

Research Article

Comprehensive Analysis of DNA Methylation and Transcriptome to Identify PD-1-Negative Prognostic Methylated Signature in Endometrial Carcinoma

Lu Cao ^{1,2}, Xiaoqian Ma,^{1,2} Pengfei Rong,¹ Juan Zhang,^{1,2} Min Yang,^{1,2} and Wei Wang ^{1,2}

¹Department of Radiology, The Third Xiangya Hospital, Central South University, 138 Tongzipo Road, Yuelu District, Changsha, China

²The Institute for Cell Transplantation and Gene Therapy, The Third Xiangya Hospital, Central South University, 138 Tongzipo Road, Yuelu District, Changsha, China

Correspondence should be addressed to Wei Wang; wang_w@csu.edu.cn

Received 24 February 2022; Revised 15 April 2022; Accepted 25 April 2022; Published 18 May 2022

Academic Editor: Yang Mi

Copyright © 2022 Lu Cao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Epigenetic mechanism plays an important role in endometrial carcinoma (EC). This study was designed to analyze the epigenetic mechanism between DNA methylation-driven genes (DEDGs) and drugs targeting DEDGs and to develop a DEDG score model for predicting the prognosis of EC. **Methods.** Expression profile and methylation profile data of PD-1-negative EC samples were obtained from TCGA. To obtain intersected DEDGs, differentially expressed genes (DEGs) and differentially methylated genes from tumor tissues and normal tissues were analyzed by limma. A linear discriminant classification model was constructed using the gene expression profile of DMDGs, methylation profile of TSS1500, TSS200, and gene body regions. Principal component analysis (PCA) and ROC analysis were conducted. The protein-drug interactions analysis of DMDGs was performed using Network Analyst 3.0 tool. Lasso Cox regression analysis was used to screen prognostic DNA methylation driving gene and to build a risk score model. The ROC curve and Kaplan-Meier survival curve were plotted to evaluate the model prediction capability. **Results.** A total of 96 DMDGs were screened from the three regions, distributed on 22 chromosomes, with consistent methylation patterns in different gene regions. Both the expression profile and methylation profile of the three regions can neatly distinguish tumor samples from normal ones, with a high classifying performance. A gene signature, which consisted of ELFN1-AS1 and ZNF132, could classify EC patients into a high-risk group and low-risk group. Prognosis of the high-risk group was significantly worse than that of the low-risk group. The risk model showed a high performance in predicting the prognosis of EC. **Conclusion.** We successfully established a risk score system with two DMDGs, which showed a high prediction accuracy of EC prognosis.

1. Introduction

Endometrial carcinoma (EC) is the fourth leading malignant tumor among women in 2021, with more than 410000 new cases being diagnosed in 2020 [1, 2]. EC usually occurs in postmenopausal women, with an average onset age of 68 years old [3]. Most endometrial cancers were diagnosed at the local stage and mostly can be cured by taking surgery, with a 5-year survival rate of about 95% [4]. Patients with

advanced and recurrent EC face great treatment challenges, and the 5-year overall survival (OS) rate of patients with distant diseases is only 17% [5]. Therefore, it is necessary to accurately diagnose EC and improve the treatment of EC. An ongoing study demonstrated that predictive biomarkers (e.g., p53 or L1 cell adhesion molecule L1CAM) may have a promising effect on improving the clinical outcome of EC [6].

Cancer development is usually driven by genetic and epigenetic changes and complex crosstalk between many different

TABLE 1: The grouping criteria of DMDGs.

Groups	Methylation cut-off	Expression cut-off
HypoUp	FDR < 0.01 and delta β value < -0.3	FDR < 0.01 and log 2FC > 1
HypoDown	FDR < 0.01 and delta β value < -0.3	FDR < 0.01 and log 2FC < -1
HyperUp	FDR < 0.01 and delta β value > 0.3	FDR < 0.01 and log 2FC > 1
HyperDown	FDR < 0.01 and delta β value > 0.3	FDR < 0.01 and log 2FC < -1

pathways [7]. These changes involve DNA methylation, histone modifiers and readers, chromatin remodelers, microRNAs, and other components of chromatin [8]. As an important epigenetic mechanism, abnormal DNA methylation is considered a biomarker for early diagnosis and prognostic monitoring of many cancers [9]. Short stature homeobox gene two (SHOX2), RAS association domain family 1A (RASSF1A), and prostaglandin E receptor 4 gene (PTGER4) methylation has been reported to be closely related to lung cancer diagnosis and prognosis [10]. According to the results of a liquid biopsy, Hassan et al. indicated that SALL1, WNT5 α , LRP1, and CDH2 have diagnostic significance for children with acute lymphocytic leukemia [11]. In high-grade serous ovarian cancer, hypermethylated NCALD and LAMA3 are associated with prognosis and chemotherapy resistance in advanced patients [12]. In addition, a number of studies have identified panels of multiple methylation driving genes to predict the prognosis of colorectal cancer [13], gastric cancer [14], and hepatocellular carcinoma [15]. However, there is a lack of a comprehensive panel of methylation markers for EC [16].

Many previous studies have shown that new gene markers are helpful to tumor prognosis and diagnostic evaluation. Previously, Lai et al. identified 9 new tumor markers [17] based on proteomic data using the LASSO method. These proteins can predict the response to immunotherapy, chemotherapy, and targeted therapy. Liu et al. [18] filtered the characteristics of focal death-related 6 lncRNAs using the method of comprehensive bioinformatics. Tan et al. [19] screened 11 genes related to lipid metabolism based on the gene set associated with lipid metabolism, which can be used as prognostic markers of endometrial cancer to predict the response to chemotherapy and immunotherapy. Ruan et al. [20] used bioinformatics methods to identify the characteristics of novel epithelial mesenchymal transformation-related genes for the prognosis of endometrial cancer. These are mainly results achieved through gene expression profiling. They focus on all patients and specific gene concentration and pay limited attention to different types of patients, especially patients with PD-1 negative.

In this study, we systematically analyzed the expression profile data of PD-1-negative EC samples from The Cancer Genome Atlas (TCGA, <https://www.tumorfusion.org/>) and DNA methylation data to screen DMDGs and predicted their targeting drugs. Further analysis on DMDGs developed a risk signature that can monitor the prognosis of EC, contributing to a better clinical management of EC.

2. Materials and Methods

2.1. Sample and Data of EC. Tissue expression data, methylation data based on Illumina Human Methylation 450Bead-Chip array, and clinical follow-up information of EC samples were all downloaded from TCGA. After excluding PD-1-positive samples (the expression of PD-1 in tumor tissues was higher than the mean value of PD-1 expression in normal tissues), a total of 208 primary tumor samples and 35 normal samples were obtained. Methylation data were obtained from 163 tumor samples and 46 normal samples.

Differential expression analysis and differential methylation analysis.

DEGs between tumor tissues and normal tissues were analyzed by R package limma [21], and the P value was adjusted by the Benjamini and Hochberg (BH) [22] method. limma was also used to analyze differential methylation of tumor tissues and normal tissues, and differential methylation CpG sites were determined with FDR < 0.05 and absolute delta β (original methylation intensity of each probe) value > 0.3 as the thresholds. The CpG locus and gene matching file came from the Illumina website (<https://www.illumina.com/>). The average β values were calculated from the region covering -200 to -1500 nucleotides upstream of Transcription Start Site (TSS1500) from Transcription Start Site to -200 nucleotides upstream of Transcription Start Site (TSS200) and gene body. The differential methylation CpG sites were calculated by limma; here, FDR < 0.01 and delta β values > 0.3 were defined as hypermethylated regions, while those with FDR < 0.01 and delta β values < -0.3 were considered as demethylated regions.

2.2. Functional Enrichment Analysis. ClusterProfiler [23] database was used to perform gene ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) path annotation analysis of DEGs and DMGs. $P < 0.05$ was set as the threshold for filtering significant functional events.

2.3. Identification of DMDGs. DNA methylation data and differentially expressed mRNA data were integrated to determine the intersection of DEGs and differentially methylated genes by Venn diagram. The concentrated genes were DMEGs, which were further divided into four groups, namely, HypoUp, HypoDown, HyperUp, and HyperDown. The grouping criteria were shown in Table 1.

2.4. Performance Evaluation of Class Energy Estimation Based on DMDGs. A linear discriminant classification model was developed using the gene expression profile of DMDGs.

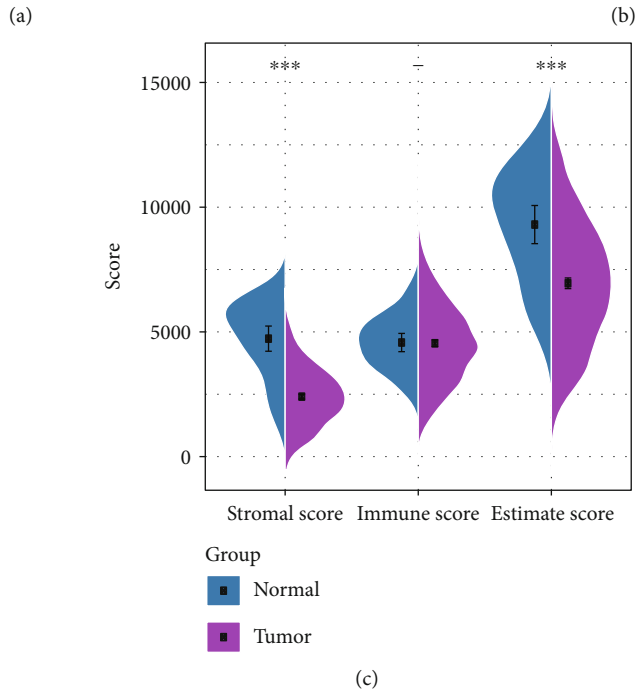
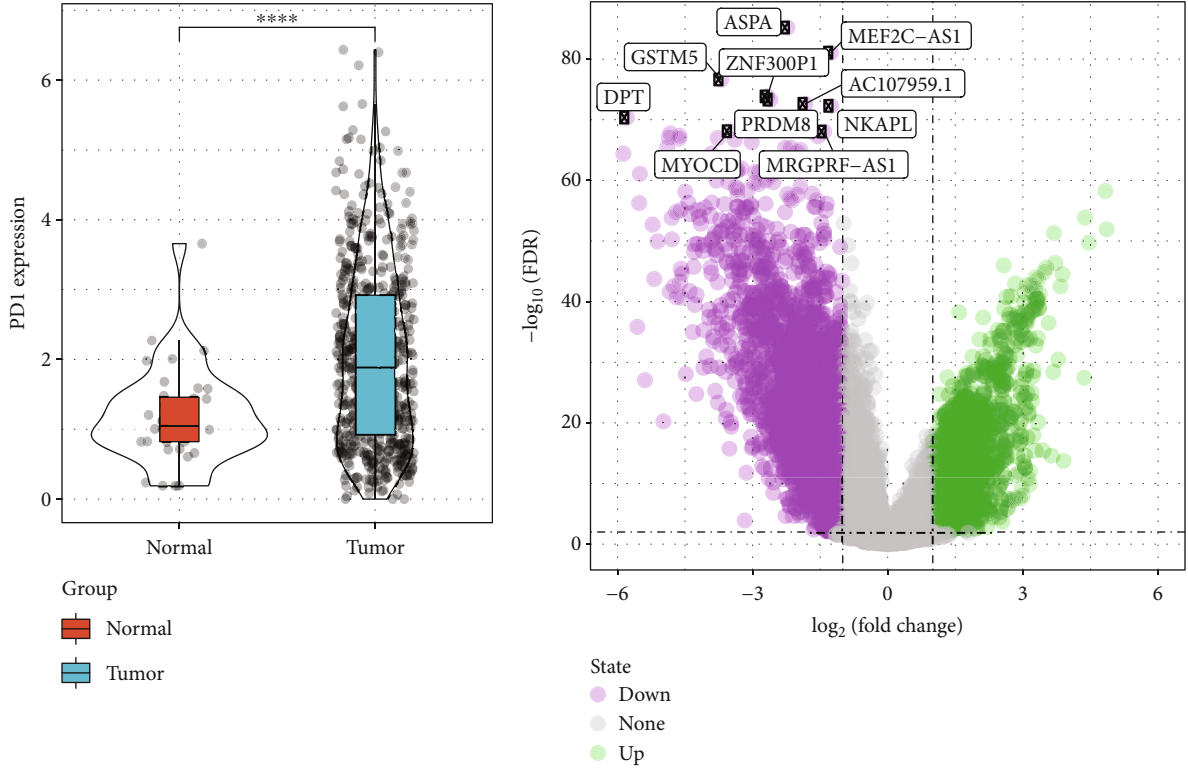
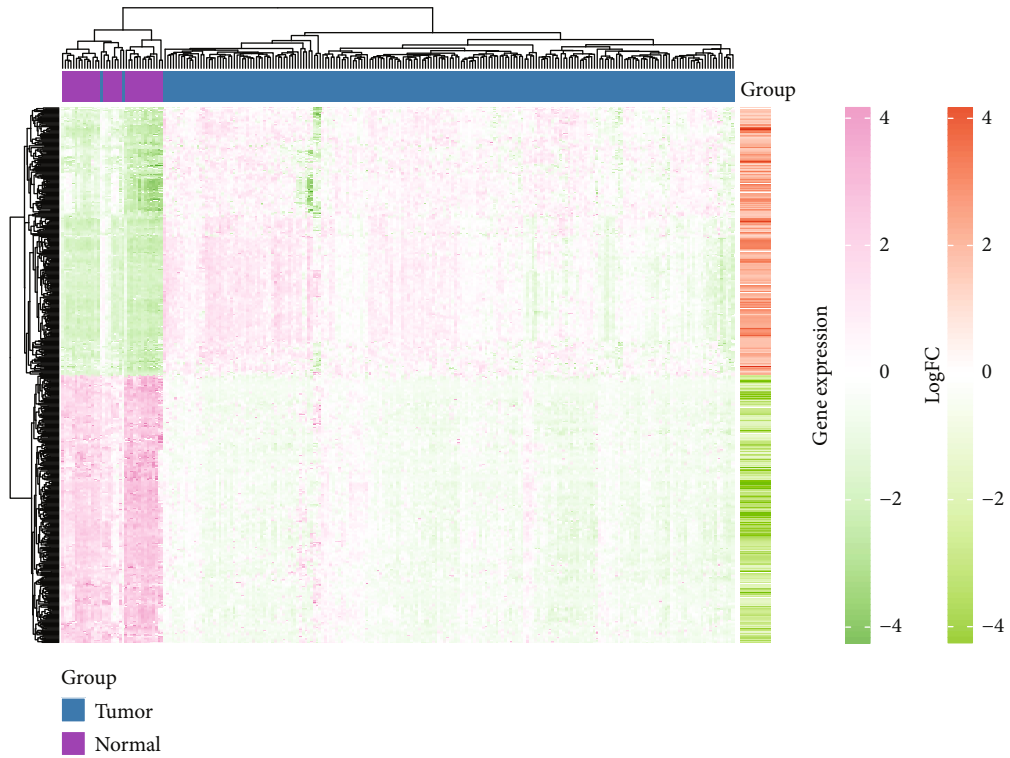
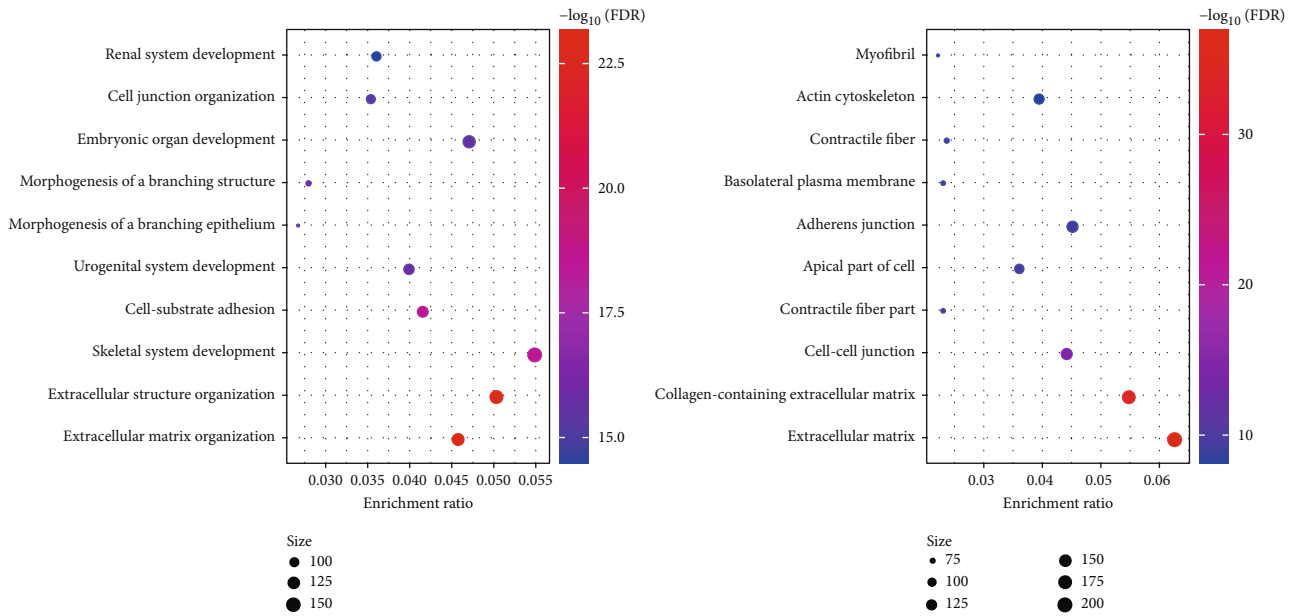


FIGURE 1: Continued.



(d)



(e)

FIGURE 1: Continued.

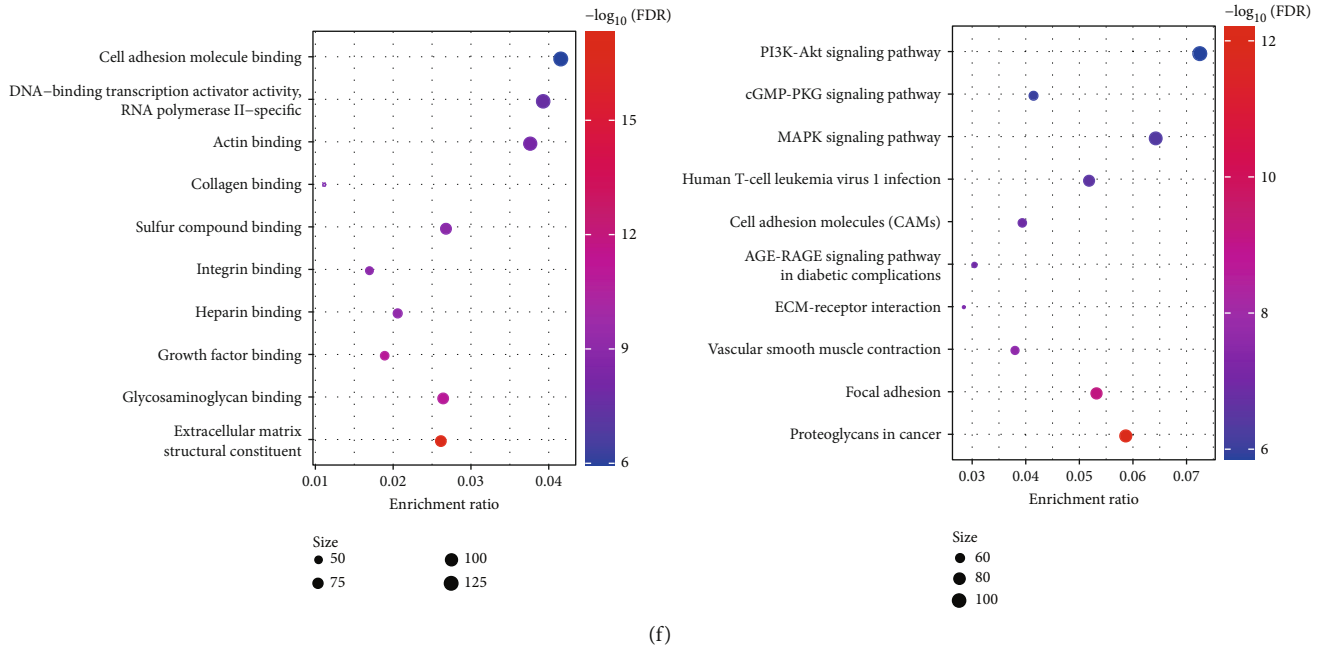


FIGURE 1: Determination and functional analysis of differentially expressed genes. (a) The difference of PD-1 expression between EC and normal tissues. (b) Volcanic map of DEGs between tumor tissue and normal tissue. (c) The difference of stromal score and immune score and ESTIMA TE score between tumor tissue and normal tissue. (d) Unsupervised hierarchical clustering heat map of differential genes. (e) Top 10 GO analysis of 4153 DEGs from the aspects of BP, CC and MF. (f) Top 10 KEGG analysis of 4153 DEGs.

Methylation profile of gene body, the methylation spectrum of TSS1500 region and the methylation spectrum of TS200 region, and PCA (principal component analysis) analysis and ROC analysis were conducted. At the same time, leave-one-out cross-validation (LOOCV) [24] method was applied to verify the effect of classification.

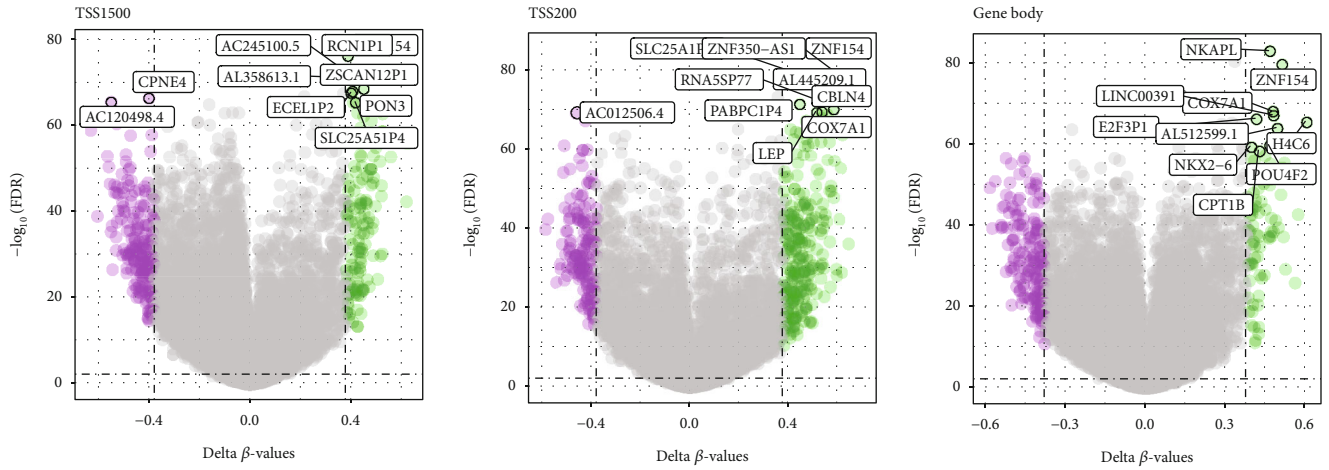
2.5. *Identification of Potential Target Therapeutic Drugs.* The DrugBank [25] database was used to identify drugs targeting upregulated DMDG. NetworkAnalyst 3.0 [26] (<http://www.networkanalyst.ca/>) is a visual analysis web platform supporting comprehensive analysis and interpretation of system-level gene expression data. In this paper, the study of the protein-drug interaction on DMDGs was carried out using NetworkAnalyst 3.0 tool, and the network map of gene-drug interaction was generated.

2.6. *Construction of Prognosis Scoring Model Based on Postestablishment Classification of DMDGs.* PD-1-negative samples were randomly divided into training set ($n = 104$) and verification set ($n = 104$). According to the expression of DMEGs and survival data, LASSO regression analysis with 1000 repetition was performed in the training set, and tenfold cross-validation was used to develop a prognosis score model. The risk score of the samples was calculated and grouped by the model. We further examined the accuracy of the model using Kaplan-Meier (KM) analysis and receiver operating characteristic (ROC) curve analysis in the internal and external verification sets.

2.7. *Statistical Analysis.* The statistical analysis of all the data in this study were carried out in R software. $P < 0.05$ as the cut-off value was considered to be statistically significant.

3. Results

3.1. *Determination and Functional Analysis of DEGs.* The difference analysis of PD-1 expression between tumor tissue and normal tissue showed that the expression of PD-1 in tumor tissue was significantly up-regulated (Figure 1(a)). According to the screening criteria of $FDR < 0.01$ and $|\log 2FC| > 1$, 4153 DEGs incorporating 1309 upregulated DEGs and 2844 downregulated DEGs were detected from 208 EC tissues and 35 normal tissues (Figure 1(b)). Between tumor tissue and normal tissue, there were significant differences in immune microenvironment scores, including matrix score and ESTIMATE score. These two scores in tumor tissue were significantly lower than those in the normal tissue (Figure 1(c)). Unsupervised hierarchical cluster analysis on 4153 DEGs demonstrated that DEGs could clearly distinguish tumor samples from normal samples (Figure 1(d)). Then, from the results of GO and KEGG enrichment analyses of differential genes, it could be found that 4153 differential genes were enriched to 107 pathways, 1582 biological process (BP), 134 cellular component (CC), and 80 molecular functions (MF). The most significantly enriched BP included renal system development, cell junction organization, and embryonic organ development. The CC showing the most significant correlation with 4153 differential genes were actin cytoskeleton, adherens junction, and cell-cell



State
 ● Down
 ● None
 ● Up

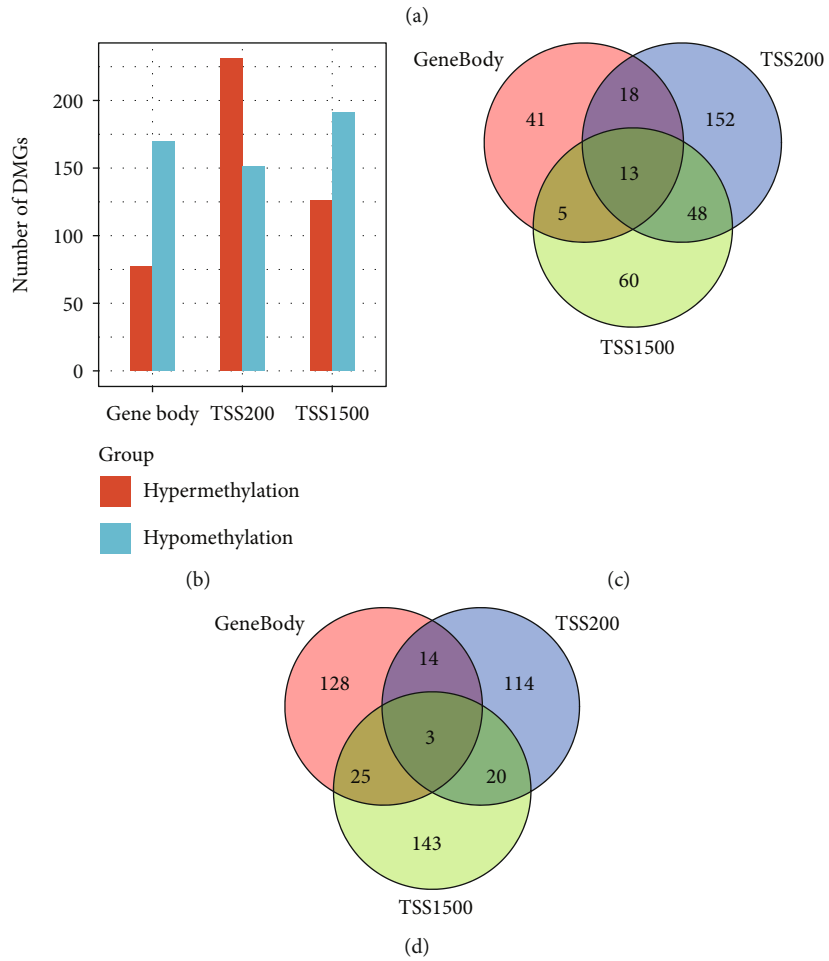
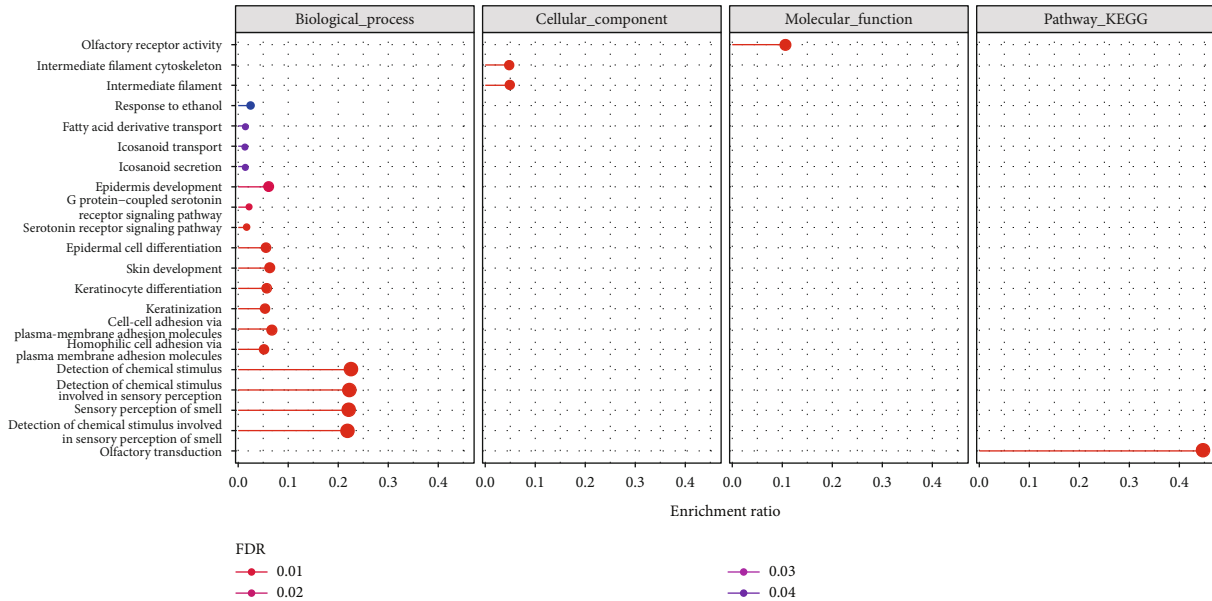


FIGURE 2: Continued.



(e)

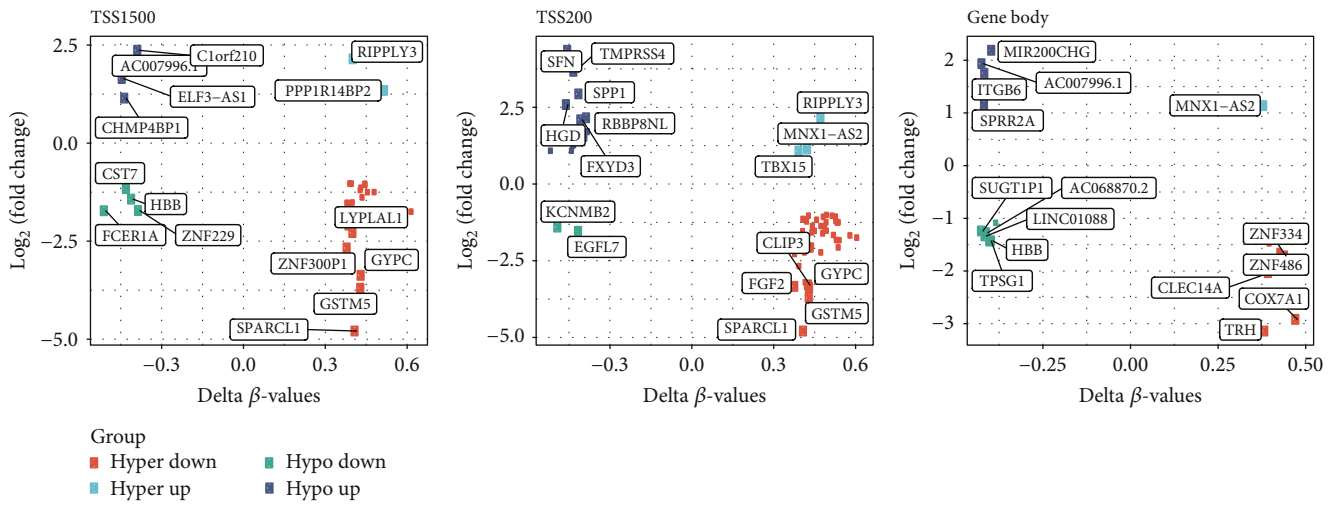
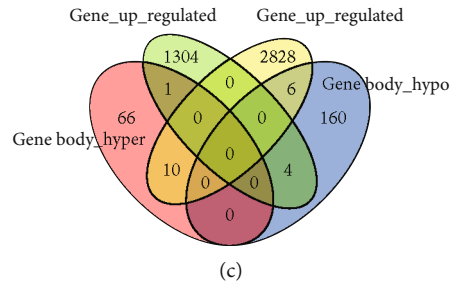
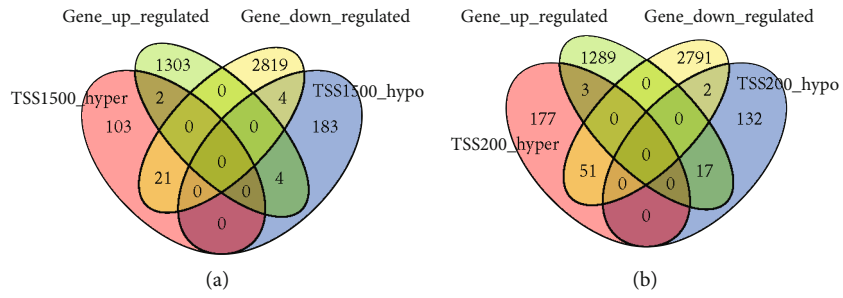
FIGURE 2: Screening and functional enrichment analysis of differential methylation genes. (a) Volcano map of differentially methylated genes in TSS1500, TSS200, and gene body regions. (b) Histogram of hypermethylation and demethylation genes in TSS1500, TSS200, and gene body regions. (c) Overlap analysis of hypermethylation genes in TSS1500, TSS200, and gene body regions. (d) Venn diagram of demethylated genes in TSS1500, TSS200, and gene body regions. (e) Differential methylation genes KEGG and GO enrichment bubble maps, in which the color from blue to red represented FDR from large to small, and dot size represented enrichment to the number of genes; the larger the dot, the more genes were enriched.

junction, and the most significantly enriched MF of these genes were cell adhesion molecule binding, DNA-binding transcription activator activity, actin binding and so on (Figure 1(e)). Furthermore, the results of KEGG demonstrated that the differential genes were mainly involved in the regulation of PI3K-Akt signaling pathway, MAPK signaling pathway, and focal adhesion and so on (Figure 1(f)).

3.2. Screening and Functional Analysis of Differentially Methylated Genes. In view of the effect of methylation on gene expression, we analyzed the methylation data of 163 tumor samples and 46 normal samples. A total of 1063 differentially methylated genes were identified from TSS1500, TSS200, and gene body regions. Specifically, the TSS1500 region contained 126 hypermethylation genes and 191 demethylation genes; the TSS200 region contained 231 hypermethylated genes and 151 demethylated genes; gene body region contained 77 hypermethylation genes and 170 demethylation genes (Figure 2(a)). After comparing the proportion of differential hypermethylation and differential demethylation genes in different regions, we found that the number of hypermethylation genes in gene body and TSS1500 regions was significantly fewer than that of demethylation genes. In the TS200 region, the number of hypermethylated genes was higher than that of demethylated genes (Figure 2(b)). Overlap analysis on the hypermethylation genes in the three regions indicated that 13 genes were present in all the three regions (Figure 2(c)). Similarly, the Venn diagram of the intersection of demethylated genes showed the presence of the three genes in all the three

regions (Figure 2(d)). Furthermore, GO and KEGG enrichment analyses were conducted to define the most significantly enriched GO terms and pathways of the differential methylation genes. Here, the data revealed that the differential methylation genes were mainly related to 17 BP (including epidermis development, cell-cell adhesion, and G protein-coupled serotonin receptor signaling pathway), 2 CC (intermediate filament cytoskeleton, intermediate filament cytoskeleton), 1 MF (olfactory receptor activity), and 1 KEGG pathway (olfactory transduction) (Figure 2(e)).

3.3. Identification of DMDGs and Biological Processes and Regulatory Mechanisms of DMDG-Related Substance Passing and Controlling. Overlap analysis was performed on differential methylated genes and DEGs in TSS1500, TSS200, and gene body region, respectively, and the genes in the crossover were found to be DMDGs. There were 31 DMDGs in TSS1500, 73 DMDGs in TSS200, and 21 DMEGs in gene body (Figures 3(a)–3(c)). Figure 3(d) showed the methylation multiple and differential expression multiple of these DMDGs, and some genes, such as GYPC, GSTM5, SPARCL1, and RIPPLY3, also appeared in different regions at the same time. Among the three regions, the proportion of DMDGs of HyperDown type was the highest (Figure 3(e)). By combining different types of DMDGs in the three regions, a total of 96 DMDGs were obtained: 23 belonged to HypoUp type, 4 belonged to HyperUp type, 10 belonged to HypoDown type, and 59 belonged to HyperDown type. The distribution of 96 DMDGs on the genome was analyzed. The results showed that they were distributed



(d)

FIGURE 3: Continued.

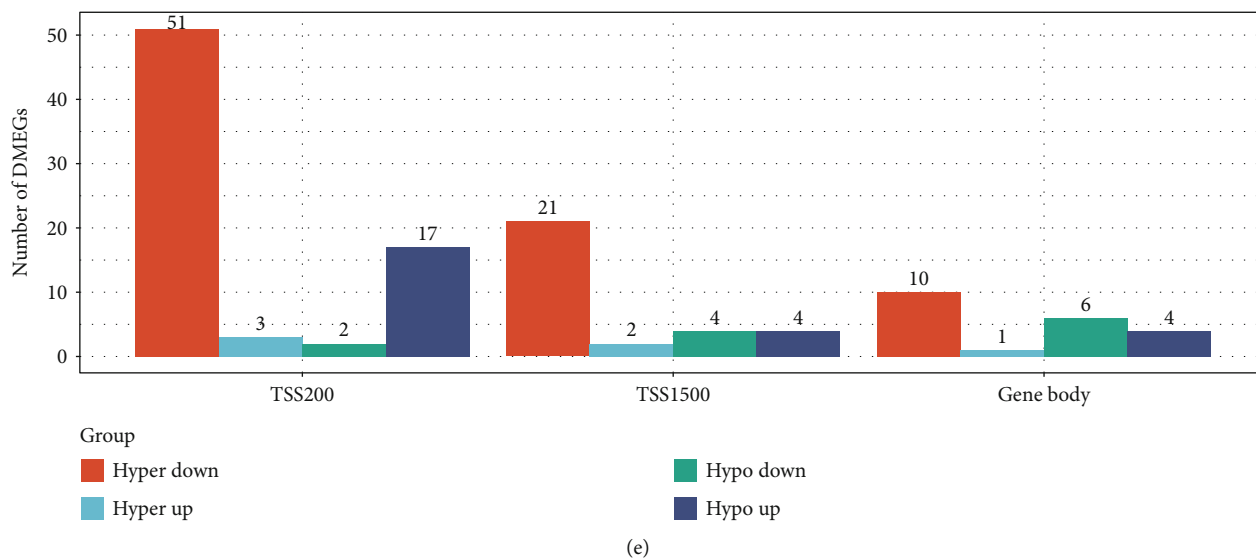


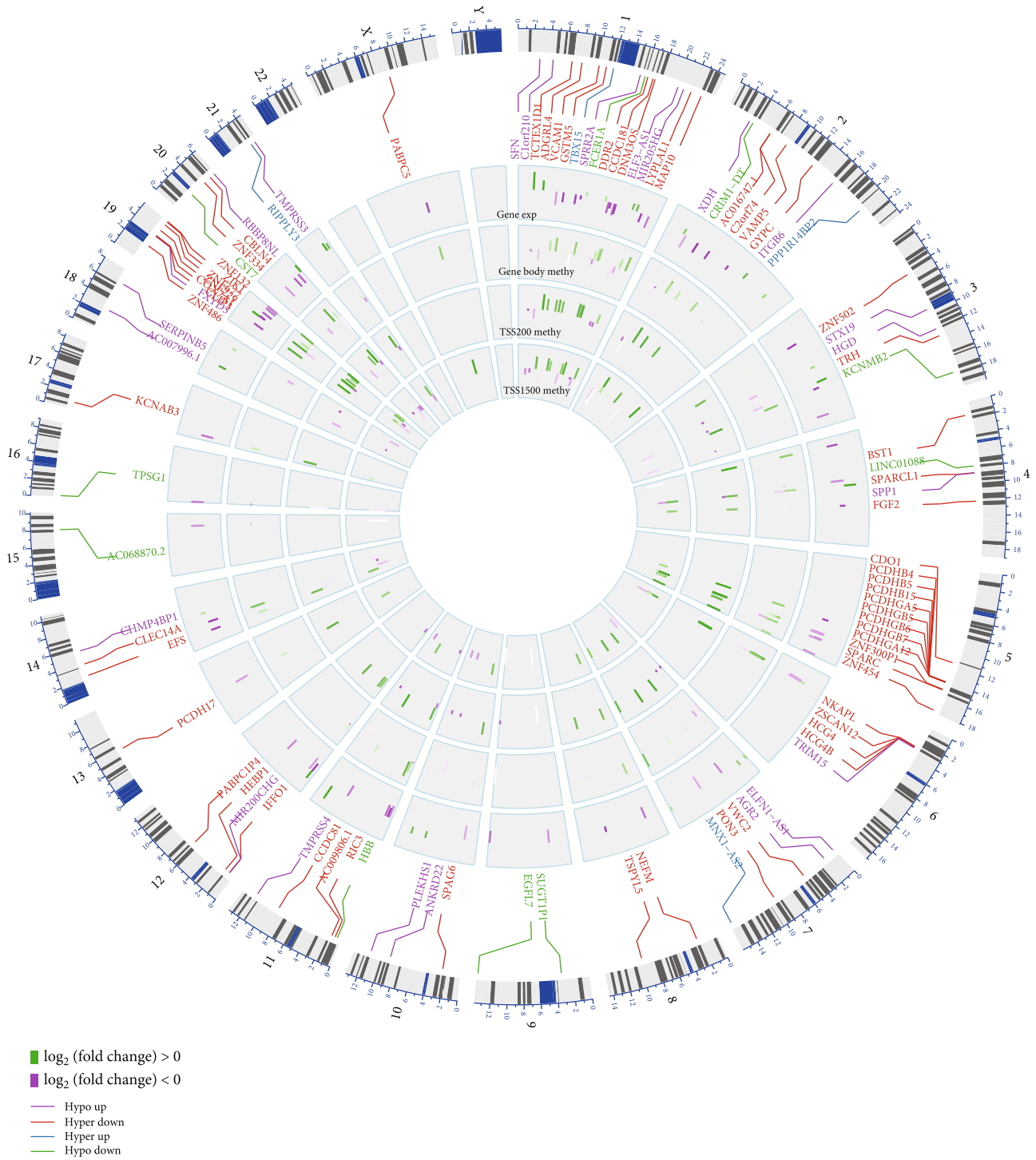
FIGURE 3: Identification of DMDGs. (a) Venn diagram of DEGs and differentially methylated genes in the TSS1500 region. (b) Venn diagram of DEGs and differentially methylated genes in TSS200 region. (c) Venn diagram of the intersection of DEGs and differential methylation genes in the gene body region. (d) Quadrant maps of DEGs and differentially methylated genes in TSS1500, TSS200, and gene body regions. (e) Histograms of four regulatory patterns of DEGs and differentially methylated genes in TSS1500, TSS200, and gene body regions.

on 22 chromosomes, in which most of the DMDGs were on chr1, chr2, chr5, and chr19 chromosomes, and the methylation patterns of DMDGs in different gene regions were consistent (Figure 4(a)). To explore the difference between gene expression and DNA methylation data in tumor and normal samples, a linear discriminant classification model was developed based on gene expression profile of DMDGs and methylation profile of the three regions, respectively. According to the results of PCA analysis, both the expression profile and the methylation profile of the three regions can neatly classify tumor samples and normal samples (Figure 4(b)). ROC analysis showed that the expression profiles and methylation profiles of the three regions had ultrahigh classification performance (Figure 4(c)). Our study on biological processes and regulatory mechanisms of 96 DMDG enrichment indicated that 14 terms were enriched, including cell-cell adhesion via plasma-membrane adhesion molecules, homophilic cell adhesion via plasma membrane adhesion molecules, and protein complex involved in cell adhesion, which were all related to cell adhesion (Figure 4(d)).

3.4. Screening and Verification of Screening and Syndromes of Potential Therapeutic Compounds in DMDGs. To identify drugs related to DMDGs, protein-chemical compound interaction analysis was carried out. Table 1 lists 13 DMDGs and compounds, including BST1, CDO1, CLEC14A, COX7A1, DDR2, FCER1A, GSTM5, HBB, RIC3, SERPINB5, SPARC, VCAM1, and XDH. The protein-chemical compound interaction network more intuitively displayed the interaction between DMDGs and drugs (Figure 5). Considering there was a lack of research on CLEC14A, molecular docking analysis was conducted by taking resveratrol-CLEC14A as

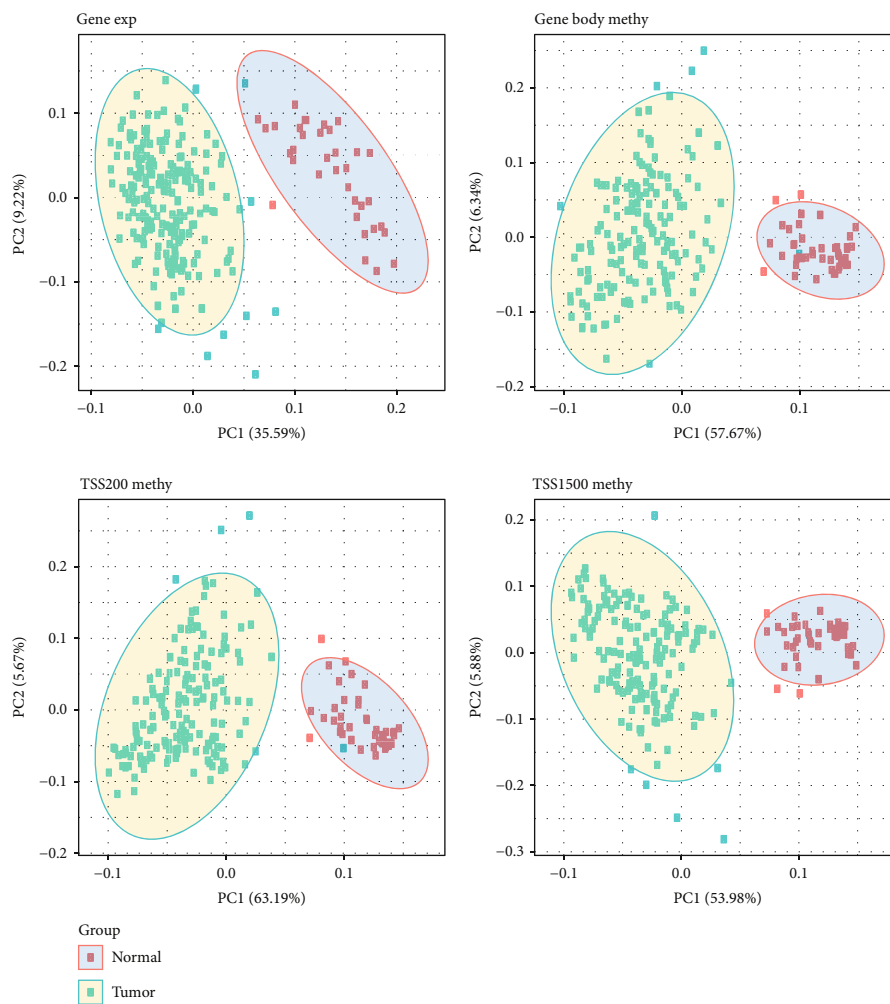
an example. Firstly, the 3D structure of CLEC14A protein was predicted by SwissModel (<https://swissmodel.expasy.org/>). There were large neutral amino acid transporters in CLEC14A, and small subunits showed a high sequence similarity, which was an optimal model plate for modeling (Figure 6(a)). The results of Ramachandran plot analysis in PDBSum (<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum>) showed that 99.1% of the residues were in the allowable area, indicating a high accuracy of predicted structure (Figure 6(b)). Stability evaluation of the protein model using root-mean-square deviation (RMSD) method demonstrated that in the process of molecular dynamics simulation of 100 ns, the RMSD value of protein skeleton (protein-backbone) only slightly increased before 10 ns, and then, the whole protein remained in a relatively stable state, which confirmed the relative stability of CLEC14A protein in molecular dynamics simulation (Figure 6(c)). Furthermore, analysis on the docking mode between resveratrol and CLEC14A protein using Autodock Vina showed that resveratrol had a strong binding to CLEC14A protein mainly through hydrogen bond interaction with amino acid residues TYR149 and ALA155 in CLEC14A protein and hydrophobic interaction with amino acid residues such as PRO154, LEU74, ARG78, and LYS88 (Figure 6(d) and 6(e)).

3.5. Establishment of a Prognostic Model Based on DMDGs. To identify genes related to the prognosis of EC from DMDGs, the expression of 96 DMDGs and survival data were analyzed by LASSO regression analysis in the training set. According to the number of times each probe appeared during 1000 times, the combination with the highest frequency contained two genes, ELFN1-AS1 and ZNF132 (Figure 7(a)). The immune infiltration score of each patient

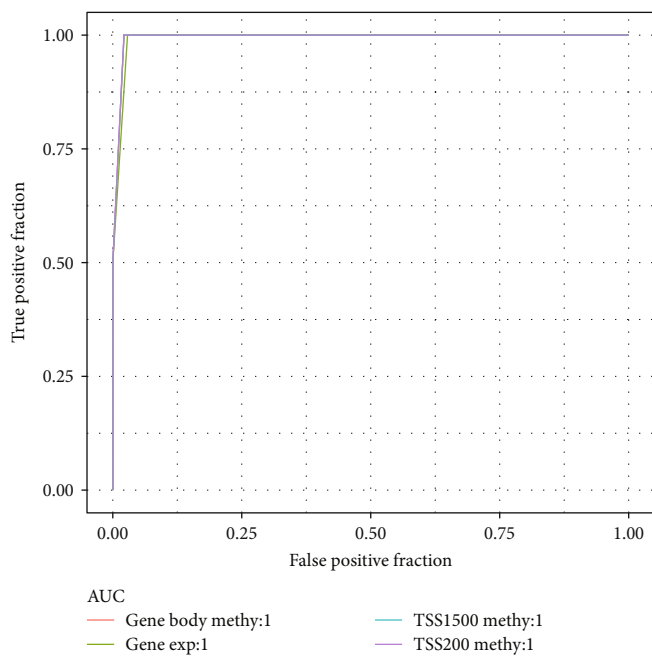


(a)

FIGURE 4: Continued.



(b)



(c)

FIGURE 4: Continued.

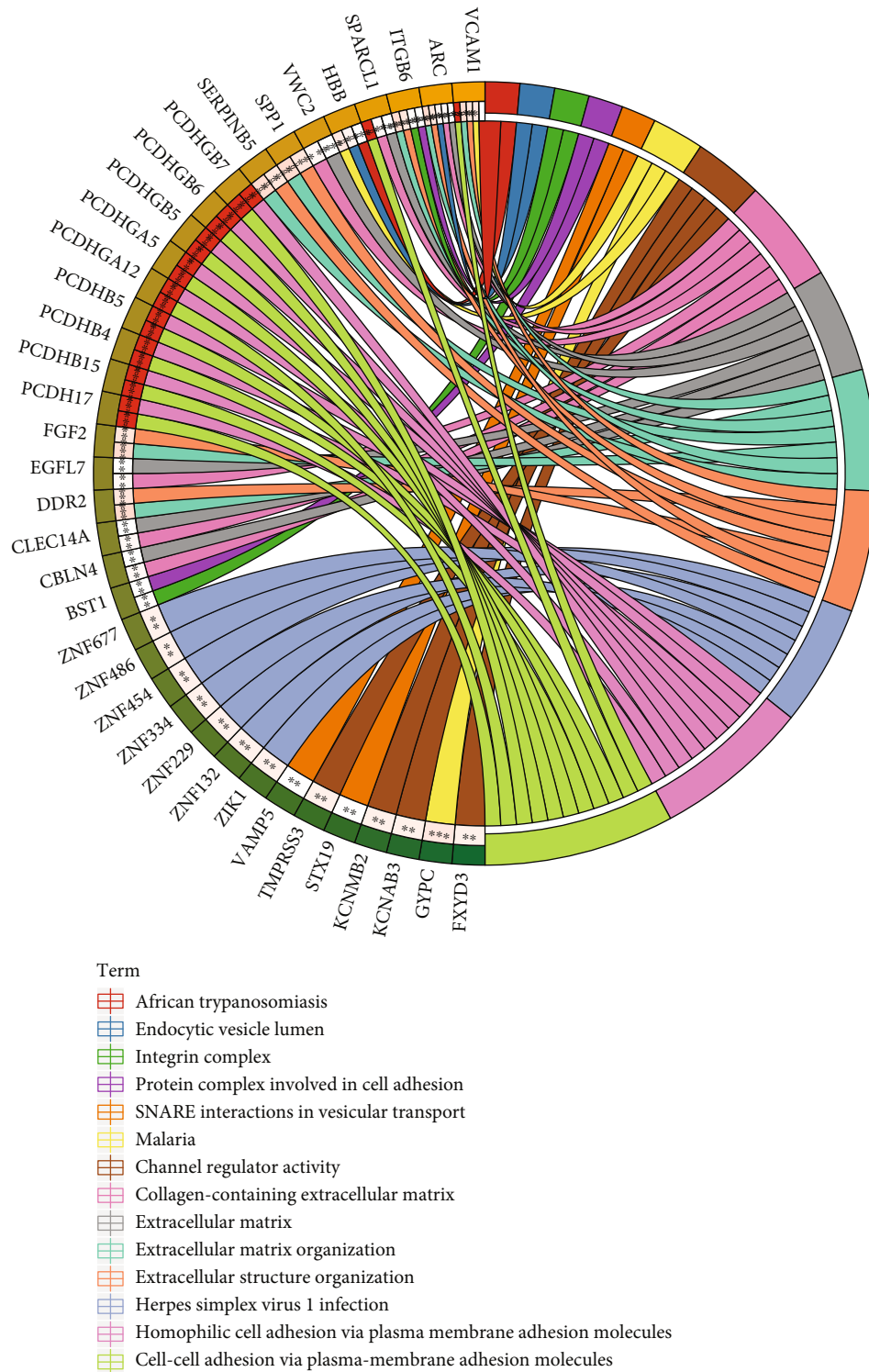


FIGURE 4: Chromosomal mapping and biological processes and mechanisms of DMDGs. (a) The location and distribution of DMDGs on chromosomes. (b) Expression of DMDGs and PCA analysis of methylation. (c) The ROC curve of classification performance of linear discriminant model based on the DMDG expression profile and methylation profile. (d) Results of KEGG pathway and GO enrichment in DMDGs, different colors represented different pathways, and wiring indicated a relationship between genes and pathways.

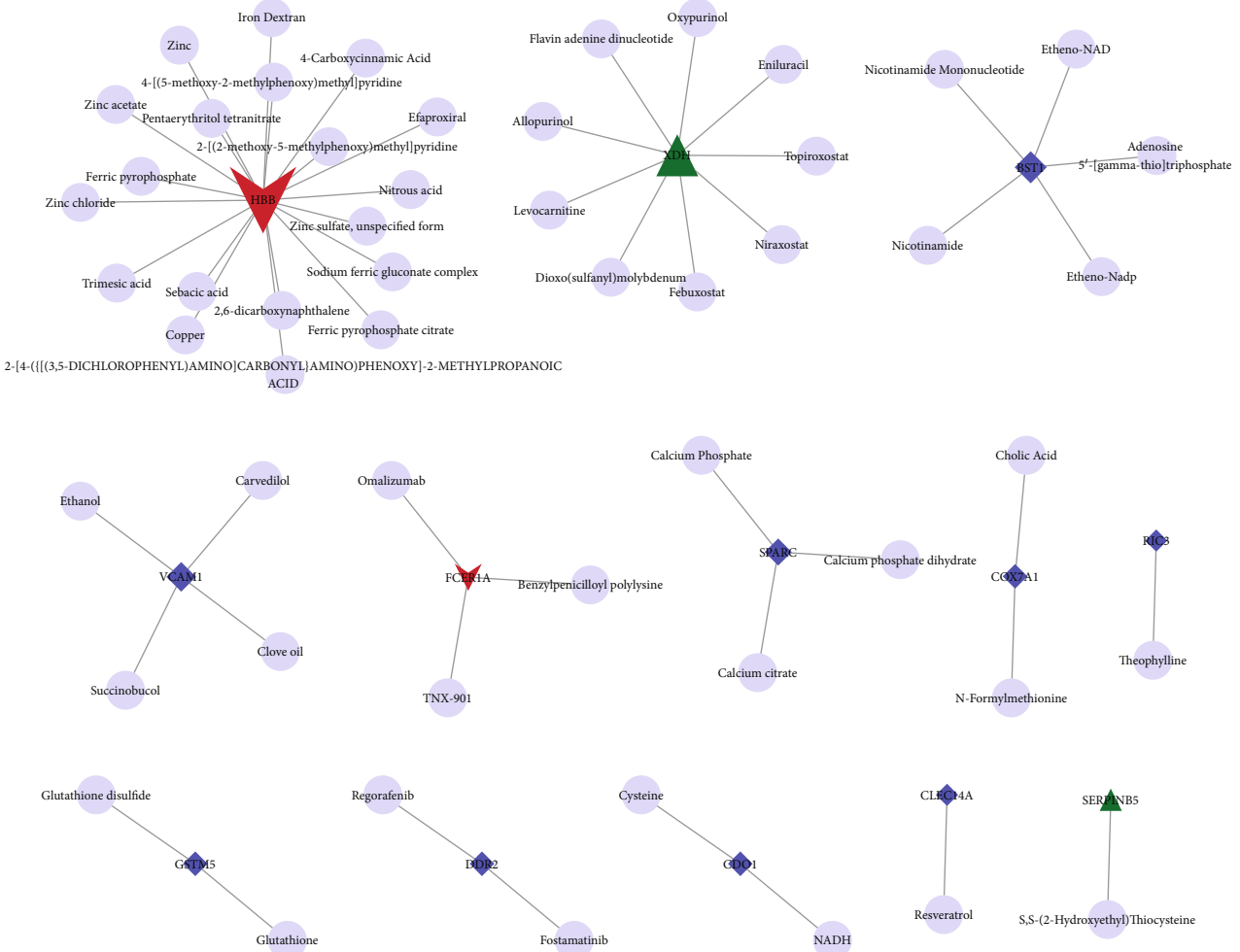


FIGURE 5: Interaction network between DMEGs and drug.

was evaluated by R software package estimate. Through analyzing the relationship between these two genes and immune infiltration, it can be observed that they had a significant negative correlation with immune microenvironment infiltration (Supplementary figure 1). We further analyzed the variation coefficient trajectories of different lambdas and the distribution of standard deviation of models under different lambdas, and the risk score model was obtained: Risk score = 0.224 * ENSG00000236081 (ELFN1 - AS1) - 1.035 * ENSG00000131849 (ZNF132) (Figure 7(b), Table 2). After using the two genes in risk score system to divide patients into groups, we found that patients with low expression of ELFN1-AS1 showed a higher survival rate and those with high expression of ZNF132 had obvious survival advantage (Figures 7(c) and 7(d)). The risk score of each sample was calculated and arranged in the training set, and the results showed that the survival time of the sample with a high-risk score was significantly shorter than that with a low-risk score, suggesting that the sample with a high-risk score may have a worse prognosis (Figure 7(e)). ROC analysis of the prognostic classification of risk score demonstrated that the AUC values of 1-, 3-, and 5-year

ROC curves were all higher above 0.7 (Figure 7(f)). The z-score cut-off value based on risk score was determined by R software package maxstat, and the samples were categorized into the high-risk group ($n = 36$) and low-risk group ($n = 68$). The survival outcome of patients in the high-risk group was found to be significantly worse than that in the low-risk group (Figure 7(g)).

3.6. Evaluation of Prognostic Model. The same method used in the training set was employed to evaluate the accuracy of the model in the verification set and the whole PD-1-negative cohort. Through the arrangement of the risk scores of the samples in each cohort, it could be observed that the number of survival patients gradually decreased with the increase of risk score whether in the verification set or the entire PD-1-negative cohort. The expression of ELFN1-AS1 was also upregulated with the increase of risk score, while ZNF132 showed an opposite trend (Figures 8(a) and 8(d)). The Kaplan-Meier curves based on the log-rank test were plotted to visualize the prognostic value of 2-gene signature in validation set and the entire PD-1-negative cohort. Consistent with the results of the verification set, high-risk

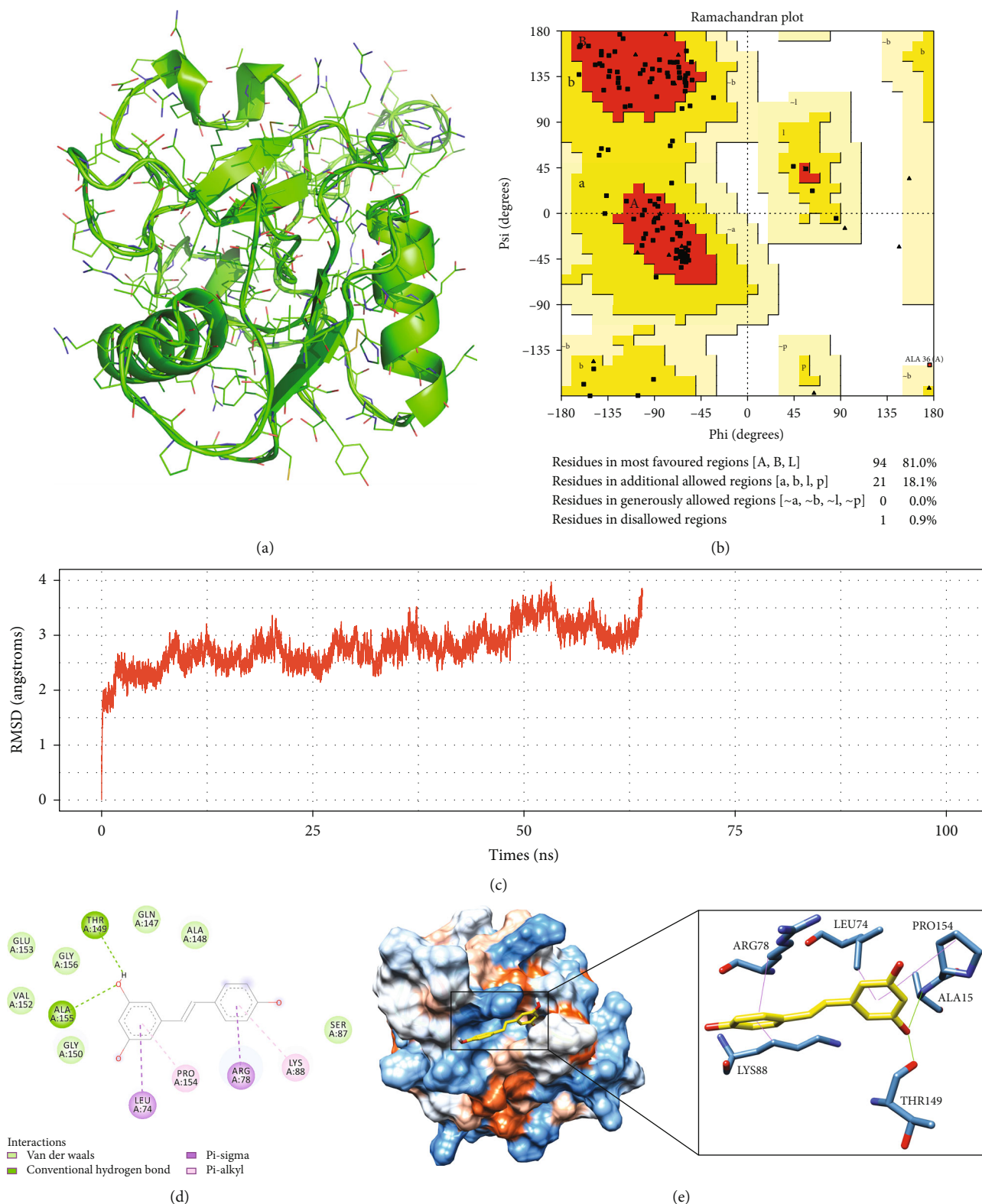


FIGURE 6: Molecular docking of resveratrol and CLEC14A. (a) 3D structure of CLEC14A protein predicted in SwissModel. (b) The Ramachandran plot of CLEC14A protein structure, and red, yellow, and white were the conformational permissible, permissible, and nonpermissible regions, respectively. (c) The change curve of the RMSD value of protein skeleton in the process of 100 ns molecular dynamics. (d) Two-dimensional interaction diagram of resveratrol and CLEC14A protein, the green dotted line represented the hydrogen bond, and the pink dotted line represented the hydrophobic interaction. (e) The 3-dimensional docking results of resveratrol and CLEC14A protein, the yellow structure represented the ligand, and the heteroatoms on it were shown as elements. The steel blue structure was an amino acid residue, and the heteroatoms on it were also shown as elements.

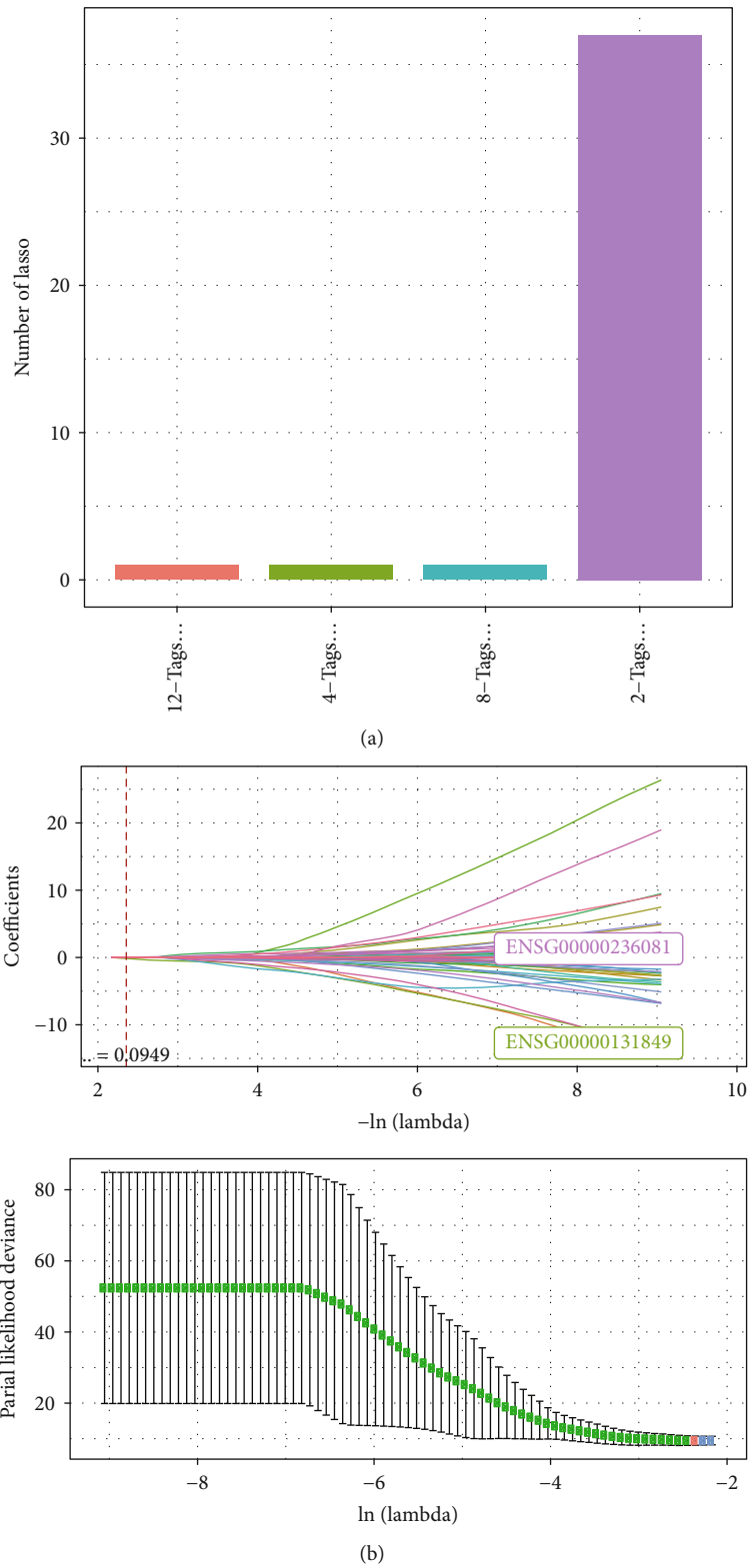
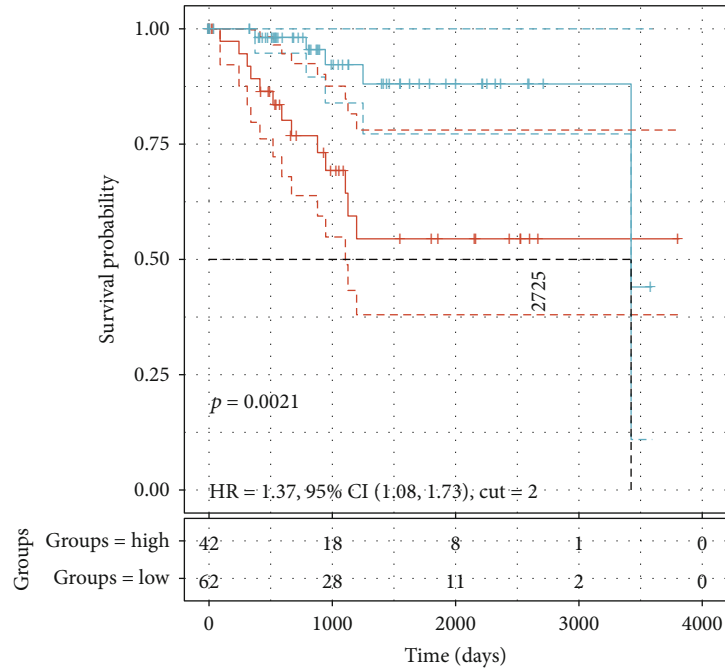
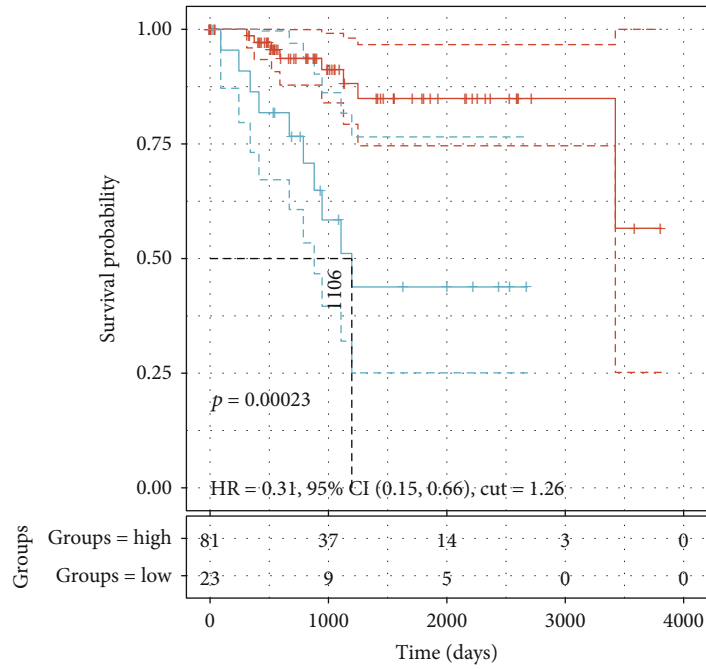


FIGURE 7: Continued.



Groups
 + Groups = high
 + Groups = low

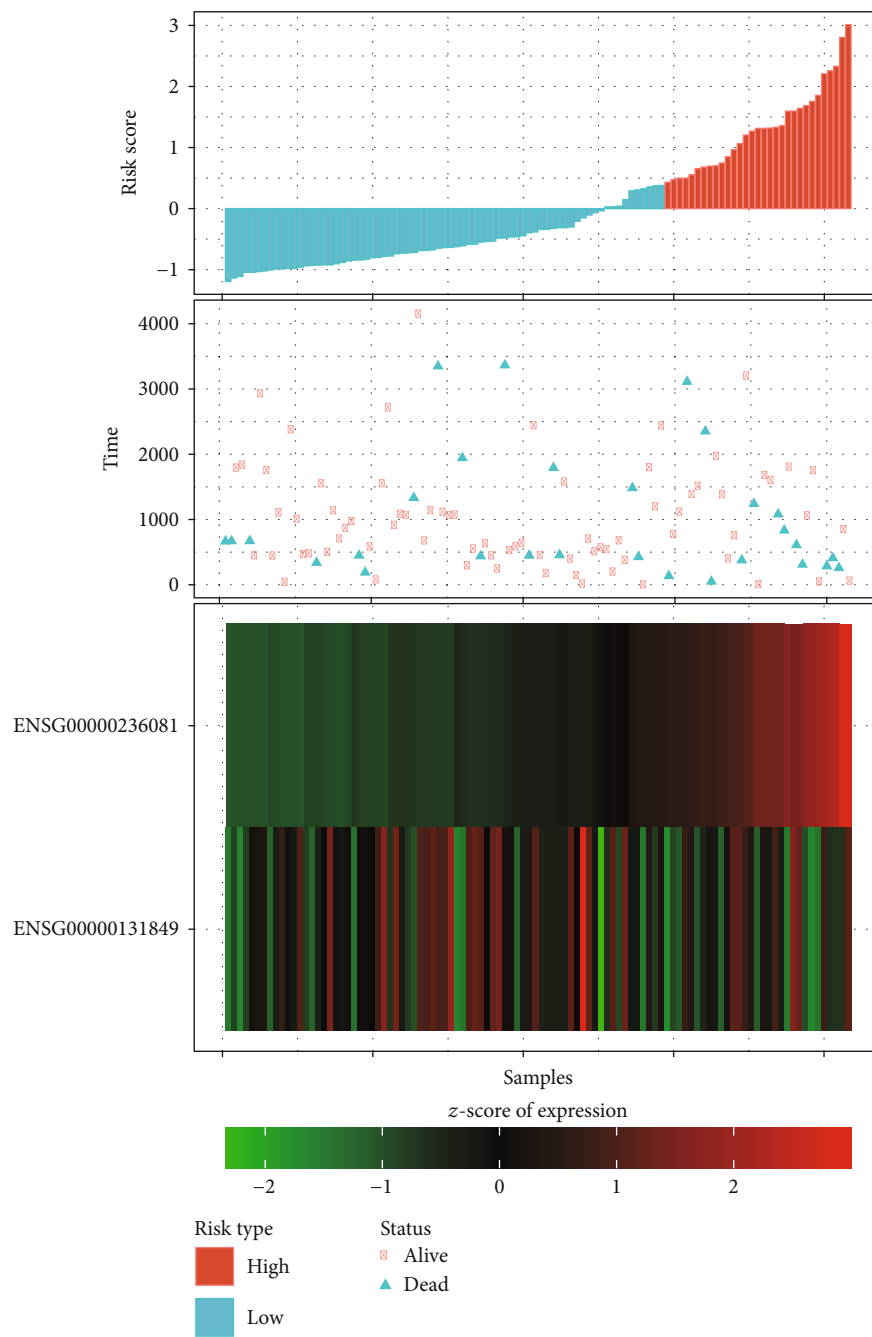
(c)



Groups
 + Groups = high
 + Groups = low

(d)

FIGURE 7: Continued.



(e)

FIGURE 7: Continued.

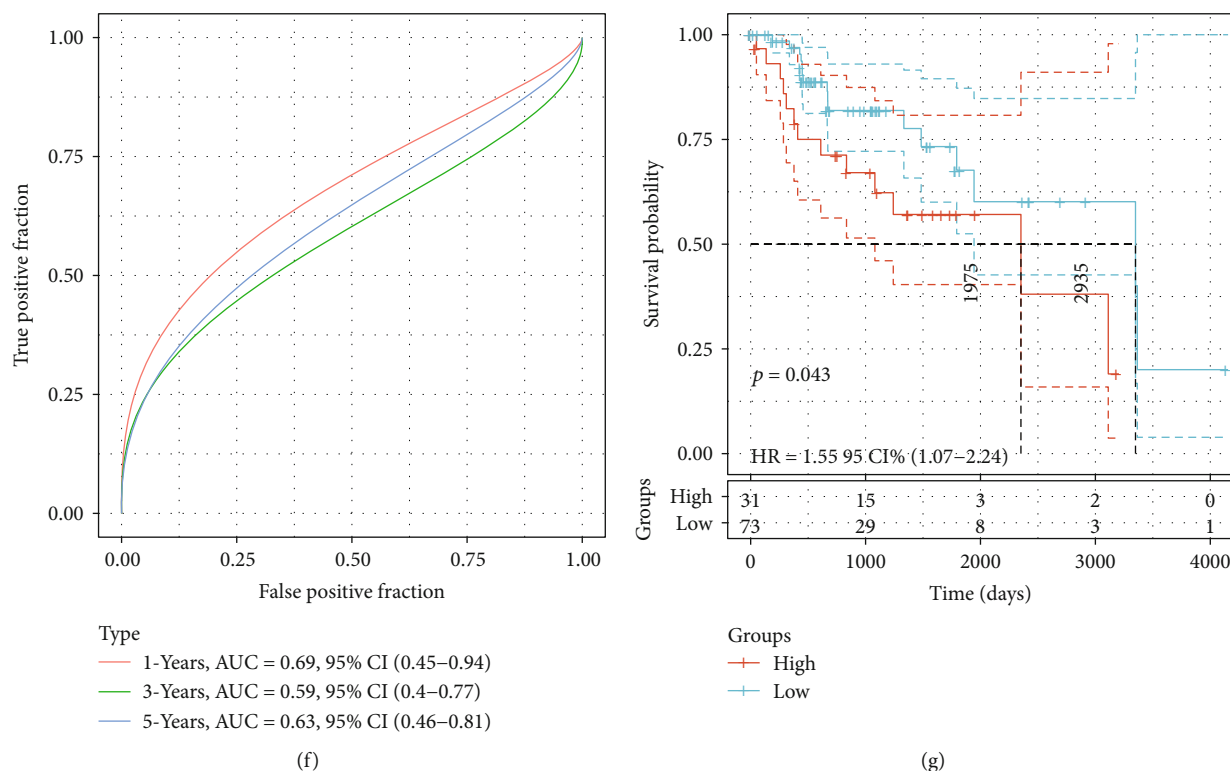


FIGURE 7: Prognostic model was established based on DMDGs. (a) The frequency of each gene combination in 1000 Lasso regression. (b) The variation coefficient locus and standard deviation distribution of different lambdas. (c) The Kaplan-Meier curve of the patients who were grouped according to the expression of ELFN1-AS1. (d) The Kaplan-Meier curve of patients who were grouped according to the expression of ZNF132. (e) The survival time, survival status, and 2-gene expression of risk score in the training set. (f) ROC curve of 2-gene signature classification. (g) Kaplan-Meier curve of samples with different risks in the training set.

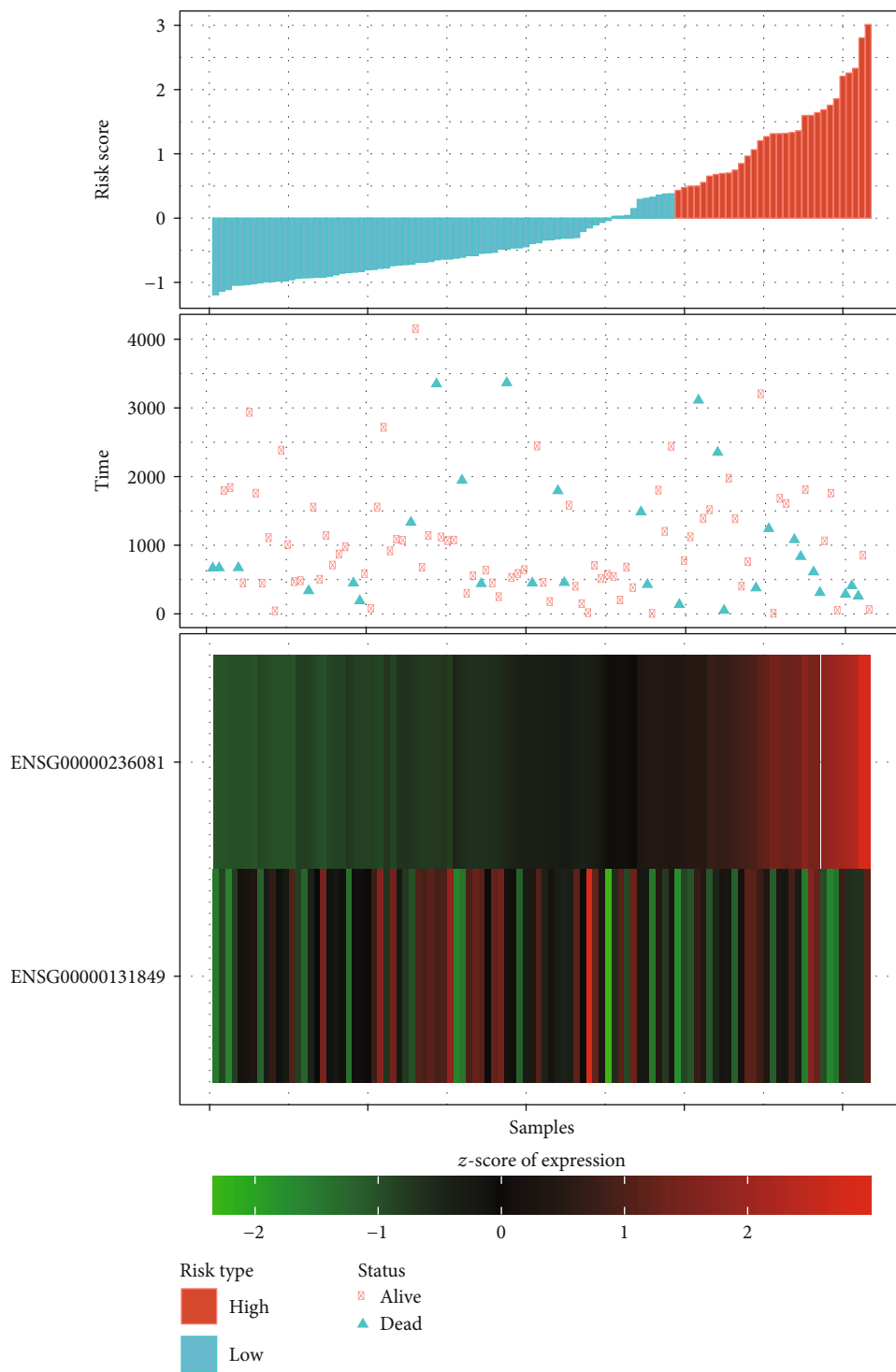
TABLE 2: 13 DMDGs and compounds that can interact with them.

Gene	GeneType	DrugCount	Drug example
BST1	HyperDown	5	Etheno-NAD, nicotinamide, adenosine 5'-[gamma-thio] triphosphate
CDO1	HyperDown	2	Cysteine, NADH
CLEC14A	HyperDown	1	Resveratrol
COX7A1	HyperDown	2	Cholic acid, N-formylmethionine
DDR2	HyperDown	2	Regorafenib, fostamatinib
FCER1A	HypoDown	3	Omalizumab, benzylpenicilloyl polylysine, TNX-901
GSTM5	HyperDown	2	Glutathione, glutathione disulfide
HBB	HypoDown	19	Iron dextran, zinc, 4-carboxycinnamic acid
RIC3	HyperDown	1	Theophylline
SERPINB5	HypoUp	1	S,S-(2-Hydroxyethyl)thiocysteine
SPARC	HyperDown	3	Calcium citrate, calcium phosphate, calcium phosphate dihydrate
VCAM1	HyperDown	4	Ethanol, carvedilol, succinobucol
XDH	HypoUp	9	Allopurinol, levocarnitine, topiroxostat

samples showed a higher death risk than low-risk samples (Figures 8(b) and 8(e)). The ROC curves in the verification set and the entire PD-1-negative cohort demonstrated that the risk system performed well in predicting 1-year and 5-year survival rates of patients with EC, with an AUC higher than 0.6 (Figures 8(c) and 8(f)).

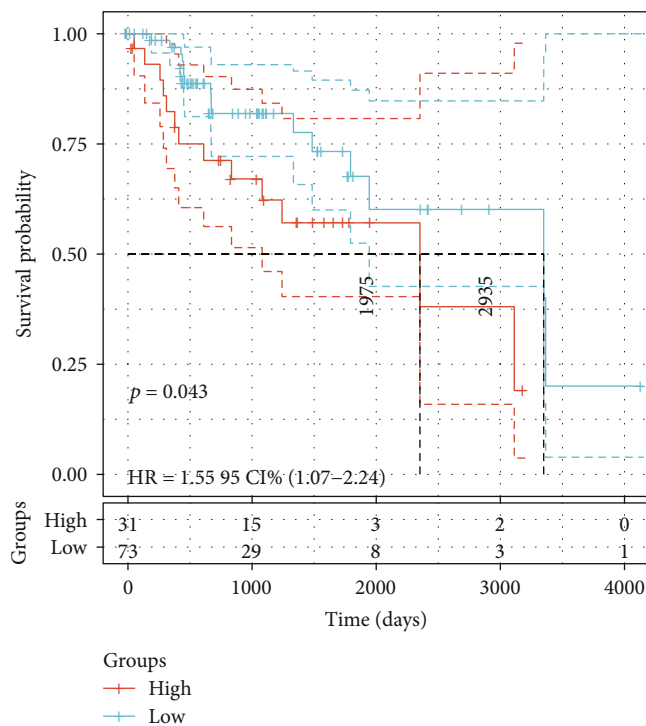
4. Discussion

At present, prediction of EC prognosis is still largely based on routine pathological indexes such as clinical stage, degree of tumor differentiation and invasion, and vascular infiltration. The development of molecular biology has enriched

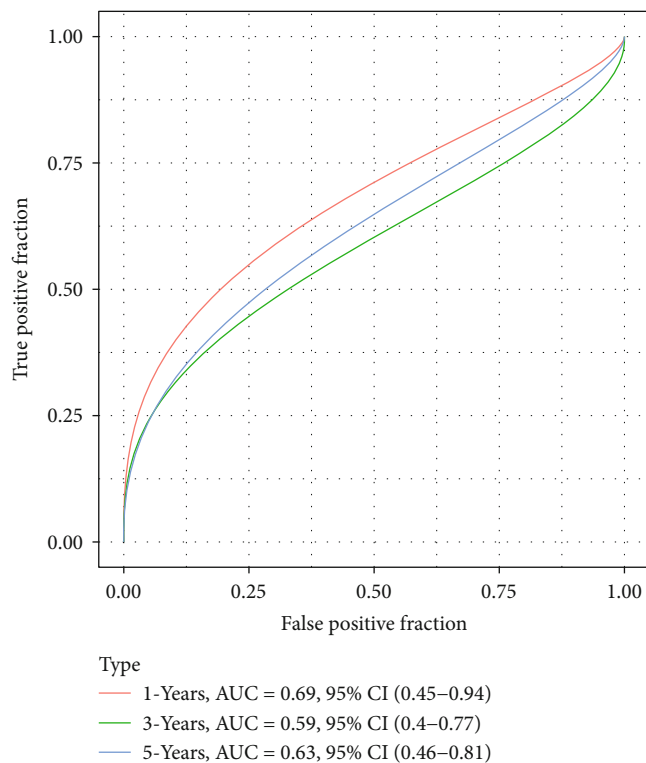


(a)

FIGURE 8: Continued.

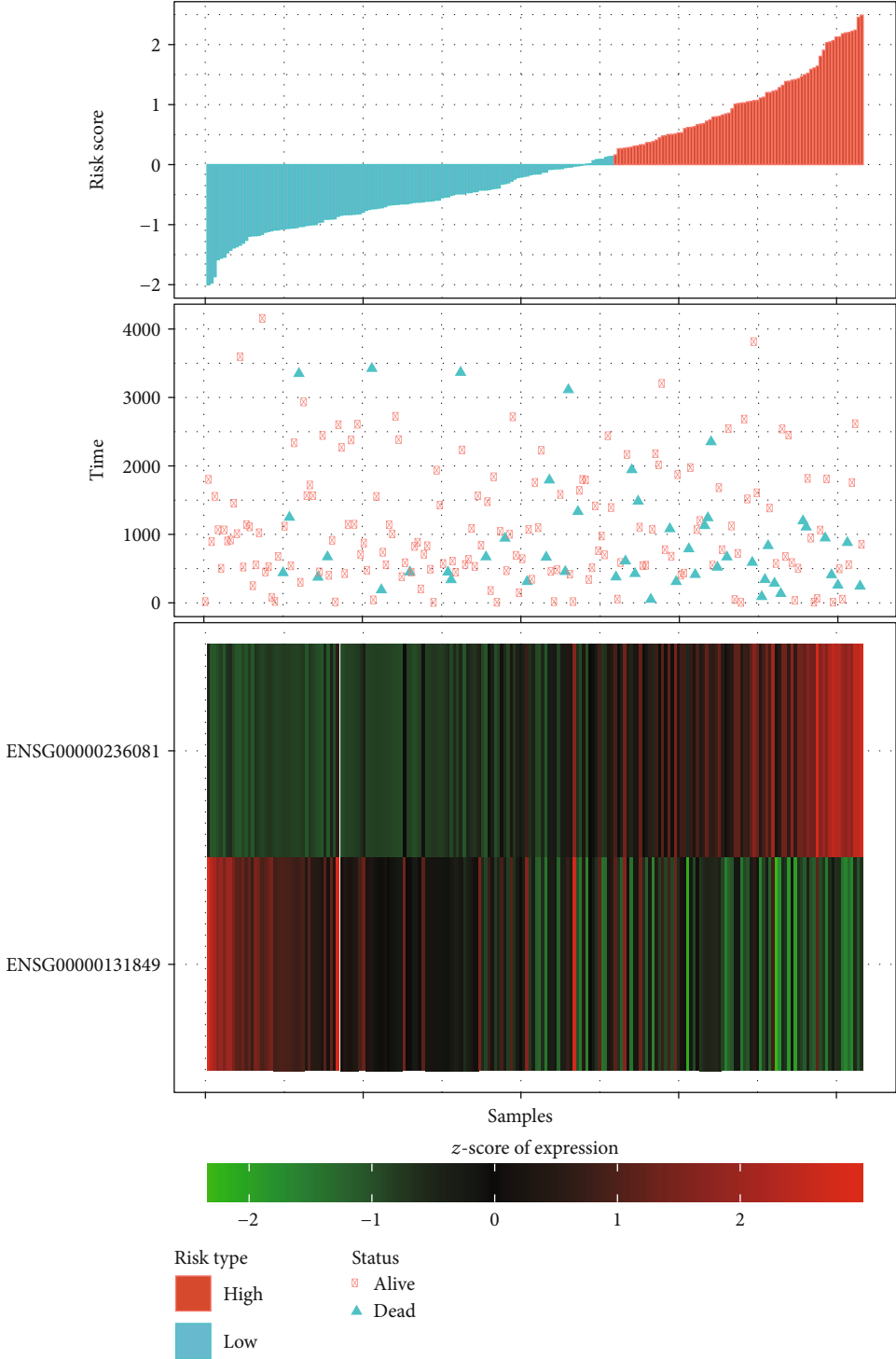


(b)



(c)

FIGURE 8: Continued.



(d)

FIGURE 8: Continued.

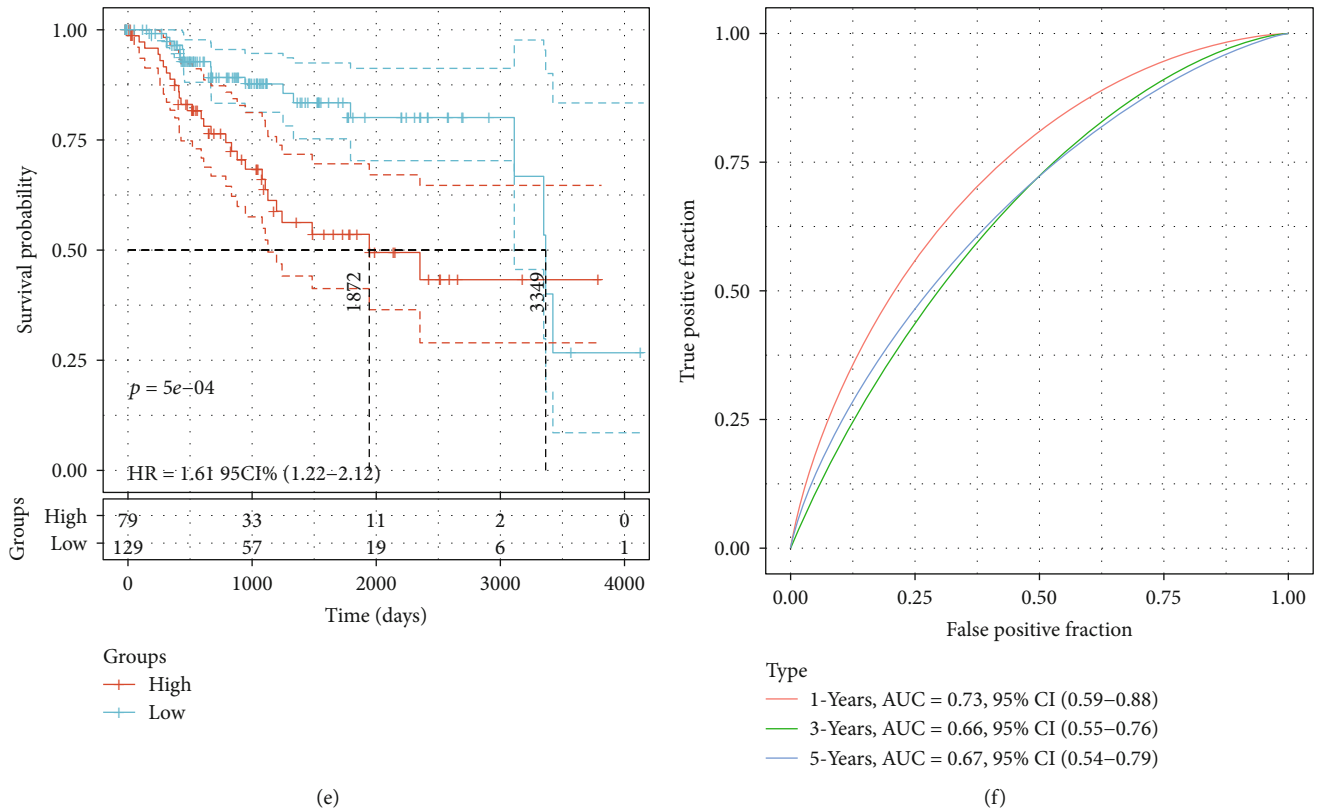


FIGURE 8: Verification of the 2-gene signature. (a) The risk score distribution, survival status scatter diagram, and expression heat map of 2 genes in the verification set. (b) Kaplan-Meier curves based on the log-rank test were created to visualize the prognostic value of 2-gene signature in the verification set. (c) The ROC curve over time in the validation set. (d) The risk score of the samples, distribution of survival time and state, and the expression heat map of ELFN1-AS1 and ZNF132 in the PD-1-negative cohort. (e) Kaplan-Meier curves based on the log-rank test were generated to visualize the prognostic value of the 2-gene signature in the entire PD-1-negative cohort. (f) ROC curve with time in PD-1-negative cohort.

the knowledge of genetic mechanism of EC, contributing to the development of new methods for diagnosis, treatment, and prognosis [27]. Antiprogrammed cell death protein 1 immunotherapy has now been approved by the U.S. Food and Drug Administration for treating the female subgroup with metastatic EC [6]. The effective response to anti-PD-L1/anti-PD-1 therapy requires the establishment of a complete immune cycle, each step of which is regulated by epigenetic modification [28]. In view of such an effect of the two, the PD-1-positive EC samples from TCGA were eliminated, and DMDGs were screened according to the joint analysis of expression data and methylation data of the remaining EC samples.

Through the overlap analysis of differential methylation genes in DEGs, TSS1500, TSS200, and gene body regions, we obtained a total of 96 DMDGs from the three regions, which were distributed on 22 chromosomes, and the most abundant genes were found on chr1, chr2, chr5, and chr19 chromosomes. Additionally, a linear discriminant classification model was developed based on the gene expression profile of DMDGs and the methylation profile of the three regions, and the results of PCA analysis confirmed that both the expression profile and methylation profile of the three regions can clearly classify tumor and normal samples. The above 96 DMGMs were enriched in 14 terms, and the most

obvious enrichment items were biological functions and mechanisms related to cell adhesion; therefore, DMGMs may be involved in regulating the pathological process of EC.

At the same time, the drug components targeting DMGMs were analyzed through the developing interactions network of protein-compounds, and it was found that the compounds targeting HBB were the most. Among several compounds targeting DMDGs, some have been shown to have potential anticancer effects. For example, Hirsh et al.'s study found that theophylline induces apoptosis of cancer cells, and they also mentioned an ongoing clinical trial using theophylline to treat patients with lung cancer [29]. Regorafenib is a multitarget kinase inhibitor approved for the treatment of metastatic colorectal cancer patients who are unresponsive to standard chemotherapy [30]. A large population-based cohort study showed that long-term treatment of carvedilol is associated with a reduced risk of upper digestive tract and lung cancer and may be a potential drug to prevent these cancers [31]. Succinobucol-loaded nanoparticles can greatly inhibit lung metastasis of breast cancer [32]. Resveratrol alleviates the deterioration of a variety of human cancers, including gynecological cancer, kidney cancer, liver cancer, bladder cancer, prostate cancer, lung cancer, thyroid cancer, esophageal cancer, gastric cancer,

colorectal cancer, and ovarian cancer. However, the potential therapeutic target and mechanism are not clear [33]. Some studies have found that CLEC14A methylation was associated with its expression and progression of lung adenocarcinoma and is a therapeutic target for solid tumors [34, 35]. Herein, we predicted that resveratrol can target methylation-driven gene CLEC14A and that the two were stably and tightly bound to each other through the formation of hydrogen bonds and hydrophobic interactions.

Finally, a risk score model composed of two methylation-driven genes was constructed through further analysis on 96 DMDGs. The progress of ELFN1-AS1 driving a variety of cancers, but whether its expression was affected by methylation was still unknown. A study has shown that ELFN1-AS1 drives progression of many cancers [36–38], but whether its expression is affected by methylation remains unknown. ZNF132 has been reported to be downregulated due to abnormally hypermethylation in prostate cancer [39] and lung cancer [40], and this is related to the prognosis of cancer. Herein, low expression of ZNF132 driven by DNA hypomethylation was also found to be significantly related to the poor prognosis of EC, and the signature composed of DMDGs showed a high accuracy in predicting the prognosis of EC. Using Sangerbox (<http://vip.sangerbox.com>), we also evaluated the prognostic relationship between ELFN1-AS1 and ZNF132 in other tumors. We filtered samples with a follow-up time of shorter than 30 days. It can be observed that ELFN1-AS1 is also significantly correlated with the poor prognosis of renal tumors, uveal melanoma, and liver cancer (Supplementary figure 2A), and ZNF132 is significantly correlated with the poor prognosis of leukemia and glioma (Supplementary figure 2B). It is a promising approach to improving diagnosis, prognosis, and treatment strategy based on methylation-driven genes, but more comprehensive researches are needed in the future for further verifying our results.

Data Availability

All data generated or analyzed during this study are included in this article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported in part by grants from the National Natural Science Foundation of China (81971721).

Supplementary Materials

Supplementary 1. Supplementary Figure 1: correlation between ELFN1-AS1, ZNF132, and immune microenvironment infiltration.

Supplementary 2. Supplementary Figure 2: prognostic correlation of ELFN1-AS1 and ZNF132 in pan-cancer. (A) Prognostic significance of ELFN1-AS1 in pan-cancer. (B) Prognostic significance of ZNF132 in pan-cancer.

References

- [1] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, "Cancer statistics, 2021," *CA: a Cancer Journal for Clinicians*, vol. 71, no. 1, pp. 7–33, 2021.
- [2] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [3] S. Rousset-Rouviere, P. Rochigneux, A. S. Chrétien et al., "Endometrial carcinoma: immune microenvironment and emerging treatments in immuno-oncology," *Biomedicine*, vol. 9, no. 6, 2021.
- [4] M. A. Clarke, B. J. Long, A. del Mar Morillo, M. Arbyn, J. N. Bakkum-Gamez, and N. Wentzensen, "Association of endometrial cancer risk with postmenopausal bleeding in women: a systematic review and meta-analysis," *JAMA Internal Medicine*, vol. 178, no. 9, pp. 1210–1222, 2018.
- [5] A. Jamieson, T. Bosse, and J. N. McAlpine, "The emerging role of molecular pathology in directing the systemic treatment of endometrial cancer," *Therapeutic Advances in Medical Oncology*, vol. 13, article 17588359211035959, 2021.
- [6] R. A. Brooks, G. F. Fleming, R. R. Lastra et al., "Current recommendations and recent progress in endometrial cancer," *CA: a Cancer Journal for Clinicians*, vol. 69, no. 4, pp. 258–279, 2019.
- [7] K. Banno, M. Yanokura, M. Iida, K. Masuda, and D. Aoki, "Carcinogenic mechanisms of endometrial cancer: involvement of genetics and epigenetics," *The Journal of Obstetrics and Gynaecology Research*, vol. 40, no. 8, pp. 1957–1967, 2014.
- [8] S. B. Baylin and P. A. Jones, "Epigenetic determinants of cancer," *Cold Spring Harbor Perspectives in Biology*, vol. 8, no. 9, 2016.
- [9] A. Koch, S. C. Joosten, Z. Feng et al., "Analysis of DNA methylation in cancer: location revisited," *Nature Reviews. Clinical Oncology*, vol. 15, no. 7, pp. 459–466, 2018.
- [10] B. Wei, F. Wu, W. Xing et al., "A panel of DNA methylation biomarkers for detection and improving diagnostic efficiency of lung cancer," *Scientific Reports*, vol. 11, no. 1, p. 16782, 2021.
- [11] S. S. Hassan, N. Maqsood, Q. Wang, S. Tao, and S. Sadaf, "A panel of epigenetically dysregulated Wnt signaling pathway genes for non-invasive diagnosis of pediatric acute lymphoblastic leukemia," *Cancer Biomarkers*, vol. 32, no. 4, pp. 459–470, 2021.
- [12] L. Y. Feng, B. B. Yan, Y. Z. Huang, and L. Li, "Abnormal methylation characteristics predict chemoresistance and poor prognosis in advanced high-grade serous ovarian cancer," *Clinical Epigenetics*, vol. 13, no. 1, p. 141, 2021.
- [13] H. Huang, J. Fu, L. Zhang et al., "Integrative analysis of identifying methylation-driven genes signature predicts prognosis in colorectal carcinoma," *Frontiers in Oncology*, vol. 11, article 629860, 2021.
- [14] Y. Bai, C. Wei, Y. Zhong et al., "Development and validation of a prognostic nomogram for gastric cancer based on DNA methylation-driven differentially expressed genes," *International Journal of Biological Sciences*, vol. 16, no. 7, pp. 1153–1165, 2020.
- [15] X. Y. Hao, A. Q. Li, H. Shi et al., "A novel DNA methylation-based model that effectively predicts prognosis in hepatocellular carcinoma," *Bioscience Reports*, vol. 41, no. 3, 2021.

- [16] S. Hutt, A. Taylor, P. Ellis, A. Michael, S. Butler-Manuel, and J. Chatterjee, "The role of biomarkers in endometrial cancer and hyperplasia: a literature review," *Acta Oncologica*, vol. 58, no. 3, pp. 342–352, 2019.
- [17] J. Lai, T. Xu, and H. Yang, "Protein-based prognostic signature for predicting the survival and immunotherapeutic efficiency of endometrial carcinoma," *BMC Cancer*, vol. 22, no. 1, p. 325, 2022.
- [18] J. Liu, R. Geng, S. Ni et al., "Pyroptosis-related lncRNAs are potential biomarkers for predicting prognoses and immune responses in patients with UCEC," *Molecular Therapy-Nucleic Acids*, vol. 27, pp. 1036–1055, 2022.
- [19] X. Tan, S. Liu, L. Yao, G. Cui, J. Liu, and J. Ding, "Comprehensive analysis of a novel lipid metabolism-related gene signature for predicting the prognosis and immune landscape in uterine corpus endometrial carcinoma," *Journal of Oncology*, vol. 2022, Article ID 8028825, 2022.
- [20] T. Ruan, J. Wan, Q. Song, P. Chen, and X. Li, "Identification of a novel epithelial-mesenchymal transition-related gene signature for endometrial carcinoma prognosis," *Genes*, vol. 13, no. 2, p. 216, 2022.
- [21] M. E. Ritchie, B. Phipson, D. Wu et al., "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, article e47, 2015.
- [22] A. Solari and J. J. Goeman, "Minimally adaptive BH: a tiny but uniform improvement of the procedure of Benjamini and Hochberg," *Biometrical Journal*, vol. 59, no. 4, pp. 776–780, 2017.
- [23] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *OMICS*, vol. 16, no. 5, pp. 284–287, 2012.
- [24] Z. Shao, M. J. Er, and N. Wang, "An efficient leave-one-out cross-validation-based extreme learning machine (ELOO-ELM) with minimal user intervention," *IEEE Transactions on Cybernetics*, vol. 46, no. 8, pp. 1939–1951, 2016.
- [25] D. S. Wishart, C. Knox, A. C. Guo et al., "DrugBank: a knowledgebase for drugs, drug actions and drug targets," *Nucleic Acids Research*, vol. 36, suppl_1, pp. D901–D906, 2008.
- [26] G. Zhou, O. Soufan, J. Ewald, R. E. W. Hancock, N. Basu, and J. Xia, "NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis," *Nucleic Acids Research*, vol. 47, no. W1, pp. W234–W241, 2019.
- [27] K. Banno, Y. Nogami, I. Kisu et al., "Candidate biomarkers for genetic and clinicopathological diagnosis of endometrial cancer," *International Journal of Molecular Sciences*, vol. 14, no. 6, pp. 12123–12137, 2013.
- [28] X. Chen, X. Pan, W. Zhang et al., "Epigenetic strategies synergize with PD-L1/PD-1 targeted cancer immunotherapies to enhance antitumor responses," *Acta Pharmaceutica Sinica B*, vol. 10, no. 5, pp. 723–733, 2020.
- [29] L. Hirsh, A. Dantes, B. S. Suh et al., "Phosphodiesterase inhibitors as anti-cancer drugs," *Biochemical Pharmacology*, vol. 68, no. 6, pp. 981–988, 2004.
- [30] H. Arai, F. Battaglin, J. Wang et al., "Molecular insight of regorafenib treatment for colorectal cancer," *Cancer Treatment Reviews*, vol. 81, article 101912, 2019.
- [31] C. S. Lin, W. S. Lin, C. L. Lin, and C. H. Kao, "Carvedilol use is associated with reduced cancer risk: a nationwide population-based cohort study," *International Journal of Cardiology*, vol. 184, pp. 9–13, 2015.
- [32] H. Cao, Z. Zhang, S. Zhao et al., "Hydrophobic interaction mediating self-assembled nanoparticles of succinobucol suppress lung metastasis of breast cancer by inhibition of VCAM-1 expression," *Journal of Controlled Release*, vol. 205, pp. 162–171, 2015.
- [33] A. Rauf, M. Imran, M. S. Butt, M. Nadeem, D. G. Peters, and M. S. Mubarak, "Resveratrol as an anti-cancer agent: a review," *Critical Reviews in Food Science and Nutrition*, vol. 58, no. 9, pp. 1428–1447, 2018.
- [34] C. Su, K. Shi, X. Cheng et al., "Methylation of CLEC14A is associated with its expression and lung adenocarcinoma progression," *Journal of Cellular Physiology*, vol. 234, no. 3, pp. 2954–2962, 2019.
- [35] J. Robinson, K. Whitworth, E. Jinks, Z. Nagy, R. Bicknell, and S. P. Lee, "An evaluation of the tumour endothelial marker CLEC14A as a therapeutic target in solid tumours," *The Journal of Pathology. Clinical Research*, vol. 6, no. 4, pp. 308–319, 2020.
- [36] C. Zhang, H. Lian, L. Xie, N. Yin, and Y. Cui, "LncRNA ELFN1-AS1 promotes esophageal cancer progression by up-regulating GFPT1 via sponging miR-183-3p," *Biological Chemistry*, vol. 401, no. 9, pp. 1053–1061, 2020.
- [37] Y. Du, Y. Hou, Y. Shi, J. Liu, and T. Li, "Long non-coding RNA ELFN1-AS1 promoted colon cancer cell growth and migration via the miR-191-5p/special AT-rich sequence-binding protein 1 Axis," *Frontiers in Oncology*, vol. 10, article 588360, 2020.
- [38] G. Ma, G. Li, A. Gou et al., "Long non-coding RNA ELFN1-AS1 in the pathogenesis of pancreatic cancer," *Annals of Translational Medicine*, vol. 9, no. 10, p. 877, 2021.
- [39] M. O. Abildgaard, M. Borre, M. M. Mortensen et al., "Down-regulation of zinc finger protein 132 in prostate cancer is associated with aberrant promoter hypermethylation and poor prognosis," *International Journal of Cancer*, vol. 130, no. 4, pp. 885–895, 2012.
- [40] K. Hamada, Y. Tian, M. Fujimoto et al., "DNA hypermethylation of the ZNF132 gene participates in the clinicopathological aggressiveness of 'pan-negative'-type lung adenocarcinomas," *Carcinogenesis*, vol. 42, no. 2, pp. 169–179, 2021.