Research article

Construction of a circadian rhythm-relevant gene signature for hepatocellular carcinoma prognosis, immunotherapy and chemosensitivity prediction

Zhiyu Ye [a,*], Ying Du [a], Wenguan Yu [a], Yunshou Lin [a], Li Zhang [a], Xiaoyu Chen [b,**]

[a] Department of Hernia and Hepatobiliary Surgery, The First Affiliated Hospital of Ningbo University, Ningbo, 315000, China
[b] Department of General Medicine, The Affiliation People's Hospital of Ningbo University, Ningbo, 315000, China

ARTICLE INFO

ABSTRACT

Aims: This study explored the molecular and biologic mechanisms underlying the association between circadian rhythm disorders (CRD) and increased risk for hepatocellular carcinoma (HCC).

Background: CRD are linked to increased risk for HCC, but the molecular and biologic mechanisms underlying this association are limited.

ObjectiveThe study constructed and validated a CRD related gene model as an independent prognostic factor for HCC, providing insight into the molecular mechanisms linking CRD to increased HCC risk and identifying potential indicators for the efficacy of immunotherapy and anticancer drugs. This helps provide important clues for personalized treatment strategies for HCC patients.

Methods: Gene sets correlated with circadian rhythm were obtained from the Molecular Signatures Database (MSigDB) to intersect with differentially expressed genes (DEGs) between tumor samples and control samples in The Cancer Genome Atlas (TCGA) and HCCDB18 from Hepatocellular Carcinoma Cell DataBase (HCCDB). The CRD related gene model was developed by univariate Cox and stepwise multivariate analysis. Immune checkpoint blockade (ICB) therapy and anticancer drugs were analyzed using the tumor immune dysfunction and exclusion (TIDE) and pRRophetic, respectively. Seurat determined the cell type of HCC by analyzing single-cell data, and malignant cells were identified using Copykat. To detect the mRNA levels of genes in the CRD related gene model, quantitative real-time polymerase chain reaction (qRT-PCR) was carried out.

Results: The activity of circadian rhythm in HCC tissue was significantly lower than that in control tissue. Subsequently, EZH2, IMPDH2, TYMS and SERPINE1 were selected to construct the CRD related gene model, which was an independent factor for HCC prognosis. Notably, low-risk patients had lower levels of immune cell infiltration and lower TIDE scores compared to high-risk patients with HCC, indicating that patients with a low risk may derive more benefit from immunotherapy. IMPDH2, TYMS and SERPINE1 expressed significantly higher in malignant cells than in benign epithelial cells.

---

* Corresponding author.
** Corresponding author.
E-mail addresses: yezhiyu2023@163.com (Z. Ye), cxy1232007@163.com (X. Chen).

Conclusions: This study presents a CRD related gene model to reveal the molecular perspective of the dependent mechanism of the association between CRD and cancer, which provides a potential indicator for understanding the preclinical efficacy of ICB and anticancer drugs.

1. Introduction

Liver disease accounts for 4 % of all deaths worldwide [1,2]. Hepatocellular carcinoma (HCC) usually develops from chronic liver disease and is a major death cause from liver disease, accounting for approximately 90 % of whole primary liver cancer cases [3–5]. Hepatitis B, cirrhosis, and liver cancer constitute the "trilogy pattern" of HCC development [6]. Only 8.1 % of HCC patients survived for more than 5 years after the date of diagnosis [7]. A variety of factors such as lack of curative treatment in patients with early detection, inadequate early detection strategies, inconsistent application of curative treatment and concurrent liver disease all contribute to a high death rate of HCC patients [8]. Further substantial reduction of HCC cases requires more widespread use of effective treatment of HBV and HCV-related chronic hepatitis and liver cirrhosis, reduction of alcohol consumption, more effective exercise, universal HBV vaccination, promotion of correct eating habits, and enhanced understanding of the molecular biology of HCC [9–11].

Suprachiasmatic nucleus in the hypothalamus regulates circadian rhythms and clock gene expressions [12]. Numerous studies has demonstrated that circadian rhythms related biological clock genes could modulate DNA repair [13] as well as inflammation related pathways to facilitate the progression of the cell cycle and cancer [14–16]. Given the essential role of circadian clock genes in tumors, certain of them have been identified as prognostic biomarkers for kinds of tumors [17–19]. It is worth noting that, although the brain detects light through the retina and regulates whole-body rhythms, the liver has an independent circadian rhythm [20,21]. Liver circadian rhythm disorders (CRDs) accelerate the development of liver diseases such as liver cirrhosis, hepatitis, fatty liver, and liver cancer, which disrupt clock function in turn [22]. Studies by Kettner et al. have shown that chronic CRD could induce spontaneous liver cancer in mice even without dietary manipulation, germline gene mutation or exogenous genotoxic stress [23]. A study showed that HCC was the second common cause to the death to the circadian gene mutant mouse models [24]. Circadian dysrhythmia is considered a hallmark of HCC, and the circadian mechanism itself has the ability to control and alter cancer markers [15]. Therefore, manipulation of the circadian rhythm may bring new hope for preventing the development of HCC and designing new strategies for the treatment of HCC [25].

In this study, we based on two public databases and obtained CRD-related genes that were analyzed by univariate and multivariate regression analyses to screen key prognostic genes for prognostic risk modeling. We assessed the level of immune infiltration and potential response to immunotherapy in patients from different HCC risk groups. Importantly, we analyzed the reliability of key CRD-associated genes using cellular experiments and at the single-cell level, respectively. In short, our study provided a new molecular perspective for comprehending the diagnosis and treatment mechanisms of HCC.

2. Materials and methods

2.1. Download and preprocess the raw data of HCC

The expression matrix and clinico-pathological data of HCC tissues were retrieved from the TCGA-LIHC cohort archived in TCGA database (https://cancergenome.nih.go) [26] and the HCCDB18 cohort archived in the HCCDB one-stop online resource (http://lifeome.net/database/hccdb/home.html). After removing the samples whose survival time and status were not recorded, 371 tumor samples and 50 cancer-adjacent control samples were retained in the TCGA-LIHC cohort, and 203 HCC samples and 177 control samples were retained in the HCCDB18 cohort.

The GSE125449 data set of the GEO database [27] provides the single-cell transcriptomic data of 9 HCC and 10 patients with intrahepatic cholangiocarcinoma, we extracted the single-cell transcriptomic data of 9 HCC, constructed the counting matrix from the original reading, and submitted the transcriptional data and transcript data for genes expressed in at least three cells, as well as for single cells with at least 200 genes to Seurat [28] for progressive analysis.

2.2. Conclusion of CRD score

The gene sets associated with CRD were acquired by visiting the Molecular Signatures Database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb), and the enrichment scores (ES) of CRD were calculated for a single sample using the single sample gene set enrichment analysis (ssGSEA) method of "GSVA" [29], in which the ranking of gene lists and the calculation of ES were dependent on the expression values of genes in the sample.

2.3. Screening of differential CRD genes

The DEGs between the control and HCC samples were analyzed by "limma" package [30,31] in R program, and the genes in the interval of adj. pval <0.05& | log2FC | > 1 were set as DEGs. Subsequently, the "ggplot2" package was employed to map the volcano [32]. The DEGs set in the HCCDB18 cohort, the DEGs set in the TCGA-LIHC cohort and the CRD gene set were uploaded to the biovenn

website to generate the Venn diagram and the intersection of the three sets was the differential CRD genes.

2.4. Development and validation in an external of CRD related gene model

Differential CRD genes were uploaded into the "survival" package [33] for univariate Cox regression analysis under p < 0.05. This step was followed by stepwise multivariate regression analysis of genes to develop a CRD-related gene model. The variables included in the model were the expression of the gene and the corresponding Cox coefficient. The CRD related gene model was used to calculate the risk scores of samples in the HCCDB18 cohort and TCGA-LIHC cohort. After ranking, the median risk score in each cohort was used as the grouping node, and the risk scores segments above the nodes were defined as the high-risk group, and the risk scores segments below the nodes were defined as the low-risk group. Overall survival (OS) was compared between risk groups using the log-rank test and visualized as Kaplan-Meier curves. The performance of the CRD related gene model was depicted by taking the true positive rate and false positive rate as the horizontal and vertical coordinates.

2.5. Drawing of nomogram

To identify independent clinical prognostic indicators of HCC from the TCGA-LIHC cohort, univariate and multivariate Cox analysis were implemented. The R package "rms" [34] was used to integrate the independent prognostic indicators to draw a Nomogram, and the predictive value of survival was calculated by generating a functional transformation relationship between the total score of the independent prognostic indicators and survival rate.

2.6. Immune infiltration analysis

27 immune-associated gene sets were obtained from the study of He et al. [35], and The Cancer Immunome Atlas was obtained from the study of Charoentong et al. [36], then the ssGSEA algorithm was utilized to calculate scores for different immune metrics based on these immunogenomic in the TCGA-LIHC cohort. For a supplement, ESTIMATE for inferring tumor purity and stromal and immune cell admixture [37] and CIBERSORT [38] evaluating the proportions 28 kinds of immune cells were also conducted.

2.7. Evaluation of therapeutic response to HCC

T cell dysfunction and T cell exclusion could be quantified by the genome-wide expression signatures of the TIDE to reflect and evaluate immune checkpoint blockade (ICB) clinical response [39]. The pRRophetic method based on ridge regression was also used to calculate the potential drug response of the sample based on the transcription profiles of the high-risk and low-risk samples [40], and the calculate the IC50 value and visualize it using the "pheatmap" package for heatmapping.

2.8. Analysis of single cell transcriptome

We further revealed cellular heterogeneity within HCC as well as pinpointed CRD-related gene expression based on single-cell level analysis [41]. For single-cell RNA sequencing (scRNA-seq) data in GSE125449 dataset, several filtering criteria were followed to ensure quality: the number of genes expressed in each cell was between 200 and 6000, the mitochondrial gene content calculated by the PercentageFeatureSet function was within 15 %, and the total UMI per cell was above 200. Seurat then applied NormalizeData function to normalize the count matrix with default parameters, and ScaleData function to convert normalized gene expression to z-score. Principal component analysis (PCA) was run using the "RunPCA" function with dim = 10. Considering the batch effect, the dimensionality reduction integration of harmony package is used for batch correction. Batch correction was performed using dimensionality reduction integration of the "harmony" package. Clusters were identified by "FindClusters" algorithm using a resolution of 1, and were annotated based on CellMarker database [42].

2.9. Cell culture and transfection

The Dulbecco's Modified Eagle Medium (DMEM, Gibco, USA) dissolved with penicillin/streptomycin and fetal bovine serum (Gibco, USA) was used for culturing the human hepatocytes (L02) and HCC cells (Huh7) at 37 °C with 5 % CO_2. The human hepatocytes (L02, Cellcook Biotech Company, Guangzhou, China) and HCC cells (Huh7, Typical Culture Reserve Center of China, Shanghai, China). Lipofectamine 2000 (Invitrogen, USA) was utilized in the transfection of negative control (si NC) and EZH2 siRNA (Sagon, China) into the cells.

2.10. Quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA separation from the cells were performing employing TRIzol (Thermo Fisher, USA) reagent. The cDNA was then synthesized from the RNA (500 ng) using HiScript II SuperMix (Vazyme, China). PCR amplification conditions were 46 cycles of 94 °C for 10 min, 94 °C for 10 s (s), and 60 °C for 45 s. The sequences of primer pairs for the genes targeted were listed in Table 1. GAPDH was an internal reference.

2.11. Statistical analysis

All data were analyzed with R program, and most common figures were plotted applying R package. Specially, the DEGs set in the TCGA-LIHC cohort and the CRD gene set were uploaded to the biovenn website to generate the Venn diagram. The ROC curve of CRD related gene model was generated by "timeROC" package [43]. Survival curves were generated using the survminer package [44], and heatmap was generated by the "pheatmap" package [45]. Differences in variables conforming to a normal distribution were tested by a two-tailed *t*-test, and differences in variables not conforming to a normal distribution were determined by a Wilcoxon test. A statistically significant p value < 0.05 was considered to be meaningful.

3. Results

3.1. Circadian rhythm was maladjusted in HCC tissues

We compared the enrichment of circadian rhythm-related pathways between HCC tissues and control tissues in the TCGA-LIHC cohort and the HCCDB18 cohort, and found that the activity of biocarta circadian pathway, circadian sleep wake cycle sleep, reactome circadian clock, circadian rhythm gene, circadian rhythm, circadian sleep wake cycle rem sleep, circadian sleep wake cycle, circadian regulation of gene expression, circadian rhythm mammal were significantly lower in HCC tissues than in normal tissues. Interestingly, dysregulated circadian clock genes have been considered as essential clinic indicators in the oncogenesis and development of HCC [46–48]. Therefore, altered "circadian regulation of gene expression" and "circadian rhythm gene" pathways in HCC groups, also drove us to study the CRD related genes with abnormal expressions. In addition, both metabolic and physiological processes are critically regulated by circadian rhythms [49]. The reduced activity of this pathway in HCC tissues suggests that circadian disorders may lead to hepatic metabolic dysfunction, which promotes tumorigenesis and progression. However, positive regulation of circadian sleep wake cycle sleep, circadian sleep/wake cycle, and non-rem sleep showed the completely opposite situation (Fig. 1). In general, we can clearly observe the obvious dysregulation of circadian rhythm-related pathways in HCC tissues, which provides important clues to our understanding of the role of CRD in the initiation and progression of HCC.

*3.2. Differential CRD-related genes and their genomic variants*

We screened 1862 DEGs between tumor tissues and paracancerous tissues in TCGA-LIHC cohort (Fig. 2A). 894 DEGs were identified between tumor samples and control samples in HCCDB18 dataset (Fig. 2B). 13 genes were in the intersection among the two DEGs sets and the integrated set of 274 CRD-related genes (Fig. 2C). TOP2A, ASS1 and NGFR were the three genes showed single nucleotide mutation (1 %) among 13 genes (Fig. 2D). We also observed only slight amplification or deletion of gene fragments occurred that these 13 genes (Fig. 2E). Despite the low frequency of mutations in genes including TOP2A, ASS1, and NGFR, they may still have a critical role in a particular setting or in a particular patient, and therefore it is advantageous to further build the comprehensiveness and predictive power of the risk model by including these genes.

3.3. Construction and optimization of CRD related gene model

4 of the 13 differential CRD genes were selected to construct the CRD related gene model by univariate Cox regression analysis and stepwise multivariate regression analysis (Fig. 3A), with the formula was: Risk score = 0.589*EZH2 +0.186*IMPDH2 -0.194*TYMS +0.075*SERPINE1. The risk score, which assigned by the CRD related gene model to each HCC sample in the TCGA-LIHC cohort showed a significant negative correlation with OS, that was, the low-risk sample had significantly lower death rate than the high-risk group. This prediction was reasonable because the area under the ROC curve (AUC) of the CRD related gene model reached the standard in 1, 3 and 5 years, which is 0.72,0.69,0.67 respectively (Fig. 3B). The ability of CRD related gene model to separate risk categories was also better in HCCDB18 dataset, with the model AUC reaching 0.75, 0.65 and 0.73 at 1, 3 and 5 years, respectively. Under the discrimination of the CRD related gene model, we observed significant OS differences between the low- and high-risk groups in the HCCDB18 dataset (Fig. 3C). Among several clinical factors for HCC provided by TCGA, only AJCC stage was an independent factor for HCC prognosis, as was risk score (Fig. 3D and E). Nomogram integrated these two factors together to predict OS. In this nomogram, a scale was used to mark the line corresponding to each variable to represent the values range of the variable, and the

Table 1
The sequences of primer pairs for target genes.

Gene	Forward primer sequence (5′-3′)	Reverse primer sequence (5′-3′)
EZH2	GACCTCTGTCTTACTTGTGGAGC	CGTCAGATGGTGCCAGCAATAG
IMPDH2	GACGACGGACTCACAGCACAG	CTTGGTCAGAGCAGAAGTCAGGTC
SERPINE1	GCAACGTGGTTTTCTCACCC	CTCTAGGGGCTTCCTGAGGT
TYMS	GCCTCGGTGTGCCTTTCA	GATGTGCGCAATCATGTACGT
BMAL1	TGGATGAAGACAACGAACCA	TAGCTGTTGCCCTCTGGTCT
CLOCK	TCTCAGACCCTTCCTCAACA	TGACCTTCTTTGCACCATCTT
CRY	TTCTTCCCATCAAAACTGG	AAACGCATCCGATTGTAACC
GAPDH	ACAACTTTGGTATCGTGGAAGG	GCCATCACGCCACAGTTTC
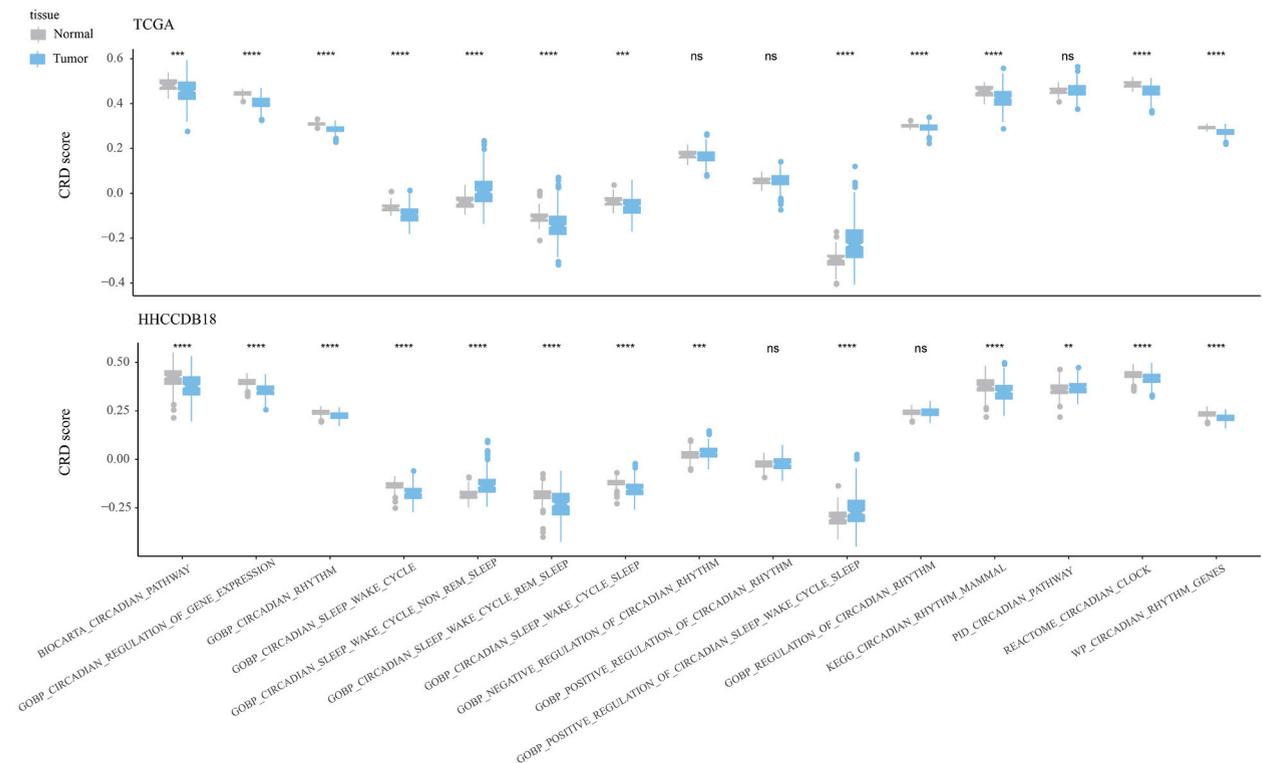
Fig. 1. Differences in circadian pathway activity between HCC tissues and control tissues from the TCGA-LIHC cohort and the HCCDB18 cohort ** <0.01, *** <0.001, **** <0.0001, ns: no significance.

contribution of the factor to clinical outcome events was reflected in the length of the line segment. The Risk score contributed much more to the nomogram than the AJCC stage (Fig. 3F). The prediction results of nomogram for 1-year, 3-year and 5-year OS in the calibration curve were close to the diagonal (Fig. 3G), indicating the good prognosis prediction ability of nomogram.

3.4. Immune indication and therapeutic response based on CRD related gene model

Immunogenomic analyses were performed according to risk groups based on the CRD related gene model. In the results, we found that compared to the low-risk group, the high-risk group significantly highly enriched activated dendritic cells (aDCs), APC co stimulation, macrophage, MHC class I, regulatory T cells (Tregs), while the levels of B cells, mast cells, cytolytic activity, NK cells, neutrophils, naive CD8 T cell, Th17, DC, monocyte, type I IFN response, NK cell and gamma delta cells were significantly lowered in the high-risk group (Fig. 4A and B). Similarly, naive CD8 T cell, Th17, Monocyte and Gamma_delta immunocytes were also had high proportions in low – risk groups (Fig. 4C). Collectively, ESTIMATE analysis also confirmed the relatively higher immune infiltration level in low-risk groups (Fig. 4D). Significant differences in TIDE score and ICB treatment response were also shown between the two samples grouped by the CRD related gene model, with lower TIDE scores and higher rates of response to ICB treatment in the low-risk group (Fig. 4E and F). The IC50 value is a measure of the potency of a drug or compound, specifically, the lower the IC50 value, the more potent the drug is as it can effectively inhibit cell growth or enzyme activity at a lower concentration. Among the drugs that showed significant IC50 differences between the high-risk and low-risk groups, Vinorelbine, Etoposide, Gemcitabine, and Doxorubicin showed significantly lower IC50 values in the high-risk group than in the low-risk group and were therefore more suitable for the treatment of high-risk samples. Compared to the high-risk group, the low-risk group had significantly lower IC50 values of Erlotinib, TGX221, CGP_60,474, and DMOG, so they might be more suitable for the low-risk group (Fig. 4G). These findings could facilitate the prediction of clinical outcomes and more personalized treatment regimens for patients in different risk groups of HCC.

3.5. Potential regulation mechanism of CRD related gene model in HCC

To explore the mechanism of the mobilization of the CRD related gene model, GSEA was performed separately for the high-risk and low-risk groups. Almost all the pathways significantly activated in the high-risk group were those driving cell proliferation, including MYC targets, PI3K-Akt-mTOR signaling, TGF-β signaling, E2F targets, DNA repair, etc (Fig. 5A). The enrichment scores of Hallmark gene sets in TCGA-LIHC cohort samples were calculated by ssGSEA, and 29 of them between high-risk and low-risk samples showed significant differences. Most of the 17 pathways positively related to risk score were carcinogenic pathways, while the 12 pathways negatively related to risk score, the strongest correlation with risk score was metabolic pathways, including xenobiotic metabolism,
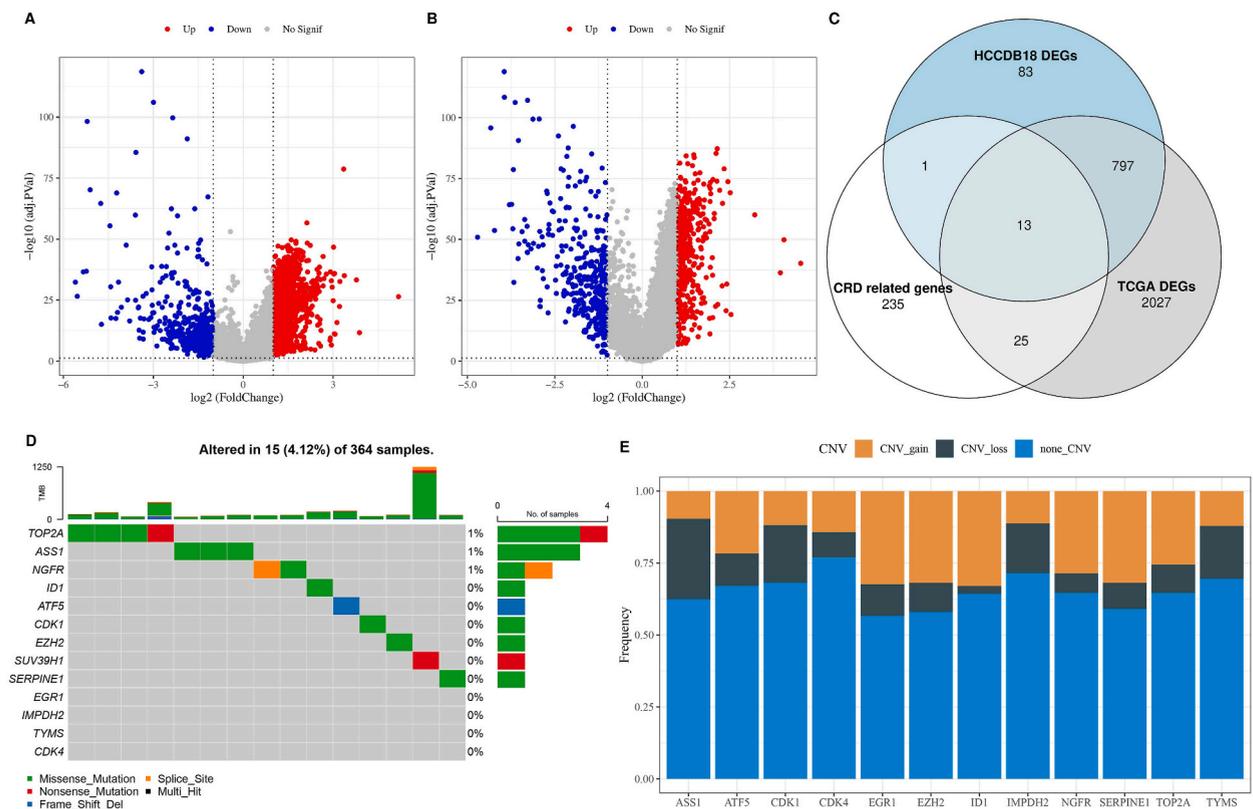
Fig. 2. Differential CRD genes and their genomic variants (A) DEGs between tumor tissues and paracancerous tissues in TCGA-LIHC cohort, where blue dots are down-regulated DEGs, red dots are up-regulated DEGs, and gray dots are non-degs. (B) DEGs between tumor samples and paracancerous samples in HCCDB18 cohort. (C) Identification of intersection among DEGs sets for TCGA-LIHC cohort, HCCDB18 cohort, CRD-related genes set. (D) The single nucleotide variation of 13 differential CRD genes. (E) Histograms show the probability of copy-number amplifications and deletions occurring in differential CRD genes.

oxidative phosphorylation, fatty acid metabolism, bile acid metabolism (Fig. 5B).

3.6. Verification of genes in CRD related gene model at single cell level

To further support the important role of the key genes we screened in the development of HCC, we verified the expression capacity of these genes in HCC by single-cell level analysis. Through clustering analysis for cells in GSE125449 dataset, we identified 19 clusters (Fig. 6A). After annotation, six cell types were identified: tumor-associated macrophages (TAMs), hepatic cells, endothelial cells, epithelial cells, T cells, and B cells (Fig. 6B). Genes that were detected to be specifically highly expressed in each type of cell were shown in Fig. 6C. Given the significant enrichment of proliferative pathways in the high-risk samples, we extracted the key cellular players in the unrestricted proliferation of HCC, namely epithelial cells, and then used copykat package to identify malignant cells among them. We found that among the four genes in the CRD related gene model, IMPDH2, TYMS and SERPINE1 were expressed at significantly higher levels in malignant cells than in normal epithelial cells (Fig. 6D), validating the important role of these genes in hepatocellular carcinoma development and progression. This suggests that the single-cell level analysis supports the plausibility of our constructed model of CRD-associated genes, revealing the specific expression patterns and potential functions of these genes in the tumor microenvironment.

3.7. Verification of the expression levels of the four model genes in HCC cells clines

The expression of EZH2, IMPDH2, SERPINE1 and TYMS was detected using qRT-PCR in HCC cells (Huh-7) and human normal hepatocytes (L02). It was observed that these four genes were all up-regulated in HCC cells lines compared with normal L02 cell lines (Fig. 7A–D). As EZH2 is essential for maintenance of circadian clock function [50], we applied siRNA to knock down the expression of EZH2 in Huh-7 cell lines and detected the levels of circadian clock genes. We could see that in comparison with si-NC cells, the expressions of BMAL1, CLOCK and CRY were distinctly upregulated in si-EZH2 cells (Fig. 7E–G), demonstrating the vital role of EZH2 in circadian clock regulation.
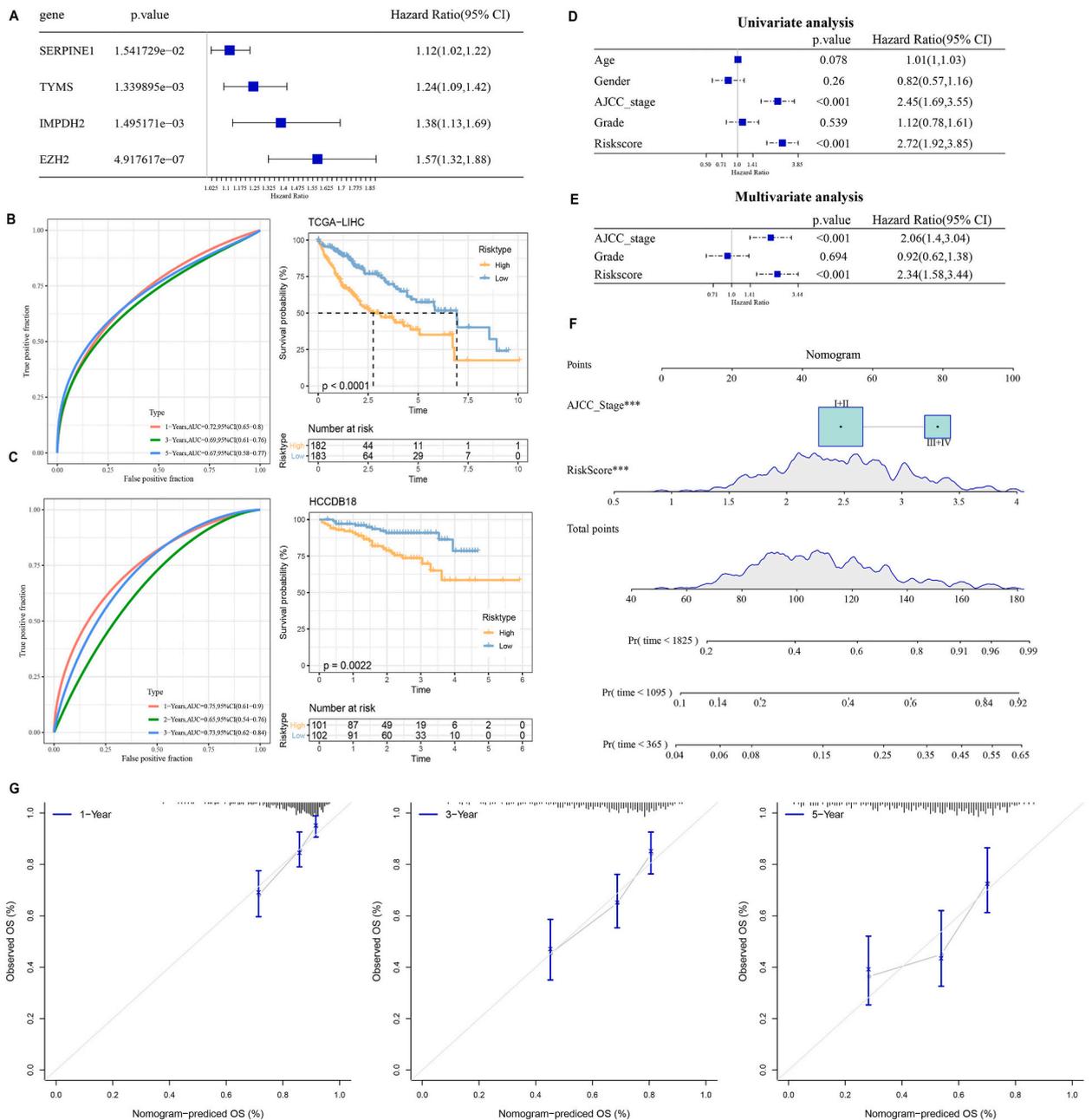
Fig. 3. Construction and optimization of CRD related gene model (A) Multivariate Cox regression analysis of the four genes in the CRD related gene model. (B) ROC curve and Kaplan-Meier curve of CRD related gene model in distinguishing the mortality risk of samples in the TCGA-LIHC cohort. (C) ROC curve and Kaplan-Meier curve of CRD related gene model in evaluating OS of samples from HCCDB18 dataset. (D, E) Univariate Cox regression analysis of several clinical factors provided by TCGA and risk scores for HCC. (F) Nomogram constructed by integrating two independent prognostic factors, the line corresponding to each variable was marked with a scale, representing the range of values of the variable, and the length of the line segment reflected the contribution of the factor to clinical outcome events. (G) Calibration curves to evaluate the performance of the nomogram in predicting 1 -, 3 -, and 5-year OS of HCC.

4. Discussion

Circadian rhythm is an evolutionarily conserved timing system, and epidemiological studies have shown that disruption of circadian rhythm in humans is associated with increased cancer risk [51]. The molecular understanding of circadian rhythm will provide a new therapeutic frontier for cancer [52]. Circadian rhythm changes in gene expression are the common features of different tumor types, and tumor genetic and genomic analysis showed that the circadian rhythm changes in gene expression more widely than
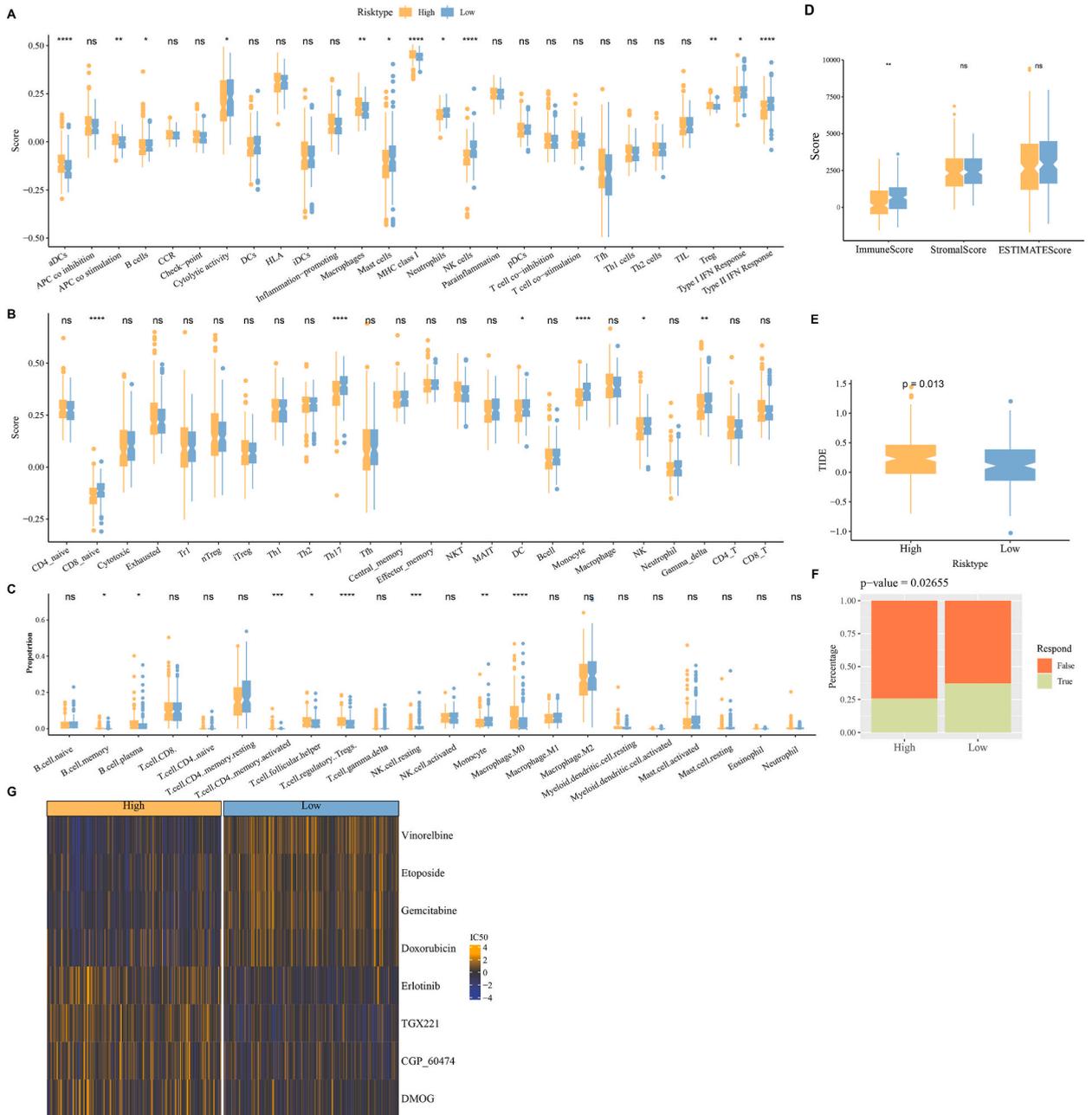
Fig. 4. Immune indication and therapeutic response based on CRD related gene model (A) Differences in enrichment scores of 27 immune-associated gene sets between the high-risk and low-risk groups grouped by the CRD related gene model. (B) The Cancer Immunome Atlas analysis based on CRD related gene model. (C) CIBERSORT analysis. (D) ESTIMATE analysis. (E) TIDE score differences between two groups of samples grouped by CRD related gene model. (F) Response to ICB treatment in samples grouped by the CRD related gene model. (G) The IC50 value of drugs in high-risk group and low-risk group samples. * <0.05 , ** <0.01 , *** <0.001 , **** <0.0001, ns: no significance.

circadian gene mutations [53]. Therefore, this study established a risk model for predicting HCC prognosis based on CRD-related genes. By analyzing the immune microenvironment of patients with different risks and evaluating their potential response to immunotherapy. Importantly, we also validated the rationality of the key genes and the model by in vitro cellular experiments and at the single-cell level. This not only strengthens the biological basis of the model, but also provides important clues for further research on the functional mechanism of CRD-related genes in HCC and personalized therapeutic strategies.

We found that the circadian rhythm was significantly suppressed in HCC tissues, which also in line with previous studies reporting that carcinogenic processes directly impair circadian rhythms [53]. We found 13 differential CRD genes between HCC tumor tissues and para-cancerous tissues, of which 4 were selected to construct CRD related gene model as well as a nomogram. Also, scRNA-seq data
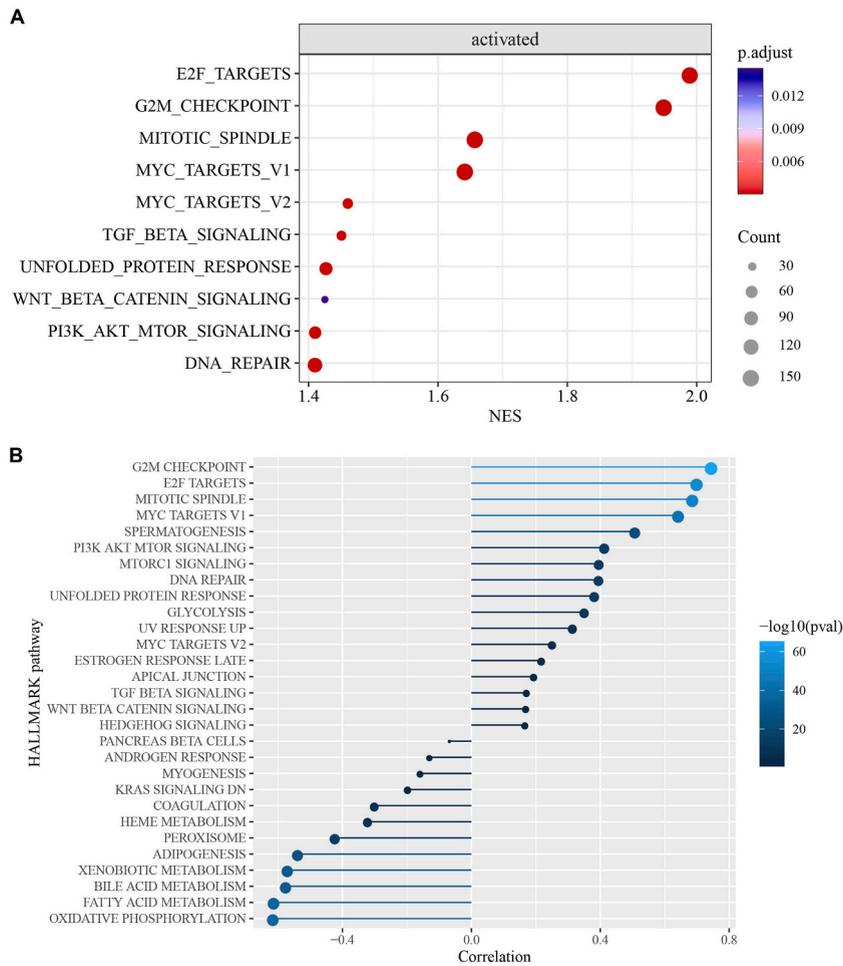
Fig. 5. Potential regulation mechanism of CRD related gene model in HCC (A) Bubble plot of the top 10 pathways significantly activated in the high-risk group. (B) The correlation between risk score and 29 pathways showing significant differences in enrichment scores between high-risk and low-risk samples.

also verified the significance of these 4 genes, 3 of which could help discriminate malignant cells, hinting a cancer promoting effect.

In detail, EZH2 (*enhancer of zeste homologue 2*) has been found to have oncogenic effects on HCC, the direction of regulation includes tumor microenvironment (TME) and cancer cell growth, and has a strong association with ICB treatment as well as circadian rhythm biological clock regulation [50]. In terms of TME regulation, EZH2 weakens PKLR expression to inactivate NK cells residing in the HCC and promotes tumor immune evasion by inducing IL-6 production [54,55]. EZH2 negatively regulates the expression of PD-L1 in HCC patients with immune-activated microenvironment [56]. Inhibition of EZH2 in HCC impaired HCC cell growth and induced cell apoptosis in a dose-dependent manner [57]. Additionally, EZH2 was demonstrated to be critical for mammalian liver circadian regulation and hematopoiesis through gene silencing [58]. Circadian clock gene BMAL1 was reported to repress cancer development and improve paclitaxel sensitivity in tongue squamous cell carcinoma via suppress EZH2 levels [59]. Here, our in vitro experiment also proved the role of EZH2 on the modulation of circadian clock gene such as BMAL1, CLOCK and CRY. The upregulated expression of IMPDH2 (*Inosine monophosphate dehydrogenase type II*) increases the proliferation and tumorigenicity of HCC in vitro by promoting cell growth rate and colony formation, and is closely related to the poor survival outcome of patients [60]. TYMS (*Thymidylate synthase*) was identified as a CD8 T cell-related genes [61] and the Silencing of TYMS in HCC significantly reduced DNA synthesis and extracellular matrix (ECM) degradation, thereby inhibiting tumor metastasis [62]. SERPINE1 (*serpin family E member 1*) encodes plasminogen activator inhibitor-1 (PAI-1), which is secreted by tumor-associated macrophages (TAMs) to enhance the malignant behavior of HCC cells [63]. All together, these references disclosed the significance of screened 4 genes on the progression of HCC with similar overexpressed phenomenon in HCC cell lines.

We also evaluated the immune infiltration level in high and low risk groups. Previous researches have pointed that CRD related genes are usually expressed in immune cells and display circadian transitions with fixed rhythms, and they involve in a series of immune regulatory processes, such as synthesis, phagocytosis, apoptosis, and release of cytolytic factors, cytokines, chemokines [64–66]. Here, higher Tregs scores and lower CD8 T cell scores were discovered in high risk groups. A pan-cancer analysis has revealed
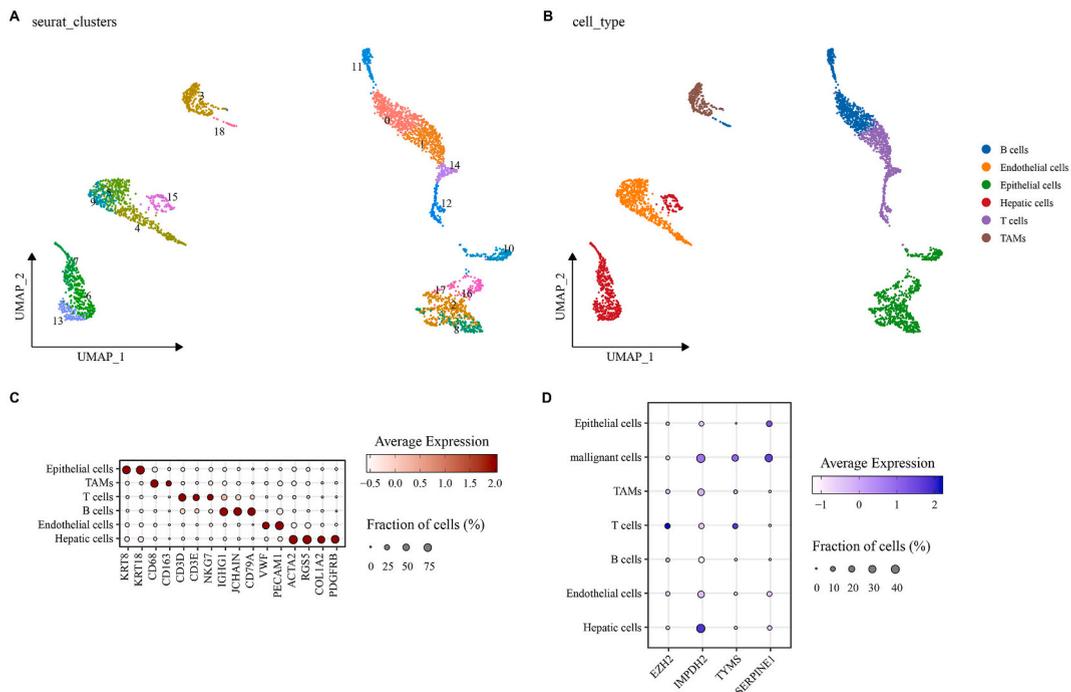
Fig. 6. Verification of genes in CRD related gene model at single cell level (A) UMAP plot shows all 19 clusters in the GSE125449 dataset. (B) The UMAP plot indicates that 19clusters were annotated as 6 cell types. (C) Genes that are specifically highly expressed in each cell subtype. (D) The expression of four genes in CRD related gene model in malignant cells and normal epithelial cells.

that Tregs were correlated with poor prognosis, while CD8 T cell were related to favorable outcomes [36]. Combined with aforemented immune or inflammatory functions of CRD related genes, we further confirmed the close connections between CRD and immune state in HCC patients.

The relationship of the CRD related gene model to immune escape and ICB treatment was reflected in the TIDE score, as TIDE evaluates the ICB efficacy of tumor samples by simulating two immune escape mechanisms. The higher TIDE score, the bigger chance for immune escape [39,67]. The TIDE score has a higher accuracy in evaluating prognosis of tumor patients treated with ICB agents [68,69] than using PD-L1 expression level and its application in assessing the ICB therapy efficacy has been confirmed in many researches [70–72]. In our research, there was a significant positive correlation between CRD related gene model and TIDE score, it can be said that high-risk samples had a higher TIDE score, a higher chance of immune escape, and a lower response rate to ICB treatment than low-risk samples as defined by the CRD related gene model. In addition, the CRD related gene model indeed also showed significant positive correlations with pathways that promote cell proliferation, including TGF-β signaling, E2F targets, PI3K-Akt-mTOR signaling, DNA repair, MYC targets, and so on. These cancer promoting pathways maybe the main causes of worsening prognosis in high-risk groups. It has been reported that CRD may promote the transformation of normal cells into malignant cells [73]. By analyzing the scRNA-seq data of HCC, we found a total of 6 major cell types. Among which, we extracted the key cellular role in the unlimited proliferation of HCC - epithelial cells, and identified the malignant cells. Through single-cell transcriptome analysis, we found that the expression levels of IMPDH2, TYMS and SERPINE1 of the four genes in CRD related gene model in malignant cells were significantly higher than those in normal epithelial cells. In particular, the high expression of EZH2 and IMPDH2 in tumor-associated macrophages suggests that they may promote tumor growth and metastasis by affecting the tumor immune microenvironment. This suggests that CRD-related genes are expressed in the immune microenvironment of HCC, providing important clues for understanding the role of CRD in tumor immunoregulation. It is worth noting that the shortcomings of this study are inevitable. The risk model and nomogram were established based on limited retrospective data. In the future, more clinical sample are needed to verify their predicting abilities. In addition, the single-cell transcriptome data used only covered a portion of the HCC samples, which will be incorporated into further studies with more HCC samples and diverse patient populations to comprehensively capture the cellular heterogeneity of the tumors and gene expression characteristics. Finally, further experimental functional validation is necessary, and we will conduct experiments *in vivo* and in vitro to confirm the specific biological functions and mechanisms of CRD-related genes, which can increase the credibility of the study results. This research constructed a prognostic model of HCC using CRD-related genes and screened four key signature genes: EZH2, IMPDH2, TYMS, and SERPINE1. The expression levels of these genes significantly affected HCC prognosis and played a critical part in tumor development and circadian rhythm regulation. In addition, our study revealed the important role of CRD-related genes in the immune microenvironment of HCC, especially the differential expression of these genes in different immune cell types, which further revealed the complex mechanism of CRD in regulating tumor immunity.
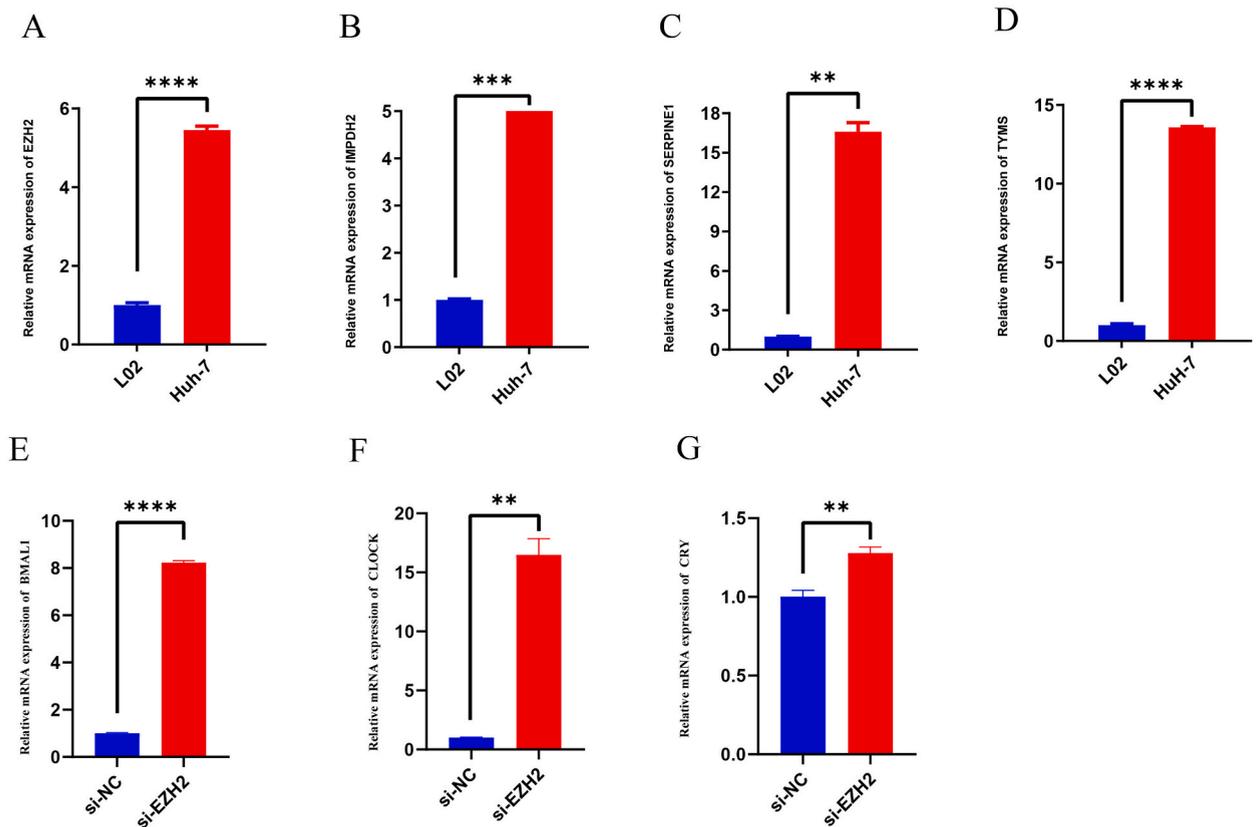
Fig. 7. Verification of genes expressions in HCC cell lines. (A–D) qRT-PCR for detecting the expression of EZH2, IMPDH2, SERPINE1 and TYMS in Huh7 and L02 cell lines. (*E*–G) qRT-PCR for detecting the expression of BMAL1, CLOCK and CRY after EZH2 inhibition. N = 3, *< 0.05, ** <0.01 , *** <0.001 , **** <0.0001. The results are presented as mean ± SEM.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

The public dataset used in this study is available in GSE125449 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE125449).

Author's contributions

All authors contributed to this present work: [ZYY] & [YD] designed the study, [WGY] acquired the data, [ZYY] applied for funding, [YSL] and [LZ] improved the figure quality, [LZ] and [XYC] drafted the manuscript, [ZYY] revised the manuscript. All authors read and approved the manuscript.

Funding

CRediT authorship contribution statement

Zhiyu Ye: Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Funding acquisition, Conceptualization. **Ying Du:** Visualization, Software, Resources, Methodology, Formal analysis. **Wenguan Yu:** Visualization, Supervision, Methodology, Data curation, Conceptualization. **Yunshou Lin:** Visualization, Supervision, Resources. **Li Zhang:** Visualization, Validation, Investigation, Formal analysis. **Xiaoyu Chen:** Writing – review & editing, Supervision, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Not applicable.

Abbreviations

CRD	Circadian rhythm disorders
HCC	Hepatocellular carcinoma
DEG	Differentially expressed gene
ICB	immune checkpoint blockade
TCGA	The Cancer Genome Atlas
MSigDB	Molecular Signatures Database
ssGSEA	Single Sample Gene Set Enrichment Analysis
PCA	Principal component analysis
ROC	RReceiver operating characteristic
AUC	Area under the curve
aDC	activated dendritic cell
Treg	Regulatory T cell
EZH2	Enhancer of zeste homologue 2
IMPDH2	Inosine monophosphate dehydrogenase type II
TYMS	Thymidylate synthase
SERPINE1	Serpin family E member 1
TAM	Tumor-associated macrophage

References

[1] M.D.M. Rigual, P. Sanchez Sanchez, N. Djouder, Is liver regeneration key in hepatocellular carcinoma development? Trends Cancer 9 (2) (2023) 140–157.
[2] H. Devarbhavi, et al., Global burden of liver disease: 2023 update, J. Hepatol. 79 (2) (2023) 516–537.
[3] A.J. Craig, et al., Tumour evolution in hepatocellular carcinoma, Nat. Rev. Gastroenterol. Hepatol. 17 (3) (2020) 139–152.
[4] X. Shu, Q. Wang, Q. Wu, The Eph/Ephrin system in hepatocellular carcinoma: functional roles and potential therapeutic targets, Oncologie 24 (3) (2022) 427–439.
[5] R. Xu, et al., A novel prognostic target-gene signature and nomogram based on an integrated bioinformatics analysis in hepatocellular carcinoma, Biocell 46 (5) (2022) 1261–1288.
[6] Z. Liao, et al., Current concepts of precancerous lesions of hepatocellular carcinoma: recent progress in diagnosis, Diagnostics (Basel) 13 (7) (2023).
[7] X. Zhang, H.B. El-Serag, A.P. Thrift, Predictors of five-year survival among patients with hepatocellular carcinoma in the United States: an analysis of SEER-Medicare, Cancer Causes Control 32 (4) (2021) 317–325.
[8] N.D. Parikh, N. Tayob, A.G. Singal, Blood-based biomarkers for hepatocellular carcinoma screening: approaching the end of the ultrasound era? J. Hepatol. 78 (1) (2023) 207–216.
[9] E. Sagnelli, et al., Epidemiological and etiological variations in hepatocellular carcinoma, Infection 48 (1) (2020) 7–17.
[10] N. Juaid, et al., Anti-hepatocellular carcinoma biomolecules: molecular targets insights, Int. J. Mol. Sci. 22 (19) (2021).
[11] X. Wang, et al., A UHPLC/MS/MS assay based on an isotope-labeled peptide for sensitive miR-21 detection in HCC serum, Oncologie 24 (3) (2022) 513–526.
[12] A. Sancar, et al., Circadian clock control of the cellular response to DNA damage, FEBS Lett. 584 (12) (2010) 2618–2625.
[13] B.S.C. Koritala, et al., Night shift schedule causes circadian dysregulation of DNA repair genes and elevated DNA damage in humans, J. Pineal Res. 70 (3) (2021) e12726.
[14] W. Xuan, et al., Circadian regulation of cancer cell and tumor microenvironment crosstalk, Trends Cell Biol. 31 (11) (2021) 940–950.
[15] F. Fagiani, et al., Molecular regulations of circadian rhythm and implications for physiology and diseases, Signal Transduct. Targeted Ther. 7 (1) (2022) 41.
[16] M.E. Geusz, A. Malik, A. De, Insights into oncogenesis from circadian timing in cancer stem cells, Crit. Rev. Oncog. 26 (4) (2021) 1–17.
[17] R. Chai, et al., Circadian clock genes act as diagnostic and prognostic biomarkers of glioma: clinic implications for chronotherapy, BioMed Res. Int. 2022 (2022) 9774879.
[18] H. Chi, et al., Circadian rhythm-related genes index: a predictor for HNSCC prognosis, immunotherapy efficacy, and chemosensitivity, Front. Immunol. 14 (2023) 1091218.

[19] H. Xian, et al., Identification of TIMELESS and RORA as key clock molecules of non-small cell lung cancer and the comprehensive analysis, BMC Cancer 22 (1) (2022) 107.

[20] K. Sato, et al., Melatonin and circadian rhythms in liver diseases: functional roles and potential therapies, J. Pineal Res. 68 (3) (2020) e12639.

[21] K.B. Koronowski, et al., Defining the independence of the liver circadian clock, Cell 177 (6) (2019) 1448–1462 e14.

[22] Y. Tahara, S. Shibata, Circadian rhythms of liver physiology and disease: experimental and clinical evidence, Nat. Rev. Gastroenterol. Hepatol. 13 (4) (2016) 217–226.

[23] N.M. Kettner, et al., Circadian homeostasis of liver metabolism suppresses hepatocarcinogenesis, Cancer Cell 30 (6) (2016) 909–924.

[24] S. Lee, et al., Disrupting circadian homeostasis of sympathetic signaling promotes tumor development in mice, PLoS One 5 (6) (2010) e10995.

[25] M. Crespo, M. Leiva, G. Sabio, Circadian clock and liver cancer, Cancers 13 (14) (2021).

[26] D. Yan, et al., Exploration of combinational therapeutic strategies for HCC based on TCGA HCC database, Oncologie 24 (1) (2022) 101–111.

[27] L. Ma, et al., Tumor cell biodiversity drives microenvironmental reprogramming in liver cancer, Cancer Cell 36 (4) (2019) 418–430.e6.

[28] A. Butler, et al., Integrating single-cell transcriptomic data across different conditions, technologies, and species, Nat. Biotechnol. 36 (5) (2018) 411–420.

[29] S. Hanzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7.

[30] M.E. Ritchie, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47.

[31] Z. Song, et al., CHDTEPDB: transcriptome expression profile database and interactive analysis platform for congenital heart disease, Congenit. Heart Dis. 18 (6) (2023) 693–701.

[32] E.K. Gustavsson, et al., ggtranscript: an R package for the visualization and interpretation of transcript isoforms using ggplot2, Bioinformatics 38 (15) (2022) 3844–3846.

[33] T.M. Therneau, T. Lumley, Package 'survival', R Top Doc 128 (10) (2015) 28–33.

[34] F.E. Harrell, F. Harrell, F.E. Harrell, rms: regression modeling strategies. R package version 4.0–0 (2013).

[35] Y. He, et al., Classification of triple-negative breast cancers based on Immunogenomic profiling, J. Exp. Clin. Cancer Res. 37 (1) (2018) 327.

[36] P. Charoentong, et al., Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade, Cell Rep. 18 (1) (2017) 248–262.

[37] K. Yoshihara, et al., Inferring tumour purity and stromal and immune cell admixture from expression data, Nat. Commun. 4 (2013) 2612.

[38] B. Chen, et al., Profiling tumor infiltrating immune cells with CIBERSORT, Methods Mol. Biol. 1711 (2018) 243–259.

[39] P. Jiang, et al., Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response, Nat. Med. 24 (10) (2018) 1550–1558.

[40] P. Geeleher, N. Cox, R.S. Huang, pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels, PLoS One 9 (9) (2014) e107468.

[41] A. Zulibiya, et al., Single-cell RNA sequencing reveals potential for endothelial-to-mesenchymal transition in tetralogy of fallot, Congenit. Heart Dis. 18 (6) (2023) 611–625.

[42] X. Zhang, et al., CellMarker: a manually curated resource of cell markers in human and mouse, Nucleic Acids Res. 47 (D1) (2019) D721–D728.

[43] P. Blanche, J.F. Dartigues, H. Jacqmin-Gadda, Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks, Stat. Med. 32 (30) (2013) 5381–5397.

[44] A. Kassambara, et al., *Package 'survminer'.* Drawing Survival Curves Using 'ggplot2'(R Package Version 03 1), 2017.

[45] R. Kolde, Pheatmap: Pretty Heatmaps, 2015.

[46] Y. Liang, et al., Dysregulation of circadian clock genes as significant clinic factor in the tumorigenesis of hepatocellular carcinoma, Comput. Math. Methods Med. 2021 (2021) 8238833.

[47] Y. Jiang, et al., The expression and function of circadian rhythm genes in hepatocellular carcinoma, Oxid. Med. Cell. Longev. 2021 (2021) 4044606.

[48] Y. Lee, A.S. Tanggono, Potential role of the circadian clock in the regulation of cancer stem cells and cancer therapy, Int. J. Mol. Sci. 23 (22) (2022).

[49] R. Costa, et al., Circadian rhythms and the liver, Liver Int. 43 (3) (2023) 534–545.

[50] J.P. Etchegaray, et al., The polycomb group protein EZH2 is required for mammalian circadian clock function, J. Biol. Chem. 281 (30) (2006) 21209–21215.

[51] S. Masri, P. Sassone-Corsi, The emerging link between cancer, metabolism, and circadian rhythms, Nat. Med. 24 (12) (2018) 1795–1803.

[52] L. Zhou, et al., Circadian rhythms and cancers: the intrinsic links and therapeutic potentials, J. Hematol. Oncol. 15 (1) (2022) 21.

[53] G. Sulli, M.T.Y. Lam, S. Panda, Interplay between circadian clock and cancer: new frontiers for cancer treatment, Trends Cancer 5 (8) (2019) 475–494.

[54] Z. Liu, et al., Extracellular vesicle-mediated communication between hepatocytes and natural killer cells promotes hepatocellular tumorigenesis, Mol. Ther. 30 (2) (2022) 606–620.

[55] M.T. Mok, et al., CCRK is a novel signalling hub exploitable in cancer immunotherapy, Pharmacol. Ther. 186 (2018) 138–151.

[56] G. Xiao, et al., EZH2 negatively regulates PD-L1 expression in hepatocellular carcinoma, J Immunother Cancer 7 (1) (2019) 300.

[57] N. Qiang, et al., Alteration of the tumor microenvironment by pharmacological inhibition of EZH2 in hepatocellular carcinoma, Int. Immunopharm. 118 (2023) 110068.

[58] Y. Zhong, et al., Ezh2 promotes clock function and hematopoiesis independent of histone methyltransferase activity in zebrafish, Nucleic Acids Res. 46 (7) (2018) 3382–3399.

[59] Q. Tang, et al., Circadian clock gene Bmal1 inhibits tumorigenesis and increases paclitaxel sensitivity in tongue squamous cell carcinoma, Cancer Res. 77 (2) (2017) 532–544.

[60] Y. He, et al., Over-expression of IMPDH2 is associated with tumor progression and poor prognosis in hepatocellular carcinoma, Am. J. Cancer Res. 8 (8) (2018) 1604–1614.

[61] J. Ren, et al., A novel prognostic model of early-stage lung adenocarcinoma integrating methylation and immune biomarkers, Front. Genet. 11 (2020) 634634.

[62] S. Li, et al., Identification and validation of TYMS as a potential biomarker for risk of metastasis development in hepatocellular carcinoma, Front. Oncol. 11 (2021) 762821.

[63] S. Chen, et al., Cancer-associated fibroblast-induced M2-polarized macrophages promote hepatocellular carcinoma progression via the plasminogen activator inhibitor-1 pathway, Int. J. Oncol. 59 (2) (2021).

[64] N. Labrecque, N. Cermakian, Circadian clocks in the immune system, J. Biol. Rhythm. 30 (4) (2015) 277–290.

[65] M.A. Venneri, et al., Circadian rhythm of glucocorticoid administration entrains clock genes in immune cells: a dream trial ancillary study, J. Clin. Endocrinol. Metab. 103 (8) (2018) 2998–3009.

[66] K. Man, A. Loudon, A. Chawla, Immunity around the clock, Science 354 (6315) (2016) 999–1003.

[67] X. Lin, et al., A novel immune-associated prognostic signature based on the immune cell infiltration analysis for hepatocellular carcinoma, Oncologie 26 (1) (2024) 91–103.

[68] T.E. Keenan, K.P. Burke, E.M. Van Allen, Genomic correlates of response to immune checkpoint blockade, Nat. Med. 25 (3) (2019) 389–402.

[69] S. Wang, et al., Antigen presentation and tumor immunogenicity in cancer immunotherapy response prediction, Elife 8 (2019).

[70] A.C. Bretz, et al., Domatinostat favors the immunotherapy response by modulating the tumor immune microenvironment (TIME), J Immunother Cancer 7 (1) (2019) 294.

[71] M. Pallocca, et al., Combinations of immuno-checkpoint inhibitors predictive biomarkers only marginally improve their individual accuracy, J. Transl. Med. 17 (1) (2019) 131.

[72] A.S. Kodous, M. Balaiah, P. Ramanathan, Single cell RNA sequencing – a valuable tool for cancer immunotherapy: a mini review, Oncologie 25 (6) (2023) 635–639.

[73] A. Shostak, Circadian clock, cell division, and cancer: from molecules to organism, Int. J. Mol. Sci. 18 (4) (2017).