

RESEARCH

Open Access



# Comprehensive bioinformatics analysis of human cytomegalovirus pathway genes in pan-cancer

Tengyue Yan<sup>1†</sup>, Xianwu Pang<sup>2†</sup>, Boying Liang<sup>3†</sup>, Qiuxia Meng<sup>4</sup>, Huilin Wei<sup>5</sup>, Wen Li<sup>6</sup>, Dahai Liu<sup>7\*</sup> and Yanling Hu<sup>1,5\*</sup>

## Abstract

**Background** Human cytomegalovirus (HCMV) is a herpesvirus that can infect various cell types and modulate host gene expression and immune response. It has been associated with the pathogenesis of various cancers, but its molecular mechanisms remain elusive.

**Methods** We comprehensively analyzed the expression of HCMV pathway genes across 26 cancer types using the Cancer Genome Atlas (TCGA) and The Genotype-Tissue Expression (GTEx) databases. We also used bioinformatics tools to study immune invasion and tumor microenvironment in pan-cancer. Cox regression and machine learning were used to analyze prognostic genes and their relationship with drug sensitivity.

**Results** We found that HCMV pathway genes are widely expressed in various cancers. Immune infiltration and the tumor microenvironment revealed that HCMV is involved in complex immune processes. We obtained prognostic genes for 25 cancers and significantly found 23 key genes in the HCMV pathway, which are significantly enriched in cellular chemotaxis and synaptic function and may be involved in disease progression. Notably, CaM family genes were up-regulated and AC family genes were down-regulated in most tumors. These hub genes correlate with sensitivity or resistance to various drugs, suggesting their potential as therapeutic targets.

**Conclusions** Our study has revealed the role of the HCMV pathway in various cancers and provided insights into its molecular mechanism and therapeutic significance. It is worth noting that the key genes of the HCMV pathway may open up new doors for cancer prevention and treatment.

**Keywords** Human cytomegalovirus pathway, Pan-cancer, Immune infiltration, Tumor mutation burden, Bioinformatics

<sup>†</sup>Tengyue Yan, Xianwu Pang and Boying Liang contributed equally to this work.

\*Correspondence:

Dahai Liu  
seansean2014@126.com  
Yanling Hu  
huyanling@gxmu.edu.cn

<sup>1</sup>Collaborative Innovation Centre of Regenerative Medicine and Medical Bioreource Development and Application Co-constructed by the Province and Ministry, Guangxi Medical University, Nanning, Guangxi 530021, China

<sup>2</sup>Guangxi Zhuang Autonomous Region Center for Disease Control and Prevention, Nanning 530028, China

<sup>3</sup>Department of Immunology, School of Basic Medical Sciences, Guangxi Medical University, Nanning, China

<sup>4</sup>School of Information and Management, Guangxi Medical University, Nanning, China

<sup>5</sup>School of Institute of Life Sciences, Guangxi Medical University, Nanning, China

<sup>6</sup>Department of Biochemistry and Molecular Biology, School of Basic Medicine, Guangxi Medical University, Nanning, China

<sup>7</sup>School of Medicine, Foshan University, Foshan, Guangdong 528000, People's Republic of China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Cancer, the predominant cause of mortality worldwide, accounts for approximately 19 million new diagnoses and close to 10 million fatalities annually [1]. Intriguingly, a noticeable shift towards a younger demographic in cancer incidence has been observed in recent years [2]. Confronted with the relentless escalation in both cancer incidence and mortality, mitigating this burden has emerged as a cardinal objective within the global public health sphere.

Human Cytomegalovirus (HCMV), a  $\beta$ -herpesvirus ubiquitous in the human population, exhibits a high infection rate of up to 90% [3, 4]. HCMV infection is prevalent worldwide, with susceptibility patterns differing between developing and developed countries. In developing nations, young people are more susceptible to HCMV, whereas in developed countries, the elderly are at a higher risk of infection [5]. Its main transmission occurs through saliva, semen, urine, breastfeeding, placental transfer, blood and organ transplantation [6, 7]. Characterized by a genome of 235 kbp, it stands out as the largest among all herpes viruses, featuring a double-stranded DNA structure, and encode at least 700 open reading frames [8, 9]. Most of the genetic products are closely related to HCMV infection and prevalence [4].

The majority of people infected with primary HCMV are asymptomatic, and only a small number of people will develop mononucleosis syndrome after infection, with fever, sweating, abnormal liver function and discomfort and other symptoms [10, 11]. The natural course of HCMV infection is complex, with primary infection occurring when an immune compromised individual is first infected, followed by an incubation period. When the infected person is exposed to HCMV again, it will be repeated infection, which is called reinfection. HCMV persists throughout the host's lifetime, and most people carry it in a latent state, but there is also a risk of it being reactivated [12–14]. In addition, HCMV can also cause a variety of diseases such as systemic lupus erythematosus, systemic sclerosis, pneumonia, atherosclerosis, mental disorders and so on [15–19]. Compare this with the diseases mentioned above, as the apoptosis disorder induced by HCMV infection is more closely related to cancer and prompts biological responses that closely mimic those supporting chronic inflammation, leukocyte dysfunction, angiogenesis, and wound healing, potentially making it a promoter of a malignant tumor [4, 20, 21]. Numerous studies have found that HCMV may be associated with various types of cancer, such as hepatocellular carcinoma [22], breast cancer [23], gastric cancer [24], cervical cancer [25], colorectal cancer [26], ovarian cancer [27], prostate cancer [28], lymphoma [29], and glioblastoma [30]. Importantly, due to the host immune evasion mechanism of HCMV, it profoundly influences the development of

the tumor microenvironment. This influence not only alters the tumor microenvironment but is also associated with poor prognosis, metastasis, and drug resistance [14, 31]. However, the specific molecular mechanism of how HCMV is involved in the regulation of tumour development remains to be further elucidated.

In this study, we comprehensively analyzed the expression of HCMV pathway genes across cancers using the Cancer Genome Atlas (TCGA) and The Genotype-Tissue Expression (GTEx) databases. Importantly, we revealed the relationships between HCMV pathway gene expression and prognosis, genomic mutations, tumor microenvironment (TME), and chemo- and immunotherapy drug sensitivity. Additionally, through comprehensive bioinformatics analysis, we identified hub genes in the HCMV pathway that may play crucial roles in regulating cancer progression.

## Methods

### Data collection and screen

We obtained pan-cancer TCGA and GTEx transcriptomic expression data ( $\log_2(\text{TPM}+0.001)$ ), as well as clinical data, from the UCSC Xena database (<https://xena.ucsc.edu/>). Additionally, we acquired pan-cancer mutation data from the TCGA database (<https://portal.gdc.cancer.gov/>). Subsequently, we procured genes associated with the Human cytomegalovirus infection pathway from the KEGG database (<https://www.genome.jp/kegg/>). In clinical data, we first eliminated missing values and then matched the tumor samples accordingly. Subsequently, we extracted pathway genes and further analyzed the differential expression between various cancerous tumors and normal tissues using the Wilcoxon test. All analyses are based on R (version: 4.2.2). We established the following inclusion and exclusion criteria: the number of normal controls for each cancer must be greater than or equal to 5, and a P-value less than 0.05 was considered statistically significant.

### Differential expression analysis

Limma-Voom: the model consistently performed well on various benchmark datasets, providing a reasonable balance between error discovery (FDR) and recall rates, which is comparable to or better than that of count-based RNA-seq methods [32, 33]. Consequently, we employed this method to further compare the gene expression levels in normal and cancer tissues. We set the threshold values at  $\text{adjust P Value} < 0.05$  and  $|\log_{2}\text{FC}| \geq 1$ , thereby identifying genes with significant differential expression.

### Immune cell infiltration

For the analysis of immune infiltration, we employed the CIBERSORT R package, setting perm to 1000. This tool, extensively utilized for investigating the proportions of

22 subtypes of human immune cells, leverages a machine learning technique known as Support Vector Regression (SVR) to enhance deconvolution performance through an amalgamation of feature selection and robust mathematical optimization techniques, notably outperforming other methods in terms of accuracy when it comes to resolving closely related cell subsets and mixtures with unknown cell types [34].

#### Tumor microenvironment

ESTIMATE: A methodology that utilizes gene expression features to infer the proportion of stromal and immune cells in tumor samples [35]. We employed the “ESTIMATER” R package to calculate the stromal score and immune score in each cancer, further estimating the tumor purity for each type of cancer (platform select “Affymetrix”).

#### Prognostic related gene screening

A univariate Cox analysis was performed on the differential genes of each cancer. To prevent the loss of some important genes, genes with a P-value less than 0.2 were screened. To minimize overfitting risk, the Lasso regression and XGBoost were then used for the feature selection. We performed feature selection and shrinkage using the “glmnet” R package. The penalty parameter ( $\lambda$ ) of the model was determined by conducting ten-fold cross-validation and following the minimum criterion, which corresponds to the  $\lambda$  value of the minimum likelihood deviation [36]. Meanwhile using the “XGBoost” R package, we set ‘survival: cox’ as the objective function and ‘cox-nloglik’ as the evaluation metric, trained the model for 100 rounds at a 0.1 learning rate with the L1 regularization hyperparameter alpha set to 0.5 to prevent overfitting. Finally, a multivariate Cox analysis was performed, selecting genes with  $P < 0.05$  as prognostic genes [37]. Patients were divided into high-risk and low-risk groups based on the median of the risk scores, followed by survival analysis using the “survminer” package in R. The “timeROC” package in R was utilized to conduct time-dependent receiver operating characteristic (ROC) curve analysis to evaluate the prognostic predictive performance of the gene features for prognosis over a 3-year period [38].

#### Tumor mutation burden

Tumor Mutation Burden (TMB), characterized as the aggregate count of somatic coding inaccuracies, base substitutions, and insertion or deletion mutations identified per million bases of DNA, serves as an effective estimator for both mutational and neoantigen loads [39]. The “maftools” R-package integrates standard analysis and visualisation modules into a single pipeline by implementing well-established statistical and computational

methods, requiring only a single and uniform input data format for the process from analysis to visualisation to annotation [40]. We further used this R package to calculate the mutations in prognosis-related genes for each cancer.

#### Screening for hub genes

The gene interaction network was established utilizing the STRING database (<https://string-db.org/>). Choose the highest confidence (0.900) as the minimum required interaction score. Further use Cytoscape software to visualize the connections between genes. For further study, after protein-protein network analysis, the MCODE plug-in was performed to screen hub genes, whose degree cutoff was =2, node score cutoff was =0.2, k core was =2, and maximum depth was =100 [41].

#### GO enrichment analysis of hub gene

To ascertain the functions of the identified hub genes, a Gene Ontology (GO) analysis was conducted, encompassing enrichment of Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) [42]. Further analysis was conducted using the “clusterProfiler”, “enrichplot” and “Goplot” R packages, with a p.adjust value of less than 0.05 set as the critical threshold.

#### Drug sensitivity analysis

The Genomics of Drug Sensitivity in Cancer (GDSC) database, as the largest public resource for information on drug sensitivity in cancer cells and molecular markers of drug response, amalgamates extensive drug sensitivity and genomic datasets to expedite the discovery of novel therapeutic biomarkers for cancer treatment [43]. And subsequently, we utilized GSCA (<http://bioinfo.life.hust.edu.cn/GSCA/#/>) [44], a tool incorporating data from the GDSC database, to investigate the association between hub genes and drug sensitivity.

## Results

#### Data set screening and preliminary analysis

We downloaded samples of 33 types of cancer from public database. After screening, we obtained a total of 18,002 samples, including 8,278 normal samples and 9,724 tumor samples. The sample information for each cancer type is shown in Fig. 1A. Using the Wilcoxon test, we excluded 7 types of cancer. Ultimately, we identified 26 types of cancer for further analysis. It’s noteworthy that gene expression in tumor tissues surpasses that in normal tissues for most cancers, as depicted in Fig. 1B.

#### Differential gene recognition

Differential expression analysis identified a total of 1,853 differentially expressed genes, including 506 up-regulated and 1,347 down-regulated genes. The top 5



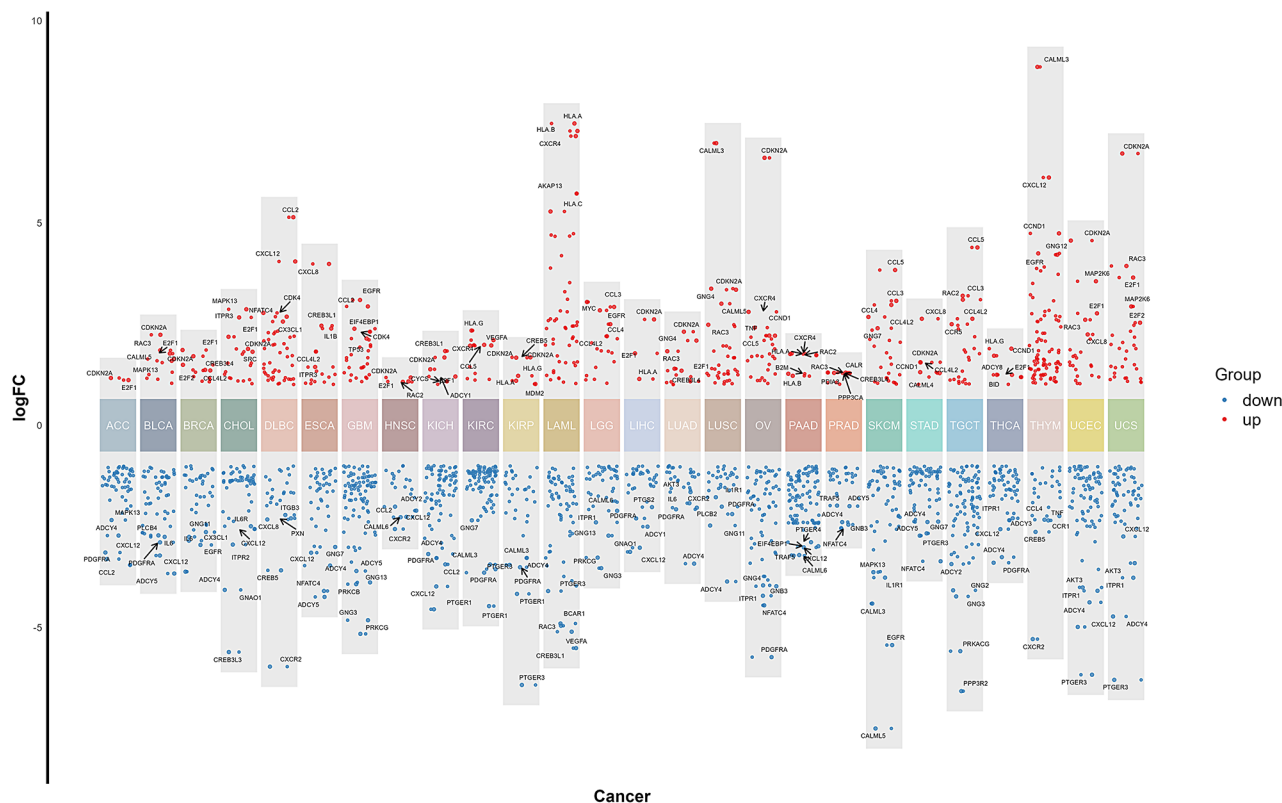
**Fig. 1** Overview of Sample Composition and Gene Expression Differences in Various Cancers. **(A)** The distribution of normal and tumor tissue samples across different cancer types: The horizontal axis represents different types of cancer, and samples are distinguished by color: blue for normal tissue and red for tumor tissue. For each cancer type, the counts of normal samples are indicated at the top, while those of tumor samples are indicated at the bottom. **(B)** The differences between normal and tumor samples across various cancers: Normal and tumor samples are depicted in blue and red, respectively. The horizontal axis represents different types of cancer, while the vertical axis indicates the level of gene expression. The symbols “\*\*\*\*”, “\*\*\*”, “\*\*”, “\*”, and “ns” correspond to  $p < 0.05$ ,  $p < 0.01$ ,  $p < 0.001$ ,  $p < 0.0001$ , and non-significance, respectively. The lack of any markers signifies the absence of normal controls for that type of cancer (MESO, UVM)

up- and down-regulated genes across the cancer samples were displayed (Fig. 2). The complete list of differentially expressed genes for each cancer sample is provided in Supplementary Table S1. Additionally, we have generated clustering heatmaps for each type of cancer to visually illustrate the differences in gene expression. The corresponding images can be found in Figure S1.

In addition, we analyzed the significant functional differences and commonalities of differential genes across various cancer types using GO enrichment analysis (Figure S2). For example, in adrenocortical carcinoma (ACC), genes primarily function in response to peptide hormones, highlighting their importance in regulating hormone balance and cell signaling. In cholangiocarcinoma (CHOL), genes play a significant role in peptidyl-serine phosphorylation, indicating their critical role in protein

modification and signal transduction. In liver hepatocellular carcinoma (LIHC), genes are crucial in responding to tumor necrosis factor, underscoring their involvement in inflammatory responses and immune regulation. In breast invasive carcinoma (BRCA), genes are particularly active in muscle cell proliferation, indicating their role in cell growth and tissue development... Notably, differential genes across multiple cancer types are involved in leukocyte migration, cellular response to abiotic stimuli, and regulation of vasculature development. This suggests that these genes play universally important roles in regulating immune responses, adapting to environmental changes, and promoting angiogenesis.





**Fig. 2** The differential gene expression across various cancer types (from left to right ACC, BLCA, BRCA, CHOL, DLBC, ESCA, GBM, HNSC, KICH, KIRC, KIRP, LAML, LGG, LIHC, LUAD, LUSC, OV, PAAD, PRAD, SKCM, STAD, TGCT, THCA, THYM, UCEC, UCS). The y axis shows the fold changes in gene expression. Gene expression levels relative to different color blocks are displayed in varying colors, with red representing up-regulated genes and blue representing down-regulated genes

**Differences in immune cell infiltration in different cancer samples**

An in-depth analysis of immune cell infiltration in a variety of cancer samples was performed using CIBERSORT, after which immune subpopulations with P-values less than 0.05 were screened out and correlation heat maps were used to show the relationship between differential genes and these subpopulations (Figure S3), the proportion of cell subpopulations in each cancer is shown in Figure S4. To further explore the differences in immune cells in tumor and normal samples, we used the Wilcoxon rank sum test for comparison. The results showed that immune infiltrating cells in the pathway mainly included T cells, macrophages, NK cells, B cells, dendritic cells, plasma cells, monocytes, and neutrophils. To visualize these findings more intuitively, we visualized the results using the “networkD3” R package (Fig. 3). These data provide us with strong evidence that these immune cell subpopulations may play a key role in the pathogenesis of various cancers.

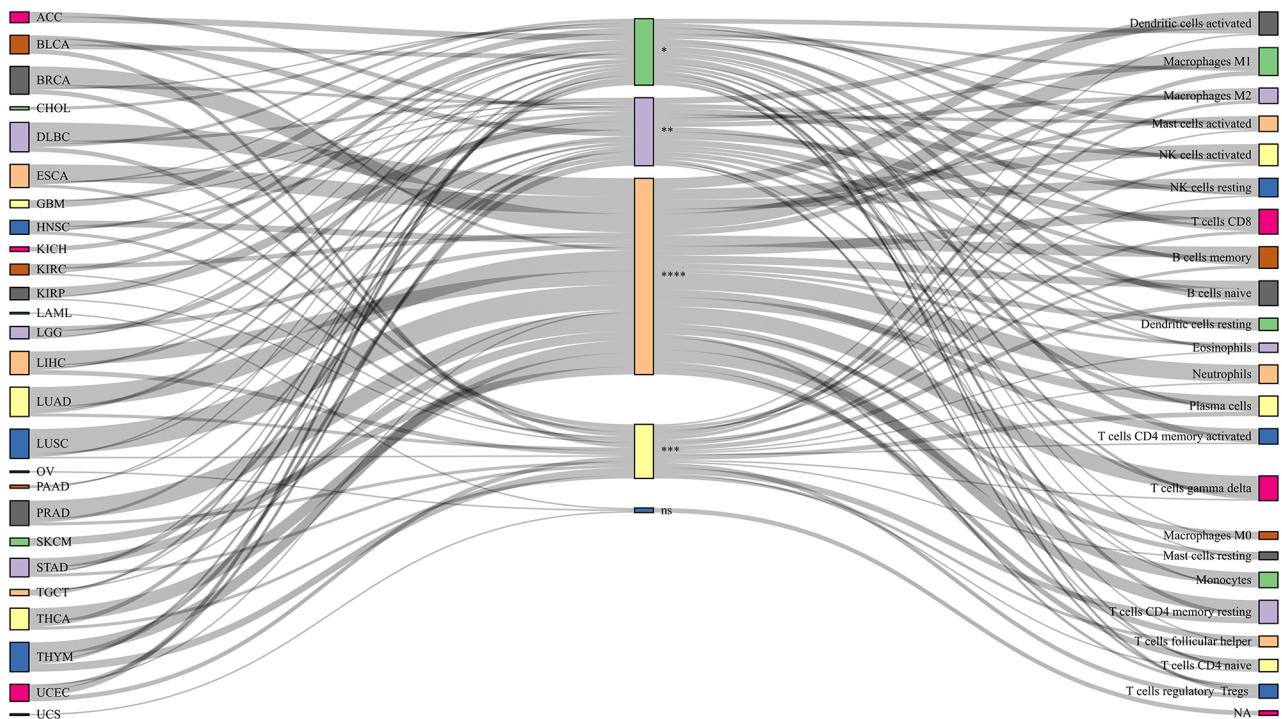
**Tumour microenvironment score**

Using the ESTIMATE algorithm, we analyzed 26 cancer types and obtained stromal scores, immune scores and

estimate scores for each cancer except LGG, which only had stromal scores available. The estimate score is calculated as the sum of stromal and immune scores [35]. As shown in Fig. 4, immune scores were increased while stromal and estimate scores were decreased across most cancer cohorts. Additionally Stromal scores, immune scores and estimate scores were all negatively correlated with tumor purity, with the estimate score showing the strongest negative correlation (Figure S5). This implies that immune cells may play an increasingly important role within the tumor microenvironment of these cancers.

**Prognostic related gene screening**

After screening 26 cancer-related genes by univariate analysis, we further screened them by LASSO regression method and successfully obtained 20 prognostic genes associated with cancer. For the six cancers (LUSC, DLBC, TGCT, THYM, UCEC, UCS) that failed the LASSO screening, we re-screened them using the XGboost method, and only TGCT was excluded (P>0.05), and finally we obtained prognostic-associated genes for these cancers (Table S2). We showed some of the genes with prognostic value for each cancer (Fig. 5A). The survival



**Fig. 3** Immune cell infiltration in various types of cancer: The squares on the left represent different cancer types, the middle squares represent salience scores, and the squares on the right represent distinct immune cell populations. The symbols “\*”, “\*\*”, “\*\*\*\*”, “\*\*\*\*\*” and “ns” correspond to  $p < 0.05$ ,  $p < 0.01$ ,  $p < 0.001$ ,  $p < 0.0001$ , and non-significance

curves of different cancers with high and low prognostic risks were clearly separated, and the prognostic assessment models constructed using these genes could effectively distinguish the prognostic status of different cancer samples (Fig. 5B–Z). These models were validated by ROC curves and showed reliable predictive efficacy (Figure S6). To further confirm the importance of these genes in the development of cancer and their clinical significance regarding expression levels, we utilized data from The Cancer Cell Line Encyclopedia (CCLE) project for external validation. The CCLE is a comprehensive resource that includes genomic and drug response data from over 1000 cancer cell lines of various tissue types, offering a valuable asset for cancer biology and precision medicine research [45]. Our analysis revealed that most genes exhibit a Chronos dependency score of less than zero, suggesting their potential role in cancer progression [46]. Furthermore, the majority of these genes show high expression levels in cancer cells (Fig. 6), suggesting their significance in specific cellular contexts.

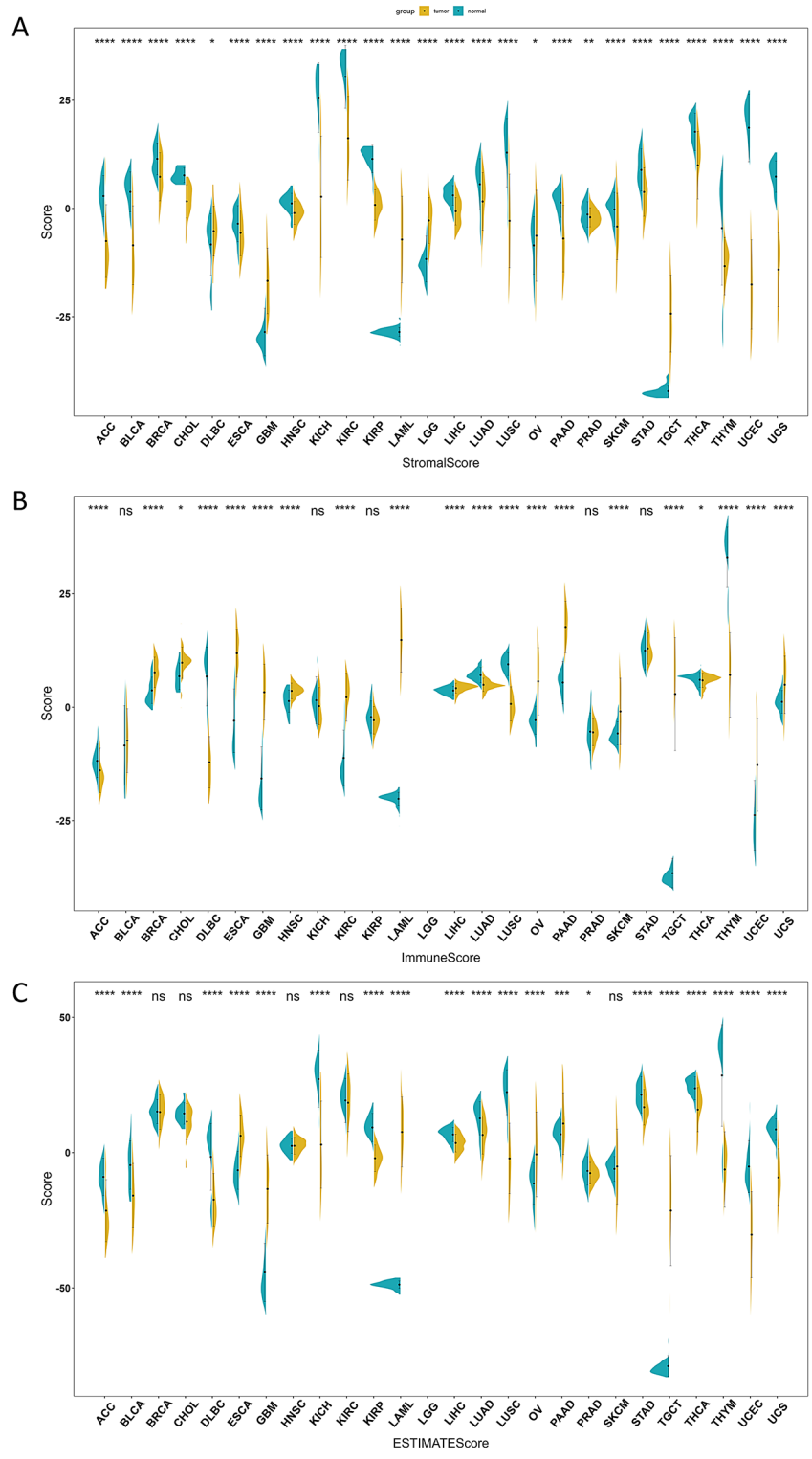
#### Tumor mutation burden

Utilizing “maftools” R package, we systematically analyzed the mutation profiles of various cancer prognostic genes. Of the 131 prognostic-related genes examined, mutations were identified in 72 (54.96%), whereas the other 59 (45.04%) did not harbor any

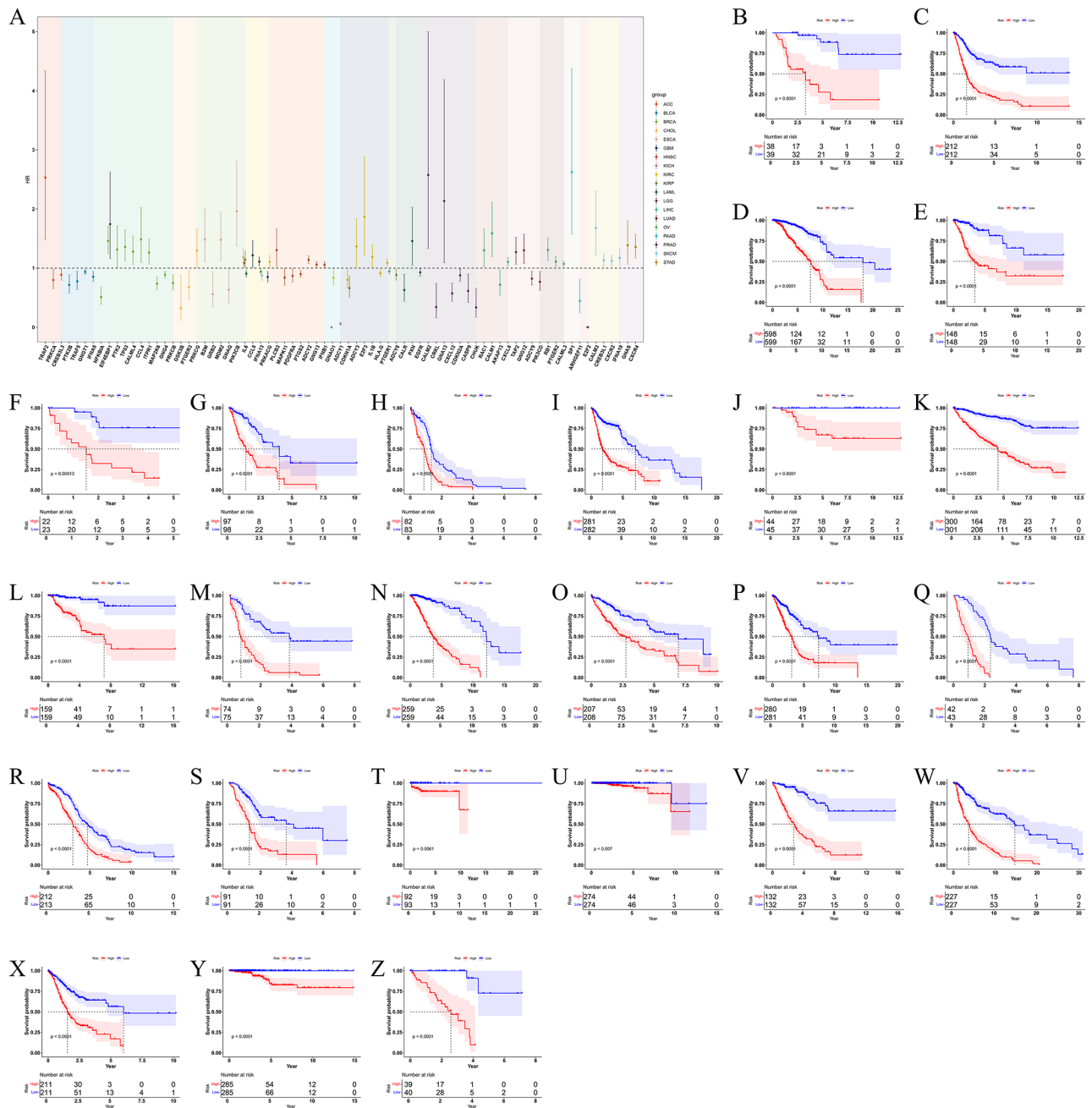
mutations (Fig. 7A–Y). Further mutation classification of these 72 mutated genes revealed a predominance of missense mutations in the majority of them. The detailed mutation spectra for each individual gene are illustrated in Figure S7.

#### Hub gene and its functional enrichment analysis

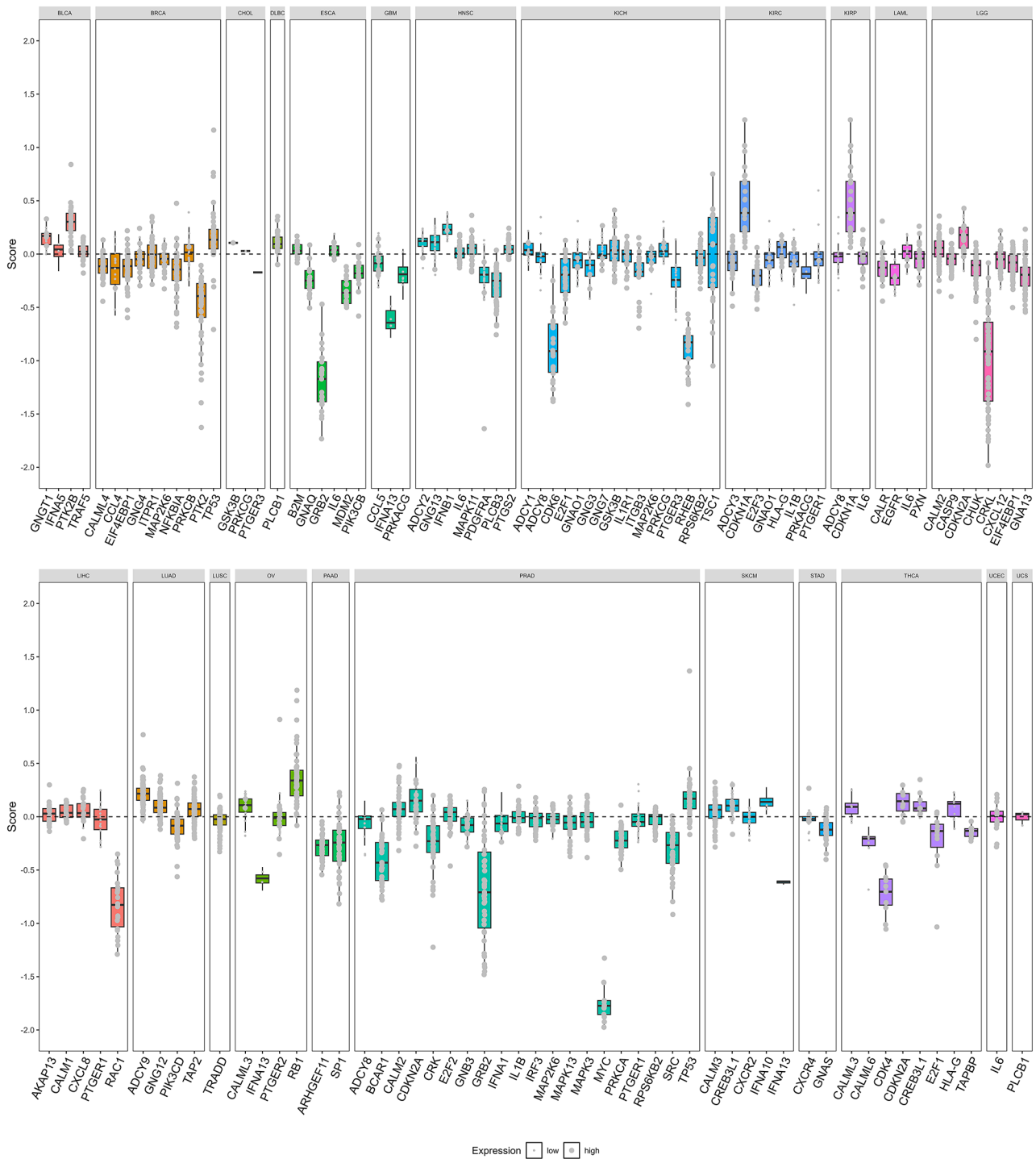
After constructing the Protein-Protein Interaction (PPI) network, MCODE analysis was further used to screen modules with scores  $\geq 6$ , highlighting the most interconnected hub genes implicated in the HCMV pathway. A total of 23 hub genes were obtained, including CALML4, GRB2, CALML3 etc. The hub gene interaction network is shown in Fig. 8A–B. GO enrichment analysis revealed that these hub genes were significantly enriched in 803 BP, 47 CC and 97 MF. The enriched BP were primarily related to cell chemotaxis and response to peptides; the CC were mostly membrane-bound and synapse-related; the MF included receptor ligand binding activities and enzyme activities (Fig. 8C), underscore the multifaceted biological implications of these genes. In addition, we visualized key genes in the HCMV pathway using the ‘pathview’ R package [47]. The results showed that most of the genes had protein-molecular interactions in the pathway. Specifically, the expression of CaM (CALML4, CALM3, CALML6) was promoted, while AC (ADCY8, ADCY3) was inhibited in the pathway (Fig. 8D), and



**Fig. 4** Assessing the tumor microenvironment via ESTIMATE algorithm: The horizontal axis represents different cancer types, and the vertical axis represents the scores. Normal samples are represented in blue, while tumor samples are indicated in yellow. The symbols “\*\*\*\*”, “\*\*\*”, “\*\*”, “\*”, and “ns” correspond to  $p < 0.05$ ,  $p < 0.01$ ,  $p < 0.001$ ,  $p < 0.0001$ , and non-significance. **(A)** Distribution of stromal scores across various cancers assessed by ESTIMATE approach. **(B)** Distribution of immune scores across multiple cancer types calculated by ESTIMATE algorithm. **(C)** Estimate scores reflecting tumor purity determined by ESTIMATE method

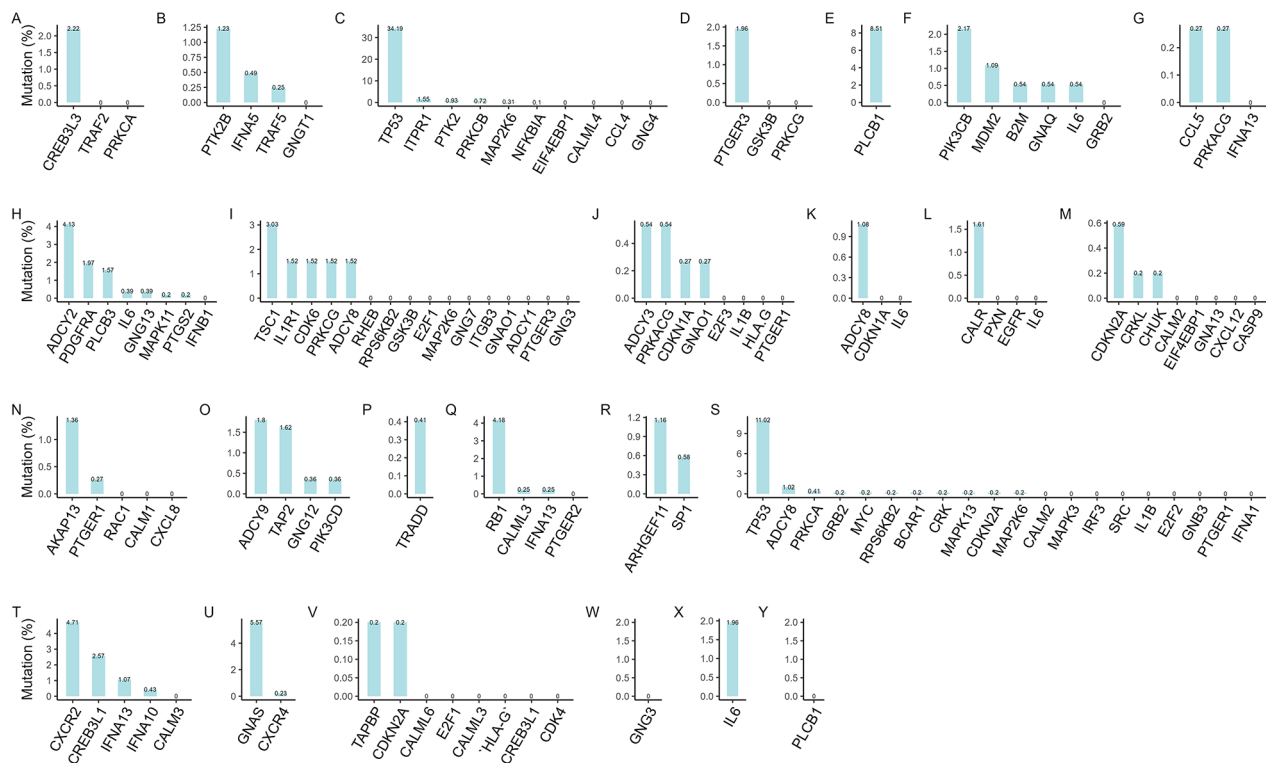


**Fig. 5** Identification and validation of prognostic gene signatures for cancers. (A) Forest plots of partial prognostic genes for each cancer: Different colors represent different cancer types, the horizontal axis represents different genes, the vertical axis represents the hazard ratio (HR) (dashed line represents HR=1) and the line segments in the figure represent confidence intervals (CI). Kaplan-Meier curves compared the prognostic situations of prognostic genes in (B) ACC, (C) BLCA, (D) BRCA, (E) CHOL, (F) DLBC, (G) ESCA, (H) GBM, (I) HNSC, (J) KICH, (K) KIRC, (L) KIRP, (M) LAML, (N) LGG, (O) LIHC, (P) LUAD, (Q) LUSC, (R) OV, (S) PAAD, (T) PRAD, (U) SKCM, (V) STAD, (W) THCA, (X) THYM, (Y) UCEC, (Z) UCS. Red lines represented high-risk group, blue represented low-risk group



**Fig. 6** Validation of pan-cancer prognostic genes in CCLE: Validation of pan-cancer prognostic genes in CCLE: Each panel represents a different cancer type, with the x-axis denoting different genes, and the y-axis representing The Chronos dependency score. The size of the points within the boxplot corresponds to the gene's expression level in cells ( $\log_2(\text{TPM} + 1)$ )





**Fig. 7** Mutations of different cancer prognostic genes. Prognostic gene mutation details in (A) ACC, (B) BLCA, (C) BRCA, (D) CHOL, (E) DLBC, (F) ESCA, (G) GBM, (H) HNSC, (I) KICH, (J) KIRC, (K) KIRP, (L) LAML, (M) LGG, (N) LIHC, (O) LUAD, (P) LUSC, (Q) OV, (R) PAAD, (S) PRAD, (T) SKCM, (U) STAD, (V) THCA, (W) THYM, (X) UCEC, (Y) UCS: The horizontal axis represents genes, and the vertical axis represents mutation rates

potentially impacting the HCMV's pathophysiological mechanisms.

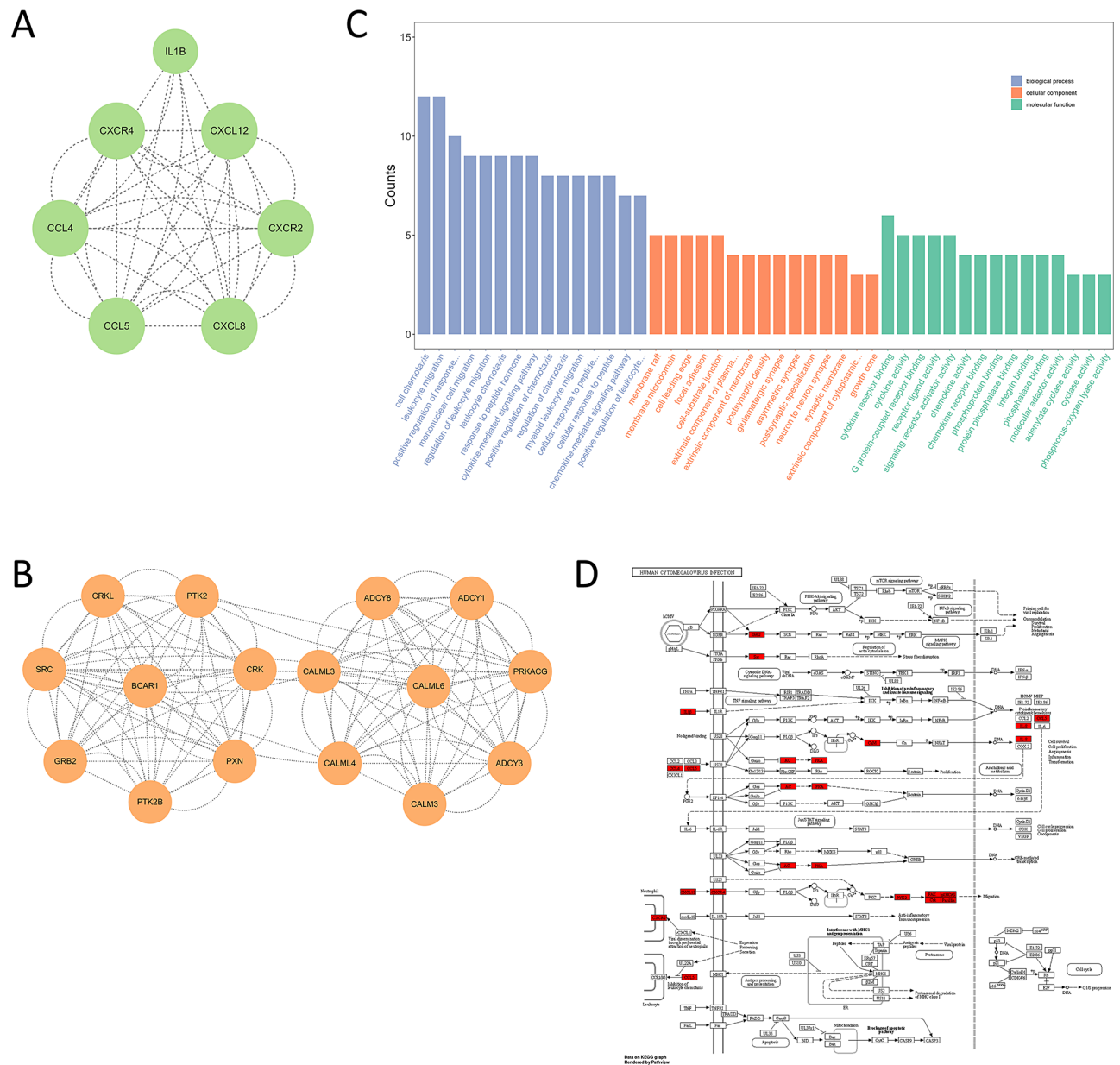
#### Analysis of the correlation between hub gene expression levels and drug sensitivity

We investigated the correlation between the expression of HCMV pathway hub genes and drug sensitivity profiles in the GDSC database. The results showed that the expression levels of ADCY1, PTK2, CRK, SRC, PXN, BRCA1, CXCL8 are mostly positively correlated with IPA-3, STF-62,247, THZ-2-49, AP-24,534, KIN001-236, QL-X-138, Y-39,983, ZSTK474, AZD8055, BX02189, CAL-101, CP466722, GSK2126458, JW-7-24-1, KIN001-244, PHA-793,887, PIK-93, SNX-2112, TG101348, AZD7762, NG-25, OSI-930, QL-XI-92, TPCA-1, Navitoclax and AICAR, these findings suggest that elevated expression of these genes may serve as a biomarker for predicting response to specific drugs, potentially guiding more tailored treatment strategies. The expression levels of ADCY1, PTK2, CRK, SRC, PXN, BRCA1, and CXCL8 were mostly negatively correlated with sensitivity to bleomycin (50  $\mu$ M), afatinib, cerutinib, and gefitinib. By contrast, the expression levels of CALML4, CALML6, CALM3, GRB2, CCR4, RTK2B, CXCR2, CCL5 to the above mentioned is the opposite (Fig. 9A), these may

provide new insights into developing combination therapies or alternative treatments to overcome drug resistance. To further validate these findings, we performed additional analyses using the Connectivity Map (CMap), which helped identify the targets of drug action [Fig. 9B]. For instance, the mechanism of action of PD-0325901 is highly similar to that of AP-24,534, AZD8055, and TG101348, as listed in the GDSC database, and all belong to kinase inhibitors. Of note, the results from CMap suggest that drugs may exert their effects by inhibiting or downregulating these target points. This not only strengthens the reliability of our results but also further confirms the potential application of these HCMV pathway hub genes in predicting drug response, demonstrating their significant value in drug response prediction.

#### Discussion

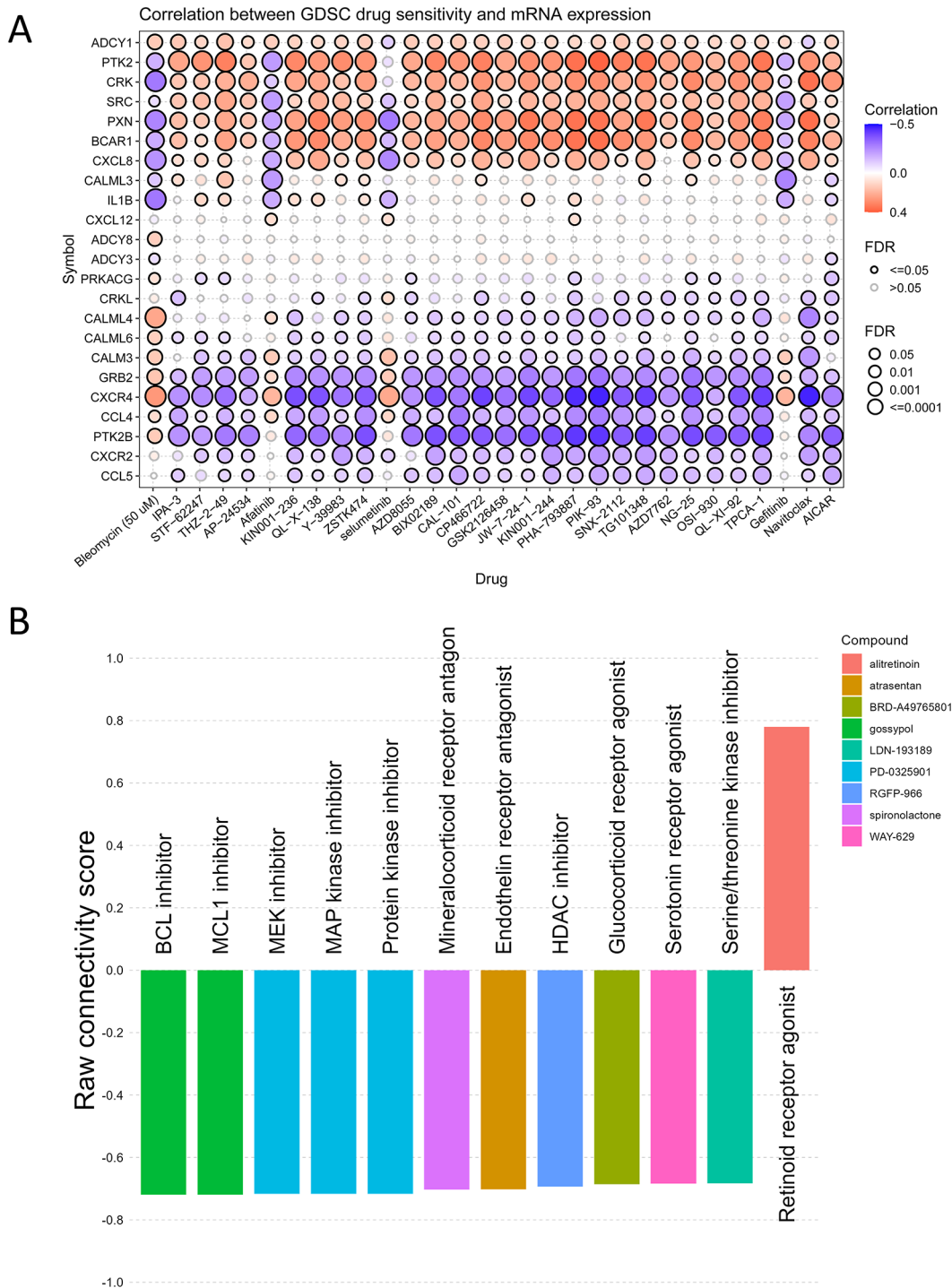
In this study, we comprehensively analysed the role of the human cell cytomegalovirus (HCMV) pathway in the development of a variety of cancers by means of bioinformatics methods. We found that the HCMV pathway has a wide range of differentially expressed genes in various cancer samples. Immune infiltration analysis demonstrated that immune cells such as T cells and macrophages play important roles in the tumor



**Fig. 8** Integrated Analyses of HCMV Pathway and Hub Genes. **(A, B)** Two significant subnetworks of PPI network. **(C)** GO enrichment analysis of Hub genes: The horizontal axis represents the potential functions enriched, while the vertical axis indicates the number of enrichments (from left to right, they are Biological Process (BP) in purple, Cellular Component (CC) in orange, and Molecular Function (MF) in green). **(D)** Pathway localization of hub genes in HCMV pathway

microenvironment, providing evidence that HCMV is involved in the regulation of tumorigenesis. Furthermore, 23 pathway key genes were identified, which provide potential targets for the development of HCMV-associated targeted therapies. Overall, our study comprehensively resolved the role of HCMV pathway in the development of various cancers, provided new insights into the relationship between HCMV and tumours, and provided a theoretical basis for the treatment of HCMV-associated tumours.

HCMV invasion induces multiple immune responses. Endothelial cells, dendritic cells, natural killer cells, monocytes, and macrophages become activated in the blood and tissues, producing abundant inflammatory mediators including IL-1 $\beta$ , IL-2, IL-6, TNF- $\alpha$ , and chemokines [48–51]. This process prompts circulating neutrophils to receive signals from antigen-presenting cells at the site of infection and prepare for a further immune response. Concurrently, the activation of local clotting factors also induces platelet aggregation, further



**Fig. 9** Integration of GDSC Analysis and CMap Validation Reveals Potential Drug Targets Associated with Hub Genes. **(A)** Gene expression-drug sensitivity correlations: Red indicates positive correlation and blue indicates negative correlation. Bubble size positively corresponds to FDR significance, with black outline highlighting correlations meeting FDR  $\leq 0.05$  threshold. **(B)** CMap validation identifies potential drug targets: The columns of various colors represent the actions of different compounds, with corresponding labels denoting the potential targets of action. The vertical axis depicts the connectivity score

enhancing inflammatory processes [52, 53]. Our analysis revealed that differential gene functions across various cancer types, including leukocyte migration, response to abiotic stimulus, and vasculature development regulation, align with the immune responses seen in HCMV invasion. These shared pathways highlight common mechanisms of immune regulation and inflammation critical to both viral infections and cancer progression. Given these findings, it is crucial to explore additional genetic regulatory layers that may influence these processes. DNA methylation, one of the most extensively studied epigenetic modifications, plays a critical role in essential biological processes such as embryonic development, genomic imprinting, and X-chromosome inactivation [54]. Aberrant DNA methylation can alter the cellular microenvironment, affect gene expression patterns, and lead to various pathological conditions, including cancer [55]. Using the MethSurv tool, we explored individual methylation CpG sites in differentially expressed cancer genes [56–58]. For instance, in adrenocortical carcinoma (ACC), we found a significant association between the methylation site cg04721825 in the PRKCA gene (HR: 6.626, 95% CI: 1.561–28.119, P: 0.010349507) and disease risk. Similarly, in bladder cancer (BLCA), the methylation site cg09825327 in the TRAF5 gene (HR: 1.682, 95% CI: 1.136–2.49, P: 0.009355414) was significantly associated with disease risk... These results suggest that variations in methylation sites may play a critical role in cancer development, highlighting an important direction for future research. However, the relationship between HCMV and the host immune system is complex, and once HCMV enters the body, it is able to establish latency in undifferentiated hematopoietic progenitor cells in the bone marrow and then reactivate, leading to a recurrence of the disease [51, 59, 60]. Latently infected monocytes disseminate the virus to various organs, and upon their reactivation, the adaptive immune system activates T helper cells and cytotoxic T cells via T cell receptors while accelerating epigenetic events. This further exacerbates the inflammatory role in pathophysiology, which may precipitate multi-organ diseases, heighten disease risk, and even cause death in severe cases [61–64].

Using the bioinformatics tools CIBERSORT and ESTIMATE, we analyzed the immune cell composition and scoring of 26 cancer types. While the degree and proportion of immune cell infiltration varied across cancers, the infiltration patterns were similar to those of cells involved in innate and adaptive immunity. Fractalkine/CX3C chemokine ligand 1 (CX3CL1) and its receptor CX3CR1 have been found to allow immature dendritic cells to migrate to cancer cells using the expression of their receptor CX3CL1 [65, 66]. The CX3CL1-CX3CR1 axis promotes NK cells to adhere to tumor cells and directly kill cancer cells [67, 68]. Cancers associated with

the CX3CL1-CX3CR1 axis include BLCA, STAD, BRCA, GBM, LGG, LIHC, LUAD, LUSC, PAAD, KICH, KIRC, KIRP, OV, PAAD, HNSC, PRAD, TGCT, UCEC, SKCM, etc [69]. . . Results from ESTIMATE revealed high tumor purity across most cancer cohorts, implying a relatively high proportion of tumor cells and low immune activity in the tumor samples. The tumor immune microenvironment represents a highly complex system that is pivotal in driving immunosuppression, distant metastasis, local drug resistance, and response to targeted therapies [70–73]. Moreover, it is closely related to the clinical prognosis of tumor patients [74]. The tumor microenvironment contains multiple immunosuppressive cell types that are induced by cancer-associated fibroblasts, such as M2 macrophages, regulatory T cells and myeloid-derived suppressor cells. These immunosuppressive cells accumulate abundantly within the tumor immune microenvironment (TIME) and play critical roles in promoting immune evasion and suppression [75–77]. , with which our results are also consistent, supporting the hypothesis that enriched immune cells reduce the proportion of tumour cells by enhancing their killing effect on tumour cells and inhibiting their proliferation. We comprehensively analysed the immune infiltration of pan-cancer, which is important for understanding the immune characteristics of tumours and developing corresponding immunotherapy strategies. Future studies can further delve into the interactions of different immune cells in the tumour immune microenvironment to more comprehensively resolve the complexity of the tumour immunoregulatory network.

We identified 25 cancer prognostic genes that are significantly associated with survival (ACC, BLCA, BRCA, CHOL, DLBC, ESCA, GBM, HNSC, KICH, KIRC, KIRP, LAML, LGG, LIHC, LUAD, LUSC, OV, PAAD, PRAD, SKCM, STAD, THCA, THYM, UCEC, UCS), these predictive models of gene composition can effectively distinguish between high and low risk groups of cancer patients. The use of large-scale public databases for biological information mining allows us to efficiently discover these potential cancer prognostic genes [78]. Our findings provide important molecular markers for prognostic assessment and risk stratification of cancer. The study by Anuraga et al. (2021), leveraging multiple database resources such as The Cancer Cell Line Encyclopedia (CCLE) and the Tumor Immune Estimation Resource (TIMER), has successfully identified prognostic biomarkers for breast cancer, contributing valuable insights to the field. Their work not only underscores the efficacy of CCLE in cancer genomics research but also inspired the application of this methodology to pan-cancer analysis, further exploring the commonalities and specificities of molecular characteristics across various cancers [79, 80]. Notably, individual cancer types, such as TGCT, did not



show significant prognostic genes in our model, which may be related to the small sample size. In follow-up studies, we will collect more samples to improve statistical efficiency and use experimental techniques to verify the expression and prognostic value of these genes in relevant cancer samples. In addition, we analyzed the mutation spectra of these prognostic genes, and found that about 55% of the genes had mutation information, and the proportion of missense mutations was relatively high. Studies have shown that tumor genomic characteristics, mutation load, and tumor-specific neoantigens are key factors in determining a patient's response to immune checkpoint blockers, and they may affect the patient's immunotherapy response. Moreover, tumor mutation load and its associated tumor-specific neoantigens appear to be key ways to predict the potential clinical efficacy of immune checkpoint blockers [81].

The present study identified 23 pivotal hub genes by constructing protein-protein interaction networks and implementing the MCODE algorithm, thus shedding light on the molecular mechanisms underlying the interplay between HCMV infection and host cells. These hub genes were implicated principally in cell chemotaxis and synaptic function modulation, intimating that viruses might harness such processes to facilitate dissemination and proliferation. Meanwhile, promotion of CaM family genes (CALM1, CALM2, CALM3, CALML3, CALML4, CALML5, CALML6) and inhibition of AC family genes (ADCY1, CDCY2, ADCY3, ADCY4, ADCCY5, ADCY6, ADCY7, ADCY8, ADCY9) were observed. These are prognosis-associated genes. Notably, while mutations prevail in AC family genes expressed in manifold cancers (e.g. HNSC, KICH, KIRC, KIRP, LUAD, PRAD), the mutation rates of CaM family genes were mostly 0% across certain expressed cancers (BRCA, LGG, LIHC, PRAD, SKCM, THCA). Additionally, we utilized the GDSC database to screen for drug sensitivities, which provides a comprehensive overview of the correlations between gene expression levels and various drug responses. To further validate these findings, we employed the Connectivity Map (CMap) database, a broad bioinformatics resource used to elucidate the connections between small molecules, biological processes, and disease states by comparing gene expression profiles to predict drug mechanisms, annotate genetic variations, and provide insights for clinical trials [82–84]. Interestingly, the results between CMap and GDSC were similar, and more importantly, CMap identified the targets of drug actions. By leveraging CMap, we were able to cross-validate our preliminary findings and gain deeper insights into drug-gene interactions, thus proffering novel theoretical foundations for therapeutic and preventive interventions against the diseases. However, inter-gene interactions are intricate and multifaceted. Whether the

pathway hub genes discovered in this study are the major factors influencing HCMV-induced disease progression remains to be further validated. Their purported capacities to exert pivotal roles in diverse HCMV-associated cancers also remain nebulous. Despite unveiling latent molecular mechanisms, inherent limitations exist as a purely in bioinformatical predictive study. Subsequent validation through experimental techniques is imperative to verify the expressional and functional alterations of these genes, alongside their precise roles during HCMV infection. Furthermore, the regulatory mechanisms of the hub genes and their downstream pathways, as well as the influences on viral infection and pathogenic mechanisms of diseases, warrant further in-depth research. Investigations on their consistencies of actions across the gamut of HCMV-induced cancers are necessitated, alongside assessments of their potentials as novel biomarkers or therapeutic targets for diseases. In summary, this study proffered valuable insights into HCMV and host interplay, and invoked multiple scientific issues for scholars across pertinent domains to pursue. We hope that future studies can further uncover the mystery of HCMV mechanism and provide theoretical basis for prevention and treatment of related diseases.

## Conclusions

Through a comprehensive and comprehensive analysis of the relationship between HCMV and pan-cancer, HCMV is significantly associated with many cancers and is involved in complex immune processes. We finally obtained prognostic genes and their mutations for 25 cancers, providing potential targets for clinical treatment. It is important to find the key genes of HCMV pathway species, which are closely related to cell chemotaxis and synaptic function, and in most tumors, CaM family genes are up-regulated and AC family genes are down-regulated, which may play an important role in the occurrence and development of cancer, and deserve further study.

## Abbreviations

HCMV	Human cytomegalovirus
TCGA	The Cancer Genome Atlas
GTEX	The Genotype-Tissue Expression
ACC	Adrenocortical carcinoma
BLCA	Bladder Urothelial Carcinoma
BRCA	Breast invasive carcinoma
CECSC	Cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL	Cholangiocarcinoma
COAD	Colon adenocarcinoma
DLBC	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma
ESCA	Esophageal carcinoma
GBM	Glioblastoma multiforme
HNSC	Head and Neck squamous cell carcinoma
KICH	Kidney Chromophobe
KIRC	Kidney renal clear cell carcinoma
KIRP	Kidney renal papillary cell carcinoma
LAML	Acute Myeloid Leukemia



LGG	Brain Lower Grade Glioma
LIHC	Liver hepatocellular carcinoma
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
MESO	Mesothelioma
OV	Ovarian serous cystadenocarcinoma
PAAD	Pancreatic adenocarcinoma
PCPG	Pheochromocytoma and Paraganglioma
PRAD	Prostate adenocarcinoma
READ	Rectum adenocarcinoma
SARC	Sarcoma
SKCM	Skin Cutaneous Melanoma
STAD	Stomach adenocarcinoma
TGCT	Testicular Germ Cell Tumors
THCA	Thyroid carcinoma
THYM	Thymoma
UCEC	Uterine Corpus Endometrial Carcinoma
UCS	Uterine Carcinosarcoma
UVM	Uveal Melanoma
CMap	Connectivity Map
HR	Hazard ratio
CCLC	The Cancer Cell Line Encyclopedia
SVR	Support Vector Regression
ROC	receiver operating characteristic
TMB	Tumor Mutation Burden
GO	Gene Ontology
CC	Cellular Component
MF	Molecular Function
GDSC	The Genomics of Drug Sensitivity in Cancer
TIME	The tumor immune microenvironment

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40246-024-00633-5>.

Supplementary Material 1  
 Supplementary Material 2  
 Supplementary Material 3  
 Supplementary Material 4  
 Supplementary Material 5  
 Supplementary Material 6  
 Supplementary Material 7  
 Supplementary Material 8  
 Supplementary Material 9

## Acknowledgements

Thanks to the Genome Center of Guangxi Medical University for your help.

## Author contributions

TY, XWP, and YLH conceived and designed the study. TY, XWP, WL and HLW collected and collated data and conducted the analysis. BYL, QXM, and DHL interpreted the data, visualized, and wrote the first draft of the manuscript. DHL and YLH reviewed and revised the manuscript and approved submission. TY and XWP contributed equally to this work and are considered co-first authors.

## Funding

This work was supported by the National Natural Science Foundation of China (No. 82160537 and 82270413), Guangxi Key Research and Development Project (Grant No. Guike AB22035027), National Key Research and Development Program of China (No. 2023YFC2605400), the Natural Science Foundation of Guangdong Province of China (Grant No. 2022A1515011368, 2023A1515011581), The Key Projects of Department of Education of Guangdong Province of China (Grant No. 2022ZDZX2057). Finally, the authors would like to

express their sincere gratitude to The Cancer Genome Atlas, The Genotype-Tissue Expression databases, and the University of California, Santa Cruz Xena browser development team.

## Data availability

No datasets were generated or analysed during the current study. The data that support the findings of this study are available from the UCSC Xena database (<https://xena.ucsc.edu/>), and The Cancer Genome Atlas database (<https://portal.gdc.cancer.gov/>).

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

All the authors declared that they had no competing interests.

Received: 23 December 2023 / Accepted: 5 June 2024

Published online: 17 June 2024

## References

- Sung H, Global Cancer Statistics. 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a cancer journal for clinicians*. 2021;71(3):209-249. [10.3322/caac.21660](https://doi.org/10.3322/caac.21660).
- Ugai T, et al. Is early-onset cancer an emerging global epidemic? Current evidence and future implications. *Nat Reviews Clin Oncol*. 2022;19(10):656-73. <https://doi.org/10.1038/s41571-022-00672-8>.
- Angelova M, et al. Human cytomegalovirus infection dysregulates the canonical Wnt/ $\beta$ -catenin signaling pathway. *PLoS Pathog*. 2012;8(10):e1002959. <https://doi.org/10.1371/journal.ppat.1002959>.
- Yu Z, et al. Apoptosis disorder, a Key Pathogenesis of HCMV-Related diseases. *Int J Mol Sci*. 2021;22(8). <https://doi.org/10.3390/ijms22084106>.
- Long X et al. Insight for Immunotherapy of HCMV Infection. *International journal of biological sciences*. 2021;17(11):2899-2911. [10.7150/ijbs.58127](https://doi.org/10.7150/ijbs.58127).
- Cannon MJ, et al. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. *Rev Med Virol*. 2010;20(4):202-13. <https://doi.org/10.1002/rmv.655>.
- Razonable RR et al. Clinical Diagnostic Testing for Human Cytomegalovirus Infections. *The Journal of infectious diseases*. 2020;221(Suppl 1):S74-s85. [10.1093/infdis/jiz601](https://doi.org/10.1093/infdis/jiz601).
- Yu X, et al. Atomic structure of the human cytomegalovirus capsid with its securing tegument layer of pp150. *Sci (New York NY)*. 2017;356(6345). <https://doi.org/10.1126/science.aam6892>.
- Balázs Z et al. Long-Read Sequencing of Human Cytomegalovirus Transcriptome Reveals RNA Isoforms Carrying Distinct Coding Potentials. *Scientific reports*. 2017;7(1):15989. [10.1038/s41598-017-16262-z](https://doi.org/10.1038/s41598-017-16262-z).
- Abgueguen P, et al. Venous thrombosis in immunocompetent patients with acute cytomegalovirus infection: a complication that may be underestimated. *Clin Microbiol Infect: Official Publication Eur Soc Clin Microbiol Infect Dis*. 2010;16(7):851-4. <https://doi.org/10.1111/j.1469-0691.2009.03022.x>.
- Wang YQ et al. Human Cytomegalovirus Primary Infection and Reactivation: Insights From Virion-Carried Molecules. *Frontiers in microbiology*. 2020;11(1511). [10.3389/fmicb.2020.01511](https://doi.org/10.3389/fmicb.2020.01511).
- Atabani SF, et al. Cytomegalovirus replication kinetics in solid organ transplant recipients managed by preemptive therapy. *Am J Transplantation: Official J Am Soc Transplantation Am Soc Transpl Surg*. 2012;12(9):2457-64. <https://doi.org/10.1111/j.1600-6143.2012.04087.x>.
- Aiello A, et al. Role of Immunogenetics in the outcome of HCMV infection: implications for Ageing. *Int J Mol Sci*. 2019;20(3). <https://doi.org/10.3390/ijms20030685>.
- Griffiths P, et al. Pathogenesis of human cytomegalovirus in the immunocompromised host. *Nat Rev Microbiol*. 2021;19(12):759-73. <https://doi.org/10.1038/s41579-021-00582-z>.

15. Attanasio U, et al. Pulmonary hypertension phenotypes in systemic sclerosis: the right diagnosis for the right treatment. *Int J Mol Sci.* 2020;21(12). <https://doi.org/10.3390/ijms21124430>.
16. Efthymiou G, et al. A comprehensive analysis of antigen-specific antibody responses against human cytomegalovirus in patients with systemic sclerosis. *Clin Immunol (Orlando Fla).* 2019;207(87–96):101016jclim201907012.
17. Chen L, et al. Up-regulation of THY1 attenuates interstitial pulmonary fibrosis and promotes lung fibroblast apoptosis during acute interstitial pneumonia by blockade of the WNT signaling pathway. *Cell Cycle (Georgetown Tex).* 2019;18(6–7):670–81. <https://doi.org/10.1080/15384101.2019.1578144>.
18. Prochnau D, et al. Human cytomegalovirus induces MMP-1 and MMP-3 expression in aortic smooth muscle cells. *Acta Microbiol Immunol Hung.* 2011;58(4):303–17. <https://doi.org/10.1556/AMicr.58.2011.4.7>.
19. Lucchese G et al. A Peptide Link Between Human Cytomegalovirus Infection, Neuronal Migration, and Psychosis. *Frontiers in psychiatry.* 2020;11(349).10.3389/fpsy.2020.00349.
20. Priel E, et al. Human cytomegalovirus viral load in tumor and peripheral blood samples of patients with malignant gliomas. *J Clin Neuroscience: Official J Neurosurgical Soc Australasia.* 2015;22(2):326–30. <https://doi.org/10.1016/j.jocn.2014.06.099>.
21. Soroceanu L, et al. Is HCMV a tumor promoter? *Virus Res.* 2011;157(2):193–203. <https://doi.org/10.1016/j.virusres.2010.10.026>.
22. Kumar A, et al. Tumor control by human cytomegalovirus in a murine model of hepatocellular carcinoma. *Mol Therapy Oncolytics.* 2016;3(16012):10–1038.mto.2016.12.
23. Richardson AK et al. Breast cancer and cytomegalovirus. Clinical & translational oncology: official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico. 2020;22(4):585–602.10.1007/s12094-019-02164-1.
24. Chen W, et al. The cytomegalovirus protein UL138 induces apoptosis of gastric cancer cells by binding to heat shock protein 70. *Oncotarget.* 2016;7(5):5630–45. <https://doi.org/10.18632/oncotarget.6800>.
25. Grce M, et al. Human papillomavirus, cytomegalovirus, and adeno-associated virus infections in pregnant and nonpregnant women with cervical intraepithelial neoplasia. *J Clin Microbiol.* 2004;42(3):1341–4. <https://doi.org/10.1128/jcm.42.3.1341-1344.2004>.
26. Bai B, et al. Human cytomegalovirus infection and colorectal cancer risk: a meta-analysis. *Oncotarget.* 2016;7(47):76735–42. <https://doi.org/10.18632/oncotarget.12523>.
27. Shanmughapriya S et al. Viral and bacterial aetiologies of epithelial ovarian cancer. *European journal of clinical microbiology & infectious diseases: official publication of the European Society of Clinical Microbiology.* 2012;31(9):2311–2317.10.1007/s10096-012-1570-5.
28. Yu C, et al. Human cytomegalovirus in cancer: the mechanism of HCMV-induced carcinogenesis and its therapeutic potential. *Front Cell Infect Microbiol.* 2023;13(1202138):103389fcimb20231202138.
29. Mehravaran H, et al. Association of Human Cytomegalovirus with Hodgkin's Disease and Non-hodgkin's lymphomas. *Asian Pac J cancer Prevention: APJCP.* 2017;18(3):593–7. <https://doi.org/10.22034/apjcp.2017.18.3.593>.
30. Harrison MAA et al. Metabolic Reprogramming of Glioblastoma Cells during HCMV Infection Induces Secretome-Mediated Paracrine Effects in the Microenvironment. *Viruses.* 2022;14(1).10.3390/v14010103.
31. Herbein GH-R. Oncogenic Human Cytomegalovirus. *Viruses.* 2022;14(11).10.3390/v14112462.
32. Buratin A et al. Systematic benchmarking of statistical methods to assess differential expression of circular RNAs. *Briefings in bioinformatics.* 2023;24(1).10.1093/bib/bbac612.
33. Law CW et al. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome biology.* 2014;15(2):R29.10.1186/gb-2014-15-2-r29.
34. Newman AM et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015;12(5):453–457.10.1038/nmeth.3337.
35. Yoshihara K et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature communications.* 2013;4(2612).10.1038/ncomms3612.
36. Simon N et al. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *Journal of statistical software.* 2011;39(5):1–13.10.18637/jss.v039.i05.
37. Chen T et al. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; San Francisco, California, USA: Association for Computing Machinery; 2016. pp. 785–794.
38. Wang W et al. The cuproptosis-related signature associated with the tumor environment and prognosis of patients with glioma. *Frontiers in immunology.* 2022;13(998236).10.3389/fimmu.2022.998236.
39. Zhang C et al. Exploration of the relationships between tumor mutation burden with immune infiltrates in clear cell renal cell carcinoma. *Annals of translational medicine.* 2019;7(22):648.10.21037/atm.2019.10.84.
40. Mayakonda A et al. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome research.* 2018;28(11):1747–1756.10.1101/gr.239244.118.
41. Bader GD et al. An automated method for finding molecular complexes in large protein interaction networks. *BMC bioinformatics.* 2003;4(2).10.1186/1471-2105-4-2.
42. The Gene Ontology (GO) project in 2006. *Nucleic acids research.* 2006;34(Database issue):D322–326.10.1093/nar/gkj021.
43. Yang W et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic acids research.* 2013;41(Database issue):D955–961.10.1093/nar/gks1111.
44. Liu CJ et al. GSCA: an integrated platform for gene set cancer analysis at genomic, pharmacogenomic and immunogenomic levels. *Briefings in bioinformatics.* 2023;24(1):bbac558.10.1093/bib/bbac558.
45. Barretina J et al. The Cancer Cell Line Encyclopedia enables predictive modeling of anticancer drug sensitivity. *Nature.* 2012;483(7391):603–607.10.1038/nature11003.
46. Dempster JM et al. Chronos: a cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects. *Genome biology.* 2021;22(1):343.10.1186/s13059-021-02540-7.
47. Luo W et al. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics.* 2013;29(14):1830–1831.10.1093/bioinformatics/btt285.
48. Kany S et al. Cytokines in Inflammatory Disease. *International journal of molecular sciences.* 2019;20(23).10.3390/ijms20236008.
49. Biolatti M et al. Modulation of the innate immune response by human cytomegalovirus. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases.* 2018;64(105–114).10.1016/j.meegid.2018.06.025.
50. Deng Y et al. A cytomegalovirus inflammasome inhibitor reduces pro-inflammatory cytokine release and pyroptosis. *Nature communications.* 2024;15(1):786.10.1038/s41467-024-45151-z.
51. Semmes EC et al. Cytomegalovirus as an immunomodulator across the lifespan. *Current opinion in virology.* 2020;44(112–120).10.1016/j.coviro.2020.07.013.
52. Mantegazza AR et al. Presentation of phagocytosed antigens by MHC class I and II. *Traffic (Copenhagen, Denmark).* 2013;14(2):135–152.10.1111/tra.12026.
53. Simmons J et al. The coagulopathy of acute sepsis. *Current opinion in anaesthesiology.* 2015;28(2):227–236.10.1097/aco.0000000000000163.
54. Skvortsova K et al. The DNA methylation landscape in cancer. *Essays in biochemistry.* 2019;63(6):797–811.10.1042/ebc20190037.
55. Liang WW et al. Integrative multi-omic cancer profiling reveals DNA methylation patterns associated with therapeutic vulnerability and cell-of-origin. *Cancer cell.* 2023;41(9):1567–1585.e1567.10.1016/j.ccell.2023.07.013.
56. Modhukur V et al. MethSurv: a web tool to perform multivariable survival analysis using DNA methylation data. *Epigenomics.* 2018;10(3):277–288.10.2217/epi-2017-0118.
57. Kao TJ et al. Prognoses and genomic analyses of proteasome 26S subunit, ATPase (PSMC) family genes in clinical breast cancer. *Aging.* 2021;13(14):17970.10.18632/aging.203345.
58. Xing C et al. Integrate analysis of the promote function of Cell division cycle-associated protein family to pancreatic adenocarcinoma. *International journal of medical sciences.* 2021;18(3):672–684.10.7150/ijms.53243.
59. Goodrum F et al. Human cytomegalovirus persistence. *Cellular microbiology.* 2012;14(5):644–655.10.1111/j.1462-5822.2012.01774.x.
60. Forte E et al. Cytomegalovirus Latency and Reactivation: An Intricate Interplay With the Host Immune Response. *Frontiers in cellular and infection microbiology.* 2020;10(130).10.3389/fcimb.2020.00130.
61. DiNardo AR et al. Postinfectious Epigenetic Immune Modifications - A Double-Edged Sword. *The New England journal of medicine.* 2021;384(3):261–270.10.1056/NEJMr2028358.
62. Surace AEA et al. The Role of Epigenetics in Autoimmune/Inflammatory Disease. *Frontiers in immunology.* 2019;10(1525).10.3389/fimmu.2019.01525.
63. Griffiths P et al. Cytomegalovirus. *Current opinion in infectious diseases.* 2014;27(6):554–559.10.1097/qco.000000000000107.

64. Leeaphorn N et al. Cytomegalovirus mismatch still negatively affects patient and graft survival in the era of routine prophylactic and preemptive therapy: A paired kidney analysis. *American journal of transplantation: official journal of the American Society of Transplantation and the American Society of Transplant Surgeons*. 2019;19(2):573-584.10.1111/ajt.15183.
65. Conroy MJ et al. CX3CL1 Signaling in the Tumor Microenvironment. *Advances in experimental medicine and biology*. 2020;1231(1-12).10.1007/978-3-030-36667-4\_1.
66. Rivas-Fuentes S et al. Regulation and biological functions of the CX3CL1-CX3CR1 axis and its relevance in solid cancer: A mini-review. *Journal of Cancer*. 2021;12(2):571-583.10.7150/jca.47022.
67. Trinh T et al. CX3CR1 deficiency-induced TIL tumor restriction as a novel addition for CAR-T design in solid malignancies. *iScience*. 2023;26(4):106443.10.1016/j.isci.2023.106443.
68. Szukiewicz D, CX3CL1 (Fractalkine)-CX3CR1 Axis in Inflammation-Induced Angiogenesis and Tumorigenesis. *International journal of molecular sciences*. 2024;25(9).10.3390/ijms25094679.
69. Korbecki J et al. Fractalkine/CX3CL1 in Neoplastic Processes. *International journal of molecular sciences*. 2020;21(10).10.3390/ijms21103723.
70. Chen F et al. New horizons in tumor microenvironment biology: challenges and opportunities. *BMC medicine*. 2015;13(45).10.1186/s12916-015-0278-7.
71. Paluskiewicz CM et al. T Regulatory Cells and Priming the Suppressive Tumor Microenvironment. *Frontiers in immunology*. 2019;10(2453).10.3389/fimmu.2019.02453.
72. Quail DF et al. Microenvironmental regulation of tumor progression and metastasis. *Nature medicine*. 2013;19(11):1423-1437.10.1038/nm.3394.
73. Schulz M et al. Microenvironmental Regulation of Tumor Progression and Therapeutic Response in Brain Metastasis. *Frontiers in immunology*. 2019;10(1713).10.3389/fimmu.2019.01713.
74. Chen Z et al. Single-cell RNA sequencing highlights the role of inflammatory cancer-associated fibroblasts in bladder urothelial carcinoma. *Nature communications*. 2020;11(1):5077.10.1038/s41467-020-18916-5.
75. Zhang A et al. Cancer-associated fibroblasts promote M2 polarization of macrophages in pancreatic ductal adenocarcinoma. *Cancer medicine*. 2017;6(2):463-470.10.1002/cam4.993.
76. Kinoshita T et al. Forkhead box P3 regulatory T cells coexisting with cancer associated fibroblasts are correlated with a poor outcome in lung adenocarcinoma. *Cancer science*. 2013;104(4):409-415.10.1111/cas.12099.
77. Chun E et al. CCL2 Promotes Colorectal Carcinogenesis by Enhancing Polymorphonuclear Myeloid-Derived Suppressor Cell Population and Function. *Cell reports*. 2015;12(2):244-257.10.1016/j.celrep.2015.06.024.
78. Desany B et al. Bioinformatics and cancer target discovery. *Drug discovery today*. 2004;9(18):795-802.10.1016/s1359-6446(04)03224-6.
79. Anuraga G et al. Potential Prognostic Biomarkers of NIMA (Never in Mitosis, Gene A)-Related Kinase (NEK) Family Members in Breast Cancer. *Journal of personalized medicine*. 2021;11(11).10.3390/jpm11111089.
80. Lazaro-Camp VJ, et al. SETDB1 in cancer: overexpression and its therapeutic implications. *Am J cancer Res*. 2021;11(5):1803-27.
81. Chabanon RM et al. Mutational Landscape and Sensitivity to Immune Checkpoint Blockers. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2016;22(17):4309-4321.10.1158/1078-0432.Ccr-16-0903.
82. Lamb J et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science (New York, NY)*. 2006;313(5795):1929-1935.10.1126/science.1132939.
83. Subramanian A et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell*. 2017;171(6):1437-1452.e1417.10.1016/j.cell.2017.10.049.
84. Wang CY, et al. Gene signatures and potential therapeutic targets of amino acid metabolism in estrogen receptor-positive breast cancer. *Am J cancer Res*. 2020;10(1):95-113.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.