Original Research *Highlight article*

Identifying immune subtypes of uterine corpus endometrial carcinoma and a four-paired-lncRNA signature with immune-related lncRNAs

Nan Li[1,2,3,*], Kai Yu[4,*], Zhong Lin[5] and Dingyuan Zeng[6]

[1]Liuzhou Maternity and Child Healthcare Hospital, Liuzhou 545001, China; [2]Affiliated Maternity Hospital and Affiliated Children's Hospital of Guangxi University of Science and Technology, Liuzhou 545001, China; [3]Guangxi Health Commission Key Laboratory of Birth Cohort Study in Pregnant Women of Advanced Age, Liuzhou 545001, China; [4]College of Animal Science and Technology, Guangxi University, Nanning 530004, China; [5]The Guangxi Zhuang Autonomous Region Reproductive Hospital, Nanning 530021, China; [6]The Department of Obstetrics and Gynecology, Liuzhou Maternity and Child Health Care Hospital, Liuzhou 545001, Guangxi
Corresponding authors: Zhong Lin. Email: uters@126.com; Dingyuan Zeng. Email: zdy13607723199@163.com
*These authors contributed equally to this article.

Impact statement

This study determined two immune subtypes and a four-paired-lncRNA signature for UCEC patients based on immune-related lncRNAs. The two immune subtypes with distinct immune features manifested differential response to immunotherapy, showing a potential to guide personalized therapy. Compared with the previous molecular subtypes, the current signature is simpler but equally effective in classifying patients. In addition, the four-paired-lncRNA signature is robust predicting UCEC prognosis. The study further demonstrated the critical role of immune-related lncRNAs in tumor progression, and these prognostic lncRNAs may be the potential targets for UCEC treatment.

Abstract

Uterine corpus endometrial carcinoma (UCEC) is the third most frequent gynecological malignancies in the female reproductive system. Long non-coding RNAs (lncRNAs) are closely involved in tumor progression. This study aimed to develop an immune subtyping system and a prognostic model based on lncRNAs for UCEC. Paired lncRNAs and non-negative matrix factorization were applied to identify immune subtypes. Enrichment analysis was conducted to assess functional pathways, immune-related genes, and cells. Univariate and multivariate Cox regression analysis were performed to analyze the relation between lncRNAs and overall survival (OS). A prognostic model was constructed and optimized by least absolute shrinkage and selection operator (LASSO) and Akaike information criterion (AIC). Two immune subtypes (C1 and C2) and four paired-prognostic lncRNAs closely associated with overall survival were identified. Some immune features, sensitivity of chemotherapy and immunotherapy, and the relation with immune escape showed variations between two subtypes. A nomogram established based on prognostic model and clinical features was effective in OS prediction. The immune subtyping system based on lncRNAs and the four-paired-lncRNA signature was predictive of UCEC prognosis and can facilitate personalized therapies such as immunotherapy or RNA-based therapy for UCEC patients.

Keywords: Uterine corpus endometrial carcinoma, lncRNAs, immune subtypes, paired-lncRNAs, immune-related lncRNAs, bioinformatics

Experimental Biology and Medicine **2022; 247: 221–236. DOI: 10.1177/15353702211053588**

Introduction

Uterine corpus endometrial carcinoma (UCEC) is a female reproductive malignancy with the third highest incidence in 2020.[1] According to the global cancer statistics, 417,367 new UCEC cases occurred worldwide in 2020, with Northern America and Eastern Europe showing a particularly high incidence and mortality.[1] Although the International Federation of Gynecology and Obstetrics (FIGO) and tumor-node-metastasis (TNM) staging system can classify patients according to the degree of tumor development, they cannot precisely guide the targeted therapies, especially for metastatic patients. Therefore, an accurate subtyping system is needed for guiding targeted treatment such as immunotherapy.

The Cancer Genome Atlas (TCGA) Research Network has previously proposed a molecular subtyping system through multi-omics analysis based on 343 UCEC samples.[2] The subtyping system can divide UCEC into four subtypes, namely, POLE-mutated group, microsatellite instability (MSI) group, low copy number group, and high copy number group.[2] For many cancer types, immunotherapy, especially programmed cell death protein 1 (PD-1)/programmed death-ligand 1 (PD-L1) inhibitors, is seen as a promising strategy in treating metastatic cancer treatment.[3,4] Pembrolizumab, also known as KEYNOTE-28, is a PD-1 inhibitor and has been approved by the U.S. Food and Drug Administration (FDA) as an effective agent in the treatment of certain UCEC patients, and in literature.[5–7] By combining pembrolizumab with other drugs such as lenvatinib, advanced patients manifest further promising antitumor activity, and this combination therapy has also been approved by FDA.[7,8] Notably, POLE-mutated and MSI groups have been found to be enriched to CD3+ and CD8+ tumor-infiltrating lymphocytes (TILs), with higher PD-1 and PD-L1 expression than other two groups, and they were positively responsive to anti-PD-1/PD-L1 therapy.[9] Although the TCGA subtyping can guide the immune blockade therapy to some extent, a number of metastatic patients still present negative response. Therefore, the realization of a personalized immunotherapy will require a more specific subtyping system based on tumor immune microenvironment.

A large amount of studies have demonstrated that long noncoding RNAs (lncRNAs) play an essential role in tumor development and regulation.[10–12] LncRNAs are key regulators in maintaining chromatin structure and modulating transcriptional activity,[13,14] but they are commonly abnormally regulated with gain or loss of copy numbers in tumors.[15,16] A series of lncRNAs, such as H19,[17] BANCR,[18] MEG3,[19] GAS5,[20] HOTAIR,[21,22] and FER1L4,[23] have been found to participate in tumor regulation of UCEC. Aberrant expression of lncRNAs appears to be characteristic and differential among various cancers; therefore, lncRNAs have the potential to serve as biomarkers and targets for personalized treatment. Previously, for example, Dong *et al.* developed a 7-lncRNA signature for predicting UCEC prognosis;[24] Liu *et al.* identified 13 immune-related lncRNAs closely related to UCEC overall survival.[25] For now, no molecular subtyping system based on immune-related lncRNAs has been reported.

Focusing on immune-related lncRNAs, the current research aimed to explore a novel immune subtyping system, so as to provide a direction for targeted immunotherapies. The prognostic lncRNAs developed in this study can predict OS of UCEC patients who are treated with a RNA-based therapy.

Materials and methods

Data source and preprocessing

TCGA-UCEC dataset was downloaded from TCGA database (version 27.0, https://portal.gdc.cancer.gov/) in 17 January 2021. Samples without follow-up data or survival data were excluded. Ensembl ID was converted to gene symbol, and median expression was taken when there were multiple gene symbols in one gene. After preprocessing, 539 samples with clinical data were included (Supplementary Table S1). Immune-related genes were obtained from ImmPort database (http://www.immport.org). The workflow of identifying immune subtypes and prognostic lncRNAs is shown in Figure 1.

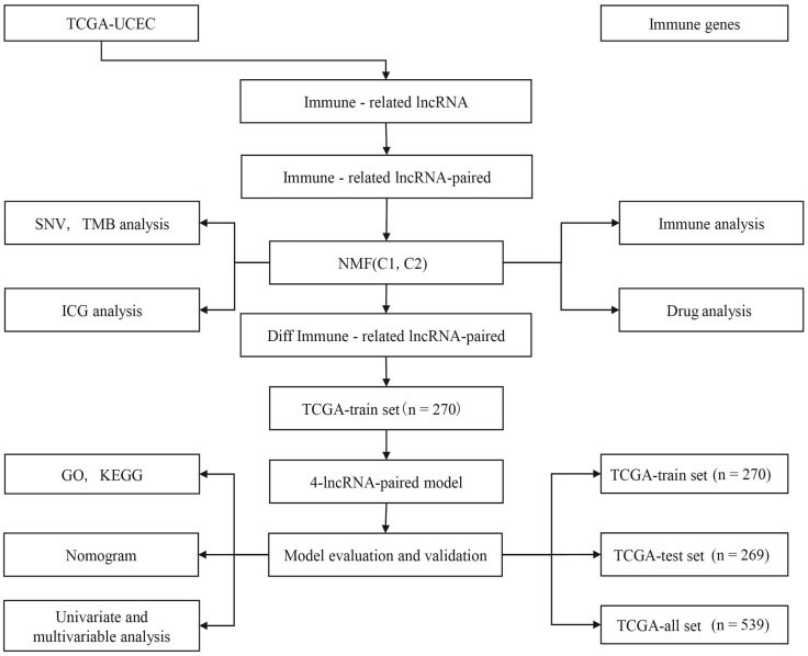

Figure 1. A workflow of constructing a prognostic model of UCEC based on TCGA-UCEC dataset and immune-related genes.

The source of immune signatures

A total of 26 classical immune biomarkers were obtained from a pan-cancer research based on immune landscape.[26] Forty-seven immune checkpoints were obtained from the work of Danilova *et al.*[27] We used Th1/IFN-γ gene signatures to calculate IFN-γ score for characterizing important cytokine.[27] Average expression of GZMA and PRF1 was used to evaluate the cytolytic activity score for studying the T-cell immune response.[28] Angiogenesis signature with a list of genes was obtained from Masiero *et al.* to evaluate the angiogenesis in tumors.[29]

Identification of immune-related lncRNAs

RNA-seq data was divided into mRNA data and lncRNA data, according to gene transfer format (GTF) file (version 05.05.21) obtained from GENCODE (https://www.gencodegenes.org/). Pearson correlation coefficient between immune-related genes and lncRNA was calculated. Under the condition of coefficient >0.8 and $p < 0.01$, 86 lncRNAs and 59 immune-related genes were screened.

Loop pairing of immune-related lncRNAs

Immune-related lncRNAs were paired as paired lncRNA-A and lncRNA-B. C was defined as the combination of lncRNA-A and lncRNA-B. C was defined as 1 if an lncRNA-A expression was higher than lncRNA-B, otherwise 0. Then a matrix was constructed based on 0 and 1. Paired lncRNAs were included for further analysis when 1 consists of 20%–80% in all samples.

Nonnegative matrix factorization (NMF) for identifying immune subtypes

Firstly, coxph function was conducted for univariate Cox regression analysis to identify 328 paired-lncRNAs associated with UCEC prognosis in TCGA-UCEC dataset ($p < 0.01$). Then, NMF R package was applied to cluster samples based on the expression of 328 paired-lncRNAs.[30] Similar to principle component analysis (PCA), NMF is an unsupervised learning technique that could extract useful information from high dimension data.[30] "Brunet" algorithm was selected and 100-time iteration was implanted. Cluster numbers (k) were set

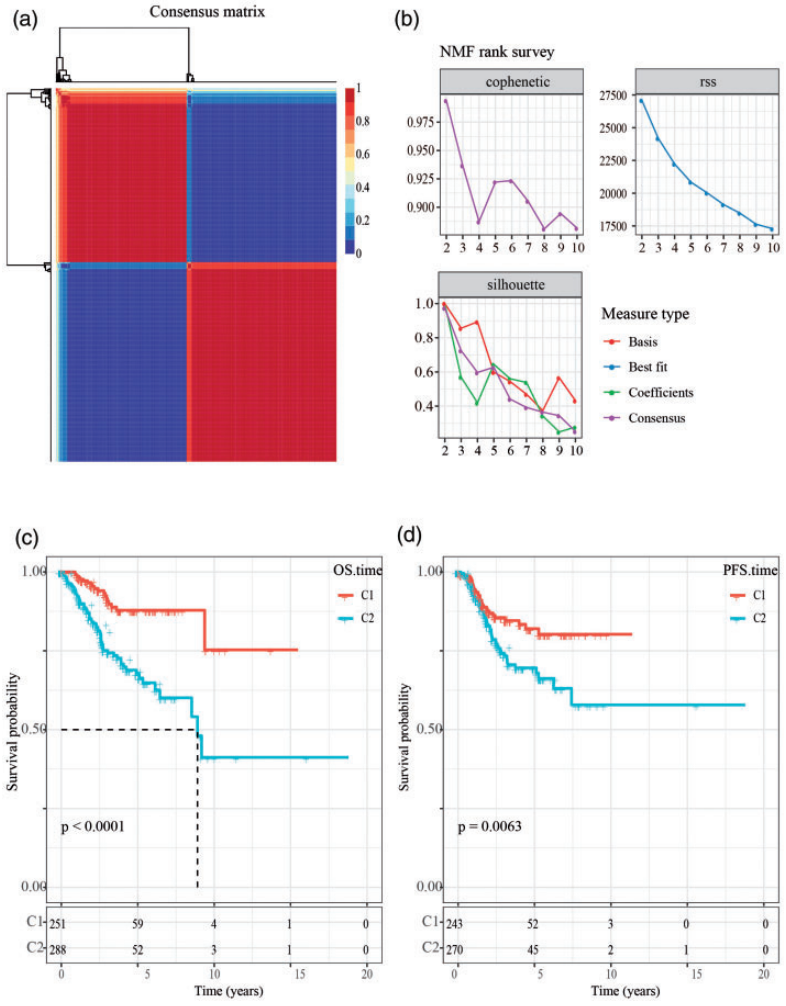

Figure 2. Identification of immune subtypes. (a) Clustering of 539 samples through NMF analysis. (b) NMF indicators of cophenetic, RSS, and silhouette with k = 2–10. (c–d) Kaplan-Meier survival curve of OS and PFS in C1 and C2 immune subtypes. Log-rank test was performed. (A color version of this figure is available in the online journal.)

from 2 to 10, with at least 10 samples in each cluster. Three indicators of cophenetic, residual sum of squares (RSS) and dispersion were used to assess the optimal cluster.

Gene mutation analysis

Tumor mutation burden (TMB), single nucleotide variations (SNVs), and mutation patterns of immune subtypes were analyzed by mutect2 software (https://software.broadinstitute.org/cancer/cga/mutect).[31] Mutect2, which can detect low allele-fractions based on a Bayesian classifier and has high sensitivity and specificity, is particularly suitable for studying sequencing data in cancers.[31] We used mutect2 to screen highly mutated genes with a mutation frequency up to 3%. *Chi*-square test was performed, and $p < 0.01$ were selected to obtain significantly mutated genes.

Single sample gene set enrichment analysis

Single sample gene set enrichment analysis (ssGSEA) in GSVA R package was conducted to calculate enrichment score of immune-related genes.[32] ssGSEA is a method widely used in analyzing the enrichment of a gene set or signature, allowing one signature to define a normalized enrichment score per sample.

CIBERSORT analysis

Enrichment score of 22 immune-related cells was determined by CIBERSORT (version 2018, http://cibersort.stanford.edu/).[33] The CIBERSORT tool could determine the proportion of immune cell types in complex tissues based on gene expression profiles, and has been applied in many studies for characterizing the immune features of tumor microenvironment.

Functional analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and gene ontology (GO) function were analyzed using WebGestaltR package.[34] GO function includes three sections, biological process (BP), molecular function (MF), and cell component (CC). WebGestalt annotates pathways and GO function, supporting complementary methods for enrichment analysis. $p < 0.05$ was considered as a

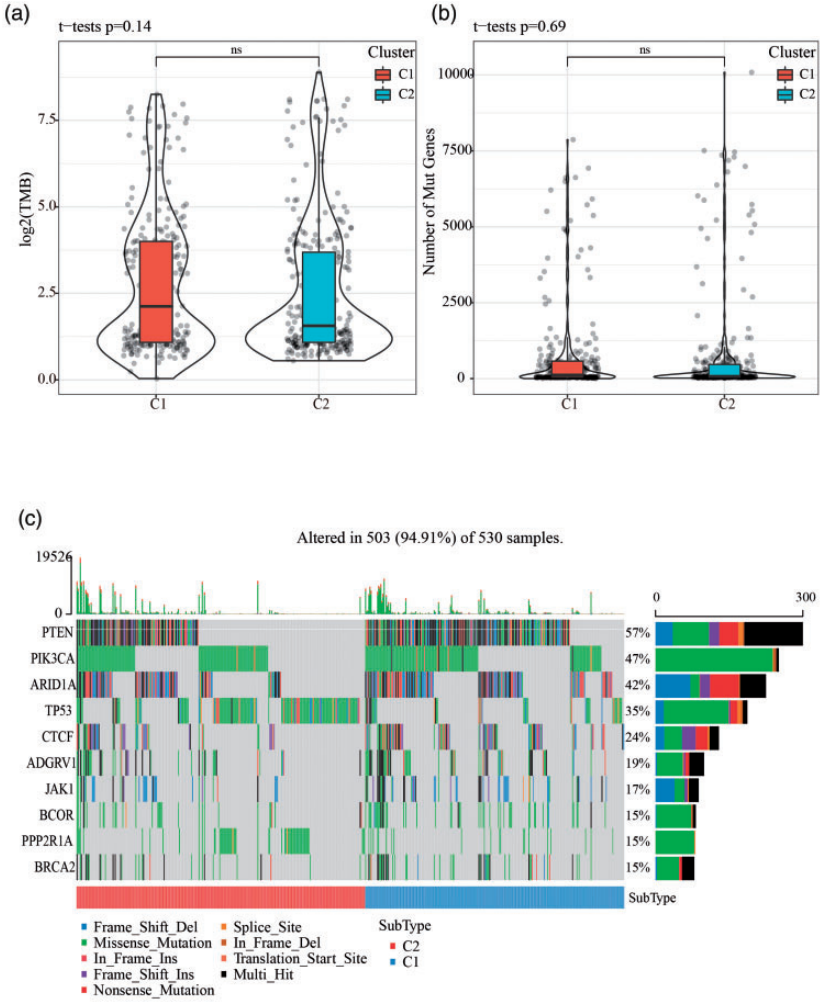

Figure 3. Mutations of TCGA-UCEC samples. (a) TMB of C1 and C2 groups. (b) Numbers of mutated genes in C1 and C2 groups. (c) Mutation patterns of the top 10 mutated genes. Student's *t* test was performed. (A color version of this figure is available in the online journal.)
ns: no significance.

significant threshold to screen annotated pathways and GO terms.

Predicting the response to immunotherapy and chemotherapy

TCGA-UCEC and IMvigor210 datasets were included to assess the response to immunotherapy and chemotherapy. Submap analysis was performed to compare the similarity of expression profiles between IMvigor210 and TCGA-UCEC datasets. If two samples from two different datasets presented similar expression features, we considered these two samples having similar response to immunotherapy. For predicting the sensitivity of chemotherapy, pRRophetic R package was applied to evaluate estimated IC50 (half maximal inhibitory concentration) of four chemotherapy drugs (cisplatin, paclitaxel, doxorubicin, and erlotinib) in each patient. Student's *t* test was performed to test the difference between two groups.

Describing genomic features

To describe the genomic features related to homologous recombination deficiency (HRD), DNA sequencing data of TCGA-UCEC dataset was used as an input to the algorithms through referring to Marquard *et al.*[35] to output the score of HRD, loss of heterozygosity (LOH), large-scale state transitions (LST), the number of sub-chromosomal regions undergoing allelic imbalance extending to the telomeres (NtAI). TCGA CNV data was used as an input to analyze somatic copy number variants (SCNVs). Student's *t* test was performed for comparing the difference of these genomic features between two immune subtypes.

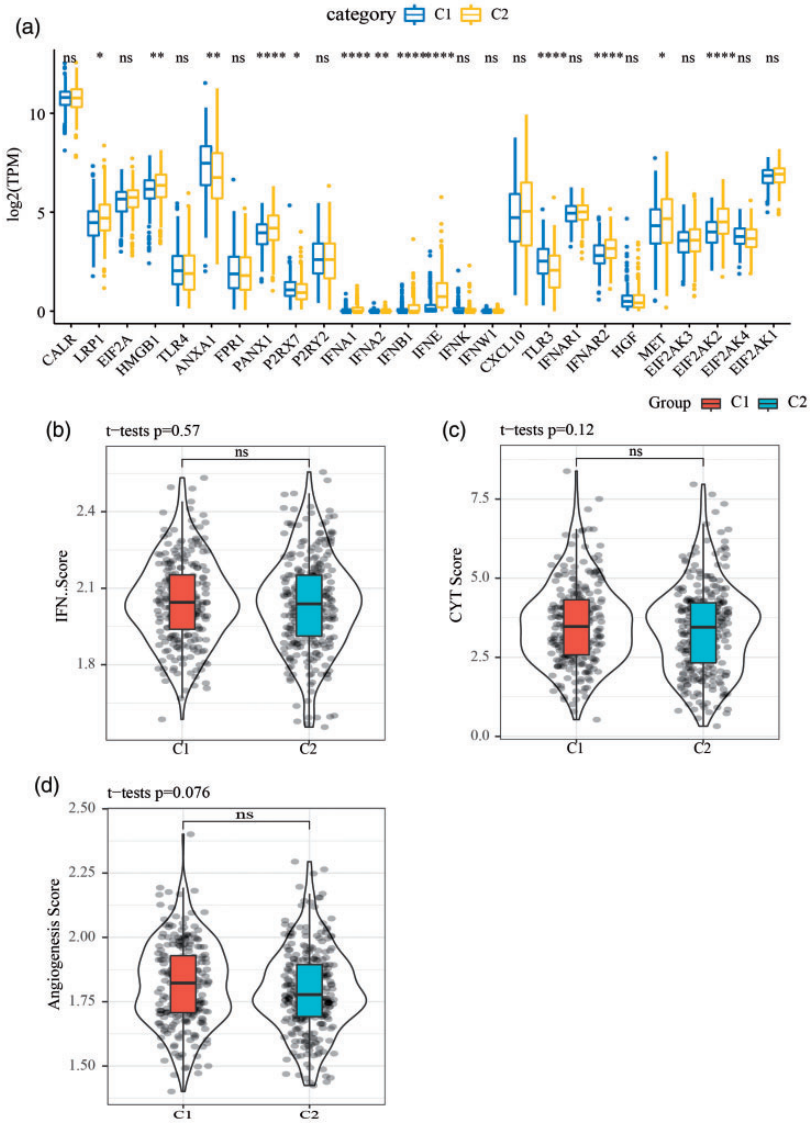

Figure 4. Immune response and tumor angiogenesis in two subtypes. (A) Log2(TPM) expression of immune biomarkers in two subtypes. (b–d) Score of IFN-γ, CYT and tumor angiogenesis in two subtypes. Student's t test was performed. ANOVA was performed. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$. (A color version of this figure is available in the online journal.)
ns: no significance; TPM: transcript per million.

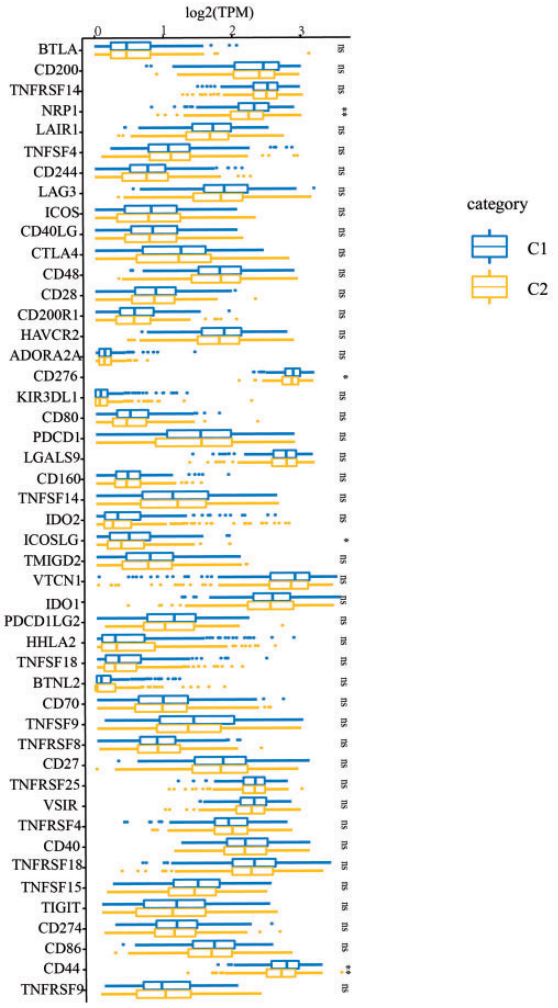

Figure 5. Log2(TPM) expression of immune checkpoints in two subtypes. ANOVA was performed. *$p < 0.05$, **$p < 0.01$. (A color version of this figure is available in the online journal.)
ns: no significance; TPM: transcript per million.

Construction of UCEC prognostic model

Differentially expressed lncRNAs between C1 and C2 groups were screened by Fisher's exact test and Benjamini & Hochberg (BH) correction. A total of 1714 lncRNA pairs were identified with false discovery rate (FDR) < 0.01. TCGA-UCEC dataset was randomly divided into training group and test group at a ratio of 1:1 (Supplementary Table S2). No significant difference between two groups was observed after performing *Chi*-square test ($p > 0.05$). Univariate and multivariate Cox regression analysis were performed through "survival coxph function" in R package.[36] Least absolute shrinkage and selection operator (LASSO) regression analysis in glmnet R package and Akaike information criterion (stepAIC) in MASS package were performed to optimize the prognostic model.[36,37] Risk score was converted to z-score; samples were assigned into high-risk and low-risk groups with set z-score = 0 as a cut-off. Kaplan-Meier (KM) survival curve was plotted to show OS and progression-free survival (PFS). Log-rank test was performed in KM plot.

Nomogram and decision curve analysis

A nomogram was constructed using rms R package (http://CRAN.R-project.org/package=rms) based on three risk signatures (HR > 1), resulting from multiple Cox regression analysis. Each signature can obtain a point and total points correspond to death rates in one-year, three-year, and five-year period can be calculated. Nomogram is frequently used for prognostic prediction and is also convenient for clinical use. The efficacy of nomogram was validated by decision curve analysis (DCA) and receiver operating characteristic (ROC) curve. DCA and ROC curve are common methods to evaluate prediction models especially for prognostic models.[38]

Statistical analysis

All statistical analyses were performed in R (v3.4.2) software. Parameters in algorithms were defined as default if there was no specific introduction. Statistical methods were presented in the corresponding sections. Bonferroni correction was used to correct p value. $p < 0.05$ was considered as significant. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

Results

Identifying immune subtypes of UCEC based on immune-related lncRNAs

Immune-related lncRNAs were paired by loop pairing. A total of 328 paired-lncRNAs associated with UCEC prognosis were screened through conducting univariate Cox regression analysis. After clustering UCEC samples through NMF algorithm, the optimal clustering (C1 and C2 groups) was generated according to cophenetic, RSS, and silhouette (Figure 2(a) and (b)). Survival analysis revealed that OS and PFS were significantly differential between C1 and C2 groups, with a more favorable prognosis in C1 group (OS, $p < 0.0001$; PFS, $p = 0.0063$, Figure 2(c) and (d)).

Mutation patterns of two immune subtypes

The TMB of all UCEC samples was calculated by mutect2 software. The results showed no difference of TMB and mutated genes between two molecular subtypes (Figure 3 (a) and (b)). Furthermore, 17,783 genes with mutation frequency up to 3% were screened. *Chi*-square test to identified 78 genes with high mutation frequency ($p < 0.01$), and the top 10 mutated genes (*PTEN*, *PIK3CA*, *ARID1A*, *TP53*, *CTCF*, *ADGRV1*, *JAK1*, *BCOR*, *PPP2R1A*, and *BRCA2*) were listed in Figure 3(c). Notably, compared with other genes, *PTEN* and *ARID1A* genes showed a high proportion of frame shift deletions and nonsense mutations, and *PIK3CA* had a high ratio of missense mutation.

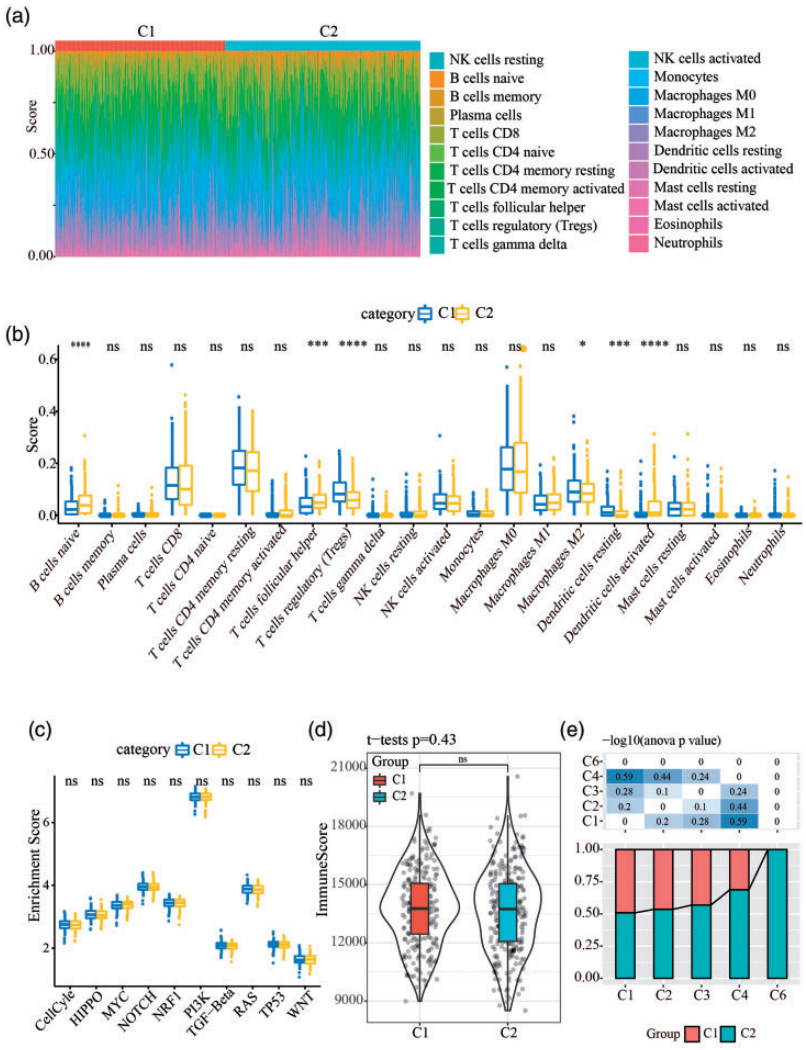

Figure 6. Immune features of C1 and C2 immune subtypes. (a–b) Enrichment score of 22 immune cells in C1 and C2 groups. (c) Enrichment score of 10 tumor-related pathways. (d) Immune infiltration score of C1 and C2 groups. (e) The relation between two immune subtypes and pan-cancer subtypes in previous study. ANOVA was performed. *$p < 0.05$, ***$p < 0.001$, ****$p < 0.0001$. (A color version of this figure is available in the online journal.)
ns: no significance.

Expression of immune biomarkers and immune checkpoints

We then evaluated whether there was a difference of immune-related gene expression between the two subtypes. After assessing 26 immune biomarkers, 14 genes were found to be differentially expressed between C1 and C2 groups ($p < 0.05$, Figure 4(a)). IFN-γ is a cytokine produced by CD8+ T cells and plays an important role in tumor microenvironment (TME). Previous study demonstrated that Th1/IFNG gene signature was correlated with PD-L1 expression.[27] ssGSEA calculation of the IFN-γ score showed no expression difference (Figure 4(b)). Cytolytic activity (CYT), defined by the expression of *GZMA* and *PRF1*, is suggested to be associated with immune response and prognosis.[28] However, no expression difference of CYT was showed between C1 and C2 groups (Figure 4(c)). As tumor angiogenesis is a potential target of anti-VEGF therapy, we selected an angiogenestic gene signature from previous research.[29] The signature was scored in C1 and C2 groups, but no expression difference between two groups was detected (Figure 4(d)). Moreover, 47 immune checkpoints were collected from the previous study,[27] and four genes (*NRP1*, *CD276*, *ICOSLG*, and *CD44*) presented differential expression between C1 and C2 groups ($p < 0.05$, Figure 5).

Immune features of two immune subtypes

Next, we scored 22 types of immune cells through CIBERSORT tool, and six of them (naive B cells, follicular helper T cells, regulatory T cells, M2 macrophages, resting dendritic cells, and activated dendritic cells) showed differential enrichment between C1 and C2 groups (Figure 6(a) and (b)). In addition, 10 oncogenic signaling pathways closely related to tumor progression were analyzed in this study;[39] there was no difference between C1 and C2 groups (Figure 6 (c)). According to the immune infiltration analysis, these two subtypes showed similar tumor microenvironment (Figure 6 (d)). A large-scale study on tumors suggested that tumors

can be classified into six subtypes based on TCGA dataset and immune features.[26] In the study, UCEC was classified into C1, C2, C3, C4, and C6 subtypes. We matched C1 and C2 groups to the five previously reported subtypes. C6 subtype was almost consisted of C2 group (Figure 6(e)), indicating the same immune feature was almost the same between C2 group and C6 subtype.

Differential response of immunotherapy between two subtypes

To explore the efficacy of immunotherapy and chemotherapy in C1 and C2 groups, submap analysis on TCGA-UCEC dataset and IMvigor210 dataset was performed. The result demonstrated that C2 group was more sensitive to immunotherapy than C1 group, with a higher ratio of CR (complete response)/PR (partial response) ($p < 0.05$, Figure 7(a)). Additionally, among four chemotherapeutic drugs of cisplatin, paclitaxel, doxorubicin, and erlotinib, C1 group showed a higher level of estimated IC50 than C2 group, indicating that these four drugs were more effective in treating UCEC patients in C2 group (Figure 7(b)).

The relation between immune escape and immune subtypes

The relation between two subtypes and immune escape was investigated here. The term "cancer immunoediting"

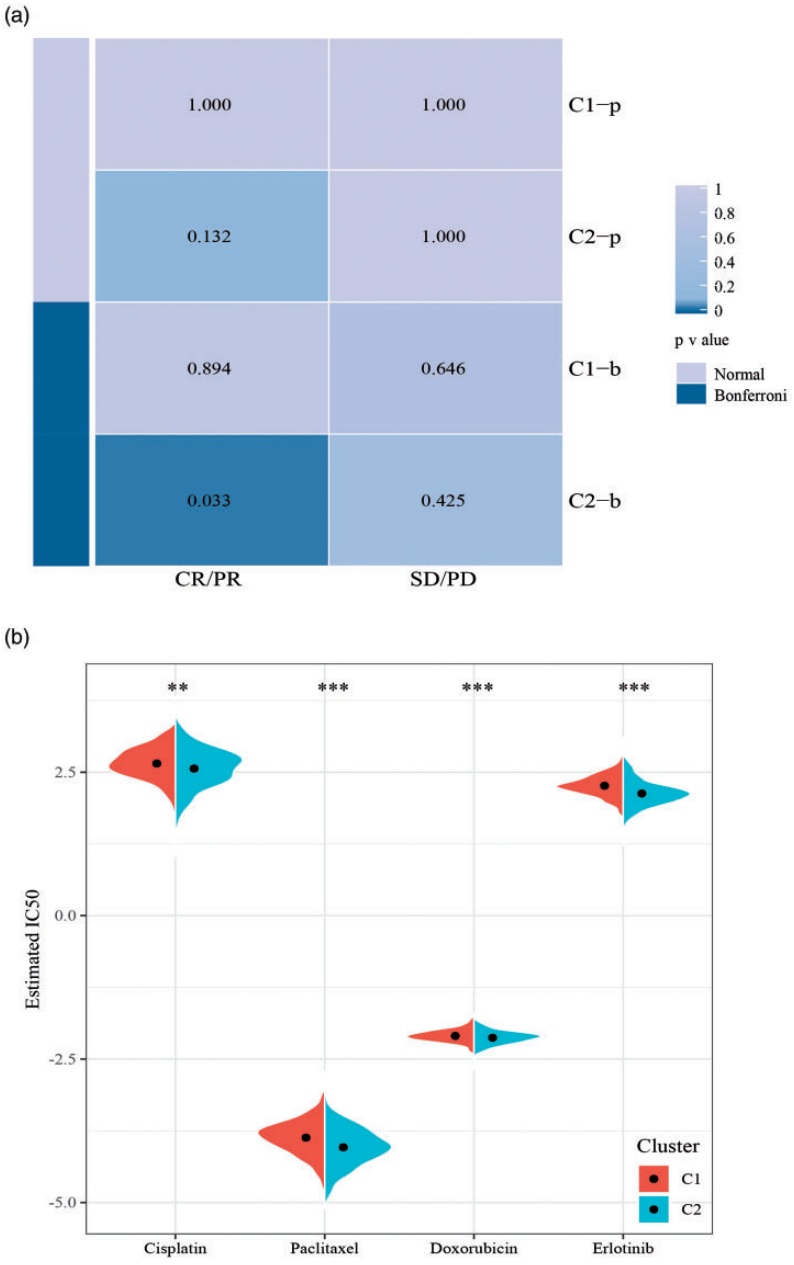

Figure 7. Sensitivity of immunotherapy and chemotherapy in C1 and C2 groups. (a) Submap analysis between immune subtypes and IMvigor210 dataset. *p* Value represents the similarity, with a lower *p* value indicative of a higher sensitivity of immunotherapy. (b) Estimated IC50 of chemotherapy drugs including cisplatin, paclitaxel, doxorubicin, and erlotinib. Student's *t* test was performed. $^{**}p < 0.01$, $^{***}p < 0.001$. (A color version of this figure is available in the online journal.)
CR: complete response; PR, partial response; SD: stable disease; PD: progressive disease.

defined by Schreiber *et al.* refers to the biological phenomenon that tumors can escape from elimination through recruiting leukocytes with anti-tumor properties.[40] Loss of antigenicity and immunogenicity are major mechanisms underlying immune escape. Therefore, we focused on some factors, including mutation load, HRD, neoantigen load, and chromosomal instability level, which may potentially affect immunogenicity. As shown in Figure 8, C2 group had higher scores in HRD, SCNVs, NtAI, LST, and LOH, which suggested genetic instability in C2 group.

Constructing the prognostic model based on immune-related lncRNA

To construct a prognostic model of UCEC, we firstly identified 1714 paired lncRNAs differentially expressed between C1 and C2 groups through performing Fisher's exact test (FDR < 0.01). TCGA-UCEC dataset was randomly divided into training group (270 samples) and test group (269 samples). In the training group, 427 paired lncRNAs were screened by univariate Cox regression analysis ($p < 0.05$). Then, we conducted LASSO regression analysis and five-fold cross-validation to construct a prognostic model with the paired lncRNAs. When lambda $= 0.0555$, the model was the optimal (Supplementary Figure S1), and seven paired lncRNAs were identified for the next optimization. Moreover, by performing stepAIC algorithm to determine a fitting degree with the less variants, we obtained a final prognostic model with four paired lncRNAs defined as

$$\begin{aligned}\text{Risk score} = {} & (0.822 \times \text{AC099329.2 vs. AC113349.1}) \\ & + (1.193 \times \text{LINC01513 vs. AL645924.1}) \\ & + (0.631 \times \text{AC068987.3 vs. LINC02345}) \\ & + (1.187 \times \text{MIR31HG vs. ZNF295} - \text{AS1})\end{aligned}$$

Survival analysis revealed that these four paired lncRNAs could accurately stratify the training group

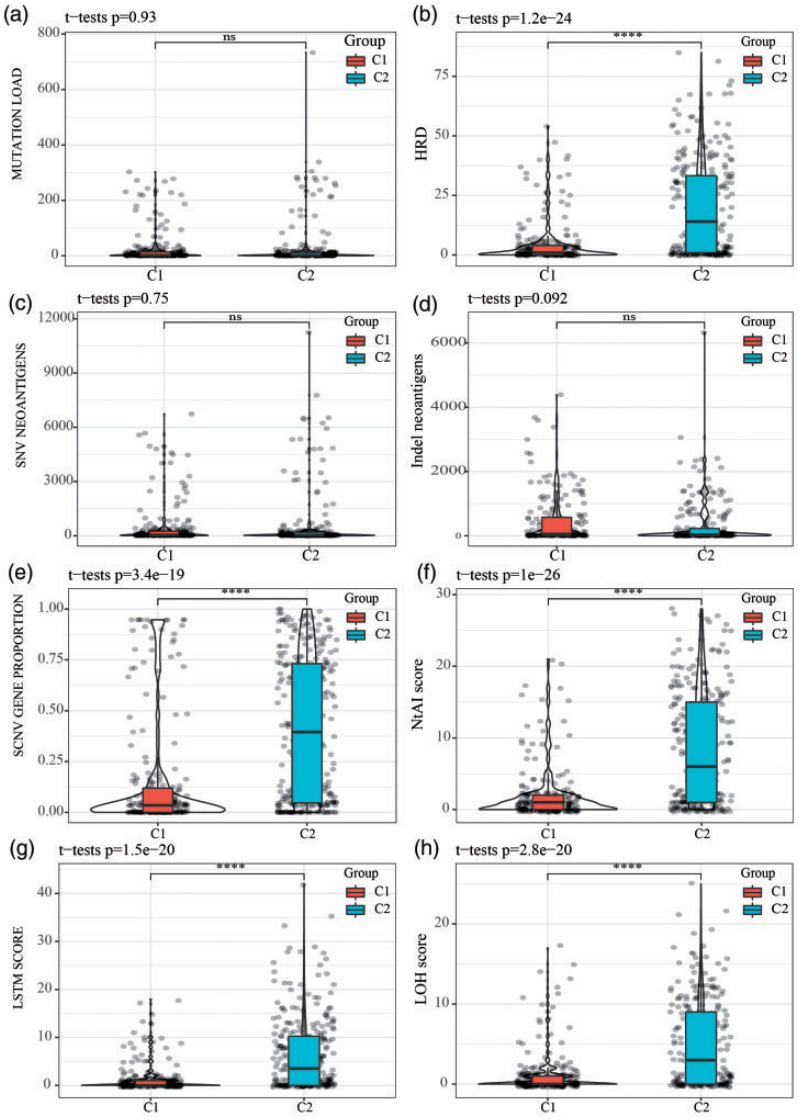

Figure 8. (a–h) Comparison of mutation loads (a), HRD (b), SNV neoantigen (c), indel neoantigens (d), SCNV gene proportion (e), NtAI score (f), LOH score (g) and LST score (h) between C1 and C2 immune subtypes. (A color version of this figure is available in the online journal.)

into two groups with differential OS ($p < 0.05$, Supplementary Figure S2(a) to (d)). Similarly, the expression of eight lncRNAs in four paired-lncRNAs was all significantly associated with prognosis ($p < 0.05$, Supplementary Figure s2(e) to (l)). We calculated the risk score of each sample and converted risk score to z-score. High-risk group and low-risk group were defined by the cutting value of z-score = 0. According to Kaplan-Meier survival curve, 139 samples were stratified into high-risk group and 131 samples were stratified into low-risk group ($p < 0.0001$, Figure 9(a)). ROC curves of one-year, three-year, five-year, and seven-year OS were shown. The prediction performance of this prognostic model was reflected by the AUC, which was 0.73, 0.71, 0.62, and 0.65 for one-year, three-year, five-year, and seven-year OS, respectively (Figure 9(b)).

The test group containing 269 samples was further used to validate the risk model. Similarly, the samples were accurately stratified into high-risk and low-risk groups, and the AUC of one-year, three-year, five-year, and seven-year OS all validated the robustness of the model (Supplementary Figure S3). A total of 539 samples of TCGA-UCEC dataset were included into analysis and classified into two groups (Supplementary Figure S4(a)). AUC of one-year, three-year, five-year, and seven-year OS was 0.80, 0.76, 0.70, and 0.71, respectively, pointing to the reliable prediction performance of the prognostic model (Supplementary Figure S4 (b)). In addition, the distribution of risk score and survival status revealed a higher density of dead status in high-risk group (Figure 10).

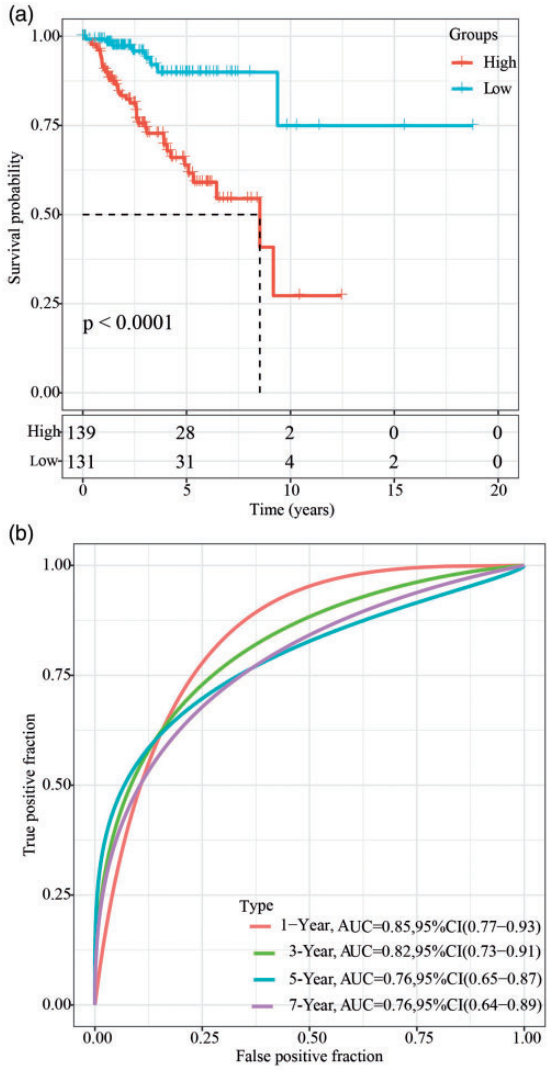

Figure 9. The performance of prognostic model in the training group. Kaplan-Meier survival curves (a) and ROC curves (b) of training group. (A color version of this figure is available in the online journal.)

Functional analysis of four paired-lncRNAs in the prognostic model

Next, the functional analysis on four paired-lncRNAs was performed through WebGestalt R package. Pearson coefficient and significance between four paired-lncRNAs and mRNAs were analyzed. A total of 1770 genes were identified under a condition of coefficient >0.25 and $p < 0.05$, and then subjected to GO and KEGG functional analysis. Numbers of enriched GO items on biological process, cellular component and molecular function were 129, 117, and 6, respectively (FDR < 0.05), and part of them were shown (Figure 11(a) to (c)). Nine KEGG pathways, including synaptic vesicle cycle, basal cell carcinoma, small cell lung cancer, cell cycle, hippo signaling pathway, microRNAs in cancers, axon guidance, human papillomavirus infection, and human papillomavirus infection, were annotated (FDR < 0.05, Figure 11(d)). Notably, most of the pathways were closely related to tumor regulation.

The relation between risk score and clinical features

The risk score of different stage, grade, immune subtype, and age all exhibited significant difference (Figure 12). Especially, C2 group showed a higher risk score than C1 group, which was consistent with the previous result of shorter OS in C2 group (Figure 12(c)). Univariate and multivariate Cox regression analysis was performed to assess the relation between risk type, clinical features, and OS using TCGA-UCEC dataset. It was found that age, grade, stage, and risk type were all positively correlated with OS (Figure 13). Particularly, the risk type was significantly associated with OS both as calculated by univariate Cox

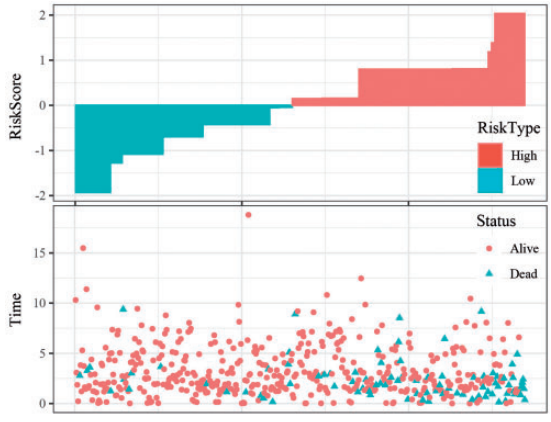

Figure 10. The distribution of survival status in high-risk and low-risk groups. (A color version of this figure is available in the online journal.)

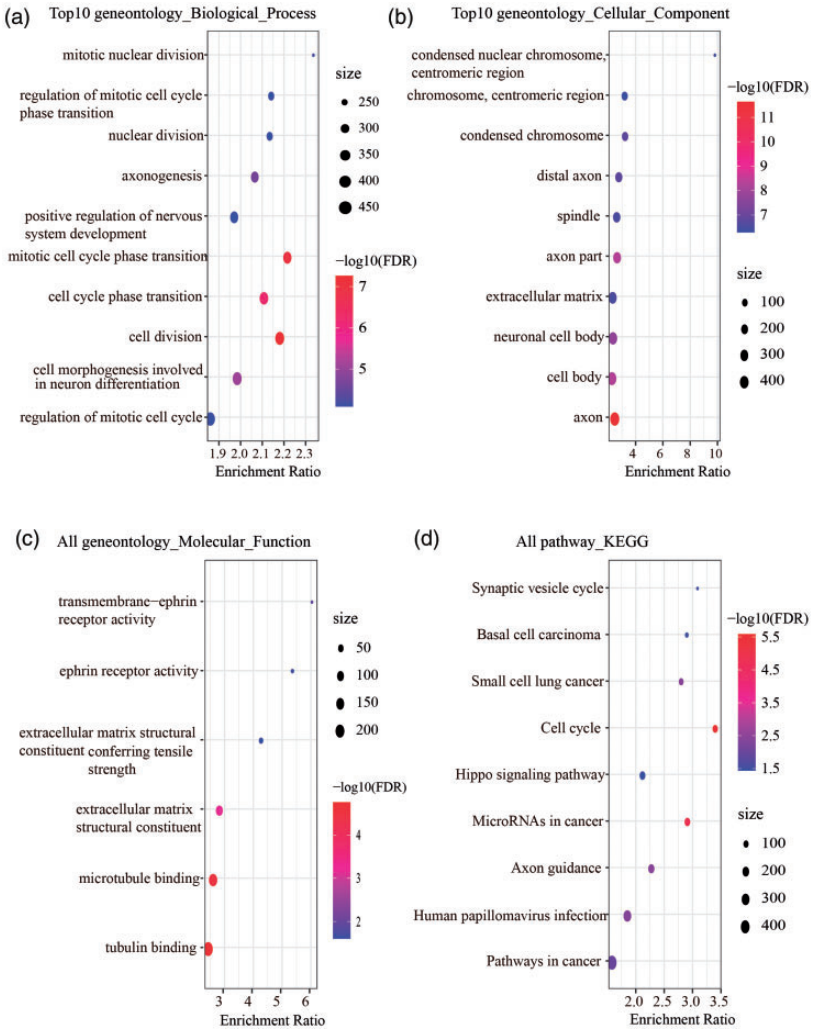

Figure 11. GO and KEGG function analysis of 1770 genes related to four paired lncRNAs. (a) The top 10 enriched biological processes. (b) The top 10 enriched cellular components. (c) Six enriched molecular function items. (d) Nine enriched KEGG pathways. Size represents the number of enriched genes. (A color version of this figure is available in the online journal.)

regression analysis (HR = 3.52, 95%CI = 2.11–5.85) and multivariate Cox regression analysis (HR = 2.67, 95% CI = 1.60–4.48).

Constructing a nomogram for predicting prognosis

Nomogram could directly exhibit and predict the prognosis in clinical practice. Risk score, grade, and stage of all samples were recruited to construct a nomogram for predicting one-year, three-year, and five-year prognosis of the patients (Figure 14(a)). The predicted OS of nomogram was compared with the actual OS (Figure 14(b)). Decision curve analysis demonstrated that nomogram was more effective than stage, grade, and risk score in predicting OS (Figure 14 (c)). Furthermore, ROC curve also showed nomogram was the optimal in the prognosis prediction, with an AUC of 0.85 (95%CI = 0.71–0.80, Supplementary Figure S5).

Discussion

LncRNAs have been increasingly found to play critical roles as transcriptional and post-transcriptional regulators in cancer progression. The dysfunction or abnormal expression of lncRNAs is strongly related to tumor development in cell proliferation, migration, and invasion.[41] Many kinds of lncRNAs act as key regulators in UCEC tumorigenesis or are associated with OS.[17–20,24,25] It has also been demonstrated that lncRNAs are involved in regulating immune responses, including in T cell development, differentiation, and activation.[42] Studied identified immune-related lncRNAs such as lncRNA-EPS as suppressors of inflammatory response[43] and lncRNA-MAF-4 as a chromatin-associated lncRNA participating in T lymphocyte differentiation.[44] The above evidence supports a concrete relation between lncRNAs and tumor immune microenvironment.

Based on the previous observations, we focused on immune-related lncRNAs to comprehensively study the correlation between immune-related lncRNAs and UCEC prognosis. In this study, we introduced a paired lncRNA strategy, which can avoid batch effect to construct the prognostic model. Different sequencing platforms can result in differences in sequencing and detecting the variations, and

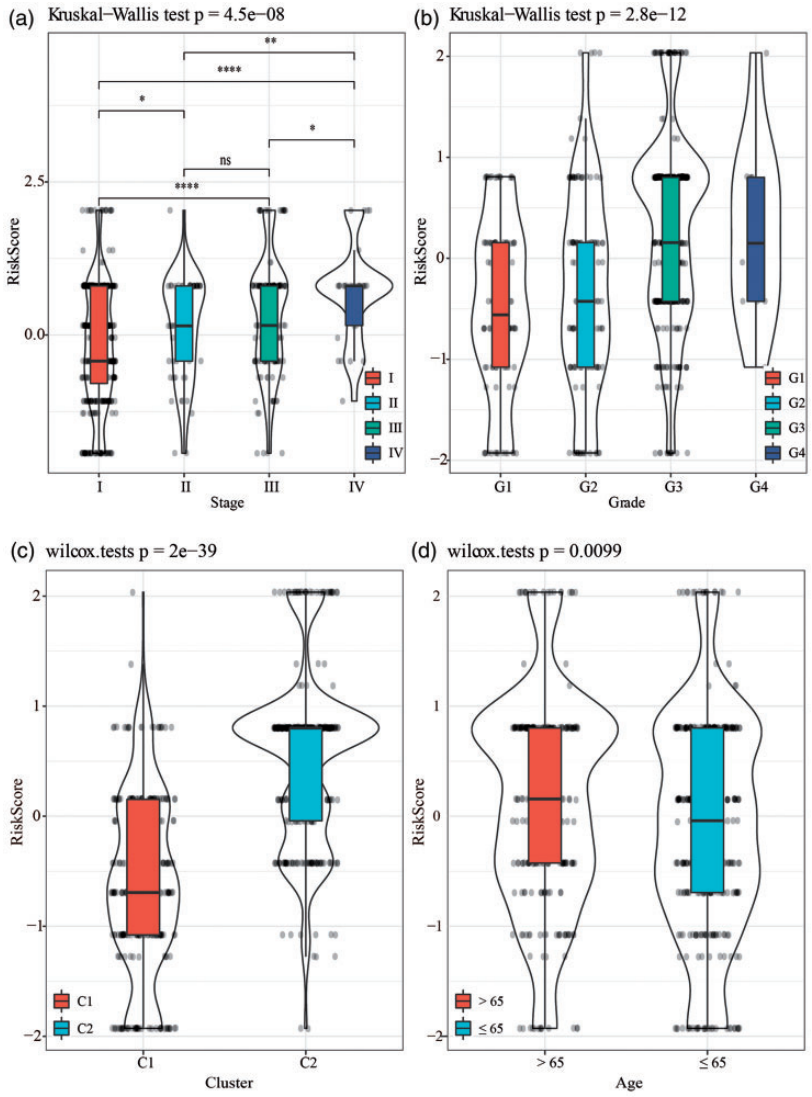

Figure 12. The relation between risk score and clinical features including stage (a), grade (b), immune subtypes (c), and age (d). Kruskal-Wallis test was performed in (a and b). Wilcoxon test was performed in (c and d). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$. (A color version of this figure is available in the online journal.)
ns: no significance.

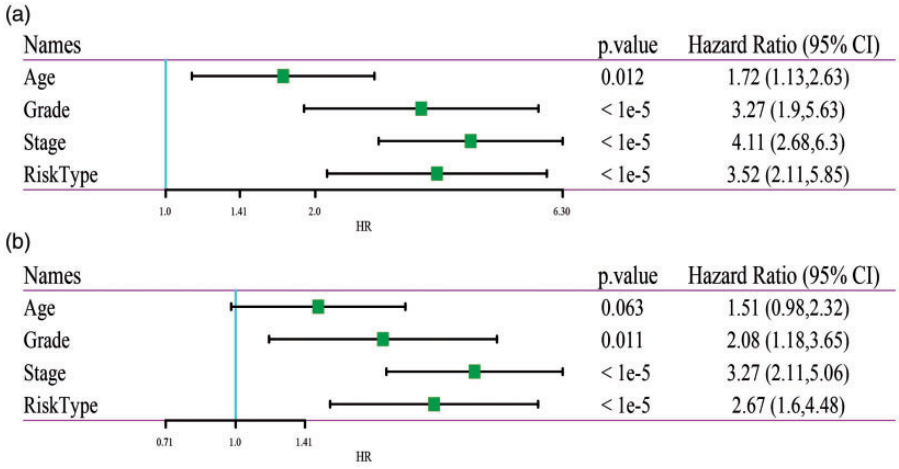

Figure 13. Univariate Cox regression analysis (a) and multivariate Cox regression analysis (b) of age, grade, stage, and risk type. (A color version of this figure is available in the online journal.)
HR: hazard ratio.

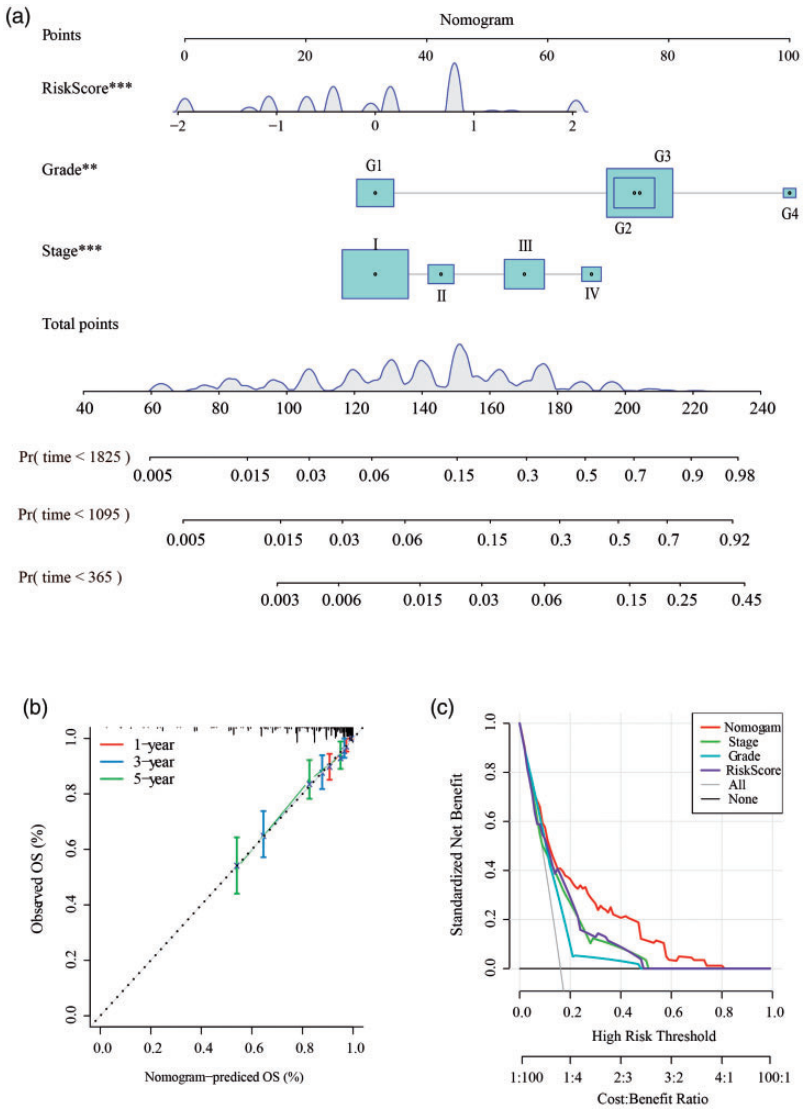

Figure 14. Construction of a nomogram for predicting prognosis. (a) A nomogram defined by risk score, grade, and stage. One-year, three-year, and five-year prognosis (possibility of mortality) were displayed. (b) The relation between observed OS and predicted OS. (c) DCA curve of nomogram, stage, grade, and risk score. (A color version of this figure is available in the online journal.)

different batches also have batch effect. Correction of normalized methods are not always effective and different normalized methodology can lead to different results; therefore, the relative ranking of paired-lncRNAs was used as a signal matrix in this study to effectively address the issue. This relative ranking method has a cross-platform advantage in decreasing the discord of data from different platforms. Based on the paired lncRNA method, we developed an immune subtyping system showing a close relation with the OS of UCEC. The UCEC samples were classified into two immune subtypes (C1 and C2), with C1 group demonstrating a more favorable prognosis than C2 group. The expression of immune biomarkers and the enrichment score of immune cells significantly varied between the two groups, pointing to the regulation of lncRNAs on immune response.

Importantly, the immune subtyping system can help to effectively select subjects with high sensitivity or positive response to chemotherapy or immunotherapy. The features of tumor microenvironment can influence the selective stress of tumor cells on the process of chemotherapy and immunotherapy.[45,46] The current findings showed that the samples in the C2 group were more sensitive to immunotherapy and chemotherapy than C1 group. Patients with C2 immune subtype may positively respond to four drugs, namely cisplatin, paclitaxel, doxorubicin and erlotinib, in chemotherapy or its combination with other therapies.

Immune escape is an event decisive of whether patients can benefit from targeted therapies. Immune system could hinder tumor progression or accelerate tumor development.[40] Tumor cells evolve to avoid elimination through anti-tumor immune response through loss of antigenicity and/or immunogenicity.[47] We observed that C2 group showed a higher proportion of homologous recombination deficiency and mutated genes, indicating a stronger possibility of loss of immunogenicity than C1 group. The loss of immunogenicity can promote tumor progression and contribute to the development of a poor prognosis, which was

consistent with the result of poor prognosis in C2 group. Previous evidence proved that lncRNAs are closely related to immune scape in tumors.[48] Recently, a study has demonstrated that LIMIT is a tumor immunogenic lncRNA correlated with immune infiltration and immune checkpoint blockade response.[49] This subtyping system based on immune-related lncRNAs also supports the close relation between lncRNAs and immune scape.

Furthermore, we screened four paired-lncRNAs based on which a four-gene prognostic signature was established to predict UCEC OS. The four-paired-lncRNA signature with a high AUC can effectively classify UCEC patients into high-risk and low-risk groups. Functional analysis revealed that tumor-related pathways, such as basal cell carcinoma, small cell lung cancer, cell cycle, hippo signaling pathway, and miRNAs in cancer, were highly enriched in these four paired-lncRNAs. Hippo signaling pathway has been found to serve as a central role in controlling cancer cell division through interacting with other cellular pathways.[50] Aberrant cell cycle activity is considered as classical characteristic of cancer development, and cell cycle proteins or regulators, especially cyclin-dependent kinases (CDKs) are mainly responsible for cancer development.[51] These enriched oncogenic pathways demonstrated that these prognostic lncRNAs were closely involved in UCEC development. In addition, a nomogram based on risk score and clinical features with a higher effectiveness was constructed for clinical use.

In conclusion, this study proposed a novel subtyping system based on immune-related lncRNAs. The subtyping system can divide UCEC patients into C1 and C2 subtypes with distinct OS, and can guide personalized therapies together with prognostic paired-lncRNAs. These lncRNAs may be potential targets for immunotherapies or RNA-based therapies.

AUTHORS' CONTRIBUTIONS

All authors participated in the design, interpretation of the studies and analysis of the data and review of the article. DYZ and ZL designed the research; KY performed the statistical analysis; NL analyzed and interpreted the data; DYZ drafted the article; ZL revised the article for important intellectual content.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

This work was supported by Key Research and Development Program of Guangxi (No. Guike AB18126056); Key Research and Development Program of Liuzhou (No. 2018BJ10301); Scientific Research and Technology Development Project of Liuzhou (No. 2018DB20501); and Guangxi Self-Finacing Research Program of Guangxi Region Health and Family Planning Commission (No. Z20190019, No. Z20200192, No. Z20210518).

ORCID iD

Dingyuan Zeng https://orcid.org/0000-0002-4271-5632

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA* 2021;**71**:209–49
2. Kandoth C, Schultz N, Cherniack AD, Akbani R, Liu Y, Shen H, Robertson AG, Pashtan I, Shen R, Benz CC, Yau C, Laird PW, Ding L, Zhang W, Mills GB, Kucherlapati R, Mardis ER, Levine DA. Integrated genomic characterization of endometrial carcinoma. *Nature* 2013;**497**:67–73
3. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Lu S, Kemberling H, Wilt C, Luber BS, Wong F, Azad NS, Rucki AA, Laheru D, Donehower R, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Greten TF, Duffy AG, Ciombor KK, Eyring AD, Lam BH, Joe A, Kang SP, Holdhoff M, Danilova L, Cope L, Meyer C, Zhou S, Goldberg RM, Armstrong DK, Bever KM, Fader AN, Taube J, Housseau F, Spetzler D, Xiao N, Pardoll DM, Papadopoulos N, Kinzler KW, Eshleman JR, Vogelstein B, Anders RA, Diaz LA Jr. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017;**357**:409–13
4. Green AK, Feinberg J, Makker V. A review of immune checkpoint blockade therapy in endometrial cancer. *Am Soc Clin Oncol Educ Book* 2020;**40**:1–7
5. Ott PA, Bang YJ, Berton-Rigaud D, Elez E, Pishvaian MJ, Rugo HS, Puzanov I, Mehnert JM, Aung KL, Lopez J, Carrigan M, Saraf S, Chen M, Soria JC. Safety and antitumor activity of pembrolizumab in advanced programmed death ligand 1-positive endometrial cancer: results from the KEYNOTE-028 study. *J Clin Oncol* 2017;**35**:2535–41
6. Marcus L, Lemery SJ, Keegan P, Pazdur R. FDA approval summary: pembrolizumab for the treatment of microsatellite instability-high solid tumors. *Clin Cancer Res* 2019;**25**:3753–8
7. Arora S, Balasubramaniam S, Zhang W, Zhang L, Sridhara R, Spillman D, Mathai JP, Scott B, Golding SJ, Coory M, Pazdur R, Beaver JA. FDA approval summary: pembrolizumab plus lenvatinib for endometrial carcinoma, a collaborative international review under project orbis. *Clin Cancer Res* 2020;**26**:5062–7
8. Makker V, Taylor MH, Aghajanian C, Oaknin A, Mier J, Cohn AL, Romeo M, Bratos R, Brose MS, DiSimone C, Messing M, Stepan DE, Dutcus CE, Wu J, Schmidt EV, Orlowski R, Sachdev P, Shumaker R, Casado Herraez A. Lenvatinib plus pembrolizumab in patients with advanced endometrial cancer. *J Clin Oncol* 2020;**38**:2981–92
9. Howitt BE, Shukla SA, Sholl LM, Ritterhouse LL, Watkins JC, Rodig S, Stover E, Strickland KC, D'Andrea AD, Wu CJ, Matulonis UA, Konstantinopoulos PA. Association of polymerase e-mutated and microsatellite-instable endometrial cancers with neoantigen load, number of tumor-infiltrating lymphocytes, and expression of PD-1 and PD-L1. *JAMA Oncol* 2015;**1**:1319–23
10. Bhan A, Soleimani M, Mandal SS. Long noncoding RNA and cancer: a new paradigm. *Cancer Res* 2017;**77**:3965–81
11. Jiang MC, Ni JJ, Cui WY, Wang BY, Zhuo W. Emerging roles of lncRNA in cancer and therapeutic opportunities. *Am J Cancer Res* 2019;**9**:1354–66
12. Li BL, Wan XP. The role of lncRNAs in the development of endometrial carcinoma. *Oncol Lett* 2018;**16**:3424–9
13. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, Sukumar S, Chang HY. Long non-coding RNA

HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 2010;**464**:1071–6
14. Lee S, Kopp F, Chang TC, Sataluri A, Chen B, Sivakumar S, Yu H, Xie Y, Mendell JT. Noncoding RNA NORAD regulates genomic stability by sequestering PUMILIO proteins. *Cell* 2016;**164**:69–80
15. Yan X, Hu Z, Feng Y, Hu X, Yuan J, Zhao SD, Zhang Y, Yang L, Shan W, He Q, Fan L, Kandalaft LE, Tanyi JL, Li C, Yuan CX, Zhang D, Yuan H, Hua K, Lu Y, Katsaros D, Huang Q, Montone K, Fan Y, Coukos G, Boyd J, Sood AK, Rebbeck T, Mills GB, Dang CV, Zhang L. Comprehensive genomic characterization of long non-coding RNAs across human cancers. *Cancer Cell* 2015;**28**:529–40
16. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, Barretina J, Boehm JS, Dobson J, Urashima M, Mc Henry KT, Pinchback RM, Ligon AH, Cho YJ, Haery L, Greulich H, Reich M, Winckler W, Lawrence MS, Weir BA, Tanaka KE, Chiang DY, Bass AJ, Loo A, Hoffman C, Prensner J, Liefeld T, Gao Q, Yecies D, Signoretti S, Maher E, Kaye FJ, Sasaki H, Tepper JE, Fletcher JA, Tabernero J, Baselga J, Tsao MS, Demichelis F, Rubin MA, Janne PA, Daly MJ, Nucera C, Levine RL, Ebert BL, Gabriel S, Rustgi AK, Antonescu CR, Ladanyi M, Letai A, Garraway LA, Loda M, Beer DG, True LD, Okamoto A, Pomeroy SL, Singer S, Golub TR, Lander ES, Getz G, Sellers WR, Meyerson M. The landscape of somatic copy-number alteration across human cancers. *Nature* 2010;**463**:899–905
17. Zhao L, Li Z, Chen W, Zhai W, Pan J, Pang H, Li X. H19 promotes endometrial cancer progression by modulating epithelial-mesenchymal transition. *Oncol Lett* 2017;**13**:363–9
18. Wang D, Wang D, Wang N, Long Z, Ren X. Long non-coding RNA BANCR promotes endometrial cancer cell proliferation and invasion by regulating MMP2 and MMP1 via ERK/MAPK signaling pathway. *Cell Physiol Biochem* 2016;**40**:644–56
19. Guo Q, Qian Z, Yan D, Li L, Huang L. LncRNA-MEG3 inhibits cell proliferation of endometrial carcinoma by repressing notch signaling. *Biomed Pharmacother* 2016;**82**:589–94
20. Guo C, Song WQ, Sun P, Jin L, Dai HY. LncRNA-GAS5 induces PTEN expression through inhibiting miR-103 in endometrial cancer cells. *J Biomed Sci* 2015;**22**:100
21. Zhou YX, Wang C, Mao LW, Wang YL, Xia LQ, Zhao W, Shen J, Chen J. Long noncoding RNA HOTAIR mediates the estrogen-induced metastasis of endometrial cancer cells via the miR-646/NPM1 axis. *Am J Physiol Cell Physiol* 2018;**314**:C690–c701
22. Łuczak A, Supernat A, Łapińska-Szumczyk S, Jachimowicz D, Majewska H, Gulczyński J, Żaczek AJ. HOTAIR in relation to epithelial-mesenchymal transition and cancer stem cells in molecular subtypes of endometrial cancer. *Int J Biol Markers* 2016;**31**:e245–51
23. Qiao Q, Li H. LncRNA FER1L4 suppresses cancer cell proliferation and cycle by regulating PTEN expression in endometrial carcinoma. *Biochem Biophys Res Commun* 2016;**478**:507–12
24. Ouyang D, Li R, Li Y, Zhu X. A 7-lncRNA signature predict prognosis of uterine corpus endometrial carcinoma. *J Cell Biochem* 2019;**120**:18465–77
25. Liu J, Mei J, Wang Y, Chen X, Pan J, Tong L, Zhang Y. Development of a novel immune-related lncRNA signature as a prognostic classifier for endometrial carcinoma. *Int J Biol Sci* 2021;**17**:448–59
26. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E, Culhane AC, Paull EO, Sivakumar IKA, Gentles AJ, Malhotra R, Farshidfar F, Colaprico A, Parker JS, Mose LE, Vo NS, Liu J, Liu Y, Rader J, Dhankani V, Reynolds SM, Bowlby R, Califano A, Cherniack AD, Anastassiou D, Bedognetti D, Mokrab Y, Newman AM, Rao A, Chen K, Krasnitz A, Hu H, Malta TM, Noushmehr H, Pedamallu CS, Bullman S, Ojesina AI, Lamb A, Zhou W, Shen H, Choueiri TK, Weinstein JN, Guinney J, Saltz J, Holt RA, Rabkin CS, Lazar AJ, Serody JS, Demicco EG, Disis ML, Vincent BG, Shmulevich I. The immune landscape of cancer. *Immunity* 2018;**48**:812–30.e14
27. Danilova L, Ho WJ, Zhu Q, Vithayathil T, De Jesus-Acosta A, Azad NS, Laheru DA, Fertig EJ, Anders R, Jaffee EM, Yarchoan M. Programmed cell death ligand-1 (PD-L1) and CD8 expression profiling identify an immunologic subtype of pancreatic ductal adenocarcinomas with favorable survival. *Cancer Immunol Res* 2019;**7**:886–95
28. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 2015;**160**:48–61
29. Masiero M, Simões FC, Han HD, Snell C, Peterkin T, Bridges E, Mangala LS, Wu SY, Pradeep S, Li D, Han C, Dalton H, Lopez-Berestein G, Tuynman JB, Mortensen N, Li JL, Patient R, Sood AK, Banham AH, Harris AL, Buffa FM. A core human primary tumor angiogenesis signature identifies the endothelial orphan receptor ELTD1 as a key regulator of angiogenesis. *Cancer Cell* 2013;**24**:229–41
30. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinform* 2010;**11**:367
31. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 2013;**31**:213–9
32. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;**14**:1–15
33. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Meth Mol Biol (Clifton, NJ)* 2018;**1711**:243–59
34. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res* 2019;**47**: W199–w205
35. Marquard AM, Eklund AC, Joshi T, Krzystanek M, Favero F, Wang ZC, Richardson AL, Silver DP, Szallasi Z, Birkbak NJ. Pan-cancer analysis of genomic scar signatures associated with homologous recombination deficiency suggests novel indications for existing cancer drugs. *Biomark Res* 2015;**3**:9
36. Ripley B, Venables B, Bates DM, Hornik K, Gebhardt A, Firth D, Ripley MB. Package 'MASS'. *Cran r* 2013;**538**:113–20
37. Hastie T, Qian J. Glmnet vignette. *Retrieved June* 2014;**9**:1–30
38. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res* 2019;**3**:18
39. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, Dimitriadoy S, Liu DL, Kantheti HS, Saghafinia S, Chakravarty D, Daian F, Gao Q, Bailey MH, Liang WW, Foltz SM, Shmulevich I, Ding L, Heins Z, Ochoa A, Gross B, Gao J, Zhang H, Kundra R, Kandoth C, Bahceci I, Dervishi L, Dogrusoz U, Zhou W, Shen H, Laird PW, Way GP, Greene CS, Liang H, Xiao Y, Wang C, Iavarone A, Berger AH, Bivona TG, Lazar AJ, Hammer GD, Giordano T, Kwong LN, McArthur G, Huang C, Tward AD, Frederick MJ, McCormick F, Meyerson M, Van Allen EM, Cherniack AD, Ciriello G, Sander C, Schultz N. Oncogenic signaling pathways in the cancer genome atlas. *Cell* 2018;**173**:321–37.e10
40. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science (New York NY)* 2011;**331**:1565–70
41. Fang Y, Fullwood MJ. Roles, functions, and mechanisms of long non-coding RNAs in cancer. *Genomics Proteomics Bioinform* 2016;**14**:42–54
42. Heward JA, Lindsay MA. Long non-coding RNAs in the regulation of the immune response. *Trends Immunol* 2014;**35**:408–19
43. Atianand MK, Hu W, Satpathy AT, Shen Y, Ricci EP, Alvarez-Dominguez JR, Bhatta A, Schattgen SA, McGowan JD, Blin J, Braun JE, Gandhi P, Moore MJ, Chang HY, Lodish HF, Caffrey DR, Fitzgerald KA. A long noncoding RNA lincRNA-EPS acts as a transcriptional brake to restrain inflammation. *Cell* 2016;**165**:1672–85
44. Ranzani V, Rossetti G, Panzeri I, Arrigoni A, Bonnal RJ, Curti S, Gruarin P, Provasi E, Sugliano E, Marconi M, De Francesco R, Geginat J, Bodega B, Abrignani S, Pagani M. The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. *Nat Immunol* 2015;**16**:318–25
45. Hirata E, Sahai E. Tumor microenvironment and differential responses to therapy. *Cold Spring Harb Perspect Med* 2017;**7**:a026781
46. Sadeghi Rad H, Monkman J, Warkiani ME, Ladwa R, O'Byrne K, Rezaei N, Kulasinghe A. Understanding the tumor microenvironment for effective immunotherapy. *Med Res Rev* 2021;**41**:1474–98
47. Beatty GL, Gladney WL. Immune escape mechanisms as a guide for cancer immunotherapy. *Clin Cancer Res* 2015;**21**:687–92

48. Jin KT, Yao JY, Fang XL, Di H, Ma YY. Roles of lncRNAs in cancer: focusing on angiogenesis. *Life Sci* 2020;**252**:117647
49. Li G, Kryczek I, Nam J, Li X, Li S, Li J, Wei S, Grove S, Vatan L, Zhou J, Du W, Lin H, Wang T, Subramanian C, Moon JJ, Cieslik M, Cohen M, Zou W. LIMIT is an immunogenic lncRNA in cancer immunity and immunotherapy. *Nat Cell Biol* 2021;**23**:526–37
50. Ehmer U, Sage J. Control of proliferation and cancer growth by the hippo signaling pathway. *Mol Cancer Res* 2016;**14**:127–40
51. Otto T, Sicinski P. Cell cycle proteins as promising targets in cancer therapy. *Nat Rev Cancer* 2017;**17**:93–115

(Received July 28, 2021, Accepted September 23, 2021)