu L, Zhu R, Chen Q, Wang Q, Chen J-Q. A four-DNA methylation biomarker is a superior predictor of survival of patients with cutaneous melanoma. *Elife.* 2019;8.
14. Tao C, Luo R, Song J, Zhang W, Ran L. A seven-DNA methylation signature as a novel prognostic biomarker in breast cancer. *J Cell Biochem.* 2019;
15. Ma X, Cheng J, Zhao P, Li L, Tao K, Chen H. DNA methylation profiling to predict recurrence risk in stage I lung adenocarcinoma: Development and validation of a nomogram to clinical management.

- J Cell Mol Med. 2020;24(13):7576–89.
16. Larsson C, Ali MA, Pandzic T, Lindroth AM, He L, Sjöblom T. Loss of DIP2C in RKO cells stimulates changes in DNA methylation and epithelial-mesenchymal transition. *BMC Cancer*. 2017;17(1):487.
 17. Frullanti E, Colombo F, Falvella FS, Galvan A, Noci S, De Cecco L, et al. Association of lung adenocarcinoma clinical stage with gene expression pattern in noninvolved lung tissue. *Int J cancer*. 2012;131(5):E643-8.
 18. Vergara D, Merlot B, Lucot J-P, Collinet P, Vinatier D, Fournier I, et al. Epithelial-mesenchymal transition in ovarian cancer. *Cancer Lett*. 2010;291(1):59–66.
 19. Chou J-L, Chen L-Y, Lai H-C, Chan MWY. TGF- β : friend or foe? The role of TGF- β /SMAD signaling in epigenetic silencing of ovarian cancer and its implication in epigenetic therapy. *Expert Opin Ther Targets*. 2010;14(11):1213–23.
 20. Cardenas H, Vieth E, Lee J, Segar M, Liu Y, Nephew KP, et al. TGF- β induces global changes in DNA methylation during the epithelial-to-mesenchymal transition in ovarian cancer cells. *Epigenetics*. 2014;9(11):1461–72.
 21. Zhou J, Jain S, Azad AK, Xu X, Yu HC, Xu Z, et al. Notch and TGF β form a positive regulatory loop and regulate EMT in epithelial ovarian cancer cells. *Cell Signal*. 2016;28(8):838–49.
 22. Ke Z, Caiping S, Qing Z, Xiaojing W. Sonic hedgehog-Gli1 signals promote epithelial-mesenchymal transition in ovarian cancer by mediating PI3K/AKT pathway. *Med Oncol*. 2015;32(1):368.
 23. Iba T, Kigawa J, Kanamori Y, Itamochi H, Oishi T, Simada M, et al. Expression of the c-myc gene as a predictor of chemotherapy response and a prognostic factor in patients with ovarian cancer. *Cancer Sci*. 2004;95(5):418–23.
 24. Hashiguchi Y, Tsuda H, Inoue T, Nishimura S, Suzuki T, Kawamura N. Alteration of cell cycle regulators correlates with survival in epithelial ovarian cancer patients. *Hum Pathol*. 2004;35(2):165–75.
 25. De Meyer T, Bijsmans ITGW, Van de Vijver KK, Bekaert S, Oosting J, Van Criekinge W, et al. E2Fs mediate a fundamental cell-cycle deregulation in high-grade serous ovarian carcinomas. *J Pathol*. 2009;217(1):14–20.
 26. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. *Genomics*. 2011;98(4):288–95.
 27. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics*. 2017;33(24):3982–4.
 28. Sandoval J, Mendez-Gonzalez J, Nadal E, Chen G, Carmona FJ, Sayols S, et al. A prognostic DNA methylation signature for stage I non-small-cell lung cancer. *J Clin Oncol*. 2013;31(32):4140–7.
 29. Chen Y, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013;8(2):203–9.

30. Wang H, Lengerich BJ, Aragam B, Xing EP. Precision Lasso: accounting for correlations and linear dependencies in high-dimensional genomic data. *Bioinformatics*. 2019;35(7):1181–7.
31. Holleczeck B, Brenner H. Model based period analysis of absolute and relative survival with R: data preparation, model fitting and derivation of survival estimates. *Comput Methods Programs Biomed*. 2013;110(2):192–202.
32. Lorent M, Giral M, Foucher Y. Net time-dependent ROC curves: a solution for evaluating the accuracy of a marker to predict disease-related mortality. *Stat Med*. 2014;33(14):2379–89.
33. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005/10/04. 2005;102(43):15545–50.
34. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.

Tables

Table 1. Five methylation sites in the DNA methylation signature.

Probe ID	Chromosomal location	Gene symbol	Gene type	CGI coordinate	Coefficient [#]	<i>P</i> value [#]
cg02218324	chr19: 45815181- 45815182	RSPH6A	Protein coding	chr19:45815232- 45816003	-0.7337	0.0002
cg04586563	chr10: 691034- 691035	DIP2C	Protein coding	chr10:688767- 690282	1.6231	0.0143
cg07172280	chr9: 137115003- 137115004	DPP7	Protein coding	chr9:137113848- 137115021	-5.1071	0.0099
cg14378057	chr6: 42820825- 42820826	BICRAL	Protein coding	chr6:42783004- 42783694	-2.2472	0.0002
cg26803305	chr18: 45724487- 45724488	SLC14A1	Protein coding	chr18:45775353- 45775733	0.8637	0.0012

[#] in multivariate Cox regression analysis.

Figures

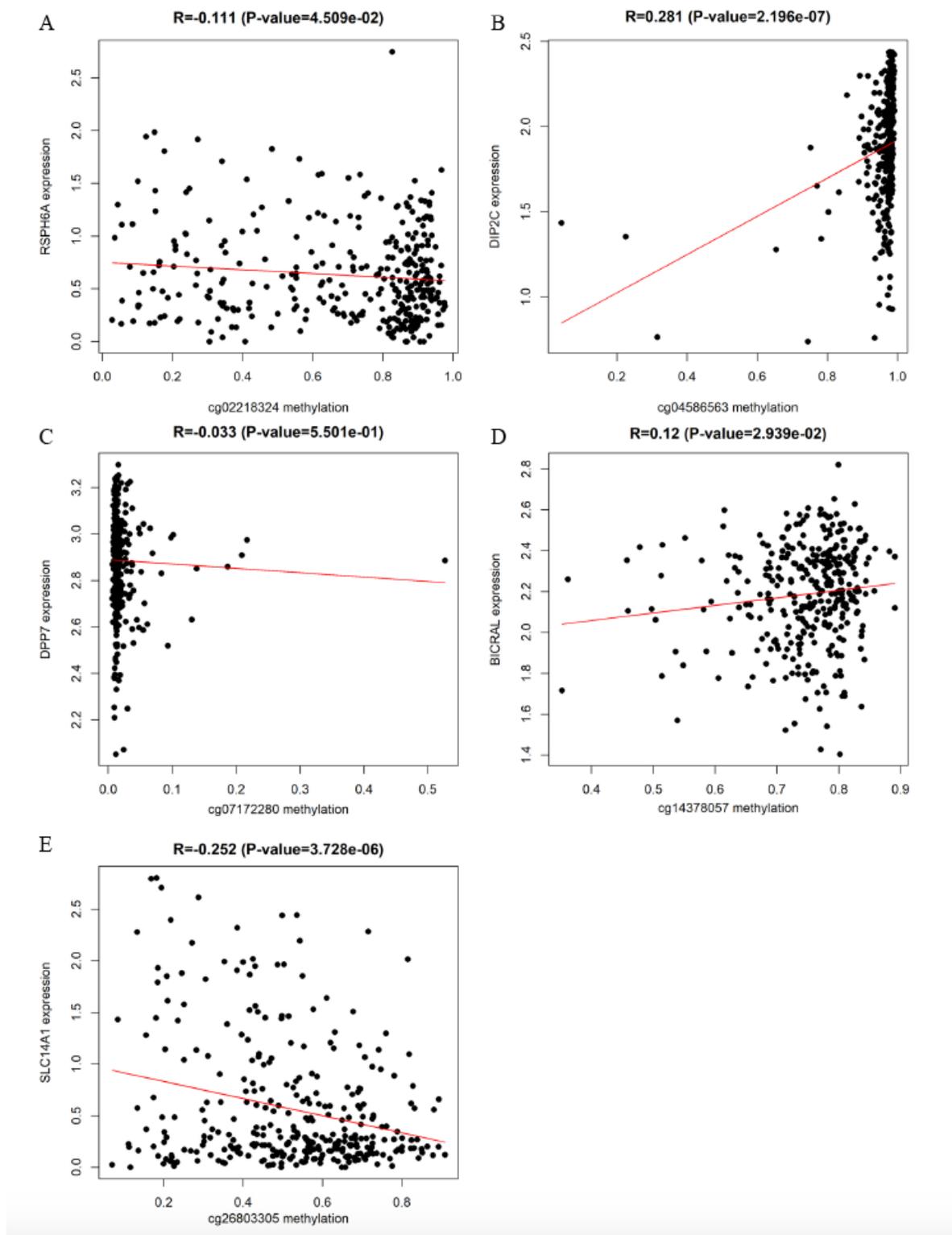


Figure 1

The correlations between the methylation levels of the five CpG sites and the expression level of the corresponding genes in the training cohort. Level of gene expression is reported as log₂-transformed FPKM, and the methylation levels of methylation sites were β -values.

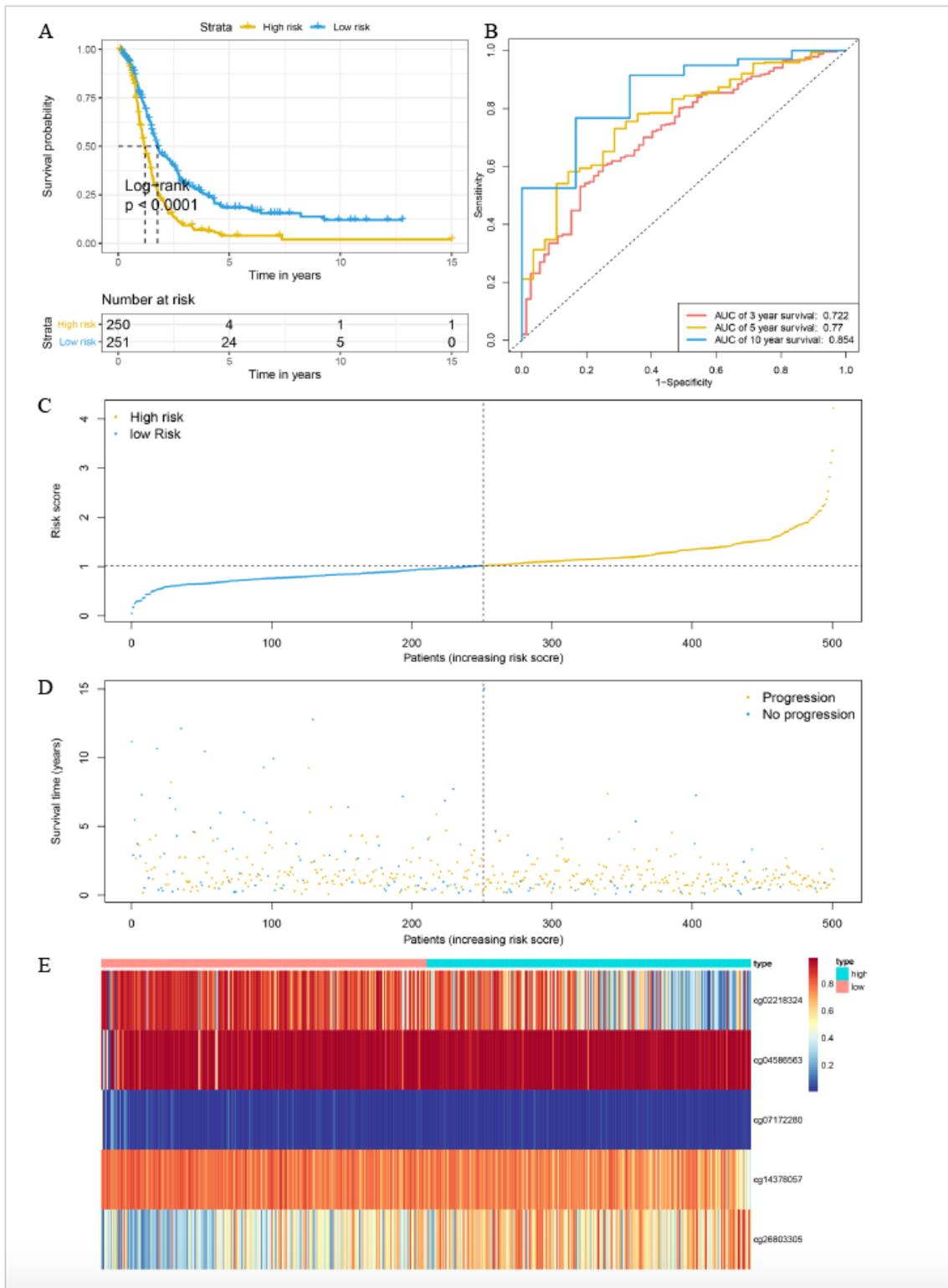


Figure 2

Performance of the DNA methylation signature in PFS prediction of ovarian cancer in the training cohort. (A) Kaplan–Meier survival curves for patients assigned to high- and low-risk groups; (B) Time-dependent ROC curves; (C) The distribution of risk score of patients with ovarian cancer in the training cohort; (D) Progression status of patients with ovarian cancer; (E) Heatmap of DNA methylation profiles of the five methylation sites. PFS: progression-free survival.

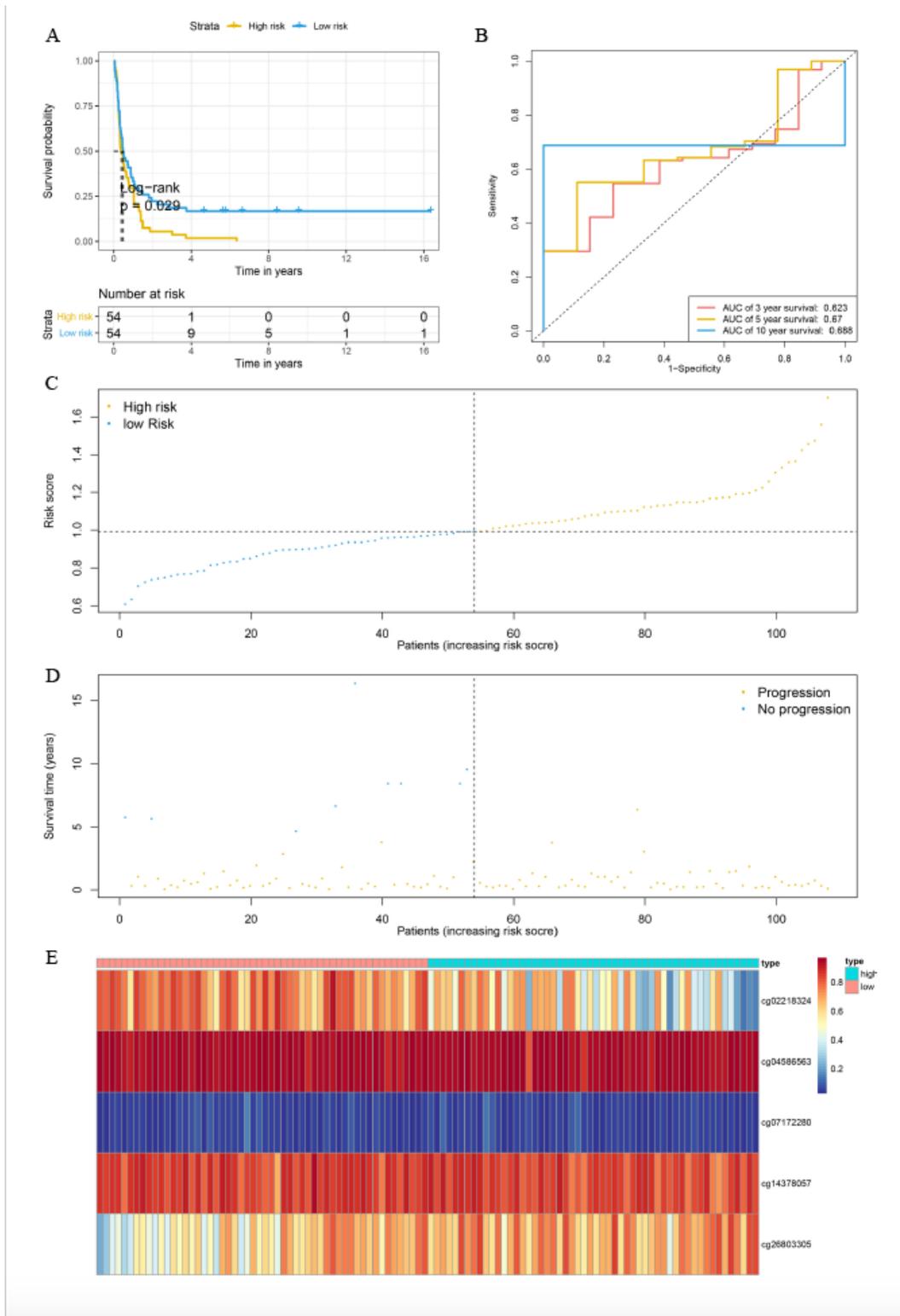


Figure 3

Performance of the DNA methylation signature in PFS prediction of ovarian cancer in the validation cohort. (A) Kaplan-Meier survival curves for patients assigned to high- and low-risk groups. (B) Time-dependent ROC curves. (C) The distribution of risk score of patients with ovarian cancer in the validation cohort. (D) Progression status of patients with ovarian cancer. (E) Heatmap of DNA methylation profiles of the five methylation sites. PFS: progression-free survival.

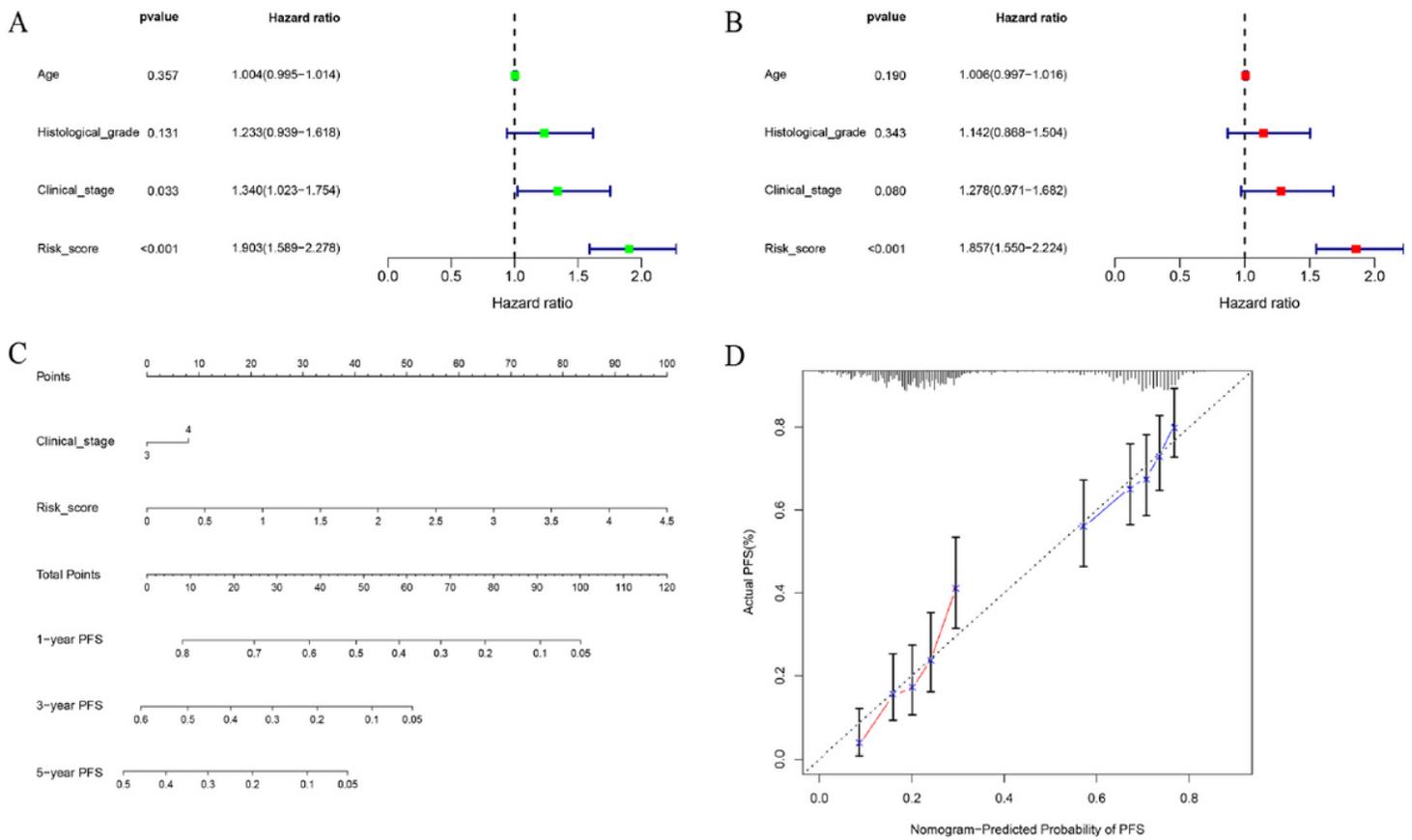


Figure 4

Integration of the DNA methylation signature risk score and clinicopathological factors. Univariate (A) and multivariate (B) Cox regression analyses of the association between clinicopathological factors, risk score and PFS of patients with ovarian cancer in the training cohort. (C) The nomogram for predicting PFS of patients with ovarian cancer. (D) The calibration curves for the 3-year PFS, 5-year PFS. Blue line represented 3-year PFS. Red line represented 5-year PFS. PFS: progression-free survival.

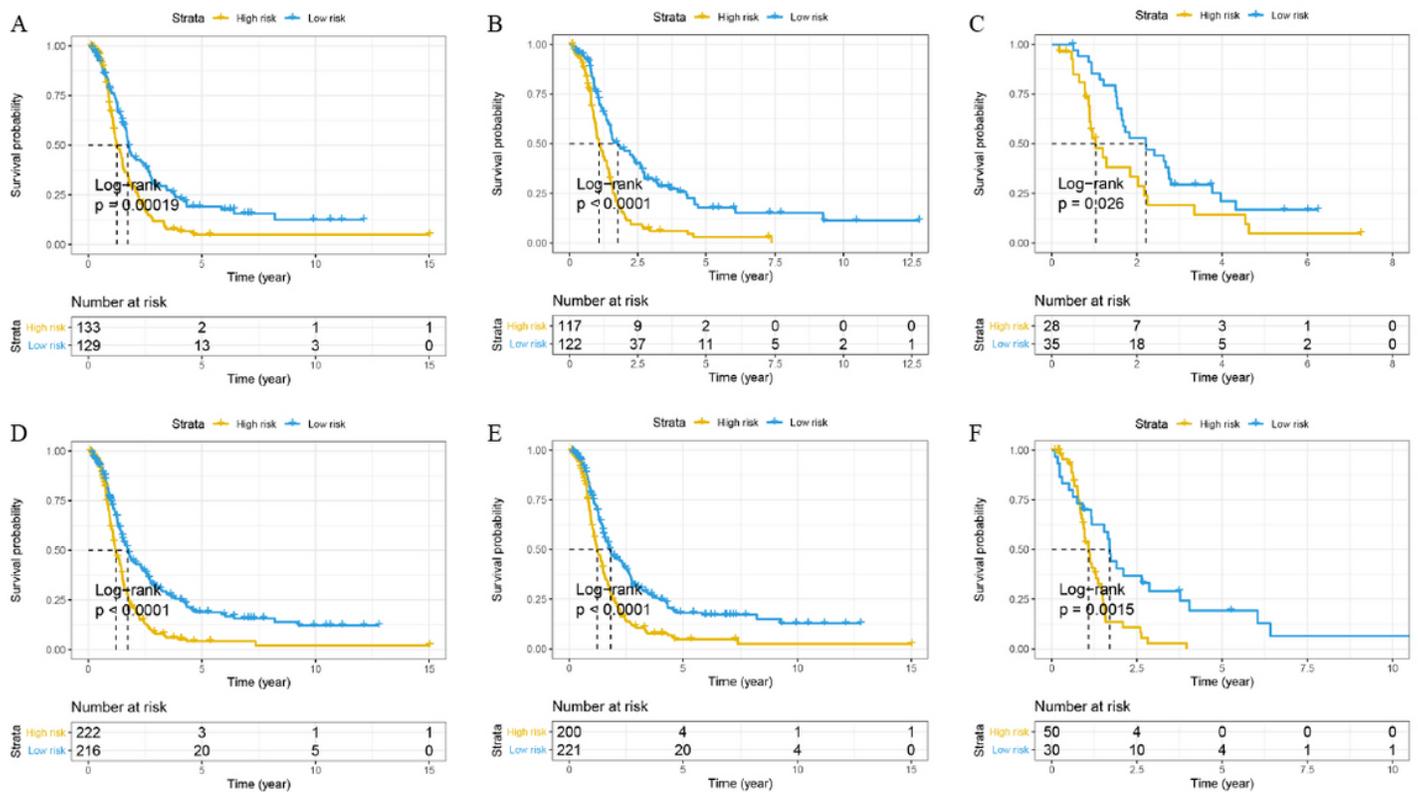


Figure 5

Evaluation of the effectiveness of the DNA methylation signature among patients in different subgroup. (A)Kaplan-Meier curves for high- and low-risk groups in patients with age < 60; (B)Kaplan-Meier curves for high- and low-risk groups in patients with age ≥ 60 ; (C)Kaplan-Meier curves for high- and low-risk groups in patients with Grade 1/2; (D)Kaplan-Meier curves for high- and low-risk groups in patients with Grade 3/4; (E)Kaplan-Meier curves for high- and low-risk groups in patients with stage III; (F)Kaplan-Meier curves for high- and low-risk groups in patients with stage IV.

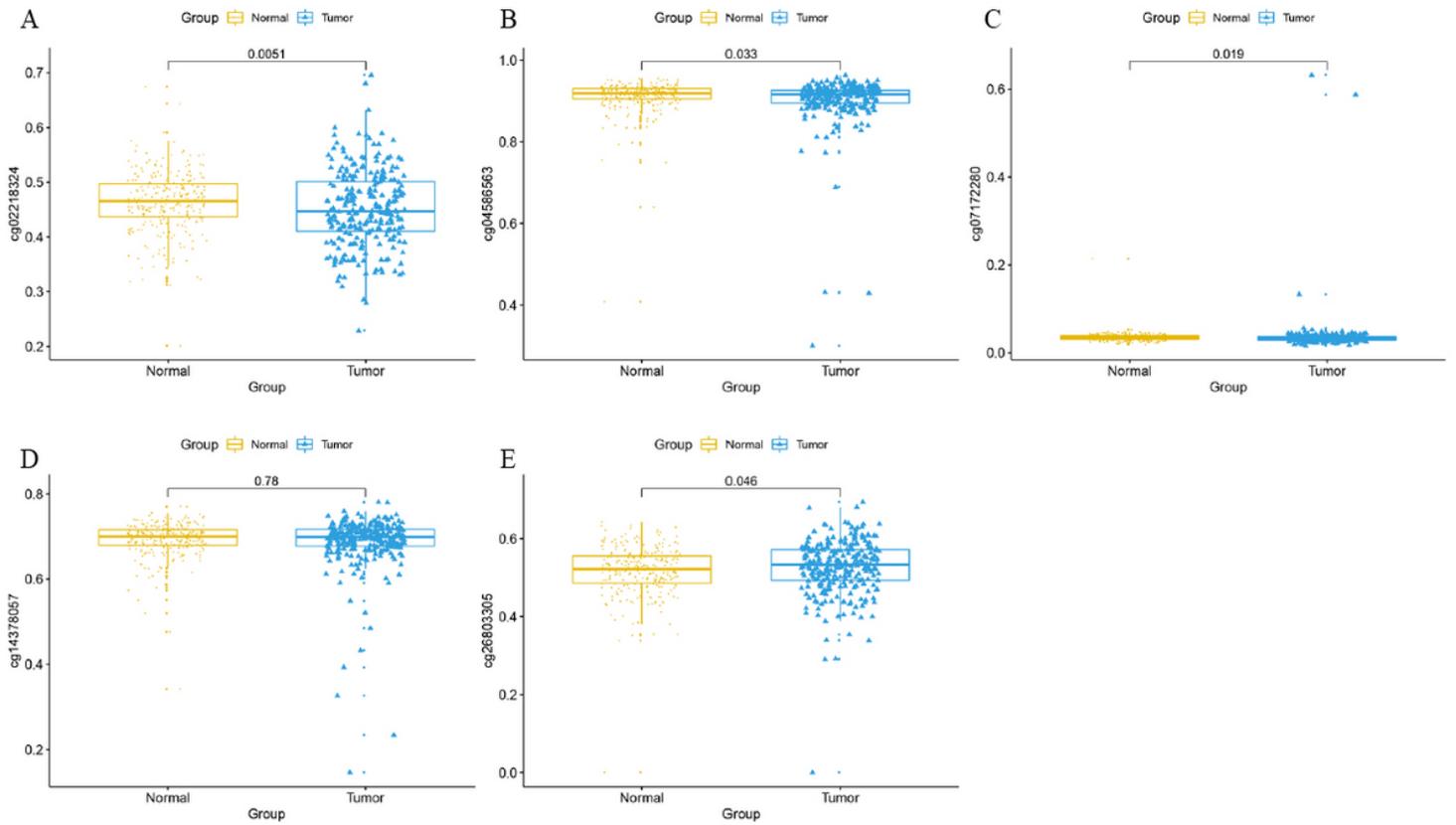


Figure 6

Boxplots of β -value in ovarian cancer patients and healthy females in the GSE19711 dataset. Wilcoxon rank test was used to determine the differences between the two groups.

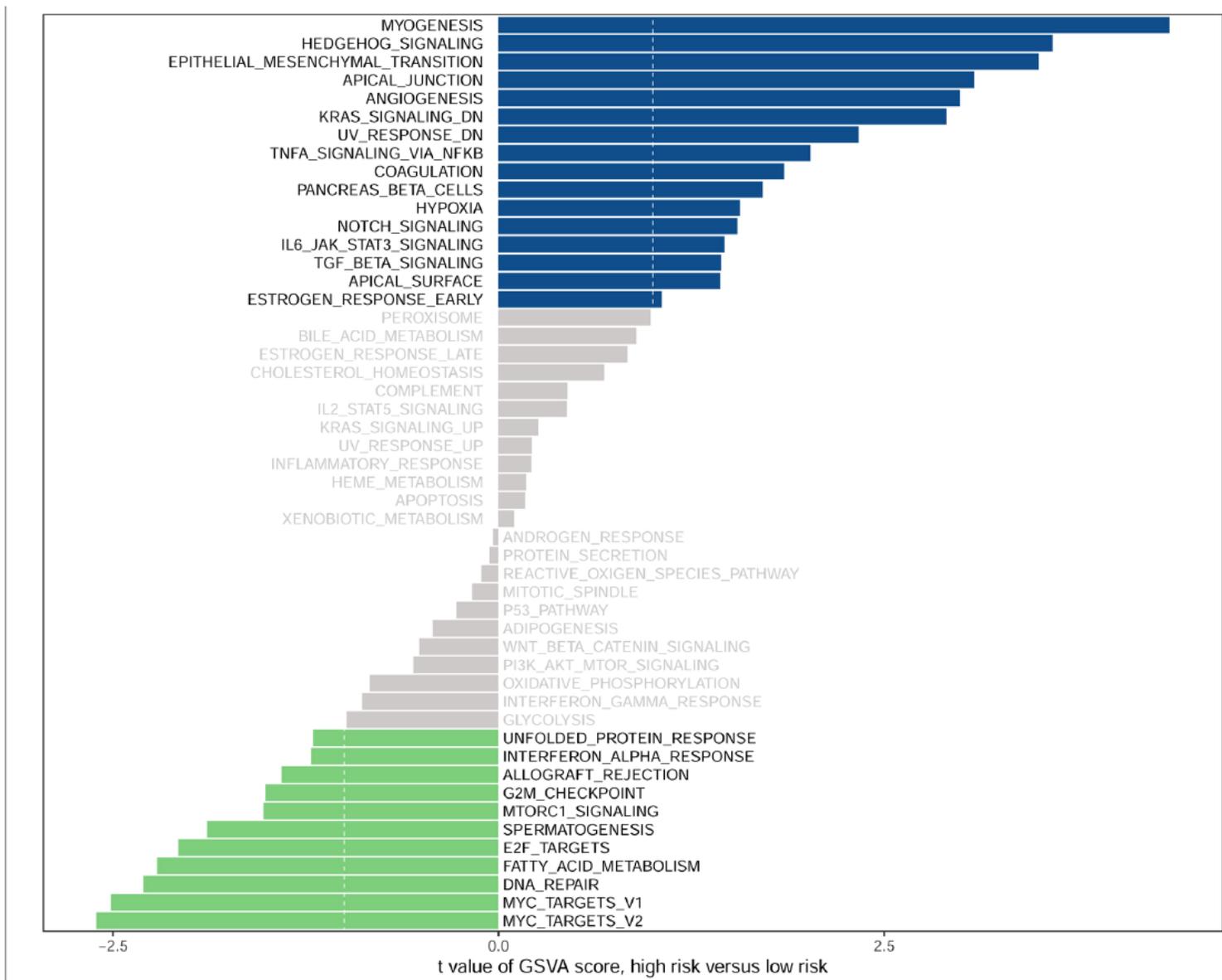


Figure 7

Summary of gene set variation analysis (GSVA) results in high-risk and low-risk groups.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.pdf](#)
- [FigureS2.pdf](#)
- [supplementTableS1.docx](#)