



## OPEN ACCESS

## EDITED BY

Yu Pan,  
Macau University of Science and Technology,  
China

## REVIEWED BY

Prajish Iyer,  
City of Hope National Medical Center,  
United States  
Xu Chen,  
Shaanxi Normal University, China  
Sen Zhang,  
University of Illinois Chicago, United States

## \*CORRESPONDENCE

Guoqing Zhu  
✉ zhuguoqing@ihcams.ac.cn  
Yuan Zhou  
✉ yuanzhou@ihcams.ac.cn

†These authors have contributed equally to  
this work

RECEIVED 25 October 2024

ACCEPTED 10 February 2025

PUBLISHED 25 February 2025

## CITATION

Liang H, Cao Z, Ren Y, Li Y, Wang H, Sun F,  
Xue M, Zhu G and Zhou Y (2025) Raman  
spectroscopy and bioinformatics-based  
identification of key genes and pathways  
capable of distinguishing between diffuse  
large B cell lymphoma and chronic  
lymphocytic leukemia.  
*Front. Immunol.* 16:1516946.  
doi: 10.3389/fimmu.2025.1516946

## COPYRIGHT

© 2025 Liang, Cao, Ren, Li, Wang, Sun, Xue,  
Zhu and Zhou. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Raman spectroscopy and bioinformatics-based identification of key genes and pathways capable of distinguishing between diffuse large B cell lymphoma and chronic lymphocytic leukemia

Haoyue Liang<sup>1,2†</sup>, Zhijie Cao<sup>1,2†</sup>, Yansong Ren<sup>1,2†</sup>, Yihan Li<sup>1,2</sup>,  
Haoyu Wang<sup>1,2</sup>, Fanfan Sun<sup>1,2</sup>, Mei Xue<sup>3</sup>, Guoqing Zhu<sup>1,2\*</sup>  
and Yuan Zhou<sup>1,2\*</sup>

<sup>1</sup>State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin, China, <sup>2</sup>Tianjin Institutes of Health Science, Tianjin, China, <sup>3</sup>School of Medical Technology, Tianjin University of Traditional Chinese Medicine, Tianjin, China

Diffuse large B-cell lymphoma (DLBCL) and chronic lymphocytic leukemia (CLL) are subtypes of non-Hodgkin lymphoma (NHL) that are generally distinct from one another, but the transformation of one of these diseases into the other is possible. Some patients with CLL, for instance, have the potential to develop Richter transformation such that they are diagnosed with a rare, invasive DLBCL subtype. In this study, bioinformatics analyses of these two NHL subtypes were conducted, identifying key patterns of gene expression and then experimentally validating the results. Disease-related gene expression datasets from the GEO database were used to identify differentially expressed genes (DEGs) and DEG functions were examined using GO analysis and protein-protein interaction network construction. This strategy revealed many up- and down-regulated DEGs, with functional enrichment analyses identifying these genes as being closely associated with inflammatory and immune response activity. PPI network analyses and the evaluation of clustered network modules indicated the top 10 up- and down-regulated genes involved in disease onset and development. Serological analyses revealed significantly higher ALB, TT, and WBC levels in CLL patients relative to DLBCL patients, whereas the opposite was true with respect to TG, HDL, GGT, ALP, ALT, and NEUT% levels. In comparison to the CLL and DLBCL groups, the healthy control samples demonstrated higher signals of protein peak positions (621, 643, 848, 853, 869, 935, 1003, 1031, 1221, 1230, 1260, 1344, 1443, 1446, 1548, 1579, 1603, 1647  $\text{cm}^{-1}$ ), nucleic acid peak positions (726, 781, 786, 1078, 1190, 1415, 1573, 1579  $\text{cm}^{-1}$ ), beta carotene peak positions (957, 1155, 1162  $\text{cm}^{-1}$ ), carbohydrate peak positions (842  $\text{cm}^{-1}$ ), collagen peak positions (1345  $\text{cm}^{-1}$ ), and lipid peak positions (957, 1078, 1119, 1285, 1299, 1437, 1443, 1446  $\text{cm}^{-1}$ ) compared to the CLL and DLBCL groups. Verification of these key genes in patient samples yielded results consistent with

findings derived from bioinformatics analyses, highlighting their relevance to diagnosing and treating these forms of NHL. Together, these analyses identified genes and pathways involved in both DLBCL and CLL. The set of molecular markers established herein can aid in patient diagnosis and prognostic evaluation, providing a valuable foundation for their therapeutic application.

#### KEYWORDS

**DLBCL, CLL, bioinformatics analysis, Raman spectroscopy, protein-protein interaction networks**

## Introduction

Diffuse large B-cell lymphoma (DLBCL) and chronic lymphocytic leukemia (CLL) are both subtypes of non-Hodgkin lymphoma (NHL). DLBCL, the most common form of NHL, is characterized by lymph node enlargement and potential organ invasion. DLBCL cases originate from mature B cells of germinal center or non-germinal center origin (1–3). The classification of DLBCL cases is based on the biological features of these cells of origin and the overall clinical disease course. ABC-DLBCL, for instance, has a poorer prognosis than GCB-DLBCL. In ABC-DLBCL tumor cells, prolonged B cell receptor (BCR) signaling leads to high levels of NF- $\kappa$ B pathway activity, whereas in GCB-DLBCL tumor cells, BCR signaling contributes to high levels of PI3K signaling activity (4, 5). The characteristics of DLBCL tumor cells influence both tumor progression and overall host immune function. Certain proteins expressed specifically within the central nervous system (CNS) can reportedly bind to the BCRs of DLBCL tumor cells, contributing to oncogenic progression (6). In one study, a bias toward IGV rearrangement was noted in PCNSL tumor cells, with the particular prioritization of IGHV4-34, enabling the recognition of galectin-3 expression on various CNS cells (7). Kawano et al. (8) described a series of primary adrenal DLBCL cases, among which 9 were positive for EBV and 3 had a history of tumors. This provides further support for potential excessive B cell infiltration in particular immunological niches in the context of chemotherapy-induced cancer-related immune dysfunction, as has been reported in EBVMCU cases (9). In their analysis of 62 primary intestinal DLBCL cases, Ishikawa et al. identified 10 cases that were EBV-positive, almost all of which presented with features of immunosuppression and clinical and pathological features overlapping with those of immunodeficient PCNSL (10). In stark contrast, cases of primary gastric EBV-positive DLBCL often present even without any apparent evidence of immunosuppression (11, 12). While the pathogenesis of lymphomas arising in immune-privileged sites appears to be linked with organ-specific immune function and systemic immune activity, the precise biological mechanisms underlying these oncogenic processes remain incompletely understood.

Unlike DLBCL, CLL tends to primarily affect elderly patients, often progresses slowly, and may be asymptomatic during its early

stages (13). CLL cases originate from the clonal expansion of abnormal populations of B cells expressing markers such as CD19 and CD20, with these abnormal lymphocyte populations ultimately accumulating in the blood, bone marrow, and lymph nodes. As a class of indolent lymphoma, tumor invasion in CLL cases tends to be relatively limited, and the overall survival of affected patients tends to be relatively long. Even with early treatment, significant improvements in clinical outcomes are rare such that conservative observation is generally adopted in most CLL cases (14). The approach to treating CLL is generally dependent on the stage of disease and overall patient health status and can include observation, immunotherapy, targeted therapy, or immunotherapeutic interventions. As a highly heterogeneous disease, certain patients present with a less aggressive form of disease, particularly in cases characterized by a lack of immunoglobulin heavy chain (IGHV) gene mutations, and mutations in the del (17p), del (11q), and TP53 genes (15). CD38 and ZAP-70 are regarded as being indicative of a poor prognosis in immunophenotyping assays (16). Specifically, CD38 is related to IGHV mutational status and may also offer some independent prognostic utility, whereas ZAP-70 is an intracellular protein normally produced in T cells that is commonly dysregulated in CLL cells from certain patients such that it is relevant to patient prognosis (17). CLL progression, in addition to being related to malignant clone characteristics, is also associated with severe immunodeficiency, as these conditions are conducive to the evasion of host immune-mediated detection and elimination. This highlights the closely interrelated, co-evolving nature of tumor cells and their microenvironment over the course of disease progression (18). Monocyte-derived nurse-like cells (NLCs), NK cells, T cells, NKT cells, and mesenchymal stromal cells are all vital environmental components that communicate with CLL cells through a series of chemokine receptors, tumor necrosis factor (TNF) family proteins, adhesion molecules, and soluble mediators (19). CLL can promote the outgrowth of immune cells that suppress immune activity such as regulatory T cells and myeloid-derived suppressor cells, thereby facilitating immune evasion (20). CLL clones often present with characteristics consistent with those of regulatory B cells (Bregs). Indeed, the phenotypes of leukemic B cells and Bregs are often similar, including CD5, CD24, and CD27 expression together with low surface IgM levels. The physiological similarities of these cells also extend to IL-10 production, consistent with their ability to serve as negative

regulators of T cell activation and immunity (21). The CD40-CD40L interactions that take place between leukemic B cells and activated CD4+ T cells can promote CLL cell proliferation and anti-apoptotic protein upregulation. T cells can also promote leukemic cell survival through the upregulation of anti-apoptotic Bcl-2 in a process mediated by secreted proteins including IL-4 and IFN- $\gamma$  (22). Peripheral T cell numbers are raised in patients with CLL, primarily owing to increased CD8+ T cell counts, resulting in decreased CD4:CD8 ratios. As a consequence, even though the overall T cell count is higher, these T cells tend to exhibit reduced functionality including impaired immune synapse formation, limited cytokine production, insufficient degranulation, and decreased cytotoxic antitumor activity (23). These CLL-related T cells also exhibit the overexpression of PD-1 and other marker proteins associated with chronic activation and exhaustion, contributing to the further impairment of their immune synapse formation and cytotoxic activity (24). Key malignant cell features and their disruption of the associated microenvironment additionally contribute to therapeutic resistance in patients. While studies of the CLL-associated microenvironment and BCR signaling have led to the design of targeted therapies including the BTK inhibitor ibrutinib, there are still challenges that remain to be addressed (25). The efficacy of novel immunotherapies has been very limited in CLL as a consequence of the lack of effector T cells in these patients (26). Additional research is thus essential to establish better approaches to treating this form of cancer.

While abnormal B cell growth is a hallmark of both CLL and DLBCL, the two diseases are not causally linked. Indeed, each is characterized by distinct genetic determinants, risk factors, and clinical features. In some instances, patients may nonetheless transition between CLL and DLBCL, as in the rare cases of CLL patients who develop Richter transformation, leading to a more invasive DLBCL diagnosis that tends to progress rapidly while exhibiting biological features and therapeutic responsiveness in line with the features of primary DLBCL. The etiological basis for Richter transformation remains incompletely understood, but it is thought to be linked to genetic factors, microenvironmental conditions, and immune escape. Clarifying the differences between these two forms of NHL and the transitions that can arise between the two is vital to optimizing the treatment of affected patients and improving their outcomes.

In this study, a bioinformatics approach was utilized for the identification of key differentially expressed genes (DEGs) between DLBCL and CLL patient samples to assist cancer diagnosis and management. These efforts are valuable, and DEGs can plausibly serve as diagnostic biomarkers suitable for accurately distinguishing between DLBCL and CLL in the clinic in the early stages of diseases, ensuring that clinicians are able to accurately select an appropriate treatment plan given that the two diseases require very different therapeutic approaches. Some of these DEGs may also offer prognostic value, supporting the precise formulation of personalized treatment plans. DEGs are also ideal candidate targets for new or existing therapeutic agents. If a given gene is overexpressed in DLBCL but absent in CLL, for instance, then any drug targeting that gene would exhibit differential efficacy between these two cancers. On the whole, studies of these DEGs will assist

the understanding of the mechanisms underlying DLBCL and CLL pathogenesis, benefitting overall research focused on the targeted treatment of these and other diseases.

The implementation of bioinformatics analyses and Raman spectroscopy-based approaches to identifying DEGs through the comparison of DLBCL and CLL samples is thus invaluable as a means of improving diagnostic accuracy for these two forms of disease, supporting personalized treatment, novel drug development, and a greater overall understanding of the mechanistic basis for these diseases.

## Materials and methods

### Data source

Data were downloaded from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>) and included the GSE57083 dataset (species: *Homo sapiens*, 20 samples collected with the GPL570 Affymetrix HG-U133-Plus\_2 Array platform) and the GSE68950 dataset (species: *Homo sapiens*, 7 total samples collected with the GPL3921 Affymetrix HG-U133A Array platform).

### Gene expression analyses

Data (CEL format) were analyzed with R (v 3.6.2) to perform the background correction and normalization of these expression data, converting the format, adding missing values, performing background correction with the MAS method, and standardizing data using quantiles. DLBCL data were separated into the DLBCL and CLL groups to screen for DEGs with unpaired t-tests performed with the limma package, using Benjamini and Hochberg (BH)-corrected P values. DEGs were defined by a  $P < 0.05$  and a  $|\log_{2}FC| > 1$ . DEG heatmaps were developed using the pheatmap package in R.

### Functional enrichment

The online DAVID tool (<https://david.ncifcrf.gov/>) was used for DEG annotation in the GO biological process (GO-BP), cellular component (GO-CC), and molecular function (GO-MF) categories. KEGG pathway enrichment was also performed.

### Protein-protein interaction networks

STRING 11.0 (<https://string-db.org/>) was utilized for the construction of the DEG PPI network, followed by the use of Cytoscape to compute the degree centrality scores for the nodes, with higher scores being indicative of more critical nodes. The functions of the 10 top DEGs in the network were further evaluated as potential key nodes.

## Sample collection

Serum biochemical analyses were performed with patients recruited from the Hematology Hospital of the Chinese Academy of Medical Sciences (Institute of Hematology, Chinese Academy of Medical Sciences). These 56 patients included 18 DLBCL patients (8 female, 10 male; age range: 25-71 years) and 38 CLL patients (13 female, 25 male; age range: 35-76 years) (Supplementary Table S1). Raman spectroscopy, ELISA and Real-time fluorescence quantitative PCR analyses were performed using blood samples from patients who visited the Hematology Hospital of the Chinese Academy of Medical Sciences from January 2024 to April 2024, including 5 men and 4 women between 31 and 73 years. For all patients, bone marrow puncture was used to assess bone marrow cell morphology, conduct flow cytometry analyses, perform electron microscopy, and conduct fusion gene, chromosome, and tissue cell chemistry analyses. Patient clinical diagnoses were based on these results and were confirmed according to FAB. These patients included 4 CLL patients, 5 DLBCL patients, and 5 controls. The Ethics Committee (KT2020016-EC-2) of the Hematology Hospital of the Chinese Academy of Medical Sciences approved this study. Routine serum biochemical testing was performed for all patients in the Clinical Testing Center of the Hematology Hospital of the Chinese Academy of Medical Sciences. A study overview is illustrated in Figure 1.

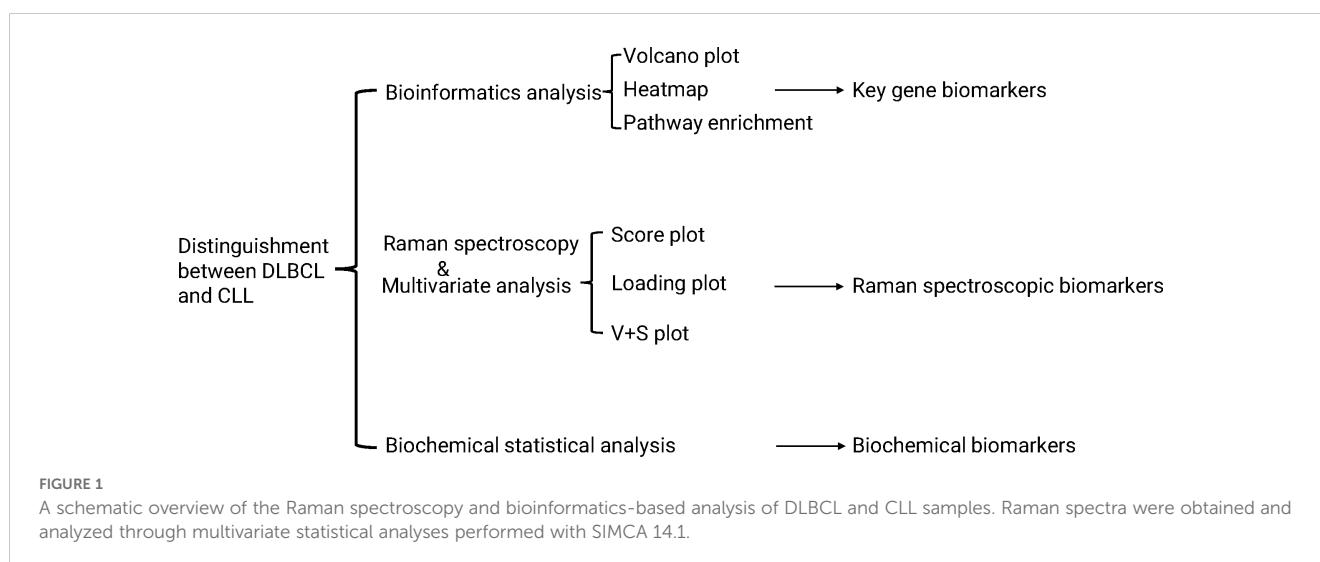
## Clinical data collection

For each participant, medical history, clinical data, and blood-related biomarkers were collected. Briefly, after 10 h of fasting, blood was collected and biochemical indices in the sera were quantified with a fully automated biochemical analyzer. Analyzed parameters included white blood cells (WBC), platelets (PLT), hemoglobin (HGB), neutrophil percentage (NEU%), absolute neutrophil count (NEU #), lymphocyte percentage (LYMPH%), absolute lymphocyte count (LYMPH #), total protein (TP), albumin

(ALB), globulin (GLB), alanine transaminase (ALT), aspartate transaminase (AST), alkaline phosphatase (ALP), glutamyl transpeptidase (GGT), total bile acid (TBA), total bilirubin (TBIL), direct bilirubin (DBIL), urea (UREA), creatinine (CREA), uric acid (UA), lactate dehydrogenase (LDH), creatine kinase (CK), CK isoenzyme (CK-MB), 5-hydroxybutyrate dehydrogenase (α-HBDH), amylase (AMY), potassium (K), sodium (Na), chlorine (Cl), calcium (Ca), phosphorus (P), magnesium (Mg), carbon dioxide combining power (CO<sub>2</sub>CP), glucose (GLU), triglycerides (TG), total cholesterol (TC), high/low-density lipoprotein (HDL and LDL), folic acid (FA), vitamin B12 (B12), ferritin (F), erythropoietin (TPO), iron (Iron), unsaturated iron (UIBC), total iron binding force (TIBC), iron saturation (ISAT), prothrombin time (PT), international standardized ratio (INR), partial thromboplastin time (APTT), thrombin time (TT), fibrinogen (FIB) antithrombin III activity (ATI I), fibrinogen decomposition products (FDB), D-dimer quantification (DD), immunoglobulin G (IgG), immunoglobulin A (IgA), immunoglobulin M (Ig), complement C3 (C3), complement C4 (C4), C-reactive protein (CRP), and rheumatoid factor Anti Streptolysin O (ASO) levels.

## Raman spectroscopy

After adding 5 μL of serum to calcium fluoride or quartz slides, samples were analyzed with a confocal Raman spectrometer XploRA Raman microscope using a 40x objective lens, a 785 nm excitation laser, and a 10 mW output power with the sample fixed on the XYZ 3D platform. Imaging was performed with a 40 0.6 NA Nikon lens, with a spot size range of  $\sim 2 \times 2 \mu\text{m}$  exposed to the laser beam. Measurements were taken in the 600-1800  $\text{cm}^{-1}$  range, with 6 measurement points per group at a 1  $\text{cm}^{-1}$  resolution. Raman spectra for quartz glass slides were analyzed as the background signal. Smoothing, background correction, and baseline correction were performed with the Labspec6 software. Intensity values for all spectra were normalized based on the 1450  $\text{cm}^{-1}$  Raman peaks, which were used as internal standards.



## Clinical and Raman spectroscopy data analyses and model development and establishment

OPLS-DA analyses of clinical data from the CLL and DLBCL sample groups and the Raman spectral data from the control, CLL, and DLBCL sample groups were performed with the SIMCA14.1 software. OPLS model performance was assessed based on the  $R^2$  and  $Q^2$  parameters. Two hundred resamplings were performed by randomly changing the  $y$  matrix. Cluster analyses were conducted and ROC curves were plotted. The Raman peaks that differed significantly in the classification model were considered candidate biomarkers using V+S analysis, selecting those peak positions with variable importance (VIP)  $> 0.5$  and a correlation coefficient (or distance from the center in the V+S plot) within the range of equivalent front positions. The significance of the candidate biomarkers was assessed using  $P < 0.05$ . Origin software was utilized for analysis.

## Measurement of serum IL-15 levels by ELISA

Blood samples were collected from normal donors, CLL patients, and DLBCL patients. Sera were obtained after centrifugation at 1000 g for 10 min and the IL-15 concentrations in the samples were measured using a human IL-15 ELISA kit (Abclonal Technology, Woburn, MA, USA). A Spark microplate reader (Tecan, Switzerland) was used to measure absorbance at 450 nm and the IL-15 concentrations in the samples were calculated from a standard curve.

## Real-time fluorescence quantitative PCR

Total RNA was extracted from peripheral blood mononuclear cells using an RNA assay kit and was reverse-transcribed to cDNA. Fluorescence quantitative PCR was used to assess the expression of IL-15 using primers designed and synthesized by Beijing Qingke Biotechnology Co., Ltd. (Supplementary Table S16). The 20  $\mu$ L reaction system was composed according to the instructions of TaKaRa's fluorescent quantitative reagent. The threshold cycle number (Ct) values of IL15 and GAPDH in the samples were obtained using the sample fitting method and, using GAPDH as the internal reference and the relative quantification (RQ) values, the relative mRNA expression of IL15 and PIK3CG was determined in the tissues. The RQ of the target gene was calculated and statistically analyzed based on  $RQ = 2^{-\Delta\Delta Ct}$ . The qPCR data was normalized using the  $2^{-\Delta\Delta Ct}$  method, setting the control group as the baseline (expression level = 1). The specificity of the PCR products was analyzed using dissolution curves to eliminate the influence of primer dimers and non-specific amplification products. The reaction conditions were pre-denaturation at 90°C for 10 seconds, followed by denaturation at 95°C for 5 seconds, and extension at 60°C for 40 seconds, for a total of 40 cycles. A blank control was used for each PCR reaction. The PCR products were analyzed using 2% agarose gel electrophoresis.

## Statistical analyses

Data were analyzed with SPSS 27.0 and GraphPad Prism 9.0. Spectral data were independent of one another. Normally distributed data are presented as means  $\pm$  SE (standard error) and were compared via one-way ANOVAs with Bonferroni's multiple comparisons test when homogeneously distributed, whereas they were compared Groups with Welch's ANOVA and Tamhane's T2 multiple comparisons test when heterogeneous variance. When non-normally distributed, the results are given as medians (interquartile range) and were compared using Kruskal-Wallis tests and Dunn's multiple comparisons tests.

Clinical data are represented as described above, respectively comparing continuous data that were normally distributed with and without equal variance using unpaired t-tests and Welch's unpaired t-test. Non-normally distributed continuous data were compared with Mann-Whitney U tests. Categorical data were compared with chi-square and Fisher's exact tests.  $P < 0.05$  was considered significant.

## Results

### DEG identification

A total of 715 DEGs (195 upregulated, 520 down-regulated) were identified in the GSE57083 dataset (Supplementary Tables S4, S5), whole 1,043 DEGs (438 upregulated, 605 down-regulated) were detected in the GSE68950 dataset (Supplementary Tables S6, S7). Of these DEGs, 179 were shared between GSE57083 and GSE68950 (39 upregulated, 140 down-regulated) (Supplementary Tables S8, S9). The distributions of these genes were reported using volcano plots (Figures 2A, C) and heatmaps (Figures 2B, D).

### Functional enrichment analyses

The functions of the overlapping DEGs were examined using the DAVID platform, leading to the identification of 99 enriched GO-BP pathways including the immune response, inflammatory response, cellular response to tumor necrosis factor, cellular response to interference gamma, and positive regulation of GTPase activity pathways. Moreover, DEG enrichment was found in 30 GO-CC terms including the cytoskeleton, plasma membrane, endothelial reticulum lumen, cytochrome, and focal adhesion terms, together with 20 GO-MF pathways of which the most strongly enriched were the cytochrome activity, protein binding, cytochrome binding, actin filament binding, and CCR chemokine receptor binding terms. These DEGs were also enriched in 11 KEGG pathways, the most strongly enriched of which were the Cytokine Cytokine Receptor Interaction, Viral Protein Interaction with Cytokine and Cytokine Receptors, Amoebiasis, NF- $\kappa$ B Signaling, and African Trypanosomatosis pathways (Figure 3, Supplementary Table S10).

When specifically focusing on upregulated DEGs, 8 enriched GO-BP terms were detected including the inflammatory response,

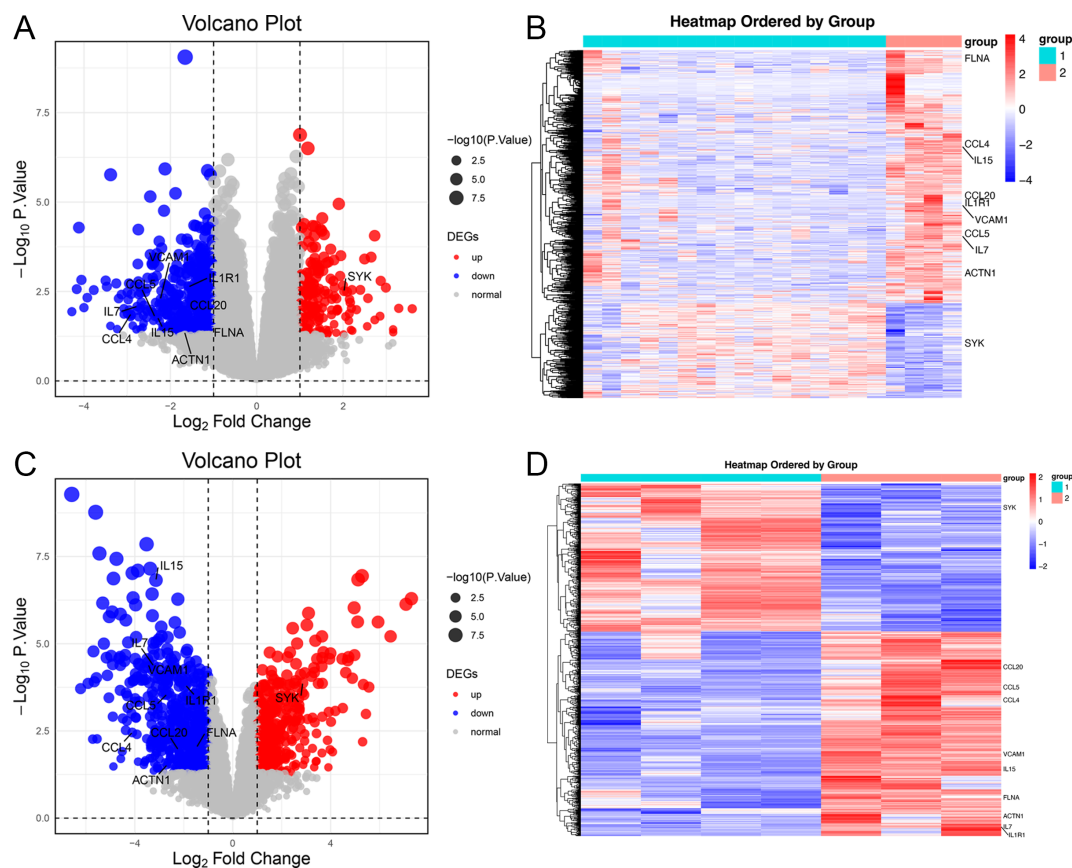


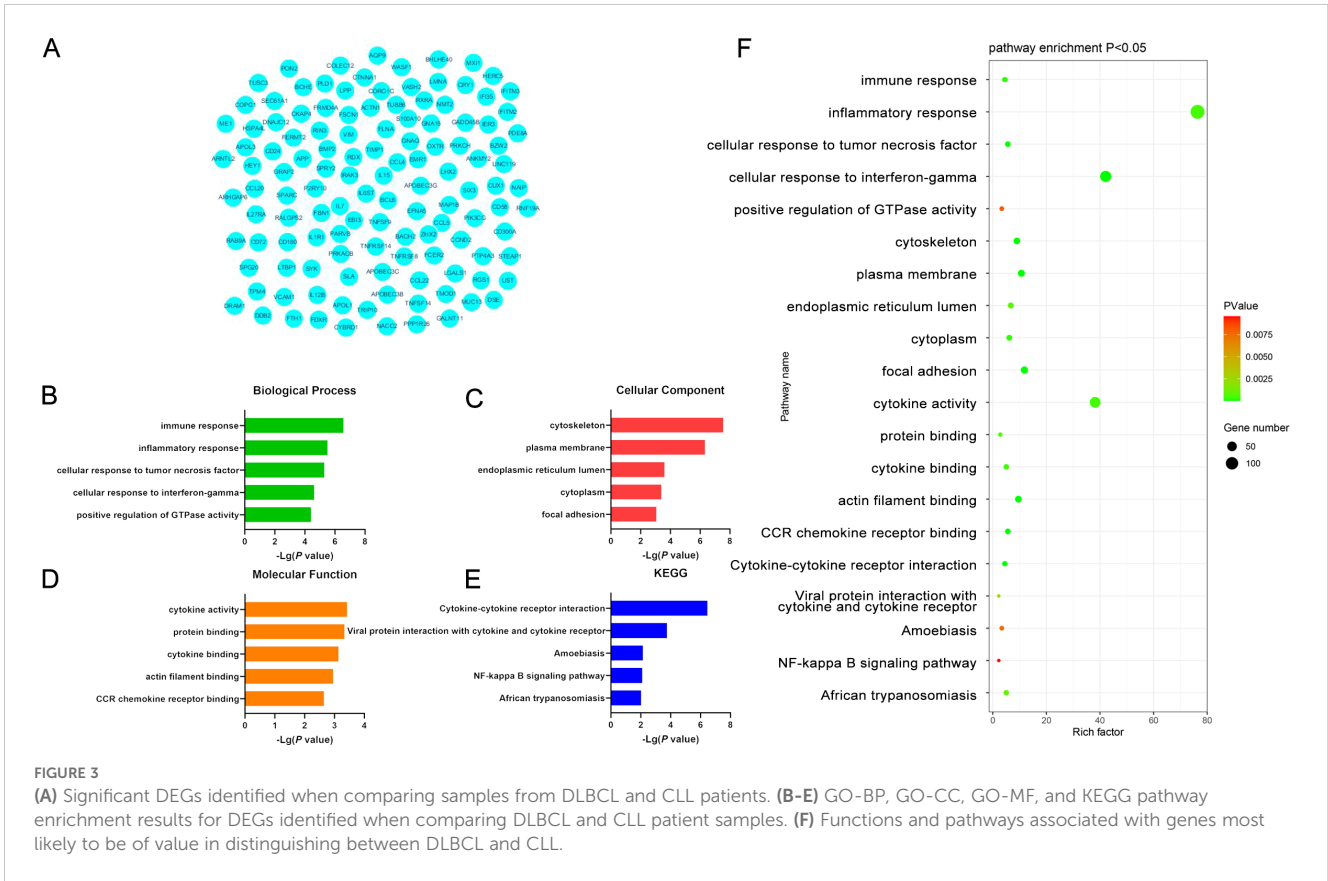
FIGURE 2

Heatmaps and volcano plots for the GSE57083 and GSE68950 datasets. (A, C) Volcano plots for DEGs in GSE57083 (A) and GSE68950 (C), with symbols representing individual genes and red or blue respectively denoting up-regulation or down-regulation.  $P < 0.05$  was the significance threshold, and a fold-change value of 1 was established as the threshold for DEG identification. Key genes such as IL15, CCL4, FLNA, CCL5, IL7, VCAM1, ACTN1, CCL20 and IL1R1 were significantly downregulated in DLBCL, indicating their role in modulating immune response and positive regulation of GTPase activity. Key genes such as SYK were significantly upregulated in DLBCL, indicating their role in protein phosphorylation process. (B, D) Heatmaps for DEGs in GSE57083 (B) and GSE68950 (D), with rows representing genes and columns representing biological samples colored based on grouping.

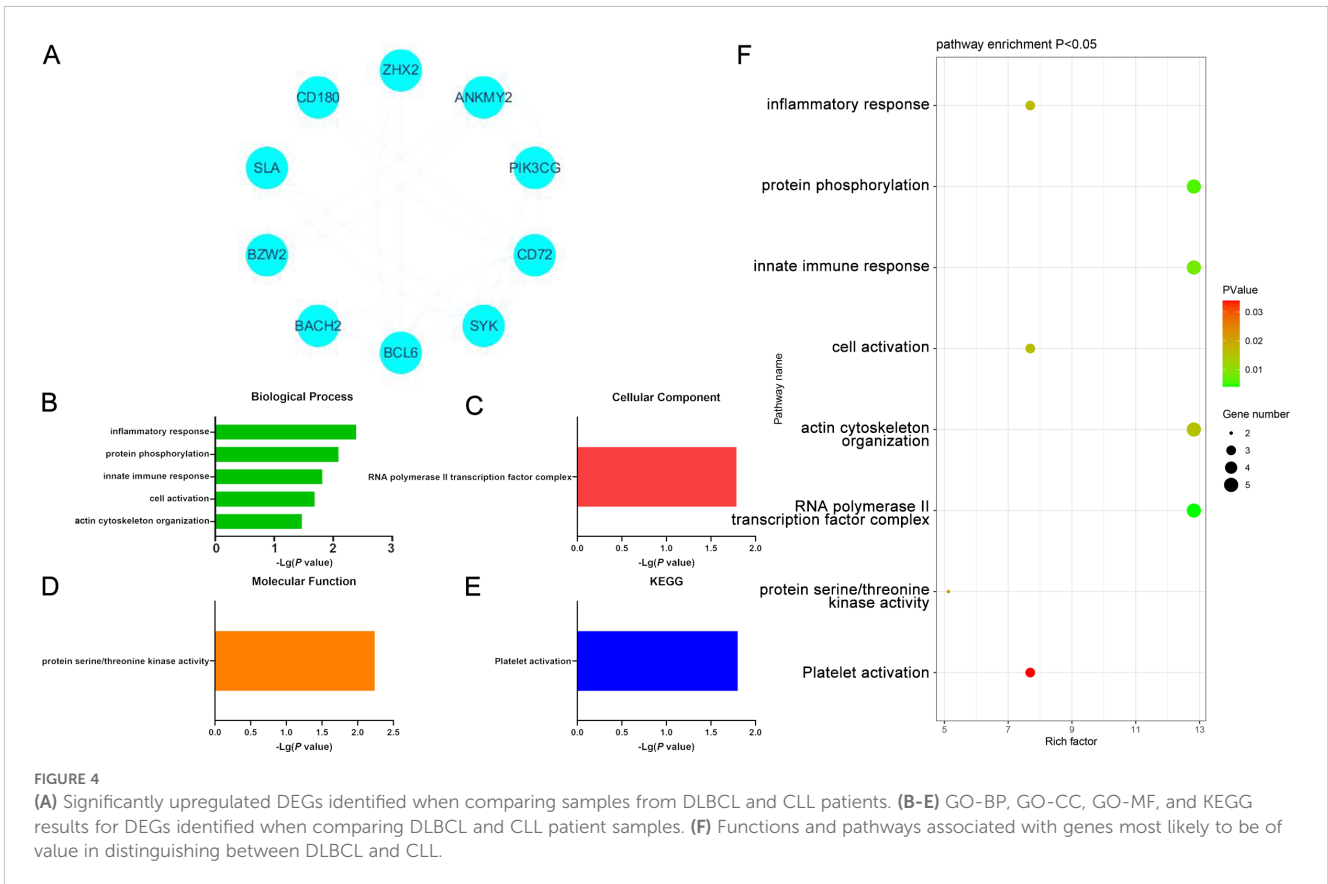
protein phosphorylation response, innate immune response, cell activation, cytoskeleton, and GTP protein-related response pathways. The only enriched GO-CC term was RNA polymerase II transcription factor complex, while the only enriched GO-MF term was the significant upregulation of protein serine/threonine kinase activity. Protein kinases are key mediators of processes like cell cycle progression, metabolic regulation, and cellular signaling, suggesting that this result may be associated with these functions. These DEGs were also enriched in the platelet activation KEGG pathway, potentially implicating it as a differential mediator of the development of these two forms of cancer. Platelets are closely associated with coagulation, immune activity, and inflammation. These upregulated DEGs may thus play a vital role in driving tumor cell activity in DLBCL (Figure 4, Supplementary Table S11).

Similar analyses were also conducted for significantly down-regulated DEGs identified when comparing DLBCL and CLL samples, yielding 90 enriched GO-BP terms including the immune response, cellular response to TNF and IFN- $\gamma$ , positive regulation of ERK1 and ERK2 signaling pathways, and lymphocyte chemotaxis pathways. Moreover, 30 GO-CC terms were enriched for these DEGs,

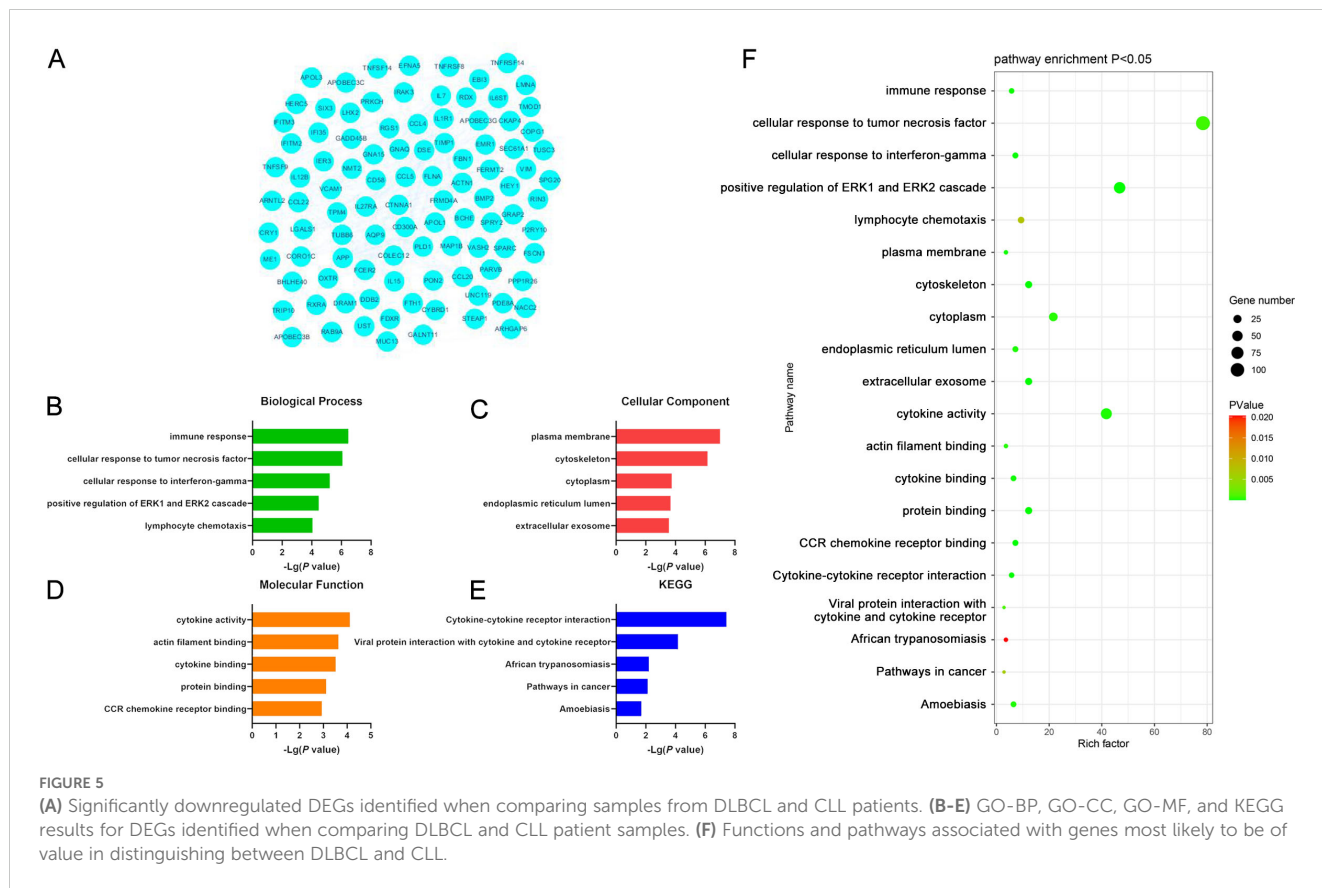
including the plasma membrane, cytoskeleton, cytoplasmic membrane, endothelial reticulum lumen, and extracellular exosome terms that may play roles related to structural preservation, signal transduction, and material transport. Additionally, 20 enriched GO-MF terms were identified including the cytokine activity, actin filament binding, cytokine binding, protein binding, and CCR chemokine receptor binding pathways potentially associated with cytoskeletal stabilization, the release of cytokines, and interactions between cells. These molecular functions are also relevant to cellular migration, immune activity, and inflammatory response regulation. These downregulated DEGs were also enriched in 9 KEGG pathways including the Cytokine Cytokine Receptor Interaction, Viral Protein Interaction with Cytokine and Cytokine Receptors, African Trypanosomatosis, Pathways in Cancer, and Amoebiasis pathways. These enriched pathways may be related to oncogenic progression, cytokine signaling, viral infections, and the pathogenesis of amoebiasis and other forms of disease. In CLL, the genes enriched in these pathways may thus be closely associated with the emergence of malignant cells, their development, and the modulation of normal physiological activity (Figure 5, Supplementary Table S12).



**FIGURE 3** (A) Significant DEGs identified when comparing samples from DLBCL and CLL patients. (B-E) GO-BP, GO-CC, GO-MF, and KEGG pathway enrichment results for DEGs identified when comparing DLBCL and CLL patient samples. (F) Functions and pathways associated with genes most likely to be of value in distinguishing between DLBCL and CLL.



**FIGURE 4** (A) Significantly upregulated DEGs identified when comparing samples from DLBCL and CLL patients. (B-E) GO-BP, GO-CC, GO-MF, and KEGG results for DEGs identified when comparing DLBCL and CLL patient samples. (F) Functions and pathways associated with genes most likely to be of value in distinguishing between DLBCL and CLL.



### PPI network analyses

A PPI network was next constructed using the DEGs overlapping between the GSE57083 and GSE68950 datasets incorporating 125 nodes and 418 edges (Figure 3A). The network had a high topology score, allowing for the identification of the key nodes within the network. The nodes with the top 10 degree values are presented in Supplementary Table S13. After visualizing this PPI network using STRING and the Cytoscape plugin, 10 key DEGs were selected including IL15, CCL4, FLNA, CCL5, IL7, VCAM1, SYK, ACTN1, CCL20, and IL1R1. A majority of these DEGs were classified as down-regulated in the above analyses, indicating markedly higher expression in CLL samples. The PPI network constructed for upregulated DEGs (10 nodes, 20 edges) is presented in Figure 4A, and the top 10 key DEGs identified in this network were ANKMY2, PIK3CG, BZW2, BACH2, BCL6, ZHX2, SYK, CD72, CD180, and SLA (Supplementary Table S14). The PPI network constructed for downregulated DEGs (100 nodes, 332 edges) is presented in Figure 5A, and the top 10 key DEGs identified in this network were IL15, VCAM1, FLNA, CCL5, CCL4, IL7, CCL20, IL1R1, ACTN1, and APP (Supplementary Table S15). qRT-PCR was performed to verify the mRNA expression levels of IL15 and PIK3CG. It was found that IL15 was strongly expressed in the peripheral blood of CLL patients, significantly higher than its expression in the peripheral blood of DCBCL patients. In contrast, PIK3CG showed significantly higher expression in the peripheral blood of DCBCL patients than in patients with CLL (Figure 6A).

The expression levels of both genes were consistent with the bioinformatics results. The IL15 concentrations in the sera of DLBCL and CLL patients were measured by ELISA, showing that its levels were significantly higher in patients with CLL compared with DCBCL patients (Figure 6B).

### Establishment of a multi-parameter approach to distinguishing between DLBCL and CLL

Next, a multi-parameter analysis approach was implemented to differentiate between DLBCL and CLL samples by combining these two groups into a single dataset. The SIMCA-P application was then used for supervised orthogonal partial least squares discriminant analysis (OPLS-DA) of the sample data, enabling detailed analyses and comparisons. Permutation plots (Figure 7A), clustering analysis plots (Figure 7B), and ROC plots (Figure 7C) were utilized to assess the discriminatory performance of the supervised OPLS-DA models. Permutation was also used to assess model validity, with permutation analyses revealing that the Q<sup>2</sup> Y-intercept was negative, consistent with a valid model free of overfitting (Figure 7A). Clustering analyses also confirmed that this model exhibits good discriminatory performance on each set of samples. The generated ROC curves yielded AUC values of exactly 1 for both DLBCL and CLL (Figure 7C), indicative of high accuracy and the ability to reliably distinguish between these two cancers.



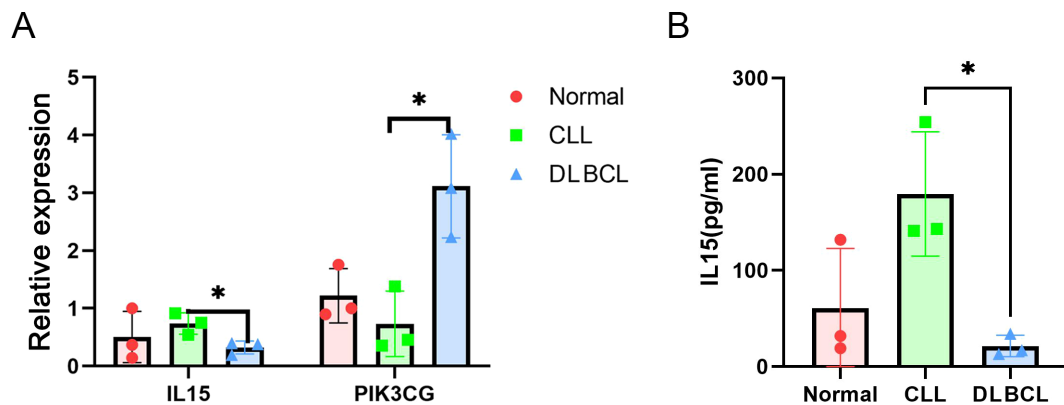


FIGURE 6

(A) Validation of the expression levels of IL15 and PIK3CG using qRT-PCR. (B) Measurement of IL15 concentrations in the serum of DLBCL and CLL patients using ELISA. \* $P \leq 0.05$ .

## Potential biochemical biomarker selection and verification

To more fully probe the links between DLBCL and CLL, the OPLS-DA model was further used to screen for potential biomarkers associated with these hematological malignancies. The OPLS-DA score plot (Figure 7D) includes a horizontal axis corresponding to score values for the main components during OSC analyses, whereas the vertical axis corresponds to scores for orthogonal components during this analysis. The clear clustering of the two sets of samples in this plot confirmed the successful differentiation between these two malignancies with this model, providing a foundation for efforts to further screen for the biomarkers that can distinguish between the two. When constructing the OPLS-DA loading plot to conduct preliminary screening for the factors that contribute to the effective differentiation between DLBCL and CLL (Figure 7E), correlations were noted between the loading and score plots such that factors with higher positive values on the vertical axis of the loading plot tended to exhibit higher positive values on the horizontal axis of the score plot, with the same also being true for negative values. The OPLS-DA VIP plot also exhibited associations between peak VIP values and model correlation coefficients (Figure 7F), with redder peak coloration being indicative of a stronger correlation coefficient for that peak position in the discrimination model. A V+S plot was also constructed incorporating VIP and correlation coefficient values in a single figure wherein individual points correspond to a given indicator of interest and the redder the coloration of that point, the stronger the correlation coefficient consistent with a greater contribution to the overall performance of the classification model. In contrast, bluer coloration is indicative of weaker contributions to this classification model. Points situated further from the center of the V+S plot also made greater contributions to the performance of the classification model. This V+S plot was thus used for further biomarker screening, testing for significance based on characteristic peak positions in subsequent analyses as a means of clarifying which features were able to reliably differentiate between DLBCL and CLL (Figure 7F). The VIP values for relevant factors in this classification model were used to assess the contributions of these factors to the model, providing a list of

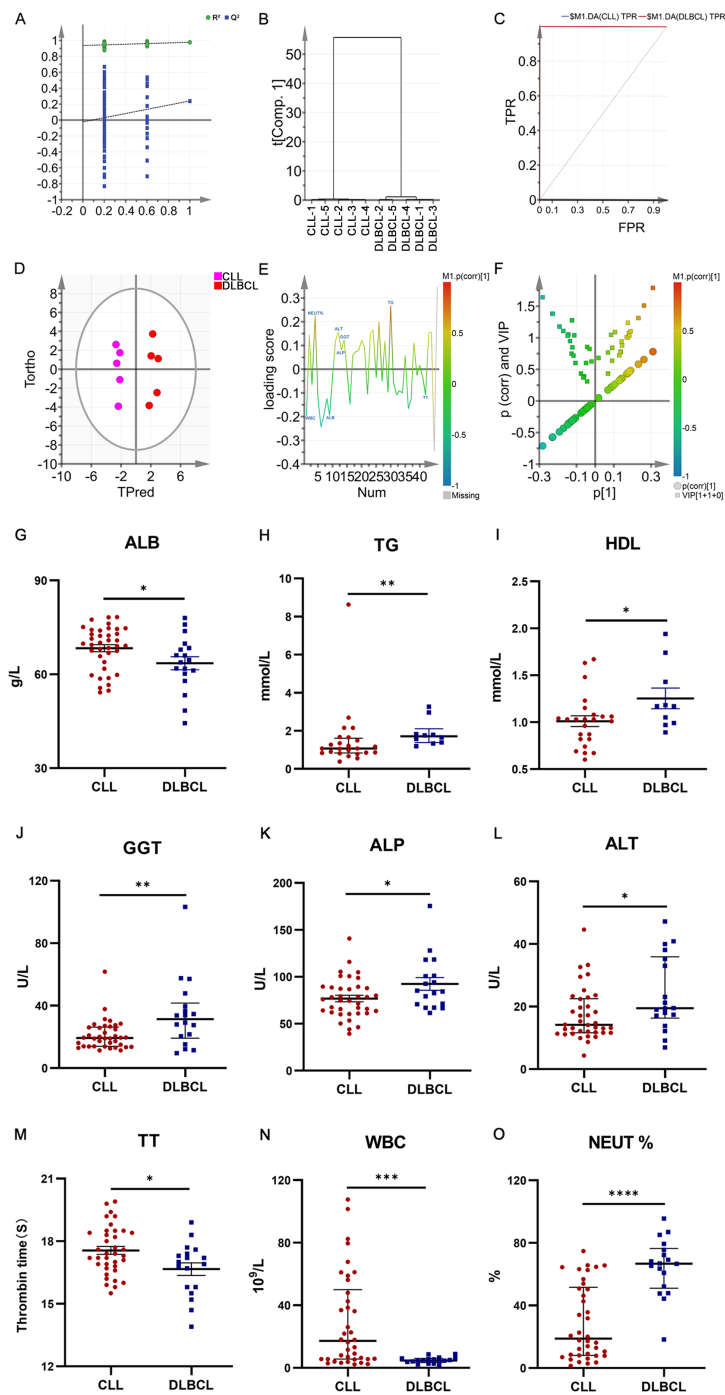
hematological indices ranked based on their VIP values, with  $VIP > 0.5$  and biological significance being used to guide biomarker selection. When the peripheral blood biochemical data from study subjects were analyzed (Figures 7G–O), ALB, TT, and WBC levels were markedly increased in the CLL group relative to the DLBCL group ( $P < 0.05$ ), whereas the opposite was true for TG, HDL, GGT, ALP, ALT, and NEUT% levels ( $P < 0.05$ ).

## Raman spectroscopy of sera from controls and CLL and DLBCL patients

In total, 25, 29, and 30 serum Raman spectra were obtained from the CLL, DLBCL, and control patient samples. The corresponding peak assignments are shown in Supplementary Table S3. Raman spectra for these groups were detectable in the 600–1800  $\text{cm}^{-1}$  range and were broadly similar, consistent with the generally comparable composition of the compounds present in samples from these three groups (Figure 8). Based on the visual inspection of these spectra alone, differentiating between these three groups is impossible, underscoring the need for a multivariate statistical classification model.

## Development of a Raman spectroscopy-based method to distinguish between CLL, DLBCL, and control samples

Raman spectral data from the control, CLL, and DLBCL patients were next used to conduct supervised OPLS-DA analyses. In the permutation analyses, no overfitting was observed as evidenced by a  $Q^2$  value with a negative Y-intercept consistent with OPLS-DA model validity (Figure 9A). Cluster analyses showed that the OPLS-DA model could differentiate among the three groups with 100% accuracy (Figure 9B). The ROC curves for all three sample types using this model were exactly 1 (Figure 9C), consistent with a high degree of discriminant performance.



**FIGURE 7**  
 (A) OPLS-DA permutation plots for the DLBCL and CLL groups. (B) Cluster analysis. (C) ROC curves for the DLBCL and CLL groups, both exhibiting AUC values of 1. (D) DLBCL and CLL discrimination OPLS-DA score plot showing Hotelling's 95% confidence ellipse. (E) CLL and DLBCL loading plot. (F) V+S plot for CLL and DLBCL (G) ALB. (H) TG (I) HDL. (J) GGT. (K) ALP. (L) ALT. (M) TT. (N) WBC. (O) NEUT%. \* $P \leq 0.05$ , \*\* $P \leq 0.01$ , \*\*\* $p \leq 0.001$ .

### Potential Raman biomarker selection and identification

To more fully clarify the utility of specific biomarkers capable of distinguishing among control, CLL, and DLBCL serum samples, the OPLS-DA model was further analyzed using an approach similar to that outlined above. In the OPLS-DA score plot (Figure 10A), the horizontal axis enabled the visualization of differences among

groups whereas the vertical axis enabled differentiation among samples. As all three sample groups were clearly differentiated in the generated plot, with control samples on the positive X-axis whereas the DLBCL and CLL samples were on the negative X-axis, these results confirmed the ability of this OPLS-DA model to effectively distinguish between these different sample types. These results provided a foundation for the further examination of metabolic differences among control, CLL, and DLBCL samples.

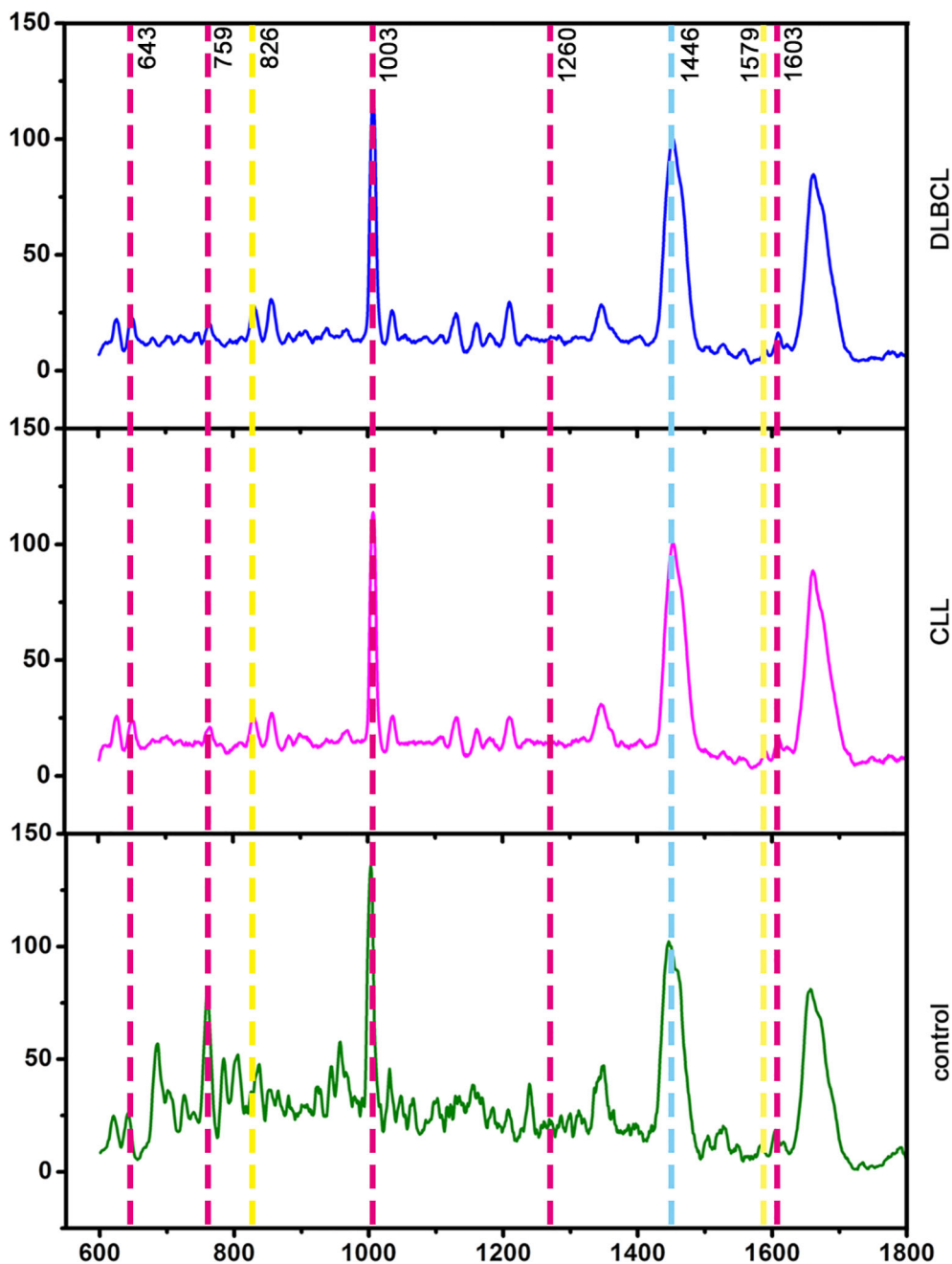


FIGURE 8  
Average serum spectra for the Control, CLL, and DLBCL groups (from bottom to top).

The OPLS-DA loading plot utilized for preliminary screening of the Raman peaks contributing to the classification model for control, CLL, and DLBCL samples is shown in Figure 9B. Peaks in red, blue, green, purple, orange, and black respectively correspond to nucleic acids, proteins,  $\beta$ -carotene, lipids, carbohydrates, and collagen. A correlation was evident between the score plot and the loading plot such that peak positions with larger values on the positive end of the loading plot vertical axis tended to be present at relatively higher levels on the positive horizontal axis of the score plot, while the same was also true for negative values. This approach revealed significantly higher peak intensity values for the peaks

corresponding to proteins ( $621, 1603 \text{ cm}^{-1}$ ), nucleic acids ( $726, 1579 \text{ cm}^{-1}$ ),  $\beta$ -carotene ( $1155 \text{ cm}^{-1}$ ), carbohydrates ( $842 \text{ cm}^{-1}$ ), and lipids ( $1446 \text{ cm}^{-1}$ ) in control samples relative to those from CLL and DLBCL patients (Figure 10B).

In the OPLS-DA V+S plot including VIP values and correlations corresponding to peak positions in the classification model (Figure 10C), individual points indicate a specific peak position and the color of these points denotes the strength of the correlation coefficient, with red and blue respectively denoting stronger and weaker contributions to the classification model. Peaks further from the plot center contributed more significantly

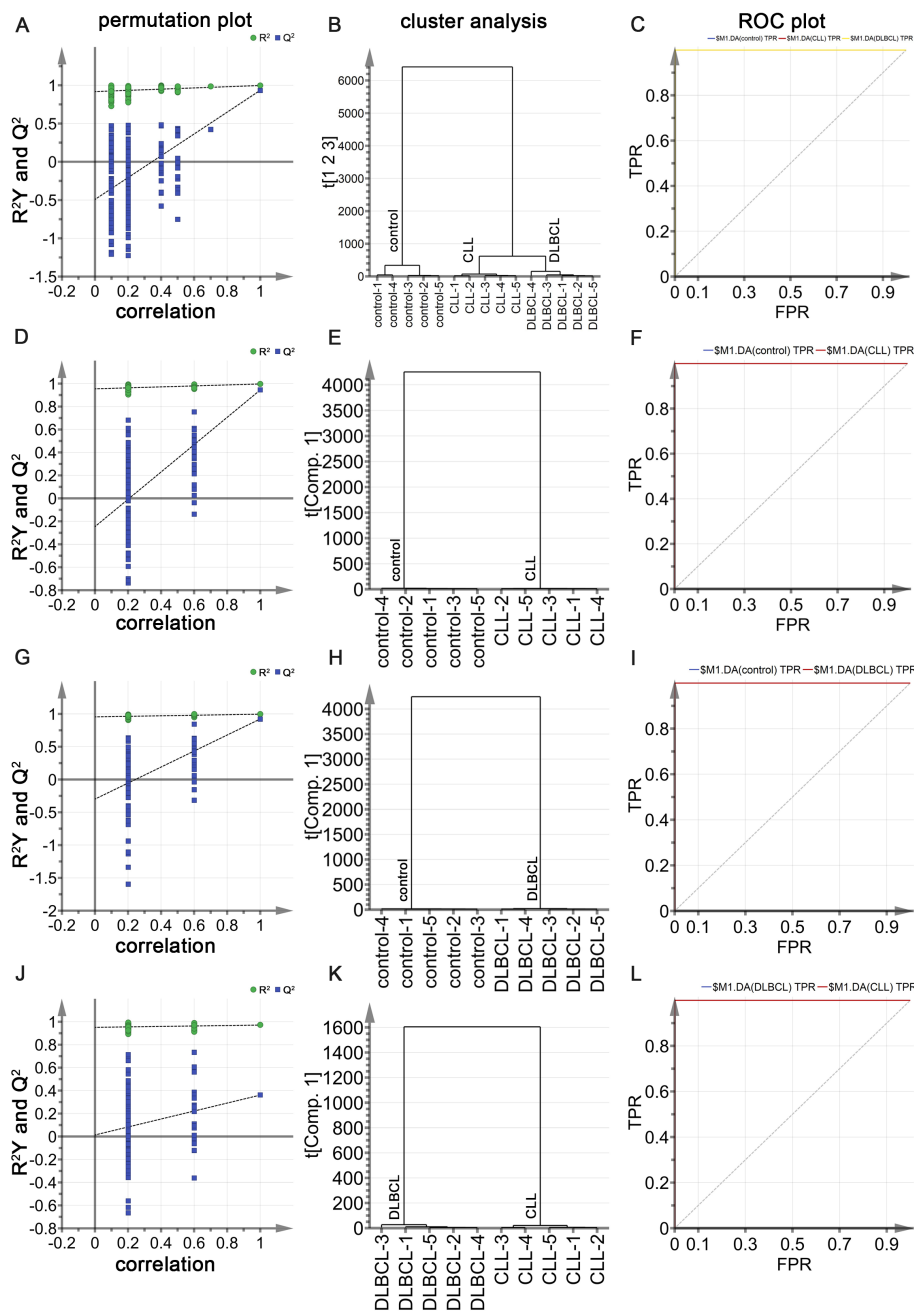


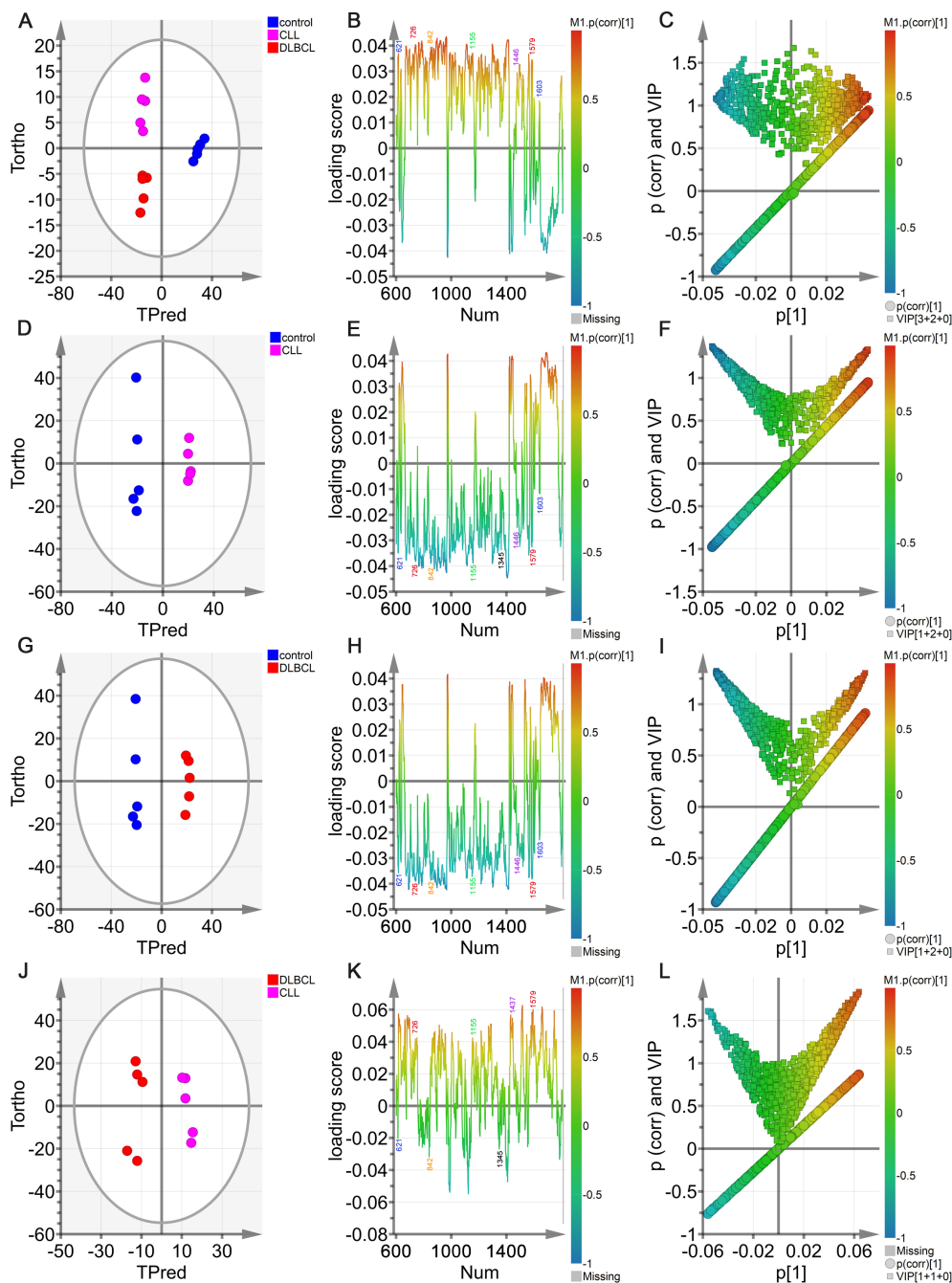
FIGURE 9

(A) OPLS-DA permutation plots for control, CLL, and DLBCL samples. (B) Cluster analyses of the groups. (C) ROC curves for the groups, each with an AUC value of 1. (D–F) Permutation plot (D), cluster analysis (E), and ROC plot (F) results for the control and CLL groups, with the latter exhibiting AUC values of 1 for both CLL and control samples. (G–I) Permutation plot (G), cluster analysis (H), and ROC plot (I) results for the control and DLBCL groups, with the latter exhibiting AUC values of 1 for both DLBCL and control samples. (J, L) Permutation plot (J), cluster analysis (K), and ROC plot (L) results for the CLL and DLBCL groups, with the latter exhibiting AUC values of 1 for both CLL and DLBCL samples.

to the model. To more fully investigate the peak positions contributing to the model, the V+S plot was also utilized to screen for promising biomarkers, performing significance testing on characteristic peak positions to clarify those peaks capable of discriminating among control, CLL, and DLBCL samples (Figure 10C). BIP values corresponding to the positions of Raman peaks in this classification model were the primary index used to assess the contribution of peak position to the model, yielding a list of Raman peak positions ranked based on VIP values, retaining

those peaks with a VIP > 0.5 and biological significance as candidate biomarkers.

Based on the model developed for the different groups, the control group was combined with the CLL and DLBCL samples for OPLS-DA analysis (Figures 10D–L). The OPLS-DA score plots for the respective control vs. CLL, control vs. DLBCL, and CLL vs. DLBCL model comparisons are illustrated in Figures 10D, G, J. Clear sample clustering was evident in the resultant scatter plots, consistent with the ability of the OPLS-DA models to effectively



**FIGURE 10**  
 OPLS-DA analyses of sera from the control, DLBCL, and CLL groups. (A) Hotelling's T2 ellipse OPLS-DA score plots showing the 95% confidence zone for the control, CLL, and DLBCL groups. (B) Loading plot for the control, CLL, and DLBCL groups. (C) V + S plot for the control, CLL, and DLBCL groups. (D-L) Hotelling's T2 ellipse score map (D, G, J), loading plot (E, H, K), and V+S plot (F, I, L) for the control and CLL (D-F), control and DLBCL (G-I), and CLL and DLBCL (J-L) groups.

extract information capable of differentiating among these groups of serum samples, reliably identifying the samples in each model (Figures 10D, G, J). The respective loading plots for the control vs. CLL, control vs. DLBCL, and CLL vs. DLBCL models are presented in Figures 10E, H, K. These results demonstrated that the control group was associated with peak positions corresponding to proteins (621, 643, 848, 853, 869, 935, 1003, 1031, 1221, 1230, 1260, 1344, 1443, 1446, 1548, 1579, 1603, and 1647  $\text{cm}^{-1}$ ), nucleic acids (726,

781, 786, 1078, 1190, 1415, 1573, and 1579  $\text{cm}^{-1}$ ), and  $\beta$ -carotene (957, 1155, and 1162  $\text{cm}^{-1}$ ). The peak intensities corresponding to sugars (842  $\text{cm}^{-1}$ ), collagen (1345  $\text{cm}^{-1}$ ), and lipids (957, 1078, 1119, 1285, 1299, 1437, 1443, 1446  $\text{cm}^{-1}$ ) was significantly increased in control samples relative to those from CLL and DLBL patients. The V+S plots for the control vs. CLL, control vs. DLBCL, and CLL vs. DLBCL models are respectively presented in Figures 10F, I, L. These plots offer a list of Raman peak plots ranked based on VIP values

suitable for the identification of candidate biomarkers for these three models (Figures 10C, F, I, L). Candidate biomarkers were defined as those peaks with a VIP > 0.5 and biological significance. In the four established models, biomarker selection was based on VIP values, correlation coefficients, loading, and the distance from the center of the V+S plots, enabling the identification of key Raman peak positions related to sample classification. During the course of biomarker validation, nonsignificant Raman peak positions were excluded from further biomarker consideration.

## Raman biomarker verification

Potential Raman biomarkers were next validated (Figures 11A, B), revealing that peak intensities associated with specific proteins in the control group were greater than those in the CLL and DLBCL groups, among which the peaks at 621, 643, 848, 853, 869, 935, 1003, 1031, 1221, 1230, 1260, 1344, 1443, 1446, 1548, 1579, 1603, and 1647  $\text{cm}^{-1}$  differed significantly relative to the CLL and DLBCL groups ( $P < 0.05$ ). Similarly, nucleic acid peak intensities in the control samples were markedly greater than those in the CLL and DLBCL groups, with significant differences for the peaks at 726, 781, 786, 1078, 1190, 1415, 1573, and 1579  $\text{cm}^{-1}$  ( $P < 0.05$ ). The peak intensities for  $\beta$ -carotene (957, 1155, 1162  $\text{cm}^{-1}$ ) were all significantly higher in control samples relative to CLL and DLBCL samples ( $P = 0.000$ ).

Peak intensity values for carbohydrates (842  $\text{cm}^{-1}$ ) and collagen (1345  $\text{cm}^{-1}$ ) in control samples were significantly greater than those for CLL and DLBCL samples ( $P < 0.001$ ). Similarly, lipid peaks (957, 1078, 1119, 1285, 1299, 1437, 1443, 1446  $\text{cm}^{-1}$ ) exhibited significantly higher intensity values in control samples relative to CLL and DLBCL samples ( $P < 0.05$ ).

To date, there has been no exploration of the utility of Raman spectroscopy as a means of distinguishing between DLBCL and CLL using serum samples, and there is a corresponding lack of information regarding glycolipid metabolism-related serological indices in these patients. This study was thus designed with the goal of developing a model capable of distinguishing among DLBCL, CLL and control samples via Raman spectroscopy combined with multivariate analyses. Screening was also

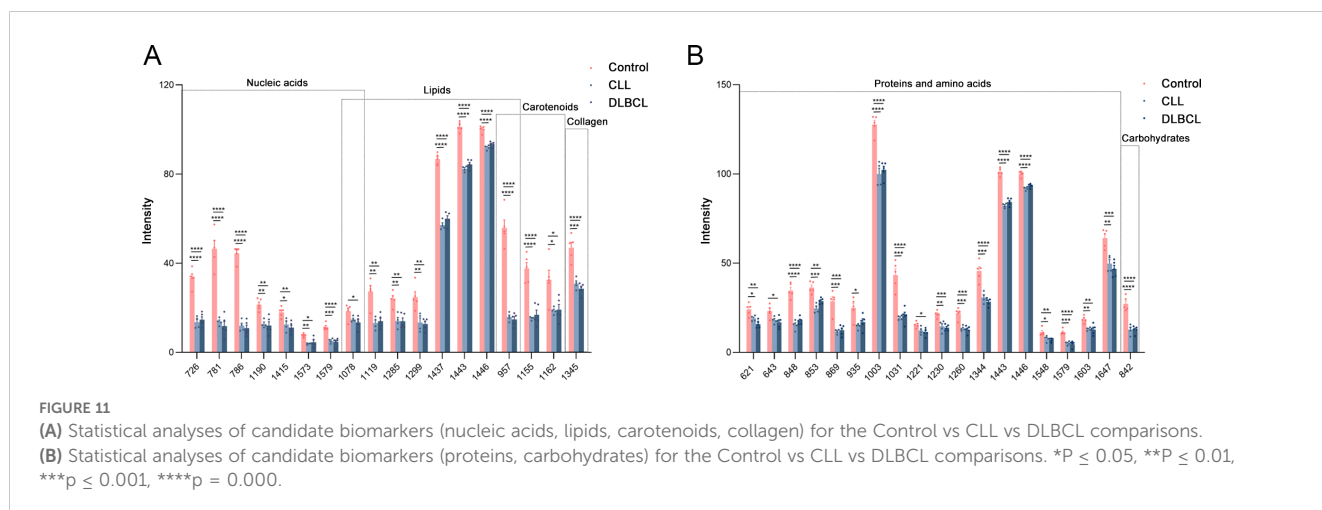
performed for Raman peaks that contributed to the DLBCL and CLL classifications for potential utilization as a biomarker for these cancers, providing a foundation for the effective and rapid detection of early-stage DLBCL and CLL.

## Discussion

DLBCL and CLL are both relatively common hematological malignancies. Intensive research efforts have informed the diagnosis of these cancers, as well as and the creation of targeted therapies, immunotherapies, and allogeneic hematopoietic stem cell transplantation (HSCT) treatment approaches. However, each of these therapeutic strategies has some risk, such as complications, recurrent disease, and death, underscoring a need for further in-depth research exploring the diagnosis, treatment, and prognostic assessment of DLBCL and CLL.

In this study, two independent datasets from the GEO dataset were normalized and analyzed, identifying 179 DEGs when comparing DLBCL and CLL samples, of which 39 and 140 were respectively upregulated and down-regulated. In functional analyses, these overlapping DEGs were enriched in pathways, including immune and inflammatory responses, cellular TNF and IFN- $\gamma$  responses, and positive regulation of GTPase activities. A core set of 10 key DEGs was selected through further bioinformatics analyses that included IL15, CCL4, FLNA, CCL5, IL7, VCAM1, SYK, ACTN1, CCL20, and IL1R1, with a majority of these genes being significantly overexpressed in CLL relative to DLBCL. To further explore the signaling pathways relevant to the treatment of these genes, DEGs were further screened to identify the top 10 key upregulated DEGs (ANKMY2, PIK3CG, BZW2, BACH2, BCL6, ZHX2, SYK, CD72, CD180, and SLA) as well as the top 10 key down-regulated DEGs (IL15, VCAM1, FLNA, CCL5, CCL4, IL7, CCL20, IL1R1, ACTN1, and APP). The top 5 DEGs in each category are the focus of the remainder of this discussion section.

Vascular cell adhesion molecule-1 (VCAM-1) is an immunoglobulin (Ig) superfamily member and type I transmembrane glycoprotein with multiple Ig-like domains containing disulfide-linked loops (27). VCAM-1 plays a role in



the inflammatory process, facilitating the interstitial adhesion of leukocytes and the migration of these cells. Endothelial exposure to inflammatory cytokines can elicit the release of soluble VCAM-1 (sVCAM-1), promoting monocyte, lymphocyte, and eosinophil recruitment, migration, and adhesion (28). VCAM-1 expression is evident in the heart, brain, kidneys, placenta, bladder, lymph nodes, and spleen. The sVCAM-1 levels in patients with atrial fibrillation, ischemic cardiomyopathy, coronary artery disease, and acute myocarditis are elevated, and both VCAM-1 and sVCAM-1 are regarded as biomarkers of autoimmune myocarditis, cancer, and immunological diseases (29–31). Under conditions of persistent inflammation, T cells, macrophages, and/or NK cells can release inflammatory mediators, including IL-1 $\beta$  and TNF- $\alpha$  (32), the latter of which promotes VCAM-1 upregulation and interacts with the receptor TNFR1 to recruit a complex consisting of TRAF2, RIP1, and cIAP1/2 that promotes TAK1 and IKK signaling pathway activity. IKK, in turn, promotes the degradation of the NF- $\kappa$ B inhibitor I $\kappa$ B, thereby driving NF- $\kappa$ B activation and VCAM-1 upregulation (33). miR-126 can promote angiogenic activity while reducing vascular inflammation, inhibiting the extravasation of leukocytes through reductions in VCAM-1 expression. When VCAM-1 is expressed on the endothelial surface, it can interact with the  $\alpha$ 1  $\beta$ 4 integrins on circulating leukocytes, thereby activating intracellular signaling that disrupts endothelial cell-cell connections and promotes actin remodeling, enabling these leukocytes to extravasate from systemic circulation into the damaged site (34).

PI3K catalytic subunit gamma (PIK3CG) is a class I catalytic PI3K subunit that, like other subunits in this class (p110- $\alpha$ , p110- $\beta$ , and p110- $\delta$ ), is capable of binding to the regulatory p85 subunit in the PI3K complex. PIK3CG is involved in the control of cellular activities through the phosphorylation of proteins and lipids (35). PIK3CG can be regulated directly by Ras and G  $\beta$  -  $\gamma$  and Ras in the G protein-coupled receptor (GPCR) pathway, controlling a diverse array of immune- and inflammation-related activities. PIK3CG has also been established as a candidate target for managing certain cancers, including Kaposi's sarcoma, medulloblastoma, and ALL (36). CLBC cell invasivity, migration, and stem cell maintenance are impaired upon the inhibition of PIK3CG activity (37). Jun et al. have also demonstrated that targeting PIK3CG with PTX can significantly influence CLBC treatment, reducing the risk of recurrent disease by 40% (38). Chung et al. further determined that PIK3CG is an essential regulator of prostate cancer activity related to KRAS activation and p53 deficiency. It may be a viable target in patients with advanced metastatic disease (39). Zhang et al. noted that PIK3CG is targeted directly by RBPJ $\kappa$ -dependent Notch signaling, and the inhibition of PIK3CG in CLBC cell lines can limit the formation of tumor balls and the migration of tumor cells such that it is a promising candidate target for CLBC subtypes.

The transcription factor basic leucine zipper and W2 domains 2 (BZW2), sometimes referred to as eIF5 mimic protein (5MP) 2, is a cytosolic member of the alkaline region leucine zipper (bZIP) superfamily that binds to cadherin and is linked to nervous system development and cellular differentiation. BZW2 is linked with various forms of cancer. For example, Wang et al. demonstrated that TEAD4 upregulation in nasopharyngeal

carcinoma, upregulated TEAD4 promotes oncogenic signaling via the AKT pathway through the activation of BZW2 transcriptional activity, driving cancer progression (40). Gao et al. (41) determined that higher levels of BZW2 expression were evident in muscle-invasive bladder cancer tissues relative to normal tissues. The knockout of BZW2 in bladder cancer cells blocked cell cycle progression and apoptosis induction. In fibrosarcoma, BZW2 can promote tumor growth, and it is also expressed at high levels in colorectal and liver cancers (42–48). The dysregulation of BZW2 may thus be a hallmark of certain forms of cancer (49). The oncogenic properties of BZW2 are associated with the translation of ATF4, another transcription factor, delaying the restart and thus promoting the survival of stress-exposed tumor cells. Li et al. conducted a systematic analysis of BZW2 in several HCC-related datasets. They documented a close link between this transcription factor and HCC patient prognostic outcomes, with BZW2 regulating eIF factors via c-Myc signaling (50).

Myc is associated with many important biological pathways. These include: (1) Metabolic pathways, where Myc is a key regulatory factor involved in metabolic reprogramming; (2) Pathways associated with the cell cycle and proliferation, where Myc promotes cell proliferation and modulates cell cycle dynamics. High levels of Myc expression induce replication pressure and genomic instability, sensitizing cells to apoptosis. Cell proliferation requires a doubling of energy and cell biomass, and replicating cells are thus particularly sensitive to deficiencies in oxidative phosphorylation (OXPHOS), and OXPHOS-deficient replicating cells are especially vulnerable to high levels of Myc, as MYC helps them evade metabolic checkpoints and accelerate cell cycle progression; (3) Pathways related to mitochondrial diseases, as upregulation of Myc and/or its typical transcriptional features have been observed in several cellular and mouse models of mitochondrial diseases. Changes in gene expression and metabolite levels associated with the mitochondrial integrated stress response (mt ISR) overlap significantly with Myc overexpression; (4) Pathways associated with the tumor microenvironment, in which Myc interacts with the tumor microenvironment through signaling pathways and molecular networks to modulate processes such as proliferation, replication pressure, and DNA repair in tumor cells, driving tumor development and drug resistance; (5) Transcriptional regulatory pathways, in which Myc acts as a global transcriptional amplifier, binding and increasing the expression of active promoters and controlling the expression of at least 15% of human genes, including genes related to cell cycle progression, metabolism, ribosome biogenesis, and translation.

The transcriptional repressor BTB domain and CNC homolog 2 (BACH2) belong to the BACH family of alkaline leucine zipper transcription factors that are found in the nucleus and cytosol and exhibit the ability to bind DNA in a sequence-specific manner (51). NK cells are key mediators of immune surveillance, protecting against cancer and infection. BACH2 is a vital mediator of adaptive immunity, but how it functions in NK cells and other innate cell populations is poorly characterized. Imianowski et al. found that BACH2 may be a negative regulator of NK cell differentiation, maturation, and function (52). BACH2 can limit NK cell maturation under conditions of weak stimulation, maintaining

numbers of undifferentiated NK cells such that knocking out BACH2 in NK cells can lead to cytotoxic NK cell accumulation within tissues, more effectively protecting against lung cancer metastasis. These results support the role of BACH2 as a global negative regulator of cytotoxic activity and NK cell-mediated tumor surveillance.

BCL6 is a transcriptional repressor involved in germinal center (GC) B cell responses, preventing premature GC B cell activation and differentiation and allowing cells to tolerate better the breakage of the DNA that occurs during Ig gene remodeling. The essential role BCL6 plays in B cell development can be coopted by malignant cells such that it serves as an oncogene in GC-derived lymphomas, with high mutation rates and translocation in DLBCL (53). BCL6 functions by binding to target DNA sequences and recruiting co-repressor complexes to suppress transcription. Specifically, its N-terminal BTB domain is responsible for recruiting the BCOR, NCOR, and SMRT co-repressor proteins to the extended groove motif of the BTB dimer interface. At the same time, its core RD2 region can inhibit gene transcription through interactions with MTA2, HDAC2, NuRD, and CBP (54). BCL6 is reportedly upregulated in glioblastoma and gastric, ovarian, and non-small cell lung cancers, and its expression is associated with malignant features, including proliferation, invasivity, migratory activity, survival, and therapeutic resistance (55).

IL-15 is a 14-15 kDa cytokine essential for NK, NKT, and memory CD8+ T cell dynamics. While IL-15 is secreted at very low levels, it can be readily delivered in trans through the unique IL-15 $\alpha$  receptor on the surfaces of cells that produce IL-15 such that it can interact with a target cell receptor consisting of IL-2R $\beta$  and common gamma chains. IL-15 and IL-15R $\alpha$  binding in solution results in the formation of the highly potent IL-15 super antagonist (IL-15 SA) complex, which can robustly activate cells responsive to IL-15, particularly NK cells, promoting a range of antitumor and antiviral effects (56). IL-15 is vital for NK and CD8+ T cell expansion and functionality (57, 58). IL-15 inhibits HIV-specific CD8+ T cell apoptosis and increases these cells' activation, expansion, cytotoxicity, and IFN- $\gamma$  production (59). Chehimi et al. showed that IL-15 therapy improved the cytotoxic functions of HIV patient-derived NK cells while significantly reducing apoptotic PBMC death (60). *In vitro*, NK cells can suppress HIV replication, much like CD8+T cells, through a mechanism potentially linked to CC chemokine secretion (61). Further studies of the potential therapeutic capability of IL-15 to remediate NK and CD8+ T cell responses in the context of HIV infection are thus warranted.

Filamin A (FLNA) possesses an N-terminal actin-binding domain (ABD) composed of two calponin homologous CH1 and CH2 domains, two rod-like domains, and 24 Ig repeat sequences that confer actin-binding activity (62). FLNA can crosslink actin filaments, linking them to membrane glycoproteins and serving as a scaffold for signaling intermediates that are involved in cytoskeletal remodeling necessary for interactions with transmembrane receptors, integrins, and second messengers, ultimately shaping cellular shape, migratory activity, motility, and signal transduction (62). Periventricular nodular ectopia (PVNH) is the most commonly associated disease, and it is characterized by

impaired neuronal migration such that ectopic neurons accumulate at ventricle edges (63, 64). CCL5 is a C-C motif chemokine family member closely associated with inflammatory and immune regulation. It can induce peripheral monocyte, memory T helper cell, and eosinophil responses, triggering activation and histamine release (65). CCL5 expression is observed in T cells, platelets, macrophages, renal tubular epithelial cells, synovial fibroblasts, and tumor cells (66). CCL5 signals through its cognate receptor CCR5 to promote tumor growth, tumor progression, tumor cell migration and metastasis (67).

In scientific research, especially in the field of biomarker research, the establishment of control groups is crucial to ensure the validity and reliability of the research results. The reasons for including a control group in the selection and identification processes of Raman biomarkers primarily include: (1) The provision of a benchmark, where the control group serves as a reference standard to assist in the confirmation of whether the observed changes represent a true association with the disease under investigation. Without this comparison, it would be difficult to determine whether the observed changes are caused by the disease itself or other factors; (2) Exclusion of bias, as factors other than the disease, such as age, sex, and lifestyle habits, may affect the expression of biomarkers. The control group helps to control these variables, ensuring the accuracy and reproducibility of the results; (3) Statistical efficacy, as statistical analysis relies on the presence of a control group to evaluate whether the observed differences are statistically significant, rather than accidental events. The most significantly enriched pathways for the DEGs between CLL and DLBCL patients were found to be associated with the immune response, cytoskeleton, cytokine activity, and cytokine-cytokine receptor interactions. The genes most significantly associated with these pathways included VCAM-1, ANKMY2, PIK3CG, BZW2, BACH2, BCL6, IL-15, FLNA, and CCL5.

After visualization of the PPI network using the STRING database and Cytoscape plugin, we identified 10 key differentially expressed genes, including IL15, CCL4, FLNA, CCL5, IL7, VCAM1, SYK, ACTN1, CCL20, and IL1R1. Most of these genes were found to be downregulated, indicating that their expression was significantly higher in CLL samples. A PPI network (consisting of 10 nodes and 20 edges) was constructed for the upregulated differentially expressed genes. The top 10 key genes identified in this network were ANKMY2, PIK3CG, BZW2, BACH2, BCL6, ZHX2, SYK, CD72, CD180, and SLA. The protein encoded by PIK3CG is one of the catalytic subunits of phosphatidylinositol 3-kinase (PI3K), and the PI3K signaling pathway plays a crucial role in various biological processes such as cell growth, proliferation, survival, metabolism, and migration. In both DLBCL and CLL, upregulation of PIK3CG may lead to abnormal activation of the PI3K signaling pathway, promoting cancer cell proliferation and survival and inhibiting cell apoptosis, and thus promoting disease progression. For example, overactivation of PI3K signaling may influence downstream target proteins such as Akt, affecting cell metabolism and survival signals, and enabling cancer cells to evade normal cell death mechanisms and continue to proliferate and form tumors. The PPI network constructed for the downregulated DEGs (including 100 nodes and 332 edges) identified the top 10 key genes



as IL15, VCAM1, FLNA, CCL5, CCL4, IL7, CCL20, IL1R1, ACTN1, and APP. Among them, IL15 is a multifunctional cytokine that plays an important role in the immune system. In CLL, overexpression of IL15 may promote disease progression through multiple pathways. In terms of promoting the proliferation of cancer cells, IL15 binds to receptors on the cell surface, leading to the activation of a series of intracellular signaling pathways and resulting in the proliferation of leukemia cells. For example, it may activate the JAK/STAT signaling pathway, regulate the expression of cell cycle-related proteins, and drive cells into the proliferation cycle. In terms of apoptosis inhibition, IL15 may also prevent apoptosis in leukemia cells by upregulating the expression of anti-apoptotic proteins or inhibiting the activity of pro-apoptotic proteins, allowing cancer cells to continue to survive and accumulate in the body. It can also affect the tumor immune microenvironment, as IL15 may regulate the functions and activities of immune cells, inhibiting the body's antitumor immune response. For example, it may affect the differentiation, proliferation, and killing activity of immune cells such as T and NK cells, making it difficult for the immune system to effectively clear leukemia cells.

In an article entitled "Immunophenotypic and genomic landscape of Richter transformation diffuse large B-cell lymphoma," Nadeu et al. explored the genetic map of Richter's Transformation, suggesting that diffuse large B-cell lymphoma (DLBCL) represents a more aggressive form of CLL and providing insights into the pathogenesis and treatment of DLBCL. Nadeu et al. mentioned ATM in their most recent article published in 2024, in which BCL2, BIRC3, BTK, CXCR4, EGR2, FBXW7, KRAS, MYD88, NFKBIE, PLCG2, POT1, SF3B1 and a series of genes were analyzed for mutations. Although no significant differences were found in the expression of all these genes, many were found to be differentially expressed in the present study. Nadeu et al. also described multiple genes associated with B-cell pre lymphocytic leukemia (B-PLL) in their article "Epigenic features support the diagnosis of B-cell prolymphatic leukemia and identify 2 clinicobiological subtypes", mainly including MYC, SF3B1, TP53, CCND1, CCND2, and CCND3. Among them, mutations in SF3B1 are associated with poor prognosis in CLL. The genes mentioned above also overlap with the DEGs in this article, indicating the credibility of our research results (68–70).

Many of the DEGs identified here were related to lipid metabolism, including CCL22, BMP2, CCL20, CCL5, and ACKR3. Among them, the role of CCL20 in lipid metabolism may be related to its function in immune cell chemotaxis and the inflammatory response. CCL20 is associated with lipid droplet accumulation-mediated macrophage survival and regulatory T cell (Treg) recruitment, indicating its potential role in lipid metabolism in the tumor microenvironment. The role of CCL5 in lipid metabolism may be related to its function in immune cell chemotaxis and the inflammatory response. Several studies have found that the expression of CCL5 is associated with the degree of T cell infiltration within tumors, indicating its potential role in lipid metabolism in the tumor microenvironment. ACKR3, the atypical chemokine receptor 3, although its primary function is that of a scavenger receptor for chemokines, plays a role in the negative

regulation of cell proliferation and the positive regulation of the ERK1/2 cascade, indirectly affecting lipid metabolism.

The Richter transformation is the process by which chronic lymphocytic leukemia (CLL) transforms into more aggressive lymphomas such as diffuse large B-cell lymphoma (DLBCL). In Iyer's articles, multiple genes and their roles in CLL and the Richter transformation are mentioned. The genes include TCL1, SF3B1, TP53, CDKN2A/CDKN2B, MYC, and NOTCH1. Abnormalities in the TCL1 gene are common in CLL patients and its overexpression is associated with the clonal expansion of B cells. SF3B1 mutations are commonly seen driver mutations in CLL, and the gene is involved in RNA splicing processes. Its mutation is related to the progression and transformation of CLL. Mutations in the TP53 gene are associated with high-risk subtypes of CLL, and TP53 mutations are typically associated with the Richter transformation, leading to disease progression. The deletion of the CDKN2A/CDKN2B genes is associated with the progression of CLL and the Richter transformation. Their absence may lead to abnormal cell cycle regulation, thereby promoting tumor development. Mutations and deletions of these genes play important roles in the occurrence, development, and response to treatment of CLL and the Richter transformation. Understanding the functions and interactions of these genes is thus crucial for developing new therapeutic strategies (71, 72).

The present study identified key genes and pathways associated with these two diseases through bioinformatics analysis followed by experimental verification. These analyses revealed many upregulated and downregulated DEGs, and determined through functional enrichment analysis that these genes are closely related to inflammation and the immune response. In addition, the author also analyzed serum samples from patients using Raman spectroscopy and found significant differences in spectral characteristics between the control, CLL, and DLBCL groups. These spectral data were combined with the DEG data to successfully distinguish between these two diseases using the supervised orthogonal partial least squares discriminant analysis (OPLS-DA) model. The author combined spectral analysis with DEG data to identify specific gene sets, providing an important foundation for the diagnosis and treatment of DLBCL and CLL. This combination helps to further the understanding of the molecular mechanisms of these two diseases and may provide support for the development of personalized treatment plans.

Experimental methods such as RT-PCR and ELISA allow assessment of the effects of genes and proteins such as PIK3CG and IL15 on cell proliferation, apoptosis, and migration. PIK3CG was found to promote the proliferation of B cells. Studies have shown that the proliferative effect of PIK3CG on B cells may be mediated by activation of the downstream AKT signaling pathway. Although its activation usually promotes cell survival, in some cases, excessive activation of PIK3CG may lead to increased cell apoptosis. This phenomenon may be related to the imbalance of intracellular signals, especially under the influence of microenvironment, which may lead to the enhancement of apoptotic signals. PIK3CG may interact with other PI3K subtypes (such as PIK3CD), which may affect cell proliferation and apoptosis. For example, excessive activation of PIK3CD in CLL cells may further enhance cell survival signaling

through interaction with PIK3CG. Interleukin-15 (IL-15) is a pleiotropic cytokine that plays a critical role in the immune system. It has complex effects on the proliferation and apoptosis of DLBCL and CLL cells, and these effects are achieved through multiple signaling pathways; for instance, IL-15 binds to receptor complexes on the cell membrane, recruiting and activating Janus kinase (JAK) family members such as JAK1 and JAK3. Activated JAK kinase phosphorylates signal transducer and activator of transcription (STAT), leading to its dimerization and translocation to the nucleus, where it regulates the transcription of related genes. In DLBCL and CLL cells, this pathway can regulate the expression of genes related to cell proliferation and apoptosis, such as the cell cycle proteins and Bcl-2 family proteins mentioned above. IL-15 stimulation can also activate PI3K, which converts phosphatidylinositol-4,5-diphosphate (PIP2) to phosphatidylinositol-3,4,5-triphosphate (PIP3), thereby recruiting and activating protein kinase B (Akt). Akt regulates cell metabolism, proliferation, and survival by phosphorylating various downstream substrates, such as glycogen synthase kinase-3  $\beta$  (GSK-3 $\beta$ ) and forkhead box protein O (FoxO). In tumor cells, activation of this pathway typically promotes cell proliferation and inhibits apoptosis. The mitogen-activated protein kinase (MAPK) signaling pathway is also one of the important pathways through which IL-15 exerts its effects. The signal activated by IL-15 can activate Raf kinase through Ras protein, which in turn activates MEK and ERK. Activated ERK can enter the nucleus where it regulates the activity of transcription factors and influences the expression of genes related to cell proliferation, differentiation, and apoptosis. In DLBCL and CLL cells, the activation status of the MAPK signaling pathway may affect the cell's response to IL-15, determining whether it proliferates or undergoes apoptosis. These studies all indicate that both PIK3CG and IL15 have complex effects on the proliferation and apoptosis of tumor cells, which are achieved through multiple signaling pathways (73–76).

Relative to the CLL and DLBCL groups, Raman spectroscopy analyses performed herein revealed significantly higher peak intensities for the peak positions corresponding to proteins (621, 643, 848, 853, 869, 935, 1003, 1031, 1221, 1230, 1260, 1344, 1443, 1446, 1548, 1579, 1603, 1647  $\text{cm}^{-1}$ ),  $\beta$ -carotene (957, 1155, 1162  $\text{cm}^{-1}$ ), carbohydrates (842  $\text{cm}^{-1}$ ), collagen (1345  $\text{cm}^{-1}$ ), and lipids (957, 1078, 1119, 1285, 1299, 1437, 1443, 1446  $\text{cm}^{-1}$ ) in the control group. Thus, leukaemia patients may be more prone to hypoalbuminemia, vitamin imbalances, dysregulated glycolipid metabolism, and immunological deficiencies. Serological results for these patients noted significantly lower TG, HDL, GGT, ALP, and ALT levels in samples from CLL patients relative to DLBCL patients ( $P < 0.05$ ), suggesting that the metabolic disorders facing DLBCL patients may be more severe such that they face a worse prognosis. The results of this study have the potential to inform the development of Raman spectroscopy-based analyses for the noninvasive screening and detection of biomarkers related to CLL and DLBCL, emphasizing the potential relevance of large volumes of biochemical data to the typing and prognostic assessment of CLL and DLBCL patients. However, the sample size was relatively small and the results may be imperfect. Given the relative rarity of CLL and DLBCL, future studies should undertake more detailed prospective analyses examining the feasibility of classifying CLL and DLBCL more

effectively and establishing relevant biomarkers to aid in this process. The present study did not perform in-depth analyses of patient medical history, treatment history, or history of smoking/alcohol intake, which may have impacted the study results. Further large-scale investigations using standardized methods for data collection will thus be vital to improve the accuracy of the analyses. Further age-based stratified analyses are warranted, with an increase in the overall sample size to control for confounding factors.

The potential application of Raman spectroscopy combined with multi-omics techniques in the diagnosis of hematological malignancies: (1) High sensitivity and specificity: Raman spectroscopy can provide molecular-level fingerprint information, distinguishing molecules with different chemical compositions and structures by detecting the vibration modes of chemical bonds in biological samples. In the diagnosis of hematological malignancies, it can detect subtle differences in molecular composition and structure between cancer cells and normal cells, such as changes in biological macromolecules such as nucleic acids, proteins, and lipids, providing a basis for early diagnosis. (2) Non-destructive testing: Raman spectroscopy does not require complex preprocessing of the sample and does not alter the integrity of the sample. It can measure biological samples (such as blood and bone marrow) directly, reducing the likelihood of errors and interference during sample processing, and also facilitating real-time and dynamic monitoring of patients. (3) Comprehensive metabolite analysis: Metabolomics can comprehensively and systematically analyze small-molecule metabolites in organisms, revealing the changes in metabolites during the onset and progression of diseases. In hematological malignancies, there are significant differences in the metabolic pathways of cancer cells compared to normal cells. Metabolomics can identify these specific metabolic markers, providing important clues for the diagnosis and classification of diseases. (4) In-depth understanding of disease mechanisms: By analyzing changes in metabolites, one can gain a deeper understanding of the pathogenesis of hematological malignancies. For example, abnormal activation or inhibition of certain metabolic pathways may be closely related to the onset and progression of tumors, which can assist in the discovery of new therapeutic targets and directions for drug development. (5) Complementary information: The information of molecular structure provided by Raman spectroscopy complements the metabolite composition information provided by metabolomics. Raman spectroscopy can rapidly identify samples with potential diagnostic value, while metabolomics can further analyze these samples in depth, to determine specific metabolic markers and changes in metabolic pathways, and improve the accuracy and reliability of diagnosis. (6) Multi-dimensional diagnostic model: The integration of Raman spectroscopy data with metabolomics data can construct a multidimensional diagnostic model. This model can consider multiple factors such as molecular structure and metabolic changes, and thus comprehensively reflect the characteristics of hematological malignancies, which helps to improve the sensitivity and specificity of diagnosis and achieve early and accurate diagnosis.

Raman spectroscopy combined with multi-omics techniques has broad application prospects in the diagnosis of hematological malignancies, and the identification of IL15 and PIK3CG as

potential therapeutic targets also provides new ideas and directions for the treatment of these diseases. Future research should explore the potential applications of these technologies and targets, and contribute to improving the diagnosis and treatment of hematological malignancies. (1) Key role in hematological malignancies: IL15 plays an important regulatory role in the immune system and often exhibits abnormal expression in hematological malignancies such as CLL and DLBCL. Overexpression of IL15 may promote tumor development through various pathways, stimulating cancer cell proliferation, inhibiting cell apoptosis, and affecting the immune microenvironment. Therefore, interventions targeting IL15 would be expected to block tumor growth and spread. (2) Targeted therapy strategy: Monoclonal antibodies, small-molecule inhibitors, and other drugs targeting IL15 or its receptors can be developed to block the binding of IL15 to its receptors or inhibit its downstream signaling pathways, thereby interfering with the proliferation and survival of tumor cells. In addition, gene-editing techniques such as CRISPR/Cas9 can be used to knock out or downregulate the expression of the IL15 gene, achieving the goal of treating tumors. (3) Challenges and prospects: Although IL15 has potential as a therapeutic target, it still faces some challenges in practical applications. For example, IL15 has a wide range of physiological functions in the immune system, and targeting IL15 may adversely affect normal immune function, leading to immune-related adverse reactions. Therefore, future research needs to further optimize the design of targeted drugs and improve their specificity and safety for the effective treatment of hematological malignancies. (4) The importance of the PI3K signaling pathway: The protein encoded by PIK3CG is one of the catalytic subunits of PI3K, and the PI3K signaling pathway plays a crucial role in various biological processes such as cell growth, proliferation, survival, metabolism, and migration. In hematological malignancies, the PI3K signaling pathway is often abnormally activated, and upregulation of PIK3CG may be one of the important reasons for this abnormal activation. Therefore, targeting PIK3CG would be expected to prevent the excessive activation of the PI3K pathway, thereby suppressing the growth and spread of tumor cells. (5) Foundation for research and drug development: Currently, some inhibitors targeting the PI3K signaling pathway have shown positive results in clinical trials, providing a reference for the study of PIK3CG as a therapeutic target. In the future, specific inhibitors targeting PIK3CG could be further developed to improve the efficacy and safety of these drugs. At the same time, the combination with other treatment methods, such as chemotherapy, radiotherapy, and immunotherapy, may achieve better therapeutic effects. (6) Potential problems and solutions: The PI3K signaling pathway also plays an important physiological function in normal cells, so targeting PIK3CG may have toxic effects on normal cells. Future research requires a deeper understanding of the differential expression and functional characteristics of PIK3CG in tumor and normal cells, and the development of more selective targeted drugs to reduce damage to normal cells. In addition, further exploration is needed to investigate the interaction between PIK3CG and other signaling pathways, as well as the resistance mechanism of tumor cells to PIK3CG targeted therapy, to provide a basis for optimizing treatment plans.

No significant differences were observed between groups for certain biochemical indicators. Even with this limitation, Raman spectroscopy could unveil significant differences when analyzing patient serum samples. This suggests that this approach offers greater sensitivity to more conventional peripheral blood analyses, providing an effective screening tool. Another limitation of this study is that patient medical history, medications, smoking status, and drinking status were not considered and may have impacted the results of these analyses. It is thus essential to improve the overall sample size and standardize the data collection approach to achieve greater sampling accuracy. Even so, the metabolomics and statistical analyses employed herein effectively compare differences among samples in a research context and may even have implications for clinical practice, providing a valuable foundation for future research.

## Conclusions

In this study, Raman spectroscopy and bioinformatics approaches were used to differentiate between DLBCL and CLL based on the differential expression of key genes, and the results of these analyses have important implications for the diagnosis and treatment of these two malignancies. Specifically, the DEGs identified herein can potentially serve as diagnostic biomarkers for DLBCL and CLL, and they may further enable the differentiation among DLBCL subtypes to support more effective patient clinical classification. They may additionally offer unique insights into the molecular processes underlying the development of these two types of cancers, as genes that are overexpressed in DLBCL but not CLL may function primarily in the development of the former NHL subtype. The results of these analyses can also help guide the development of personalized targeted treatments, as small molecule inhibitors or antibodies targeting proteins vital to DLBCL cell survival and proliferation may aid in the treatment of this form of cancer. These DEGs can also be leveraged to guide the prognostic assessment of patients with these cancers, aiding clinicians in the selection of the most appropriate treatment plan. Lastly, the bioinformatics analyses performed herein will help unveil the mechanisms through which drugs function in these cancers, potentially facilitating the selection of combination drug regimens. Through comparisons of gene expression patterns between these diseases, genes capable of impacting drug efficacy may be identified, allowing for the optimization of the use of drugs that already exist or the selection of novel drug combinations.

In summary, the bioinformatics analyses of genes differentially expressed between DLBCL and CLL conducted herein can both improve the accuracy of efforts to diagnose these two diseases while also supporting new disease treatment strategies that will contribute to better patient quality of life and treatment outcomes.

## Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#), further inquiries can be directed to the corresponding author/s.

## Ethics statement

The studies involving humans were approved by the Ethics Committee of the Hematology Hospital of the Chinese Academy of Medical Sciences. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

HL: Conceptualization, Formal analysis, Methodology, Project administration, Resources, Writing – original draft, Writing – review & editing. ZC: Methodology, Writing – review & editing. YS: Resources, Writing – review & editing. YL: Methodology, Writing – review & editing. HW: Formal analysis, Methodology, Writing – review & editing. FS: Methodology, Writing – review & editing. MX: Methodology, Writing – review & editing. ZG: Resources, Supervision, Writing – review & editing. YZ: Funding acquisition, Project administration, Supervision, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by grants from National Natural Science Foundation of China (82270148), CAMS Innovation Fund for Medical Sciences (2023-I2M-2-007), Haihe Laboratory of Cell Ecosystem Innovation Fund (24HHXBSS00008) and Tianjin Health and Wellness Industry National Smart Social Governance Experimental Feature Base 2023 Annual Project for Leading Talents in Scientific Research (TJHIA-2023-022).

## References

1. Takahara T, Nakamura S, Tsuzuki T, Satou A. The immunology of DLBCL. *Cancers (Basel)*. (2023) 15:835. doi: 10.3390/cancers15030835
2. Campo E, Jaffe ES, Cook JR, Quintanilla-Martinez L, Swerdlow SH, Anderson KC, et al. The international consensus classification of mature lymphoid neoplasms: A report from the clinical advisory committee. *Blood*. (2022) 140:1229–53. doi: 10.1182/blood.2022015851
3. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. (2000) 403:503–11. doi: 10.1038/35000501
4. Davis RE, Ngo VN, Lenz G, Tolar P, Young RM, Romesser PB, et al. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature*. (2010) 463:88–92. doi: 10.1038/nature08638
5. Chen LF, Monti S, Juszczynski P, Ouyang J, Chapuy B, Neuberg D, et al. Syk inhibition modulates distinct pi3k/akt- dependent survival pathways and cholesterol biosynthesis in diffuse large B cell lymphomas. *Cancer Cell*. (2013) 23:826–38. doi: 10.1016/j.ccr.2013.05.002
6. Montesinos-Rongen M, Terrao M, May C, Marcus K, Blümcke I, Hellmich M, et al. The process of somatic hypermutation increases polyreactivity for central nervous system antigens in primary central nervous system lymphoma. *Haematologica*. (2021) 106:708–17. doi: 10.3324/haematol.2019.242701
7. Montesinos-Rongen M, Küppers R, Schlüter D, Spieker T, Van Roost D, Schaller C, et al. Primary central nervous system lymphomas are derived from germinal-center

## Acknowledgments

All authors kindly thank the core facilities of the State Key Laboratory of Experimental Hematology, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College for the technical assistance.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2025.1516946/full#supplementary-material>

8. Kawano T, Tsuyuki Y, Suzuki Y, Shimada K, Kato S, Takahara T, et al. Clinicopathologic Analysis of Primary Adrenal Diffuse Large B-Cell Lymphoma: A Reappraisal of 23 Japanese Patients based on EBV Association and Pd-L1 Expression in Tumor cells. *Am J Surg Pathol*. (2021) 45:1606–15. doi: 10.1097/pas.0000000000001809
9. Hiam-Galvez KJ, Allen BM, Spitzer MH. Systemic immunity in cancer. *Nat Rev Cancer*. (2021) 21:345–59. doi: 10.1038/s41568-021-00347-z
10. Ishikawa E, Tanaka T, Shimada K, Kohno K, Satou A, Eladl AE, et al. A prognostic model, including the EBV status of tumor cells, for primary gastric diffuse large B-cell lymphoma in the rituximab era. *Cancer Med*. (2018) 7:3510–20. doi: 10.1002/cam4.1595
11. Ishikawa E, Kato S, Shimada K, Tanaka T, Suzuki Y, Satou A, et al. Clinicopathological analysis of primary intestinal diffuse large B-cell lymphoma: prognostic evaluation of cd5, pd-L1, and epstein-barr virus on tumor cells. *Cancer Med*. (2018) 7:6051–63. doi: 10.1002/cam4.1875
12. Miyagi S, Ishikawa E, Nakamura M, Shimada K, Yamamura T, Furukawa K, et al. Reappraisal of primary Epstein-Barr virus (EBV)-positive diffuse large B-cell lymphoma of the gastrointestinal tract: comparative analysis among immunosuppressed and nonimmunosuppressed stage I and II-IV patients. *Am J Surg Pathol*. (2020) 44:1173–83. doi: 10.1097/pas.0000000000001499

13. Cardoso Borges F, Ramos A, Lourenço A, Gomes da Silva M, Miranda A. Detailing the epidemiological and clinical characteristics of chronic lymphocytic leukaemia in Portugal—results from a population-based cancer registry cohort study. *PLoS One*. (2021) 16:e0258423. doi: 10.1371/journal.pone.0258423
14. Weil C, Chodick G, Shalev V, Kan I, Afik R, Cohen R, et al. Epidemiology, longitudinal treatment patterns and outcomes of chronic lymphocytic leukemia in Israel. *Leuk Lymphoma*. (2021) 62:1136–45. doi: 10.1080/10428194.2020.1858293
15. Bewarder M, Stilgenbauer S, Thurner L, Kaddu-Mulindwa D. Current treatment options in CLL. *Cancers (Basel)*. (2021) 13:2468. doi: 10.3390/cancers13102468
16. González-Gascón YMI, Muñoz-Novas C, Rodríguez-Vicente AE, Quijada-Álamo M, Hernández-Sánchez M, Pérez-Carretero C, et al. From biomarkers to models in the changing landscape of chronic lymphocytic leukemia: evolve or become extinct. *Cancers (Basel)*. (2021) 13:1782. doi: 10.3390/cancers13081782
17. Hotinski AK, Best OG, Kuss BJ. The future of laboratory testing in chronic lymphocytic leukaemia. *Pathology*. (2021) 53:377–84. doi: 10.1016/j.pathol.2021.01.006
18. Griggio V, Perutelli F, Salvetti C, Boccellato E, Boccadoro M, Vitale C, et al. Immune dysfunctions and immune-based therapeutic interventions in chronic lymphocytic leukemia. *Front Immunol*. (2020) 11:594556. doi: 10.3389/fimmu.2020.594556
19. Ten Hacken E, Burger JA. Microenvironment interactions and B-cell receptor signaling in chronic lymphocytic leukemia: implications for disease pathogenesis and treatment. *Biochim Biophys Acta*. (2016) 1863:401–13. doi: 10.1016/j.bbmc.2015.07.009
20. Zarobkiewicz M, Kowalska W, Chocholska S, Tomczak W, Szymańska A, Morawska I, et al. High M-dmsc percentage as a negative prognostic factor in chronic lymphocytic leukaemia. *Cancers (Basel)*. (2020) 12:2614. doi: 10.3390/cancers12092614
21. Forconi F, Moss P. Perturbation of the normal immune system in patients with CLL. *Blood*. (2015) 126:573–81. doi: 10.1182/blood-2015-03-567388
22. Peters FS, Strefford JC, Eldering E, Kater AP. T-cell dysfunction in chronic lymphocytic leukemia from an epigenetic perspective. *Haematologica*. (2021) 106:1234–43. doi: 10.3324/haematol.2020.267914
23. Burger JA, Gribben JG. The microenvironment in chronic lymphocytic leukemia (CLL) and other B cell malignancies: insight into disease biology and new targeted therapies. *Semin Cancer Biol*. (2014) 24:71–81. doi: 10.1016/j.semcancer.2013.08.011
24. Vlachonikola E, Stamatopoulos K, Chatzidimitriou A. T cell defects and immunotherapy in chronic lymphocytic leukemia. *Cancers (Basel)*. (2021) 13:3255. doi: 10.3390/cancers13133255
25. Bond DA, Woyach JA. Targeting btk in CLL: beyond ibrutinib. *Curr Hematol Malign Rep*. (2019) 14:197–205. doi: 10.1007/s11899-019-00512-0
26. Fraietta JA, Lacey SF, Orlando EJ, Pruteanu-Malinici I, Gohil M, Lundh S, et al. Determinants of response and resistance to CD19 chimeric antigen receptor (CAR) T cell therapy of chronic lymphocytic leukemia. *Nat Med*. (2018) 24:563–71. doi: 10.1038/s41591-018-0010-1
27. Cook-Mills JM, Marchese ME, Abdala-Valencia H. Vascular cell adhesion molecule-1 expression and signaling during disease: regulation by reactive oxygen species and antioxidants. *Antioxid Redox Signal*. (2011) 15:1607–38. doi: 10.1089/ars.2010.3522
28. Dessein PH, Joffe BI, Singh S. Biomarkers of endothelial dysfunction, cardiovascular risk factors and atherosclerosis in rheumatoid arthritis. *Arthritis Res Ther*. (2005) 7:R634–43. doi: 10.1186/ar1717
29. Wilhelmi MH, Leyh RG, Wilhelmi M, Haverich A. Upregulation of endothelial adhesion molecules in hearts with congestive and ischemic cardiomyopathy: immunohistochemical evaluation of inflammatory endothelial cell activation. *Eur J Cardiothorac Surg*. (2005) 27:122–7. doi: 10.1016/j.ejcts.2004.09.027
30. Semaan HB, Gurbel PA, Anderson JL, Muhlestein JB, Carlquist JF, Horne BD, et al. Soluble VCAM-1 and E-selectin, but not ICAM-1 discriminate endothelial injury in patients with documented coronary artery disease. *Cardiology*. (2000) 93:7–10. doi: 10.1159/00006995
31. Willeit K, Pechlaner R, Willeit P, Skrobilin P, Paulweber B, Schernthaner C, et al. Association between vascular cell adhesion molecule 1 and atrial fibrillation. *JAMA Cardiol*. (2017) 2:516–23. doi: 10.1001/jamacardio.2017.0064
32. Kong DH, Kim YK, Kim MR, Jang JH, Lee S. Emerging roles of vascular cell adhesion molecule-1 (Vcam-1) in immunological disorders and cancer. *Int J Mol Sci*. (2018) 19:1057. doi: 10.3390/ijms19041057
33. Boulanger E, Wautier MP, Wautier JL, Boval B, Panis Y, Wernert N, et al. Ages bind to mesothelial cells via rage and stimulate VCAM-1 expression. *Kidney Int*. (2002) 61:148–56. doi: 10.1046/j.1523-1755.2002.00115.x
34. Zhong L, Simard MJ, Huot J. Endothelial microRNAs regulating the NF- $\kappa$ B pathway and cell adhesion molecules during inflammation. *FASEB J*. (2018) 32:4070–84. doi: 10.1096/fj.201701536R
35. Malek M, Kielkowska A, Chessa T, Anderson KE, Barneda D, Pir P, et al. Pten regulates PI(3,4)P<sub>2</sub> signaling downstream of class I PI3K. *Mol Cell*. (2017) 68:566–80.e10. doi: 10.1016/j.molcel.2017.09.024
36. Kameda MM, Messer KS, Ralainirina N, Li HY, Leem CJ, Gorjestani S, et al. PI3K $\gamma$  is a molecular switch that controls immune suppression. *Nature*. (2016) 539:437–42. doi: 10.1038/nature19834
37. Zhang SB, Chung WC, Wu GM, Egan SE, Miele L, Xu KL. Manic fringe promotes a claudin-low breast cancer phenotype through notch-mediated PI3K $\gamma$  induction. *Cancer Res*. (2015) 75:1936–43. doi: 10.1158/0008-5472.CCR-14-3303
38. Chang J, Hong L, Liu YZ, Pan YW, Yang H, Ye WR, et al. Targeting PI3K $\gamma$  in combination with paclitaxel as a potential therapeutic regimen in claudin-low breast cancer. *Cancer Manag Res*. (2020) 12:2641–51. doi: 10.2147/cmar.S250171
39. Chung WC, Zhou XC, Atfi A, Xu KL. PI3K $\gamma$  is a potential therapeutic target in androgen receptor-indifferent metastatic prostate cancer. *Am J Pathol*. (2020) 190:2194–202. doi: 10.1016/j.ajpath.2020.07.013
40. Wang YQ, Wu DH, Wei DH, Shen JY, Huang ZW, Liang XY, et al. Tead4 is a master regulator of high-risk nasopharyngeal carcinoma. *Sci Adv*. (2023) 9:eadd0960. doi: 10.1126/sciadv.add0960
41. Gao HF, Yu GH, Zhang X, Yu S, Sun Y, Li YH. Bzw2 gene knockdown induces cell growth inhibition, G1 arrest and apoptosis in muscle-invasive bladder cancers: A microarray pathway analysis. *J Cell Mol Med*. (2019) 23:3905–15. doi: 10.1111/jcmm.14266
42. Kozel C, Thompson B, Hustak S, Moore C, Nakashima A, Singh CR, et al. Overexpression of eIF5 or its protein mimic 5mp perturbs eIF2 function and induces ATF4 translation through delayed re-initiation. *Nucleic Acids Res*. (2016) 44:8704–13. doi: 10.1093/nar/gkw559
43. Jin X, Liao MM, Zhang LH, Yang MY, Zhao JF. Role of the novel gene BZW2 in the development of hepatocellular carcinoma. *J Cell Physiol*. (2019) 234:16592–600. doi: 10.1002/jcp.28331
44. Sato K, Masuda T, Hu QJ, Tobo T, Gillaspie S, Niida A, et al. Novel oncogene 5mp1 reprograms C-myc translation initiation to drive malignant phenotypes in colorectal cancer. *EBioMedicine*. (2019) 44:387–402. doi: 10.1016/j.ebiom.2019.05.058
45. Obayashi E, Luna RE, Nagata T, Martin-Marcos P, Hiraishi H, Singh CR, et al. Molecular landscape of the ribosome pre-initiation complex during mRNA scanning: structural role for eIF3c and its control by eIF5. *Cell Rep*. (2017) 18:2651–63. doi: 10.1016/j.celrep.2017.02.052
46. Li GX, Lu AQ, Chen AN, Geng S, Xu Y, Chen X, et al. BZW2/5mp1 acts as a promising target in hepatocellular carcinoma. *J Cancer*. (2021) 12:5125–35. doi: 10.7150/jca.53282
47. Igarashi K, Kurosaki T, Roychoudhuri R. Bach transcription factors in innate and adaptive immunity. *Nat Rev Immunol*. (2017) 17:437–50. doi: 10.1038/nri.2017.26
48. Yao C, Lou GH, Sun HW, Zhu Z, Sun Y, Chen ZY, et al. Bach2 enforces the transcriptional and epigenetic programs of stem-like CD8<sup>+</sup> T cells. *Nat Immunol*. (2021) 22:370–80. doi: 10.1038/s41590-021-00868-7
49. Imanowski CJ, Whiteside SK, Lozano T, Evans AC, Benson JD, Courreges CJF, et al. Bach2 restricts NK cell maturation and function, limiting immunity to cancer metastasis. *J Exp Med*. (2022) 219:e20211476. doi: 10.1084/jem.20211476
50. Basso K, Dalla-Favera R. Roles of BCL6 in normal and transformed germinal center B cells. *Immunol Rev*. (2012) 247:172–83. doi: 10.1111/j.1600-065X.2012.01112.x
51. Ghetu AF, Corcoran CM, Cerchietti L, Bardwell VJ, Melnick A, Privé GG. Structure of a BCOF corepressor peptide in complex with the BCL6 BTB domain dimer. *Mol Cell*. (2008) 29:384–91. doi: 10.1016/j.molcel.2007.12.026
52. Huang CX, Gonzalez DG, Cote CM, Jiang YW, Hatzi K, Teater M, et al. The BCL6 RD2 domain governs commitment of activated B cells to form germinal centers. *Cell Rep*. (2014) 8:1497–508. doi: 10.1016/j.celrep.2014.07.059
53. Louwen F, Kreis NN, Ritter A, Friemel A, Solbach C, Yuan J. BCL6, a key oncogene, in the placenta, pre-eclampsia and endometriosis. *Hum Reprod Update*. (2022) 28:890–909. doi: 10.1093/humupd/dmac027
54. Ren ZL, Gao YF, Gao Y, Liang GM, Chen Q, Jiang SJ, et al. Distinct placental molecular processes associated with early-onset and late-onset preeclampsia. *Thrombostasis*. (2021) 11:5028–44. doi: 10.7150/thno.56141
55. Guo Y, Luan LM, Patil NK, Sherwood ER. Immunobiology of the IL-15/IL-15R $\alpha$  complex as an antitumor and antiviral agent. *Cytokine Growth Factor Rev*. (2017) 38:10–21. doi: 10.1016/j.cytogfr.2017.08.002
56. Waldmann TA, Tagaya Y. The multifaceted regulation of interleukin-15 expression and the role of this cytokine in NK cell differentiation and host response to intracellular pathogens. *Annu Rev Immunol*. (1999) 17:19–49. doi: 10.1146/annurev.immunol.17.1.19
57. McMichael AJ. Environment, sustainability and health: the learning curve steepens. *Med J Aust*. (2001) 175:569–70. doi: 10.5694/j.1326-5377.2001.tb143728.x
58. Mueller YM, Bojczuk PM, Halstead ES, Kim AH, Witek J, Altman JD, et al. IL-15 enhances survival and function of HIV-specific CD8<sup>+</sup> T cells. *Blood*. (2003) 101:1024–9. doi: 10.1182/blood-2002-07-1957
59. Mavilio D, Lombardo G, Kinter A, Fogli M, La Sala A, Ortolano S, et al. Characterization of the defective interaction between a subset of natural killer cells and dendritic cells in HIV-1 infection. *J Exp Med*. (2006) 203:2339–50. doi: 10.1084/jem.20060894
60. Chehimi J, Marshall JD, Salvucci O, Frank I, Chehimi S, Kawecky S, et al. IL-15 enhances immune functions during HIV infection. *J Immunol*. (1997) 158:5978–87. doi: 10.4049/jimmunol.158.12.5978
61. Kottlilil S, Chun TW, Moir S, Liu SY, McLaughlin M, Hallahan CW, et al. Innate immunity in human immunodeficiency virus infection: effect of viremia on natural killer cell function. *J Infect Dis*. (2003) 187:1038–45. doi: 10.1086/368222

62. Rosa JP, Raslova H, Bryckaert M. Filamin A: key actor in platelet biology. *Blood*. (2019) 134:1279–88. doi: 10.1182/blood.2019000014
63. Lange M, Kasper B, Bohring A, Rutsch F, Kluger G, Hoffjan S, et al. 47 patients with flna associated periventricular nodular heterotopia. *Orphanet J Rare Dis*. (2015) 10:134. doi: 10.1186/s13023-015-0331-9
64. Bandaru S, Ala C, Salimi R, Akula MK, Ekstrand M, Devarakonda S, et al. Targeting filamin a reduces macrophage activity and atherosclerosis. *Circulation*. (2019) 140:67–79. doi: 10.1161/circulationaha.119.039697
65. Aldinucci D, Gloghini A, Pinto A, Colombatti A, Carbone A. The role of cd40/cd40l and interferon regulatory factor 4 in hodgkin lymphoma microenvironment. *Leuk Lymphoma*. (2012) 53:195–201. doi: 10.3109/10428194.2011.605190
66. Jiao XM, Nawab O, Patel T, Kossenkov AV, Halama N, Jaeger D, et al. Recent advances targeting ccr5 for cancer and its role in immuno-oncology. *Cancer Res*. (2019) 79:4801–7. doi: 10.1158/0008-5472.Can-19-1167
67. Aldinucci D, Borghese C, Casagrande N. The ccl5/ccr5 axis in cancer progression. *Cancers (Basel)*. (2020) 12:1765. doi: 10.3390/cancers12071765
68. El Hussein S, Medeiros LJ, Lyapichev KA, Fang H, Jelloul FZ, Fiskus W, et al. Immunophenotypic and genomic landscape of richter transformation diffuse large B-cell lymphoma. *Pathology*. (2023) 55:514–24. doi: 10.1016/j.pathol.2022.12.354
69. Pineyro JA, Lopez-Oreja I, Nadeu F, Martinez-Farran A, Arostegui JI, Lopez-Guerra M, et al. Association of genomic alterations with the presence of serum monoclonal proteins in chronic lymphocytic leukemia. *Cells*. (2024) 13:1839. doi: 10.3390/cells13221839
70. Charalampopoulou S, Chapiro E, Nadeu F, Zenz T, Bea S, Martinez-Farran A, et al. Epigenetic features support the diagnosis of B-cell prolymphocytic leukemia and identify 2 clinicobiological subtypes. *Blood Adv*. (2024) 8:6297–307. doi: 10.1182/bloodadvances.2024013327
71. Bertilaccio MTS, Chen SS. Mouse models of chronic lymphocytic leukemia and richter transformation: what we have learnt and what we are missing. *Front Immunol*. (2024) 15:1376660. doi: 10.3389/fimmu.2024.1376660
72. Ten Hacken E, Gruber M, Hernandez-Sanchez M, Hoffmann GB, Baranowski K, Redd RA, et al. Generation of mouse models carrying B cell restricted single or multiplexed loss-of-function mutations through crispr-cas9 gene editing. *STAR Protoc*. (2023) 4:102165. doi: 10.1016/j.xpro.2023.102165
73. Pongas G, Cheson BD. Pi3k signaling pathway in normal B cells and indolent B-cell Malignancies. *Semin Oncol*. (2016) 43:647–54. doi: 10.1053/j.seminoncol.2016.11.011
74. Dunleavy K, Erdmann T, Lenz G. Targeting the B-cell receptor pathway in diffuse large B-cell lymphoma. *Cancer Treat Rev*. (2018) 65:41–6. doi: 10.1016/j.ctrv.2018.01.002
75. Sindaco P, Pandey H, Isabelle C, Chakravarti N, Brammer JE, Porcu P, et al. The role of interleukin-15 in the development and treatment of hematological Malignancies. *Front Immunol*. (2023) 14:1141208. doi: 10.3389/fimmu.2023.1141208
76. Cai MH, Huang X, Huang XT, Ju DW, Zhu YZ, Ye L. Research progress of interleukin-15 in cancer immunotherapy. *Front Pharmacol*. (2023) 14:1184703. doi: 10.3389/fphar.2023.1184703