Observational Study

Medicine

OPEN

Screening and identification analysis of core markers for leukemia and cervical cancer

Calmodulin 3 as a core target

Danfeng Shao, MD[a,*], Huaiying Yu, MD[a], Xiaoqing Zhu, MD[b]

Abstract

Leukemia is a type of malignant tumor affecting hematopoietic system. Cervical cancer is a malignant tumor of female reproductive tract. The relationship between Calmodulin 3 (CALM3) and leukemia, cervical cancer remains unclear. Leukemia dataset GSE26294 and cervical cancer dataset GSE173097 profiles were downloaded from gene expression omnibus. Principal component analysis, differentially expressed genes (DEGs) screening, weighted gene co-expression network analysis (WGCNA), functional enrichment analysis, gene set enrichment analysis, immune infiltration analysis, protein–protein interaction network construction and analysis, survival analysis were performed. Gene expression heatmaps were plotted. Comparative Toxicogenomics Database (CTD) was used to find diseases most related to core genes. 77 DEGs were identified. According to gene ontology, in biological process category, they were mainly enriched in cell proliferation, immune response, and apoptotic process. In cellular component category, they were mainly enriched in nucleus and Golgi apparatus. In molecular function category, they were mainly enriched in DNA binding, protein binding, transcription factor activity. In Kyoto encyclopedia of gene and genome analysis, they were mainly enriched in cell adhesion molecules, Wnt signaling pathway, cAMP signaling pathway, cGMP-PKG signaling pathway, and basal cell carcinoma. The soft-threshold power in WGCNA was set to 1, generating 6 modules. Finally identifying 3 core genes (CALM3, secreted frizzled-related protein 4, plasminogen). CTD analysis revealed that core genes were related to leukemia, coagulation disorders, vaginal tumors, cervical tumors, autoimmune diseases, and inflammation. CALM3 is lowly expressed in leukemia samples and highly expressed in cervical cancer samples.

Abbreviations: ALL = acute lymphoblastic leukemia, AML = acute myeloid leukemia, BP = biological process, CALM3 = calmodulin 3, CC = cellular component, CTD = Comparative Toxicogenomics Database, DEGs = differentially expressed genes, GEO = gene expression omnibus, GO = gene ontology, GSEA = gene set enrichment analysis, KEGG = Kyoto Encyclopedia of Gene and Genome, MAD = median absolute deviation, MF = molecular function, PCA = principal component analysis, PPI = protein–protein interaction, STRING = search tool for the retrieval of interacting genes, WGCNA = weighted gene co-expression network analysis.

Keywords: CALM3, cervical cancer, differentially expressed genes, leukemia

1. Introduction

Leukemia and cervical cancer are 2 malignant tumors that pose serious threats to human health globally. Leukemia is a malignant disease originating from hematopoietic stem cells in the bone marrow, characterized by the abnormal proliferation of white blood cells that invade the blood and other tissues.[1] According to related studies, leukemia results in tens of thousands of deaths annually, with its incidence and mortality rates continually rising worldwide.[2] Cervical cancer, primarily caused by human papillomavirus (HPV) infection, is one of the most common malignant tumors of the female reproductive system.[3] Statistics show that in high-income countries, the incidence and mortality of cervical cancer have declined due to widespread screening and vaccination programs. However, in low- and middle-income countries, the incidence and mortality rates of cervical cancer remain high.[4] The pathogenesis of these 2 diseases is not completely understood, but existing research suggests that genetic factors, chromosomal abnormalities, and gene fusions may play critical roles in their development. For instance, the occurrence of leukemia is often closely associated with specific chromosomal translocations,[5] such as the BCR-ABL fusion gene. Similarly, the development of cervical cancer involves various genetic mutations and signaling pathway

The authors have no funding and conflicts of interest to declare.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

The data in this article are from public databases and are exempt from ethical review.

[a] Department of Gynaecology, Hangzhou Fuyang first people's Hospital, Hangzhou, China, [b] Department of Gynaecology, The Second Affiliated Hospital of Medical College of Zhejiang University, Hangzhou, China.

** Correspondence: Danfeng Shao, Department of Gynaecology, Hangzhou Fuyang First People's Hospital, Hangzhou, China (e-mail: shaodanfeng55@163.com).*

How to cite this article: Shao D, Yu H, Zhu X. Screening and identification analysis of core markers for leukemia and cervical cancer: Calmodulin 3 as a core target. Medicine 2025;104:14(e41665).

Received: 30 July 2024 / Received in final form: 1 February 2025 / Accepted: 7 February 2025

http://dx.doi.org/10.1097/MD.0000000000041665

abnormalities, including mutations in the p53 and RB genes.[6] Therefore, in-depth research into the molecular mechanisms of leukemia and cervical cancer is crucial for understanding their pathophysiological foundations and developing new therapeutic strategies.

The rapid development of bioinformatics provides powerful tools for studying the molecular mechanisms of diseases. The application of high-throughput sequencing and gene chip technologies allows for comprehensive analysis of genomic, transcriptomic, and proteomic data, thereby uncovering key genes and signaling pathways related to diseases. The advantage of bioinformatics lies in its ability to process and analyze large volumes of biological data, integrating various data types to offer more comprehensive biological insights. Using high-throughput sequencing techniques like RNA-seq, we can analyze gene expression differences between cancer patients and healthy controls, identifying potential pathogenic genes and pathways.[7] In summary, bioinformatics supports the efficient processing and analysis of extensive biological data, providing strong support for revealing the molecular mechanisms of diseases and laying the foundation for developing new diagnostic and therapeutic strategies.[8]

Currently, the role of Calmodulin 3 (CALM3) in leukemia and cervical cancer is unclear. Therefore, this study aims to utilize bioinformatics techniques to identify core genes between leukemia and cervical cancer and normal tissues, and to conduct enrichment and pathway analyses. By validating the significant role of CALM3 in leukemia and cervical cancer using public datasets, and conducting fundamental cell experiments, we seek to uncover the potential mechanisms of CALM3 in these 2 cancers. The goal of this research is to elucidate the potential mechanisms of CALM3 in these 2 diseases, providing new targets and strategies for future diagnosis and treatment.

2. Methods

2.1. Leukemia and cervical cancer datasets

In this study, the leukemia dataset GSE26294 and the cervical cancer dataset GSE173097 profiles were downloaded from the gene expression omnibus database (http://www.ncbi.nlm.nih.gov/geo/) generated from GPL6102 and GPL14951. GSE26294 includes 8 leukemia and 8 normal samples, while GSE173097 includes 5 cervical cancer and 6 normal samples. These datasets were used to identify differentially expressed genes (DEGs) in leukemia and cervical cancer.

2.2. Principal component analysis

Principal component analysis (PCA) was performed to reduce the dimensionality of these data samples and visualize them, revealing the underlying structure of the data and discovering differential biological characteristics. We used the R package stats (version 3.6.0) for the analysis. Specifically, we first performed z-score normalization on the expression profiles and further used the prcomp function for dimensionality reduction to obtain the reduced matrix. This was used to show the distribution and differences between healthy and diseased samples.

2.3. Screening of DEGs

We first performed log2 transformation on the GSE26294 and GSE173097 datasets, used the lmFit function for multiple linear regression, and calculated the adjusted t-statistics, adjusted F-statistics, and log-fold changes through empirical Bayesian adjustment of standard errors approaching a common value. This provided the significance of differential expression for each gene, resulting in a volcano plot. The DEGs from GSE26294 and GSE173097 were intersected to obtain the final DEGs.

2.4. Weighted gene co-expression network analysis

First, we calculated the median absolute deviation (MAD) for each gene in the GSE173097 gene expression matrix, removing the top 50% of genes with the smallest MAD. Using the goodSamplesGenes method from the R package weighted gene co-expression network analysis (WGCNA), we removed outlier genes and samples. Further, we constructed a scale-free co-expression network using WGCNA. To classify genes with similar expression profiles into gene modules, average linkage hierarchical clustering was performed based on TOM dissimilarity measures, with a minimum module size (gene group) of 30. The sensitivity was set to 3. For further module analysis, we calculated the dissimilarity of module eigengenes, selected a cut line for the module dendrogram, and merged some modules. Additionally, modules with a distance <0.25 were also merged. It is important to note that the gray module is considered a collection of genes that cannot be assigned to any module.

2.5. Functional enrichment analysis

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses assess gene function and biological pathways. The intersected DEGs were input into the KEGG API (https://www.kegg.jp/kegg/rest/keggapi.html) to get the latest KEGG pathway gene annotations. The R package clusterProfiler (version 3.14.3) was used for enrichment analysis to obtain enriched gene sets. Additionally, the GO annotations from the R package org.Hs.e.g..db (version 3.1.0) were used as the background. The minimum gene set was set to 5, and the maximum to 5000. A *P*-value of <.05 and an FDR of <.25 were considered statistically significant.

The Metascape database (http://metascape.org/gp/index.html) was also used for comprehensive gene list annotation and analysis resources, and visualization was exported.

2.6. Gene set enrichment analysis

For gene set enrichment analysis (GSEA), we obtained the GSEA software (version 3.0) from the GSEA website (DOI:10.1073/pnas.0506580102, http://software.broadinstitute.org/gsea/index.jsp). Disease and normal tissue samples were divided into 2 groups, and the c2.cp.kegg.v7.4.symbols.gmt subset was downloaded from the Molecular Signatures Database (DOI:10.1093/bioinformatics/btr260, http://www.gsea-msigdb.org/gsea/downloads.jsp) to assess related pathways and molecular mechanisms. Based on gene expression profiles and phenotype grouping, the minimum gene set was set to 5, the maximum to 5000, and 1000 permutations were performed. A *P*-value of <.05 and an FDR of <.25 were considered statistically significant. GO and KEGG analyses of the whole genome were also conducted by GSEA.

2.7. Immune infiltration analysis

CIBERSORT (http://CIBERSORT.stanford.edu/) is a widely used method for estimating the abundance of immune cells in a mixed cell population using gene expression data. The LM22 gene file defines 22 immune cell subtypes. We applied an integrated bioinformatics approach using the CIBERSORT package to analyze the gene expression matrix of the leukemia dataset GSE26294, deconvolving the expression matrix of immune cell subtypes using linear support vector regression, estimating immune cell abundance, and using a confidence *P*-value of <.05 as a cutoff standard to select samples with sufficient confidence.

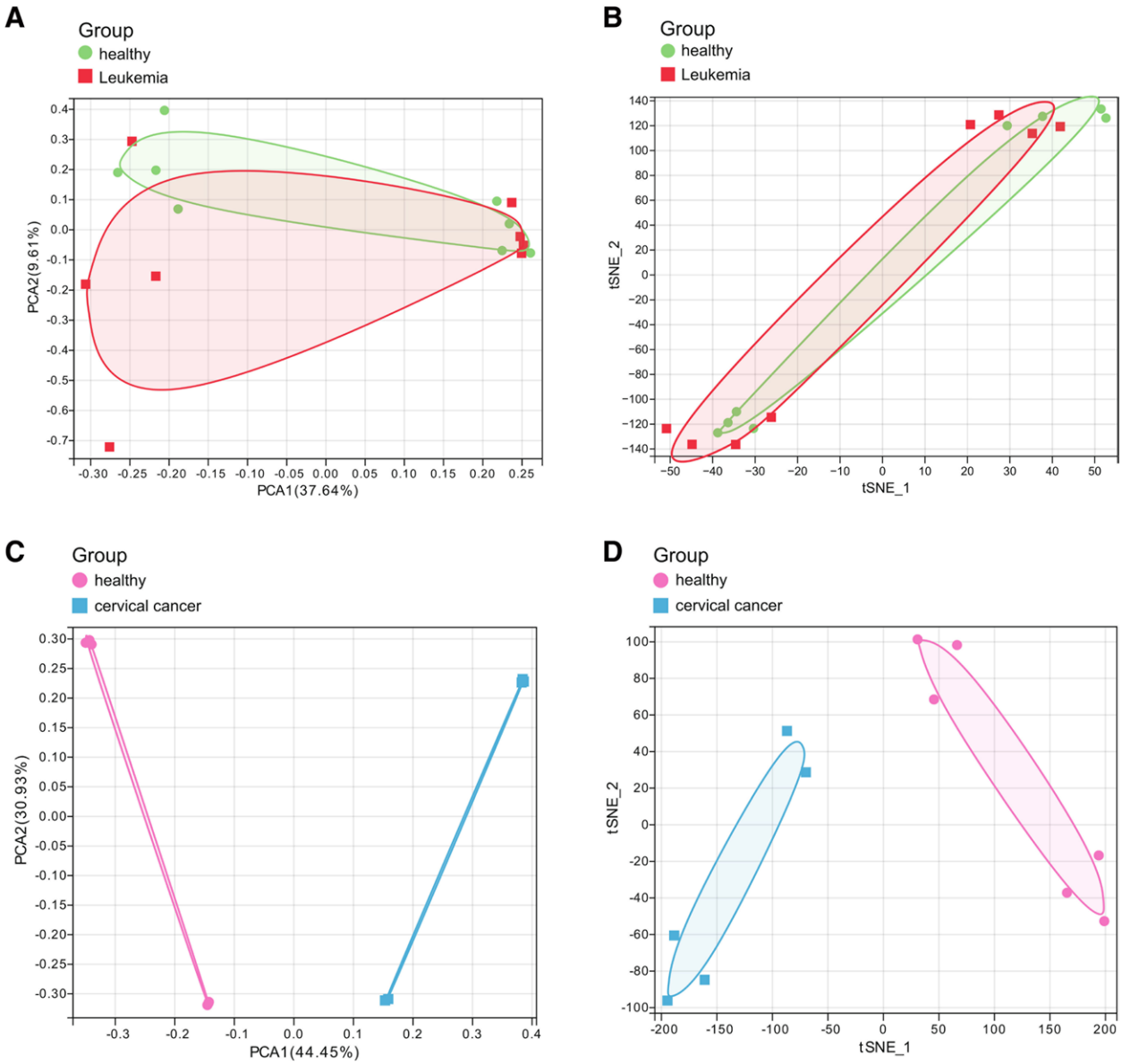

Figure 1. (A) Principal component analysis (PCA) results of leukemia. (B) tSNE analysis results of leukemia. (C) PCA results of cervical cancer. (D) tSNE analysis results of cervical cancer. tSNE = t-Distributed Stochastic Neighbor Embedding.

2.8. Protein–protein interaction network construction and analysis

The search tool for the retrieval of interacting genes (STRING) database (http://string-db.org/) collects, scores, and integrates all publicly available protein–protein interaction (PPI) information sources, supplemented by computational predictions. The DEGs were input into the STRING database to construct a PPI network with a confidence score >0.4. Cytoscape software provides biological network analysis and 2D visualization for biologists. In this study, the PPI network formed by the STRING database was visualized and core genes were predicted using Cytoscape software. The PPI network was imported into Cytoscape, and the top 10 related genes were calculated using 4 algorithms (Maximal Clique Centrality, Maximum Neighborhood Component, Density of Maximum Neighborhood Component, BottleNeck), with the intersection visualized and core gene lists exported.

2.9. Survival analysis

We obtained clinical survival data for cervical cancer from TCGA and used the R package maxstat (version 0.7-25) to calculate the optimal cutoff value for the RiskScore of core genes, setting the minimum group sample size >25% and the maximum group sample size <75%. The optimal cutoff value was calculated, and patients were divided into high and low groups. The survfit function from the R package survival was used to analyze the prognostic differences between the 2 groups, and the log-rank test method was used to assess the significance of prognostic differences between different groups.

2.10. Gene expression heatmap

We used the R package heatmap to plot heatmaps of the expression levels of core genes identified in the PPI network in GSE26294 and GSE173097, visualizing the expression differences of core genes between leukemia, cervical cancer, and normal samples.

2.11. Comparative toxicogenomics database analysis

The CTD integrates data on interactions between chemicals, genes, functional phenotypes, and diseases, providing significant

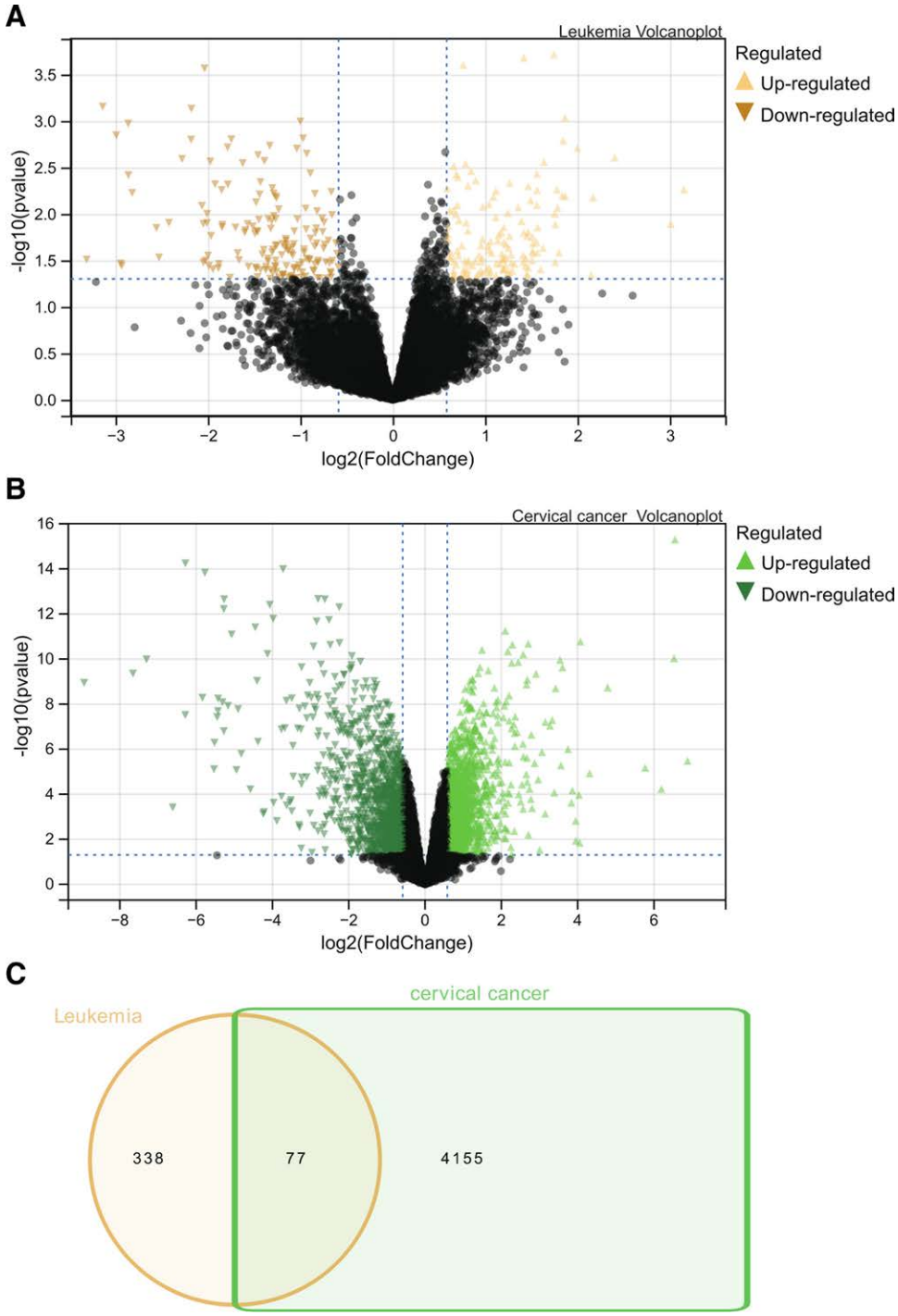

Figure 2. Differentially expressed gene (DEG) Analysis. (A) Leukemia dataset GSE26294. (B) Cervical cancer dataset GSE173097. (C) Intersection of differential genes in leukemia and cervical cancer. A total of 77 DEGs.

convenience for research on disease-related environmental exposure factors and potential mechanisms of drug action. The core genes were input into the CTD website to find the most related diseases, and radar charts of expression differences for each gene were drawn using Excel.

2.12. miRNA

We used miRNA prediction websites to find miRNA information for target genes. In our study, the online databases TargetScan (www.targetscan.org) and miRTarBase (mirtarbase.cuhk.edu.cn) were used to screen miRNAs regulating core genes.

3. Results

3.1. Principal component analysis

We performed dimensionality reduction on leukemia and cervical cancer data samples to guide subsequent differential analysis. Using PCA and t-Distributed Stochastic Neighbor Embedding methods for dimensionality reduction and visualization, we revealed the underlying structure of the data. The results showed differentiation between leukemia and corresponding normal samples (Fig. 1A and B), as well as between cervical cancer and corresponding normal samples (Fig. 1C and D). Despite some similarity due to the same tissue origin, a portion of the samples

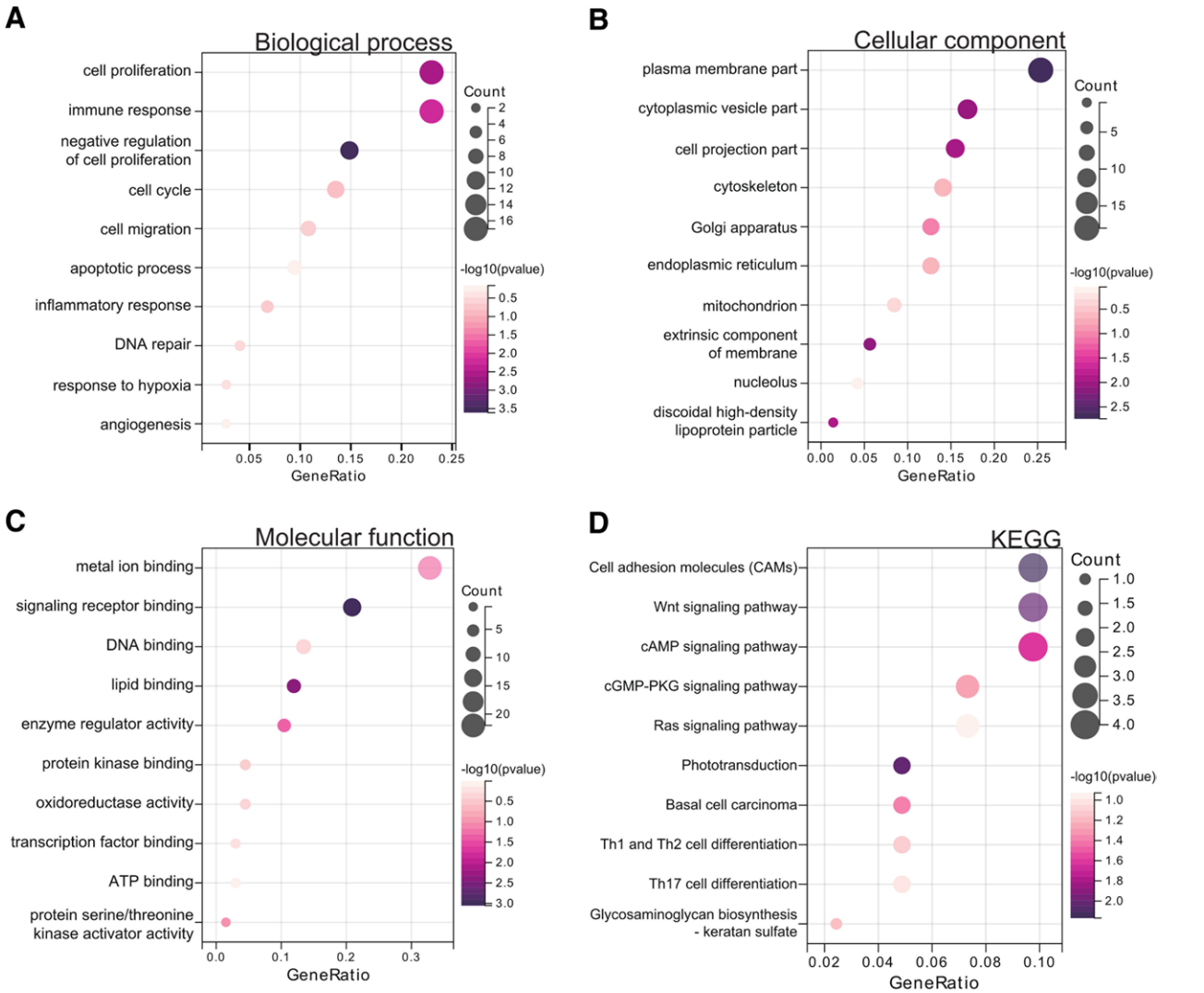

Figure 3. GOKEGG enrichment analysis of differentially expressed genes. (A) Biological process analysis. (B) Cellular component analysis. (C) Molecular function analysis. (D) KEGG enrichment analysis. KEGG = Kyoto encyclopedia of gene and genome.

showed good differentiation, highlighting differences between disease and control samples.

3.2. Differentially expressed gene analysis

In this study, based on predefined cutoff values, we identified DEGs for the leukemia dataset GSE26294 (Fig. 2A) and the cervical cancer dataset GSE173097 (Fig. 2B). We then used a Venn diagram to intersect the DEGs from both datasets, yielding a total of 77 common DEGs (Fig. 2C).

3.3. Functional enrichment analysis

3.3.1. DEGs. We performed GO and KEGG analyses on the DEGs. According to the GO analysis, in the biological process category, the genes were primarily enriched in cell proliferation, immune response, and apoptotic processes (Fig. 3A). In the cellular component category, the genes were mainly enriched in the nucleus and Golgi apparatus (Fig. 3B). In the molecular function category, the genes were primarily involved in DNA binding, protein binding, and transcription factor activity (Fig. 3C).

In the KEGG pathway analysis, the genes were mainly enriched in cell adhesion molecules, the Wnt signaling pathway, the cAMP signaling pathway, the cGMP-PKG signaling pathway, and basal cell carcinoma (Fig. 3D).

3.3.2. Gene set enrichment analysis. Additionally, we performed GSEA on the whole genome to identify potential enrichment terms among non-DEGs and to validate the results of the DEGs. The intersection of enrichment terms with GO and KEGG enrichment terms of DEGs revealed similar results between the leukemia dataset GSE26294 (Fig. 4A–D) and the cervical cancer dataset GSE173097 (Fig. 4E–H). DEGs were mainly enriched in cell proliferation, apoptosis processes, nucleus, transcription factor activity, Wnt signaling pathway, and basal cell carcinoma.

3.3.3. Metascape enrichment analysis. In Metascape enrichment items, GO enrichment results included important pathways such as the Wnt signaling pathway, growth regulation, cell adhesion molecules, regulation of epithelial cell proliferation, calcium-mediated signaling, and negative regulation of cell population proliferation (Fig. 5A). We also generated enrichment networks colored by enrichment items and *P*-values (Fig. 5B and C). Metascape enrichment results further validated the reliability of the above GO and KEGG enrichment results.

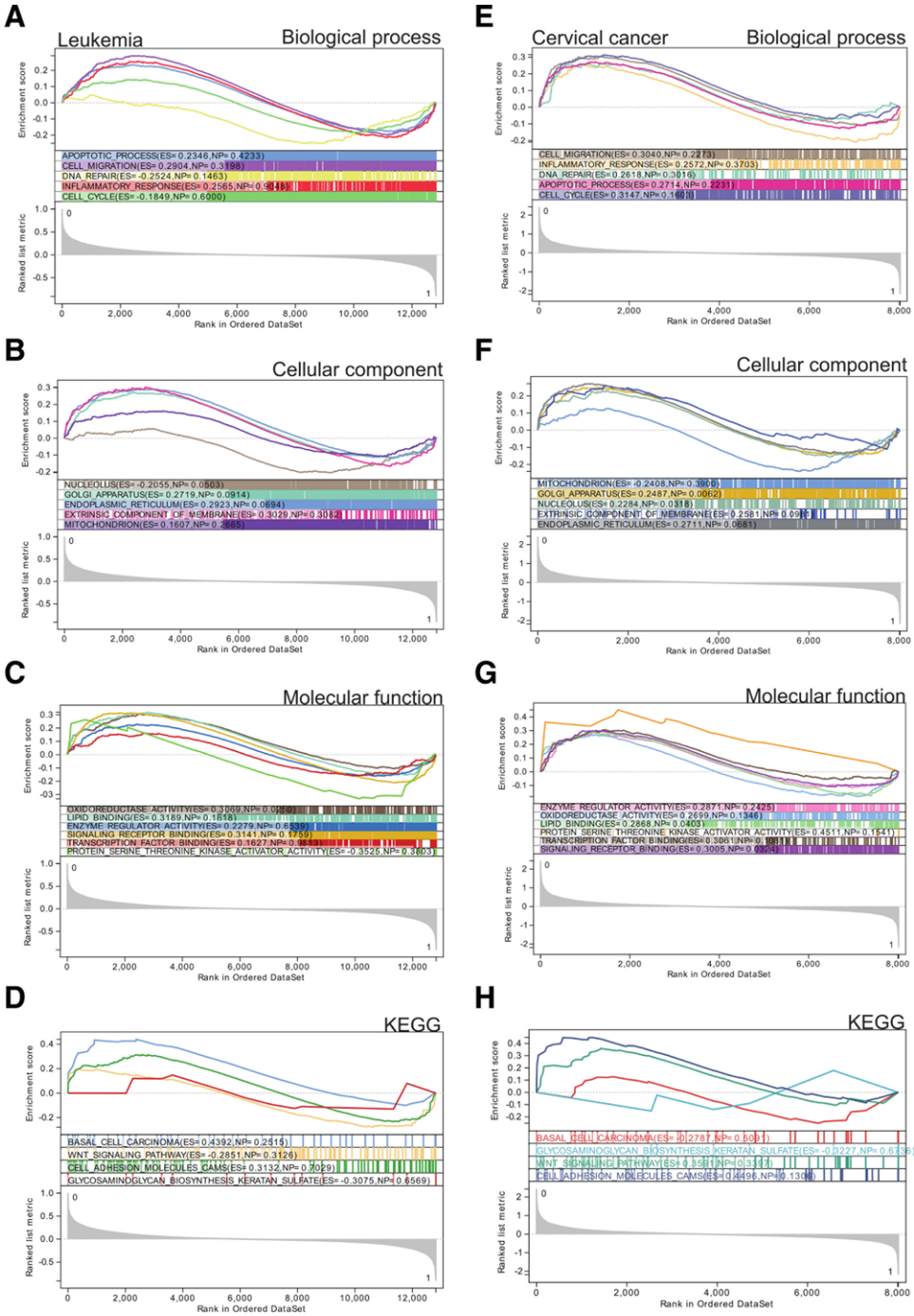

Figure 4. (A–D) GSEA of leukemia. (A) Biological process analysis. (B) Cellular component analysis. (C) Molecular function enrichment analysis. (D) KEGG enrichment analysis. (E–H) GSEA enrichment analysis of cervical cancer. (E) Biological process analysis. (F) Cellular component analysis. (G) Molecular function enrichment analysis. (H) KEGG enrichment analysis. GSEA = gene set enrichment analysis, KEGG = Kyoto Encyclopedia of Gene and Genome.

3.4. Immune infiltration analysis

Using the CIBERSORT software package, we analyzed the gene expression matrix of the leukemia dataset GSE26294 and obtained the proportion of immune cells in the entire gene expression matrix with 95% confidence. The analysis showed that dendritic cells (resting) and macrophages (M0) had higher proportions in the samples (Fig. 6A). Further correlation analysis of infiltrating immune cells revealed a co-expression pattern among immune cell components. The results indicated a strong positive correlation between monocytes and CD8 + T cells (Fig. 6B). This suggests that the interaction between monocytes and CD8 + T cells may play an important role in the progression of leukemia, providing new perspectives for leukemia treatment, especially targeting the interaction between these cell types.

3.5. Weighted gene co-expression network analysis

The selection of soft-threshold power is a critical step in WGCNA. We performed network topology analysis to determine the soft-threshold power, which was set to 1 for the WGCNA (Fig. 7A). We constructed a hierarchical clustering tree of all genes, generating 6 modules (Fig. 7B) and analyzed

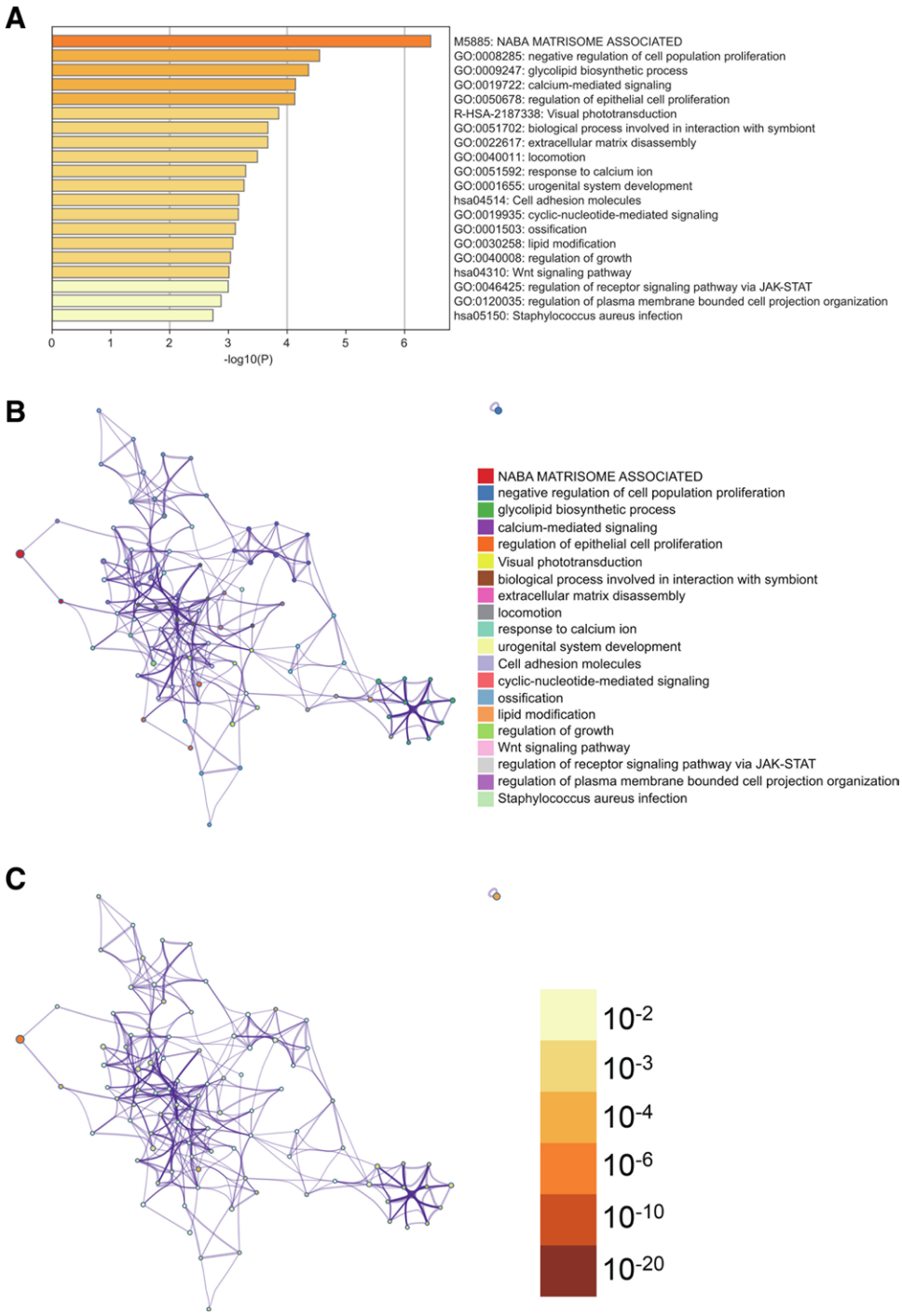

Figure 5. Metascape enrichment analysis. (A) Bar graph of enriched terms across input gene lists, colored by *P*-values. (B) Network of enriched terms: colored by cluster ID, where nodes that share the same cluster ID are typically close to each other. (C) Colored by *P*-value, where terms containing more genes tend to have a more significant *P*-value.

the interactions among important modules (Fig. 7C). We also generated heatmaps of module-trait relationships (Fig. 8A) and scatter plots of the GS vs MM correlation for related hub genes (Fig. 8B–D).

We calculated the module eigengene-gene expression correlation to obtain MM. According to the cutoff standard ($|MM| > 0.8$), highly connected genes in clinically significant modules were identified as hub genes. By extracting these hub genes from important modules, we intersected them with DEGs and plotted a Venn diagram for subsequent PPI network construction (Fig. 8E).

3.6. Protein–protein interaction network construction and analysis

The PPI network of DEGs was constructed using the STRING online database and analyzed with Cytoscape software (Fig. 9A). Four algorithms (Maximal Clique Centrality, Maximum Neighborhood Component, Density of Maximum Neighborhood Component, BottleNeck) were used to identify central genes, and a Venn diagram was used to find the intersection (Fig. 9B). The results of the 4 algorithms are shown (Fig. 9C–F), and we ultimately identified 3 core genes: CALM3, secreted frizzled-related protein 4 (SFRP4), and plasminogen (PLG).

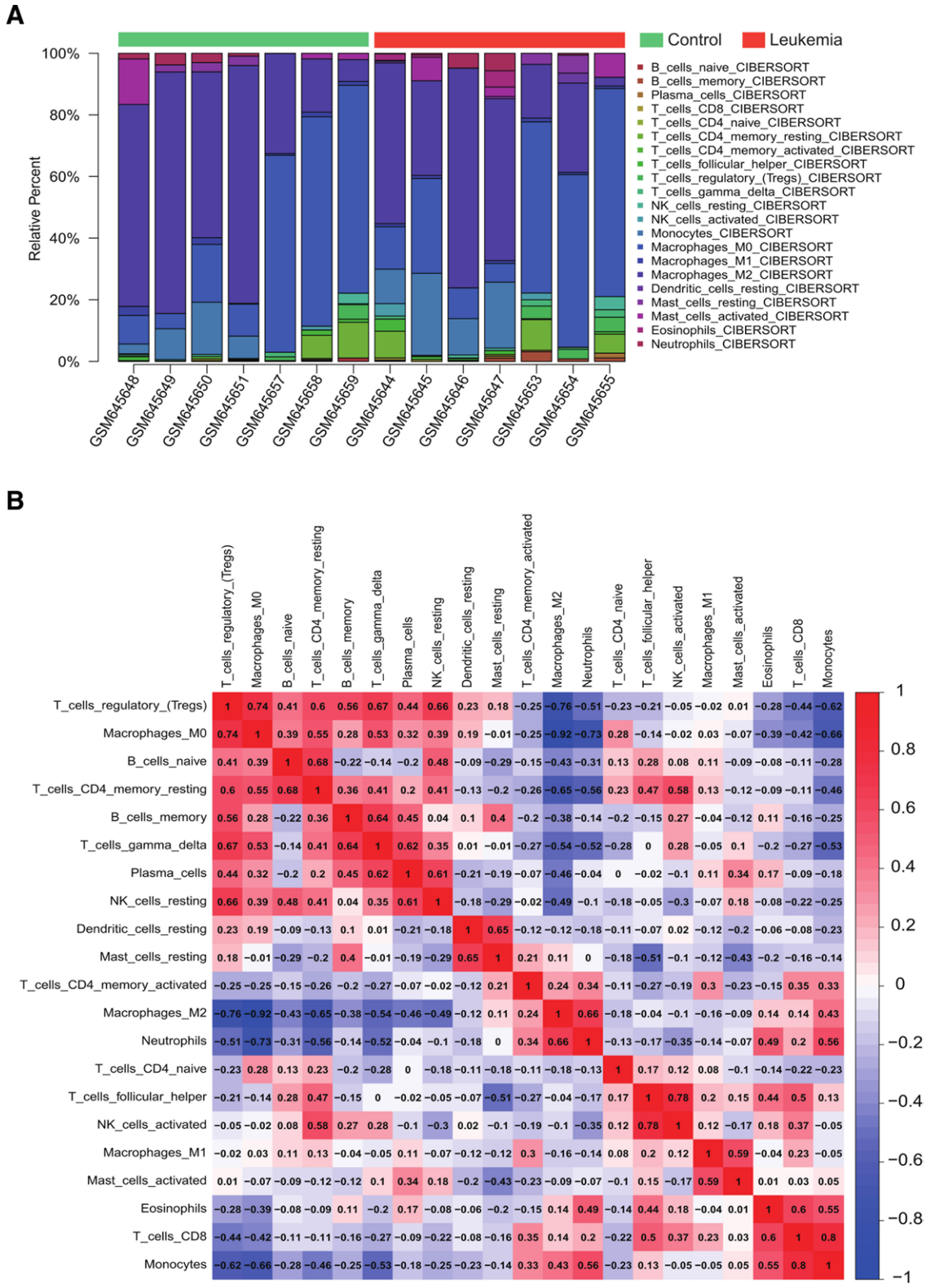

Figure 6. Immunoinfiltration analysis. (A) Whole gene expression matrix results in proportion of immune cells. (B) Map of co-expression patterns between immune cell components.

3.7. Core gene expression heatmap

We visualized and plotted heatmaps of the expression levels of core genes in the leukemia dataset GSE26294 and the cervical cancer dataset GSE173097. CALM3 was found to be lowly expressed in leukemia samples but highly expressed in normal samples (Fig. 10A). In contrast, CALM3 was highly expressed in cervical cancer samples and lowly expressed in normal samples (Fig. 10B). This opposite expression pattern suggests that CALM3 may have different biological functions and regulatory mechanisms in different types of cancer. In leukemia, low CALM3 expression may be related to disease onset and progression, while in cervical cancer, high CALM3 expression may promote cancer cell growth and invasion.

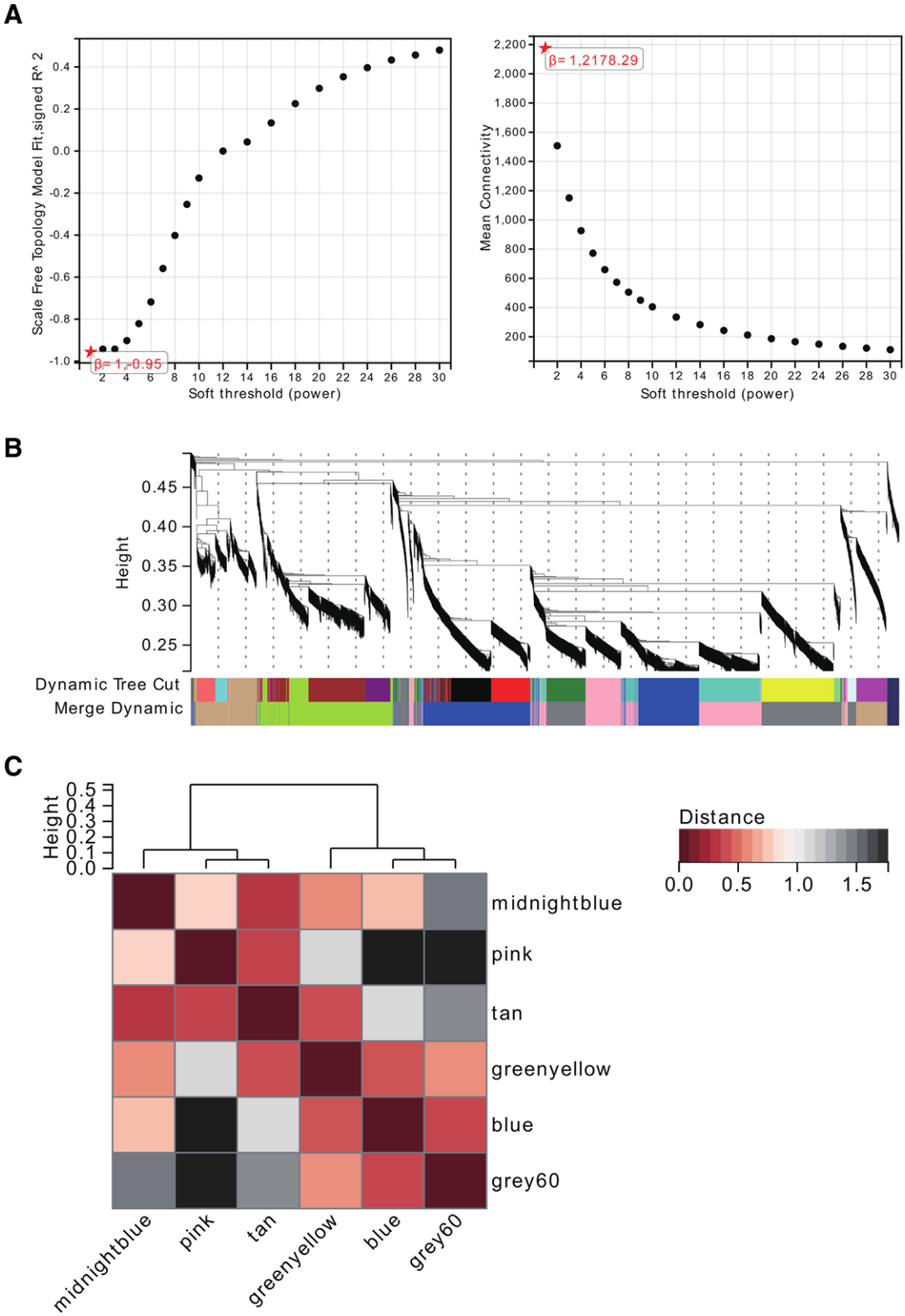

Figure 7. WGCNA. (A) β = 1, –0.95. β = 1, 2178.29. (B, C) The hierarchical clustering tree of all genes was constructed, and 6 important modules were generated. WGCNA = weighted gene co-expression network analysis.

3.8. Comparative toxicogenomics database analysis

We input the hub gene list into the CTD website to find diseases related to the core genes, enhancing our understanding of gene-disease associations. The core genes (CALM3, SFRP4, PLG) were found to be related to leukemia, coagulation disorders, vaginal tumors, cervical tumors, autoimmune diseases, and inflammation (Fig. 11).

3.9. Survival analysis

Using cervical cancer survival data downloaded from TCGA, we obtained prognosis score relationship plots and heatmaps of core gene expression differences between cervical cancer and normal tissue samples. We found that the survival time and rate of the low-risk group were significantly higher than those of the high-risk group (Fig. 12A). Visualizing the

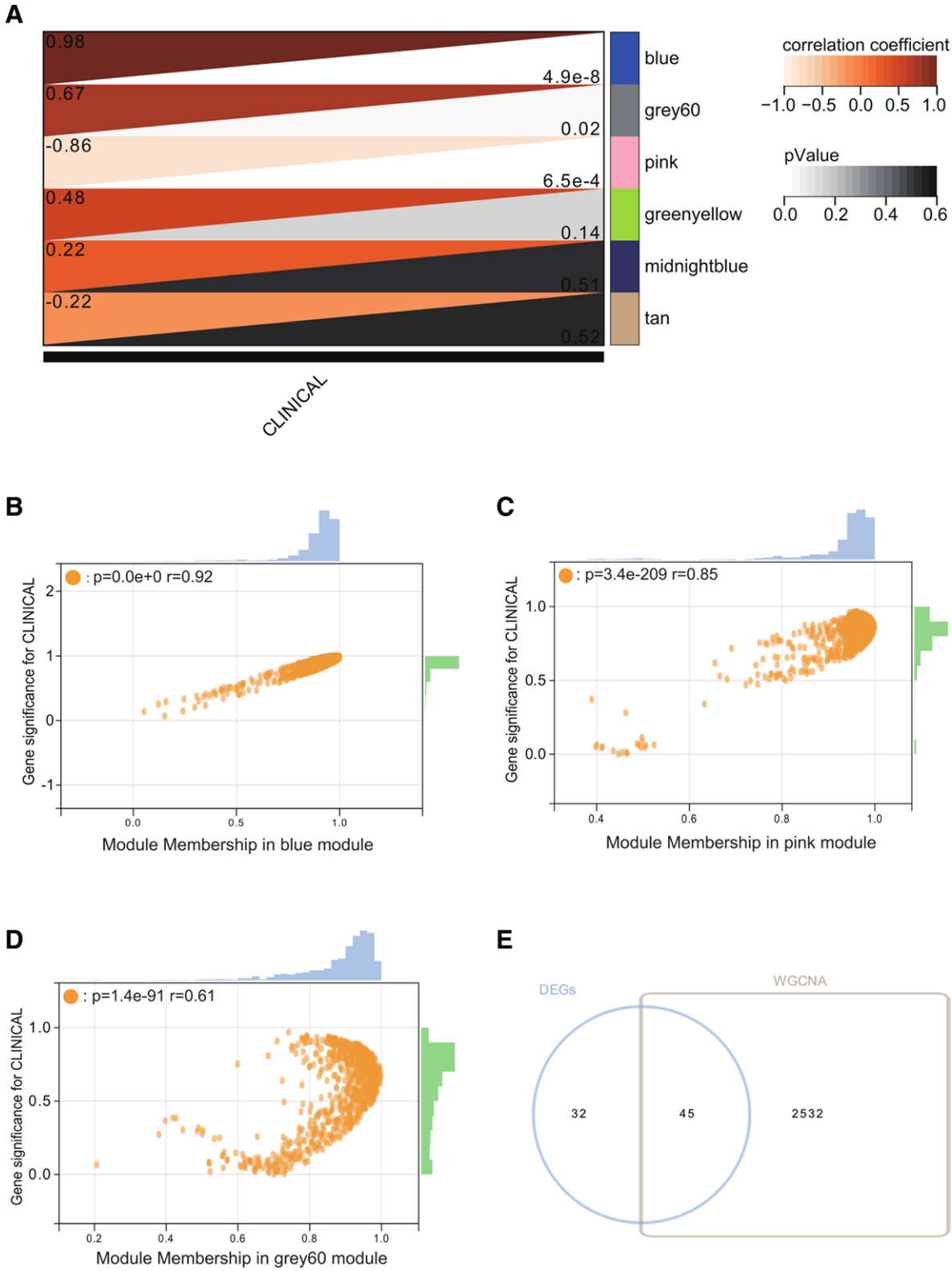

Figure 8. WGCNA. (A) The heat map of correlation between modules and phenotypes. (B, C, D) The scatter map of correlation between GS and MM of related hub genes. (E) The differentially expressed genes (DEGs) screened by WGCNA and DEGs was used to obtain venn map. WGCNA = weighted gene co-expression network analysis.

expression heatmap of core genes in cervical cancer survival data showed that CALM3 is a risk factor, with expression increasing with risk score (Fig. 12B). We also plotted overall survival rates related to core genes (CALM3, SFRP4, PLG) in cervical cancer (Fig. 12C), indicating that high CALM3 expression is associated with poor overall survival. Although the Log-rank *P*-value and p (HR) value were slightly above .05, both were close to significance, suggesting that high CALM3 expression may be an important factor influencing survival rates.

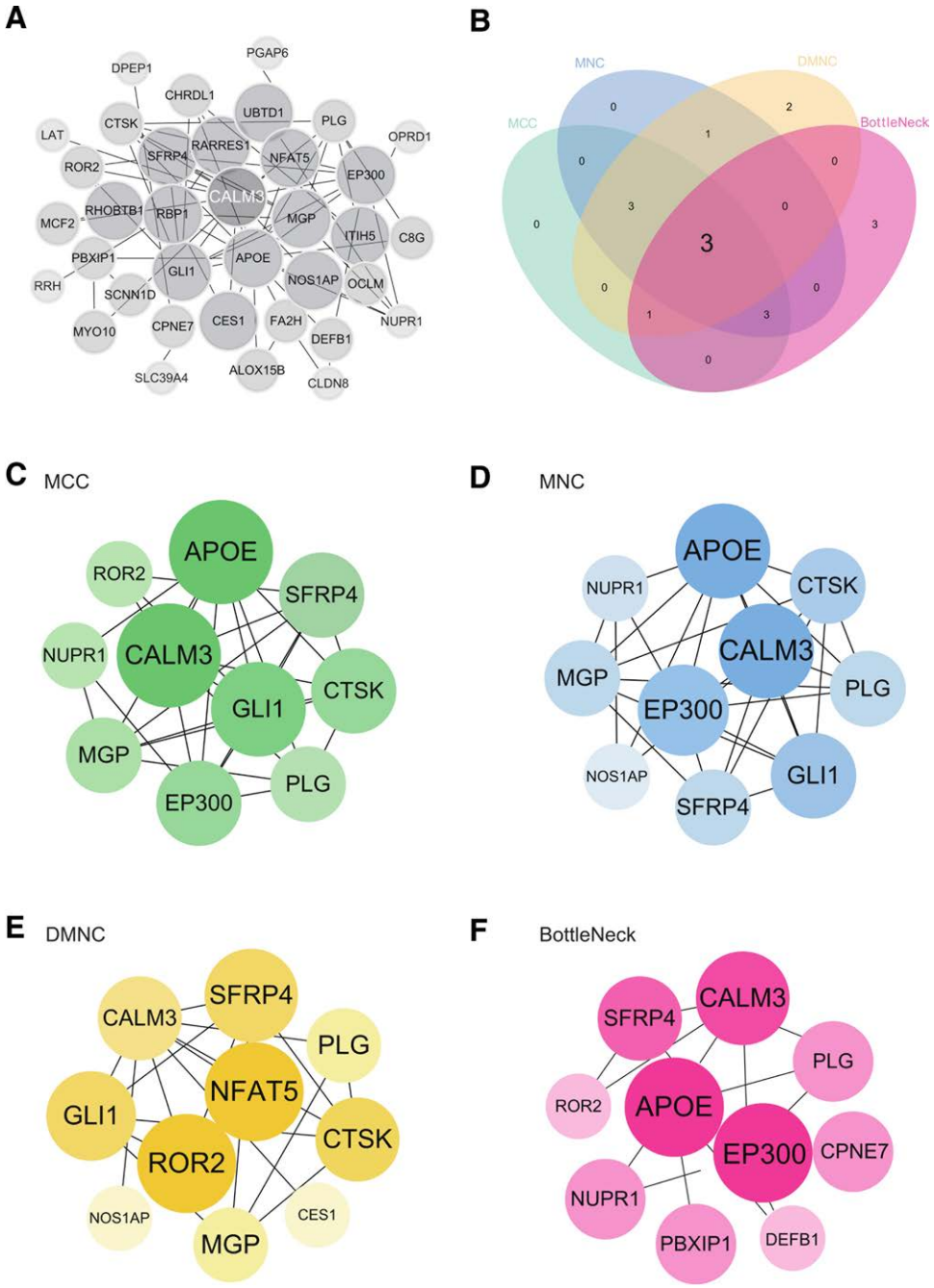

Figure 9. Construction and analysis of protein–protein interaction (PPI) networks. (A) PPI network of DEGs. (B) Core genes (CALM3, SFRP4, PLG) were obtained by merging using Venn diagrams. (C) MCC was used to identify the central gene. (D) MNC was used to identify the central gene. (E) DMNC was used to identify the central gene. (F) BottleNeck was used to identify the central gene. CALM3 = Calmodulin 3, DEGs = differentially expressed genes, DMNC = Density of Maximum Neighborhood Component, MCC = Maximal Clique Centrality, MNC = Maximum Neighborhood Component, PLG = plasminogen, SFRP4 = secreted frizzled-related protein 4.

3.10. miRNA prediction and functional annotation related to target genes

In this study, we input the target gene CALM3 into miRNA prediction websites to find related miRNAs, enhancing our understanding of gene expression regulation (Table 1). The TargetScan website predicted that the miRNA related to CALM3 is hsa-miR-122-5p. The miRTarBase website predicted that the miRNAs related to CALM3 are hsa-miR-1-3p, hsa-miR-7-5p, and hsa-miR-320a.

4. Discussion

Leukemia is a malignant tumor of the hematopoietic system, primarily characterized by the abnormal proliferation of hematopoietic stem cells in the bone marrow, leading to a large number of immature white blood cells entering the bloodstream.[9] Leukemia can be classified into acute and chronic types based on its progression speed and cell type. Acute leukemia progresses rapidly, with patients usually showing severe symptoms within weeks, while chronic leukemia progresses more slowly.[10] Epidemiological data on leukemia show varying incidence rates in different regions and populations. Globally, children and the elderly are the main affected groups. According to a study in the International Journal of Hematology, the epidemiological characteristics of leukemia exhibit significant differences worldwide, especially across different age groups and genders. In the United States, acute lymphoblastic leukemia (ALL) is one of the most common cancers in children, whereas acute myeloid leukemia

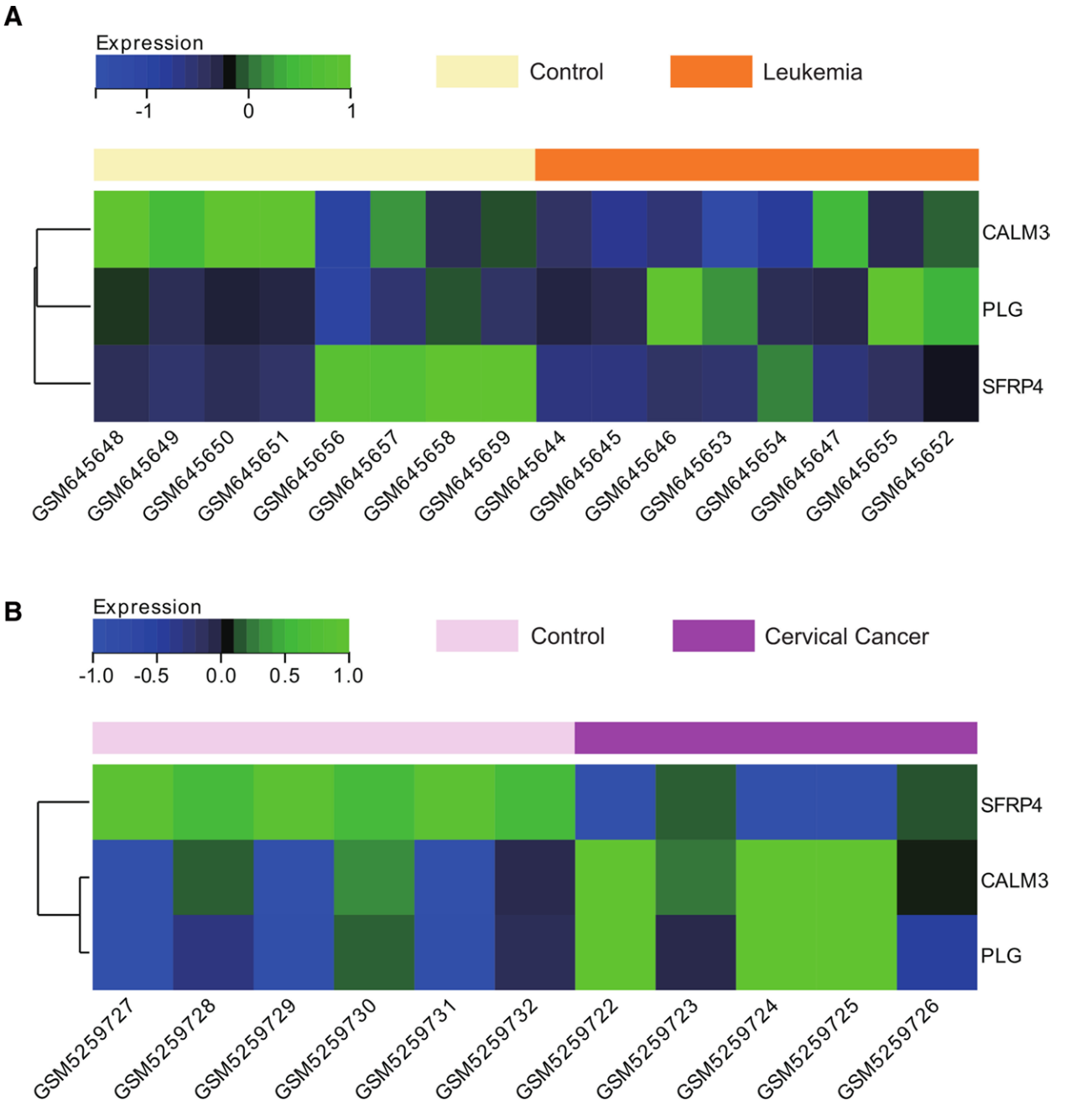

Figure 10. Heat map of gene expression. (A) Heat map of expression of core genes in the matrix of leukemia datasets. (B) Expression heat map of core genes in the matrix of cervical cancer dataset.

(AML) is more common in adults. According to the American Cancer Society, ALL accounts for a high proportion of childhood cancers, while AML mainly affects adults, particularly the elderly.[11,12] Leukemia patients typically present with symptoms such as anemia, bleeding, infection, and bone pain. Pathological features include a significant proliferation of immature white blood cells in the bone marrow and peripheral blood. These abnormal cells not only interfere with the production of normal blood cells but also invade other organs, leading to multisystem damage.[13] According to a study in the Chinese Journal of Hematology, the symptoms and pathological features of leukemia patients include significant proliferation of immature white blood cells in the bone marrow, leading to a series of systemic symptoms.[14] Leukemia is highly hazardous; without timely treatment, the survival period of acute leukemia patients may be only a few months. Studies have shown that patients with acute leukemia (including ALL and AML) have a very short survival period if they do not receive timely and appropriate treatment, with the high aggressiveness of acute leukemia being a major cause of its fatality.[15]

In recent years, significant progress has been made in the study of the molecular mechanisms of leukemia, thanks to advances in molecular biology techniques. Gene mutations, chromosomal translocations, and epigenetic changes are key factors leading to leukemia.[16] In-depth research into the molecular mechanisms of leukemia is crucial for the development of targeted therapies. The discovery of the BCR-ABL fusion gene in chronic myeloid leukemia (CML) has provided an important basis for targeted therapy, with tyrosine kinase inhibitors like imatinib significantly improving the prognosis of CML patients.[17]

Cervical cancer is one of the most common malignancies of the female reproductive system, mainly caused by persistent

Figure 11. CTD analysis. Three core genes (CALM3, SFRP4, PLG) are associated with leukemia, coagulation disorders, vaginal tumors, cervical tumors, autoimmune diseases, and inflammation. CALM3 = calmodulin 3, CTD = Comparative Toxicogenomics Database, PLG = plasminogen, SFRP4 = secreted frizzled-related protein 4.

infection with high-risk human papillomavirus (HPV).[18] After HPV infection, the viral genome integrates into the host cell DNA, leading to abnormal expression of cell cycle regulatory genes such as E6 and E7, which in turn induces malignant transformation of the cells.[19,20] Epidemiological data show higher incidence rates in developing countries, whereas countries implementing HPV vaccination and screening programs have seen a significant decrease in cervical cancer incidence.[21] Cervical cancer patients may present with symptoms such as abnormal vaginal bleeding, contact bleeding, and increased vaginal discharge.[22] The pathological features mainly include cervical intraepithelial neoplasia and invasive cancer, with most cases being squamous cell carcinomas. A key challenge in cervical cancer is that it often presents asymptomatically in its early stages, leading to late diagnosis, complicating treatment, and increasing mortality rates.[23] Regular screening (such as Pap smears and HPV testing) for early detection is crucial for improving survival rates.[24] These screenings can identify precancerous lesions, allowing for intervention before the disease progresses. Therefore, early diagnosis and treatment are critical

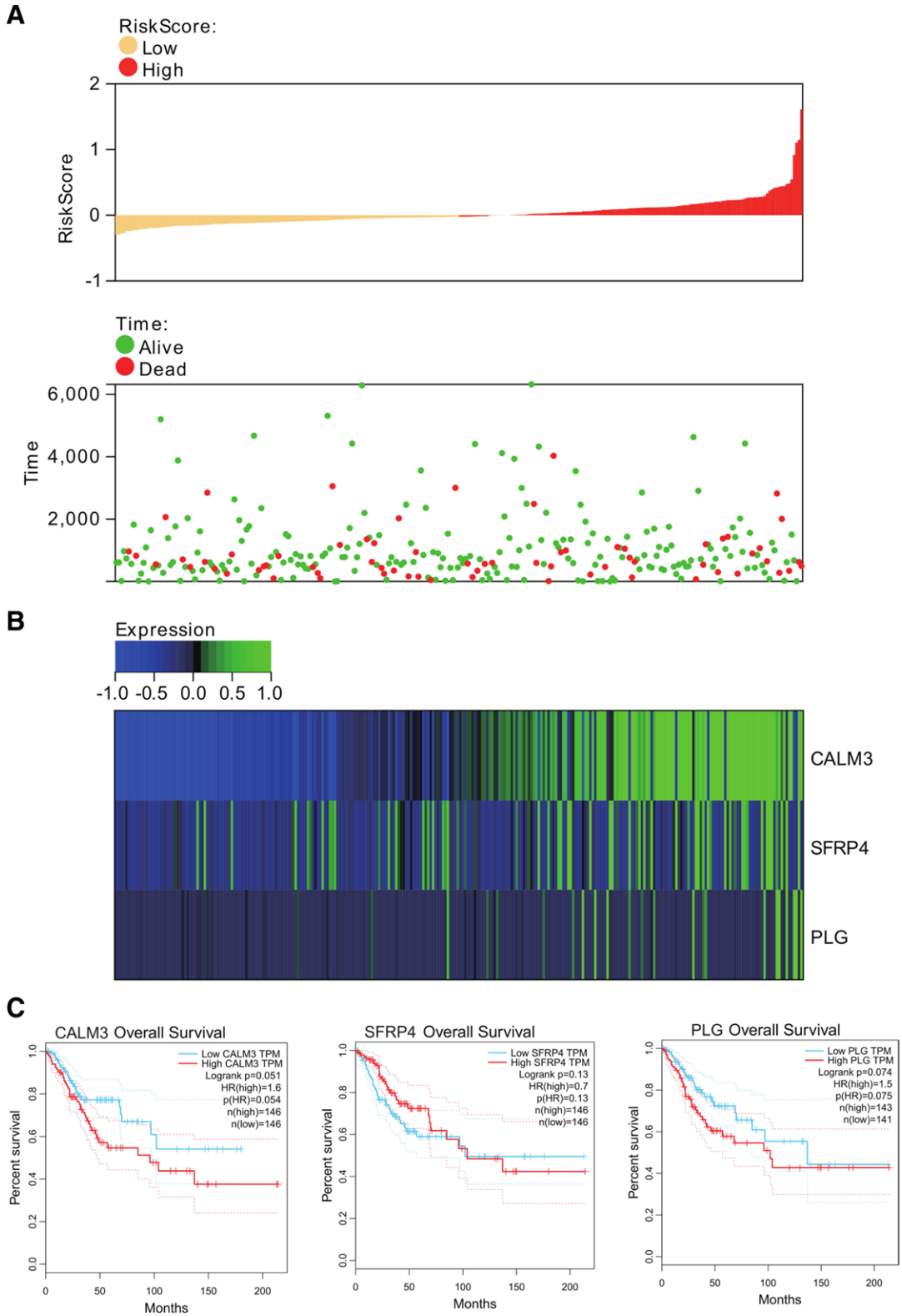

Figure 12. Prognostic survival analysis. (A) Trend graph of risk score in relation to survival time and survival rate. (B) Expression calorimetric map of core genes in liver cancer survival data. (C) Overall survival of core genes in cervical cancer.

for improving the survival rate of cervical cancer patients, and studying its molecular mechanisms is important for discovering new diagnostic markers and therapeutic targets.

The occurrence of cervical cancer is closely related to infection with high-risk types of HPV, particularly HPV16 and HPV18. These viruses integrate their genomes into the host cell DNA, leading to the inactivation of tumor suppressor genes such as p53 and pRb through the expression of E6 and E7 proteins, respectively, promoting uncontrolled cell proliferation and malignant transformation. The E6 protein forms a complex with E6AP (E6-associated protein), inducing ubiquitination and degradation of the p53 protein. p53, known as the "guardian

Table 1

A summary of miRNAs that regulate hub genes.

Databases	Gene	miRNA		
TargetScan	CALM3	hsa-miR-122-5p		
miRTarBase	CALM3	hsa-miR-1-3p	hsa-miR-7-5p	hsa-miR-320a

CALM3 = calmodulin 3.

of the genome," plays a crucial role in responding to DNA damage and other stress conditions; its loss of function allows cells to evade apoptosis and promotes carcinogenesis.[25,26] The E7 protein binds to pRb (retinoblastoma protein), releasing its inhibition of the cell cycle. Normally, pRb prevents cells from entering the S phase (DNA synthesis phase) by binding to E2F transcription factors. The E7 protein promotes pRb degradation, releasing E2F and activating genes associated with the S phase, leading to continuous cell proliferation.[27,28]

Our study shows that CALM3 is downregulated in leukemia and upregulated in cervical cancer. The differential expression of CALM3 in leukemia and cervical cancer suggests its potential role in the development and progression of these diseases. Further investigation into the molecular mechanisms of CALM3 could help elucidate its specific functions in tumor biology and provide important insights for developing new targeted therapeutic strategies.

CALM3 is part of the calmodulin gene family, including CALM1 and CALM2, which encode the same calmodulin protein. Calmodulin, encoded by CALM3, is crucial for regulating various enzymes, ion channels, and other proteins. It interacts with many targets and regulates their activity, including kinases, phosphatases, and cytoskeletal proteins. For example, calmodulin regulates the activity of the Kv7.2 potassium channel, which is essential for maintaining neuronal excitability.[29] Calmodulin is involved in key cellular processes such as muscle contraction, cell division, and signal transduction pathways, mediating processes like inflammation, immune response, apoptosis, and metabolic pathways. By binding to calcium ions, calmodulin undergoes conformational changes that enable it to interact with and regulate its target proteins.[30] CALM3 has been found to be associated with the progression of various diseases. Research indicates that CALM3 plays a key role in the pathological processes of cardiovascular diseases. For example, CALM3 regulates calcium signaling in cardiomyocytes, affecting cardiac contractility and cell survival in conditions like cardiac hypertrophy and heart failure.[31] In Alzheimer disease, altered calmodulin activity affects synaptic plasticity and memory formation.[32] Additionally, one study suggests that CALM3 may play a significant role in the pathogenesis of leukemia.[33]

These studies indicate that CALM3 plays important roles in various physiological and pathological processes in multiple diseases, and its functional abnormalities may serve as potential therapeutic targets. This further supports our findings, which show that CALM3 is downregulated in leukemia and upregulated in cervical cancer. This expression difference highlights CALM3's potential roles in these diseases, providing new clues for further research into its specific mechanisms in disease development and progression, which could aid in developing new diagnostic and therapeutic approaches.

4.1. Limitation and future research directions

Although this study revealed the differential expression of CALM3 in leukemia and cervical cancer through bioinformatics analysis, there are still some limitations. This study did not further validate the specific functions of CALM3 through gene overexpression or knockout experiments in animal models. Therefore, future research should combine in vivo and in vitro experiments to explore the mechanisms of CALM3 in greater depth and its potential applications in disease treatment and prognosis, providing new ideas and evidence for targeted therapies.

5. Conclusion

CALM3 is downregulated in leukemia and upregulated in cervical cancer, potentially acting as a key target in both diseases. This study provides valuable insights for further research on the CALM3 gene in leukemia and cervical cancer, laying a solid foundation for future investigations.

Author contributions

Conceptualization: Danfeng Shao.
Data curation: Huaiying Yu, Xiaoqing Zhu.
Formal analysis: Huaiying Yu, Xiaoqing Zhu.
Methodology: Danfeng Shao.
Software: Huaiying Yu.
Validation: Huaiying Yu.
Visualization: Danfeng Shao, Huaiying Yu, Xiaoqing Zhu.
Writing – review & editing: Danfeng Shao.
Writing – original draft: Xiaoqing Zhu.

References

[1] Juliusson G, Hough R. Leukemia. Prog Tumor Res. 2016;43:87–100.
[2] Whiteley AE, Price TT, Cantelli G, Sipkins DA. Leukaemia: a model metastatic disease. Nat Rev Cancer. 2021;21:461–75.
[3] Schiffman M, Doorbar J, Wentzensen N, et al. Carcinogenic human papillomavirus infection. Nat Rev Dis Primers. 2016;2:16086.
[4] Rahangdale L, Mungo C, O'Connor S, Chibwesha CJ, Brewer NT. Human papillomavirus vaccination and cervical cancer risk. BMJ. 2022;379:e070115.
[5] Vakana E, Platanias LC. AMPK in BCR-ABL expressing leukemias. Regulatory effects and therapeutic implications. Oncotarget. 2011;2:1322–8.
[6] Wrede D, Tidy JA, Crook T, Lane D, Vousden KH. Expression of RB and p53 proteins in HPV-positive and HPV-negative cervical carcinoma cell lines. Mol Carcinog. 1991;4:171–5.
[7] Ji F, Sadreyev RI. RNA-seq: basic bioinformatics analysis. Curr Protoc Mol Biol. 2018;124:e68.
[8] Jorge NA, Ferreira CG, Passetti F. Bioinformatics of cancer ncRNA in high throughput sequencing: present state and challenges. Front Genet. 2012;3:287.
[9] Greaves M. A causal mechanism for childhood acute lymphoblastic leukaemia. Nat Rev Cancer. 2018;18:471–84.
[10] Jurgec S, Jezernik G, Gorenjak M, Büdefeld T, Potočnik U. Meta-analytic comparison of global RNA transcriptomes of acute and chronic myeloid leukemia cells reveals novel gene candidates governing myeloid malignancies. Cancers (Basel). 2022;14:4681.
[11] Yao Y, Lin X, Li F, Jin J, Wang H. The global burden and attributable risk factors of chronic lymphocytic leukemia in 204 countries and territories from 1990 to 2019: analysis based on the global burden of disease study 2019. Biomed Eng Online. 2022;21:4.
[12] Ou Y, Long Y, Ji L, et al. Trends in disease burden of chronic lymphocytic leukemia at the global, regional, and national levels from 1990 to 2019, and projections until 2030: a population-based epidemiologic study. Front Oncol. 2022;12:840616.
[13] Estey EH. Acute myeloid leukemia: 2019 update on risk-stratification and management. Am J Hematol. 2018;93:1267–91.
[14] Lin K, Jia H, Cao M, et al. Epidemiological characteristics of leukemia in China, 2005-2017: a log-linear regression and age-period-cohort analysis. BMC Public Health. 2023;23:1647.
[15] Yi M, Li A, Zhou L, Chu Q, Song Y, Wu K. The global burden and attributable risk factor analysis of acute myeloid leukemia in 195 countries and territories from 1990 to 2017: estimates based on the global burden of disease study 2017. J Hematol Oncol. 2020;13:72.
[16] Ley TJ, Miller C, Ding L, et al.; Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368:2059–74.

[17] Druker BJ, Guilhot F, O'Brien SG, et al.; IRIS Investigators. Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. N Engl J Med. 2006;355:2408–17.
[18] Bosch FX, de Sanjosé S. The epidemiology of human papillomavirus infection and cervical cancer. Dis Markers. 2007;23:213–27.
[19] Hong D, Liu J, Hu Y, et al. Viral E6 is overexpressed via high viral load in invasive cervical cancer with episomal HPV16. BMC Cancer. 2017;17:136.
[20] Yeo-Teh NSL, Ito Y, Jha S. High-risk human papillomaviral oncogenes E6 and E7 target key cellular pathways to achieve oncogenesis. Int J Mol Sci. 2018;19:1706.
[21] Arbyn M, Weiderpass E, Bruni L, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. Lancet Glob Health. 2020;8:e191–203.
[22] Plaisy MK, Boni SP, Coffie PA, et al.; IeDEA West Africa Collaboration. Barriers to early diagnosis of cervical cancer: a mixed-method study in Côte d'Ivoire, West Africa. BMC Womens Health. 2023;23:135.
[23] Kakotkin VV, Semina EV, Zadorkina TG, Agapov MA. Prevention strategies and early diagnosis of cervical cancer: current state and prospects. Diagnostics (Basel). 2023;13:610.
[24] Gates A, Pillay J, Reynolds D, et al. Screening for the prevention and early detection of cervical cancer: protocol for systematic reviews to inform Canadian recommendations. Syst Rev. 2021;10:2.
[25] László B, Antal L, Gyöngyösi E, et al. Coordinated action of human papillomavirus type 16 E6 and E7 oncoproteins on competitive endogenous RNA (ceRNA) network members in primary human keratinocytes. BMC Cancer. 2021;21:673.
[26] Beaudenon S, Huibregtse JM. HPV E6, E6AP and cervical cancer. BMC Biochem. 2008;9(suppl 1):S4.
[27] Pal A, Kundu R. Human papillomavirus E6 and E7: the cervical cancer hallmarks and targets for therapy. Front Microbiol. 2019; 10:3116.
[28] Tomaić V. Functional roles of E6 and E7 oncoproteins in HPV-induced malignancies at diverse anatomical sites. Cancers (Basel). 2016; 8:95.
[29] Alaimo A, Villarroel A. Calmodulin: a multitasking protein in Kv7.2 potassium channel functions. Biomolecules. 2018;8:57.
[30] Vetter SW, Leclerc E. Novel aspects of calmodulin target recognition and activation. Eur J Biochem. 2003;270:404–14.
[31] Bers DM. Calcium cycling and signaling in cardiac myocytes. Annu Rev Physiol. 2008;70:23–49.
[32] Ghosh A, Giese KP. Calcium/calmodulin-dependent kinase II and Alzheimer's disease. Mol Brain. 2015;8:78.
[33] He J, Ni Z, Li Z. CALM3 affects the prognosis of leukemia and hemorrhoids. Medicine (Baltimore). 2023;102:e36027.