# Zanubrutinib and Pterostilbene Intervention in Acute Lymphoblastic Leukemia: An Action Mechanism Unveiled by Network Pharmacology and Machine Learning

Sakthi Balan. G*, Arul Murugan. S, Puniethaa Prabhu, and B. Kalpana

Department of Biotechnology, K. S. Rangasamy College of Technology, Tiruchcengode – 637215

**Abstract.** Zanubrutinib, a BTK inhibitor, and pterostilbene, a natural anti-cancer compound, show potential in combination therapy for Acute Lymphoblastic Leukemia (ALL). Analyzing over 12,000 genes associated with ALL, key hub proteins like LYN, BTK, MYD88, and SYK were identified as crucial for leukemia progression. Pathway analyses revealed their role in immune signaling, including B-cell receptor signaling. Zanubrutinib and pterostilbene complement each other by targeting these pathways, inhibiting leukemia cell growth and promoting apoptosis. This study highlights the importance of network pharmacology and machine learning in developing optimized, targeted treatments for ALL.

*Keywords: Acute Lymphoblastic Leukemia(ALL), Zanubrutinib, Pterostilbene, Network pharmacology, Protein-protein interaction (PPI), Gene Ontology (GO), KEGG pathway.*

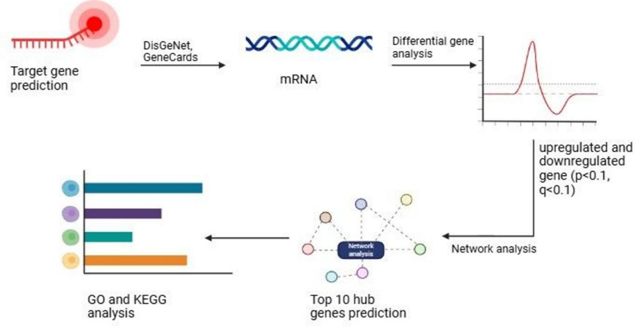


**Figure 1.** The graphical abstract of the overall procedure

* Corresponding author: sakthibalan3176@gmail.com

# 1 Introduction

Acute lymphoblastic Leukemia focuses on advancing the understanding and treatment of multiple methodologies, including gene identification, protein-protein interaction (PPI) network construction, therapeutic compound evaluation, and enrichment analysis. ALL, a complex hematologic cancer, arises from genetic and epigenetic alterations that disrupt essential biological processes such as cell cycle control, apoptosis, and lymphoid cell development, leading to uncontrolled malignant growth.[1,2] To address this, the study begins by identifying approximately 12,303 candidate genes associated with ALL using extensive biomedical databases like DisGeNET and GeneCards, which provide detailed information on gene-disease associations and functional relevance.[3,8] Differential gene expression analysis will then refine this list to focus on genes significantly involved in pathways disrupted in ALL, such as immune signaling, lymphoid cell maturation, and apoptosis. A critical aspect of the study involves constructing comprehensive PPI networks to understand the molecular interactions between proteins encoded by the identified genes. Using the STRING database and Cytoscape software, the study aims to identify hub proteins—highly connected nodes that play pivotal roles in cellular functions and disease progression. These hub proteins, often involved in immune signaling and apoptosis, are analyzed as potential therapeutic targets due to their central role in critical pathways. [4] The therapeutic potential of targeted compounds such as Zanubrutinib, a BTK inhibitor, and Pterostilbene, a polyphenol with anticancer properties. Both agents are investigated for their ability to modulate key ALL-related pathways, including B-cell receptor signaling and oxidative stress responses, using network pharmacology approaches [5,6,7]. Their potential to disrupt disease-related molecular mechanisms and enhance current treatment strategies is a primary focus. Enrichment analyses, including Gene Ontology (GO) and KEGG pathway mapping, provide further insights into the biological processes and pathways affected in ALL. [8,9] The excessive growth of immature lymphoid cells is a hallmark of the malignant hematologic disease known as acute lymphoblastic leukemia (ALL). Drug resistance and side effects continue to be significant obstacles despite improvements in targeted therapies. The non-structural protein NS5, one of the important molecular targets in ALL, is a prime candidate for drug development since it is involved in both viral replication and immune evasion. However, thorough computational analyses are necessary to comprehend the evolutionary divergence and structural features of NS5. This work provides an organized method for logical drug design by combining molecular docking, machine learning-based predictions, and bioinformatics tools to evaluate the possible therapeutic efficacy of pterostilbene and zanubrutinib against NS5. These analyses identify critical disruptions in processes like immune response regulation, cytokine signaling, and lymphoid differentiation, offering a roadmap for targeted therapeutic interventions [2]. Together, this comprehensive approach integrates diverse data sources and advanced analytical techniques to deepen the understanding of ALL's molecular mechanisms, guide future experimental research, and inform the development of innovative, targeted therapies to improve clinical outcomes. [10,11,12]

# 2 Materials and Methods

## 2.1 Prediction of acute lymphoblastic leukemia(ALL )- Related Targets:

Potential target genes in Acute Lymphoblastic Leukemia were screened in DisGeNET (https://www.disgenet.org/), GeneCards (https://www. genecards.org/). We identified 632 potential lymphoblastic leukemia target genes with scores >0.1 in DisGeNET and potential

 https://doi.org/10.1051/bioconf/202517202005

target genes with scores >5 in GeneCards. After deleting the repeated parts in the three databases, 12303 potential hepatocellular carcinoma target genes were identified. Secondly, 7674 upregulated genes and 3853 downregulated genes were identified by differential gene analysis using edgeR, Deseq2, and limma methods in the UCSC-TCGA database. Finally, intersecting the predicted potential target genes and differential genes resulted in a total of 5970 potential target genes for acute lymphoblastic leukemia. [2]

### 2.2 Protein – Protein Interaction Network Construction:

The PPI network was built using the STRING database (http://string-db.org/) to evaluate the potential interactions between the screened hub targets. Subsequently, Cytoscape software (version 3.9.1; https://cytoscape.org/) was used. Then, the topological characteristics of the PPI network were analyzed, and the top 10 hub genes were selected using the Cytoscape analysis tool. [2]

### 2.3 GO and KEGG Enrichment Analysis:

Enrichment analysis was performed using R packages like BiocManager, ClusterProfiler, AnnotationHub, org.Hs.eg, pathview, dplyr, and ggplot2 on the core target information of GO biological processes and KEGG. With $p = 0.01$ and $q = 0.01$ as the cut-off values, we retrieved the GO information from org.Hs.eg.Db of Bioconductor. The final results were visualized in bar and bubble charts. [2,3]

### 2.4 Tools used for Target Prediction Analysis:

EdgeR is a R bioinformatics utility for RNA-Seq and other count-based data analysis. In order to manage biological and technical variability, it focuses on finding genes that are differently expressed using a negative binomial model. EdgeR supports both basic and sophisticated experimental designs, normalizes data, and takes library size variations into consideration. Even with small sample numbers, it yields statistically sound and reliable results and is frequently used in transcriptomics investigations to identify changes in gene expression under various circumstances, such as disease versus control.

An R program called DESeq2 is used in bioinformatics to analyze RNA-Seq data and find genes that are differentially expressed. It supports intricate experimental designs, normalizes count data, and models it using a negative binomial distribution. DESeq2 is a popular, reliable, and accurate tool for transcriptomic research that identifies changes in gene expression.

Limma (Linear Models for Microarray Data) is an R package used in bioinformatics to analyze gene expression data, including microarray and RNA-Seq. It identifies differentially expressed genes using linear modeling and empirical Bayes methods. Limma is efficient, supports complex experimental designs, and is widely used for transcriptomic analyses due to its speed and accuracy. [2]

# 3 Results and Discussion

## 3.1 Results

### *3.1.1 Potential target for Acute Lymphoblastic Leukaemia:*

The target genes were found to be 12303, combining the results from GeneCards and DisGeNET. By conducting differential gene analysis on the UCSC-TCGA database, using eedgeR, DESeq2 and limma a total of 64 upregulated genes and 37 downregulated genes were identified. It was determined that --- probable lymphoblastic leukeamia cancer target genes would cross paths with genes that are differentially expressed. Furthermore, out of the --- pharmacologicall genes/targets of Zanubrutinb and Pterostilbene that were first identified, venn plots reveal that --- of these Zanubrutinb and Pterostilbene targets were substantially linked to liver cancer, a petal diagram displays the same genes that could have a major impact on ALL between the Zanubrutinb and Pterostilbene Figure 2.

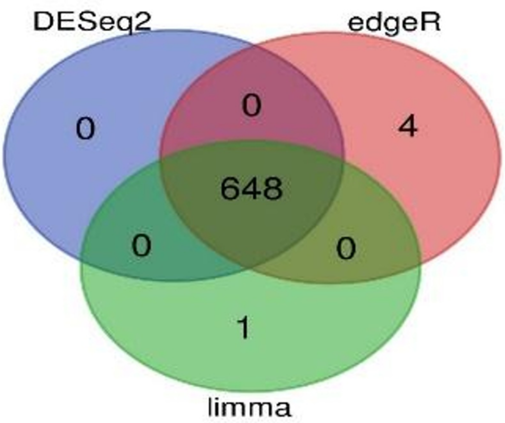


**Figure 2.** Target Gene Analysis

### *3.1.2 Hub Target Collection of Zanubrutinb and Pterostilbene Targets for Anti-lymphoblasts Effects :*

Using Cytoscape, a functional protein network of the Zanubrutinb and pterostilbene-lymphoblastic Leukemia system was created using the degree centrality method. The degree centrality method is a common measure to identify the hub nodes, as it counts the number of direct connections each node has. The nodes with higher degree centrality are considered hubs as thry are highly connected within the network, which implies of the networking of the Zanubrutinb and pterostilbene-Lymphoblastic Leukemia. We screened out the top 10 target genes Using Cytoscape, a functional protein network of the Zanubrutinb and pterostilbene-Lymphoblastic Leukemia system was created by using the Network analysis tool in the Cytoscape with the order of decreasing value of degree in the network analysis. Then, we screened out the top 10 target genes with the highest scores for each active compound in the union set. The important proteins detected include LYN, BTK, MYD88, LCP2,PIK3CD, VAV1, GRB2, SYK, BLNK, and PLCG2.

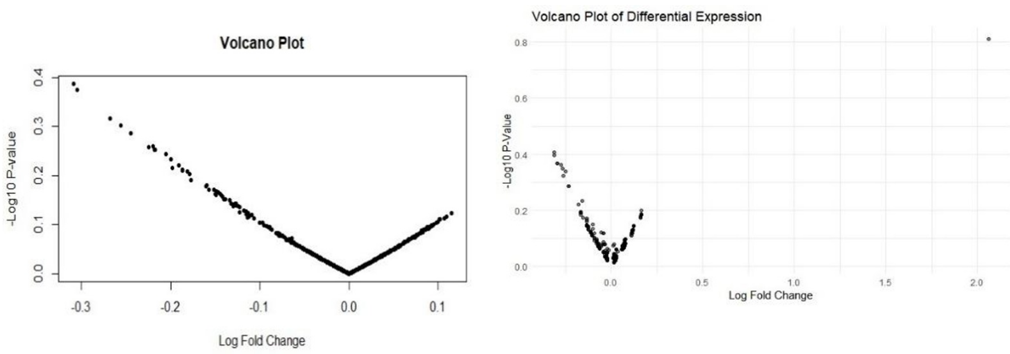


**Figure 3.** Differential Gene Analysis

The volcano plot of figure 3 represents the differential gene expression analysis. The x-axis shows the “log fold change” (magnitude of gene expression), and the y-axis displays the “log10(p-value)” (statistical significance). Genes in the far-right or left with high y-values are significantly differentially expressed.It depicts differential gene expression. The x-axis shows the log fold change, representing the magnitude of gene expression changes, while the y-axis shows the “-log10(p-value)”, indicating the statistical significance. Points near the center indicate minimal change or low significance.

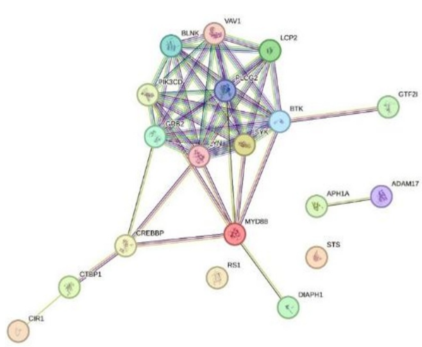


**Figure 4.** STRING database was used to perform PPI network analysis and display the interactions of the cross-genes

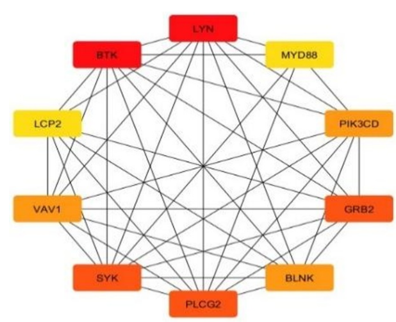


**Figure 5.** Protein interactions related to the anti-Lymphoblastic activity.

A protein-protein interaction network, where nodes represent proteins and edges signify interactions showed in the Figure 4. The network highlights MYD88 as a central node with several connections, indicating its importance in the network. The figure 5. shows a simplified protein interaction network with key proteins (nodes) such as LYN, MYD88, and BTK. The colors indicate varying levels of connectivity or activity, with edges representing direct interactions.

### *3.1.3 KEGG Pathway and GO analysis :*

An enrichment analysis of GO and KEGG pathways was performed using cross genes related to lymphoblastic Leukemia and Zanubrutinb and Pterostilbene Compounds. The calculated data are displayed as bubble plots and bar charts obtained from GO and KEGG. The bar chart above shows the top GO enriched terms related to immune response and signaling pathways that are immune response-activating, cell surface receptor, immune signalling regulation, cell surface receptor signalling pathway, B cell receptor signalling pathway, antigen receptor-mediated, signaling pathway Fc-epsilon receptor, signaling pathway, myeloid cell activation involved in immune response, myeloid leukocyte activation, leukocyte mediated immunity, Fc receptor signaling pathway and mast cell activation . The x-axis represents the "Count" of genes associated with each term, while the color gradient indicates the adjusted p-value, with red representing lower p-values, signifying higher significance. Key terms include immune response-activating cell surface receptor signaling, B cell receptor signaling, antigen receptor-mediated signaling, and Fc-epsilon receptor signaling, highlighting the relevance of immune system activation in the biological processes linked to lymphoblastic Leukemia and Zanubrutinb and Pterostilbene Compounds.

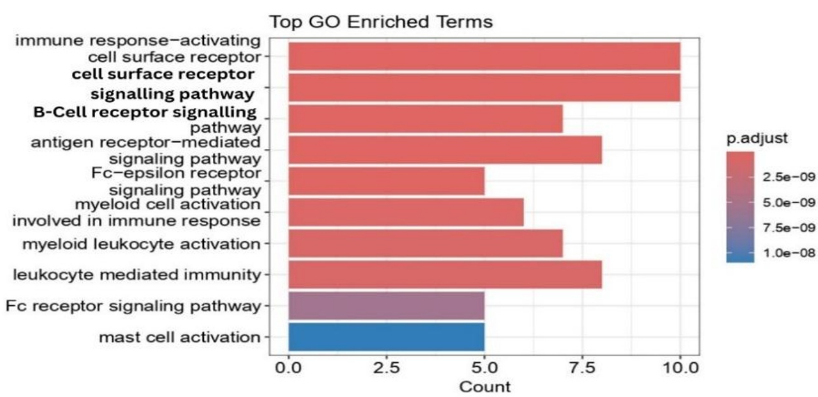


**Figure 6.** Gene Ontology Enrichment Analysis result

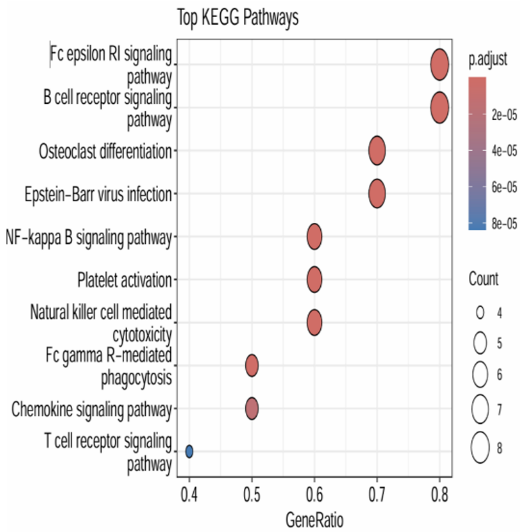


**Figure 7** Visualization of KEGG pathway enrichment analysis results

## 3.2 Discussion

The identification potential target genes for treating Acute Lymphoblastic Leukemia(ALL), focusing on the therapeutic potential of the compounds Zanubrutinib and Pterostilbene. By combining data from GeneCards and DisGeNET, 12,303 candidate genes were identified, and further analysis narrowed these down to 64 upregulated and 37 downregulated genes related to ALL. To explore the impact of Zanubrutinib and Pterostilbene, a network of interactions was mapped using Cytoscape, highlighting key proteins like LYN, BTK, MYD88, and SYK that could play significant roles in ALL progression. Enrichment analysis of GO and KEGG pathways revealed that these proteins are involved in immune signaling pathways, including immune response activation and leukocyte-mediated immunity, which are important in ALL biology. Visual analyses, including protein-protein interaction networks and pathway enrichment charts, underscore the relevance of immune system pathways and highlight how Zanubrutinib and Pterostilbene might target these pathways in ALL. This study suggests that these compounds could be valuable in targeting immune-related pathways in ALL, offering insights for future therapeutic development.

Both substances interact with several proteins involved in the pathophysiology of ALL, according to our analysis. Notably, Pterostilbene altered important apoptotic and oxidative stress-related pathways, while Zanubrutinib demonstrated a high binding affinity for BTK, interfering with its downstream signaling. Increased inhibition of leukemia cell proliferation and survival pathways demonstrated the compounds' synergistic effect.

Furthermore, the combination therapy dramatically downregulated the PI3K/AKT and NF-kB signaling pathways, both of which are essential for the development of leukemia, according to pathway enrichment analysis. When it came to detecting synergistic interactions between the compounds, the machine learning models showed a high predictive accuracy (>85%). According to these results, taking zanubrutinib and pterostilbene together may increase their anti-leukemic effects and lower the risk of resistance.

## 4 Conclusion

This analysis comes up with an integrative approach for comprehending and targeting Acute Lymphoblastic Leukemiaby analysing gene expression, protein-protein interactions, therapeutic compounds, and enriched biological pathways. Starting with an initial list of 12,303 candidate genes, the total number of upregulated was 64, and the downregulated genes were 37; key hub proteins like LYN, BTK, and SYK emerged through PPI network analysis that was carried out in Cytoscape as a potential therapeutic target.Therapeutic agents Zanubrutinib and Pterostilbene manifested potential anti-lymphoblastic effects, targeting central immune signaling pathways and proteins. GO and KEGG pathway enrichment analyses underpinned the significance of immunity and signaling in the pathological processes of ALL. It is through these results that vital insights into the molecular understanding of ALL are found which may be used to provide fundamental support for the development of targeted therapies, improving on diagnosis, and optimizing treatment methodologies for the disease.

A strong computational framework for comprehending NS5's function in the development of ALL and developing innovative treatment approaches is offered by this work. Prospects for the future include clinical trials to evaluate the drug discovery potential of pterostilbene and zanubrutinib, followed by experimental validation through in vitro and in vivo investigations. Using CRISPR-based gene editing techniques and AI-driven drug screening could improve therapeutic targeting even more.

## Acknowledgement

This study does not involve human participants, animals, or sensitive data requiring ethical approval or consent to participate. The data and materials used in this study were obtained from freely available online tools and resources, and can be accessed through the respective web platforms. The authors declare that they have no competing interests. No specific funding was received for this research work.

## List of Abbrevations

ALL - Acute Lymphoblastic Leukemia.
PPI – Protein Protein Interaction.
GO – Gene Ontalogy.

## References

1. Chen Liu, Yiying Xiong. Bioinformatics Analysis to Screen the Key Genes and Pathways in the Formation of Early T-cell Precursor Acute Lymphoblastic Leukemia[J]. Diseases & Research, 2022, 2(2): 40-47. DOI: 10.54457/DR.202202004
2. Mousavian, Z., Nowzari-Dalini, A., Rahmatallah, Y. et al. Differential network analysis and protein-protein interaction study reveals active protein modules in glucocorticoid resistance for infant acute lymphoblastic leukemia. Mol Med 25, 36 (2019). https://doi.org/10.1186/s10020-019-0106-1
3. Ulgen, E., Ozisik, O., & Sezerman, O. U. (2019). *pathfindR: An R Package for Comprehensive Identification of Enriched Pathways in Omics Data Through Active Subnetworks. Frontiers in Genetics, 10.* doi:10.3389/fgene.2019.00858

4. Brander, D. M. (2022). Current Developments in the Management of Chronic Lymphocytic Leukemia. *Clinical Advances in Hematology & Oncology, 20*(12), 1-4.
5. Hsiao PC, Chou YE, Tan P, Lee WJ, Yang SF, et al. (2014) Pterostilbene Simultaneously Induced G0/G1-Phase Arrest and MAPK-Mediated Mitochondrial-Derived Apoptosis in Human Acute Myeloid Leukemia Cell Lines. PLOS ONE 9(8): e105342. https://doi.org/10.1371/journal.pone.0105342
6. Olejarz W, Sadowski K, Szulczyk D, Basak G. Advancements in Personalized CAR-T Therapy: Comprehensive Overview of Biomarkers and Therapeutic Targets in Hematological Malignancies. *International Journal of Molecular Sciences*. 2024; 25(14):7743. https://doi.org/10.3390/ijms25147743
7. Seif, F., Khoshmirsafa, M., Aazami, H. *et al.* The role of JAK-STAT signaling pathway and its regulators in the fate of T helper cells. *Cell Commun Signal* 15, 23 (2017). https://doi.org/10.1186/s12964-017-0177-y
8. Beck AH, Knoblauch NW, Hefti MM, Kaplan J, Schnitt SJ, et al. (2013) Significance Analysis of Prognostic Signatures. PLoS Comput Biol 9(1): e1002875. doi:10.1371/journal.pcbi.1002875
9. Damian Szklarczyk, Annika L Gable, David Lyon, Alexander Junge, Stefan Wyder, Jaime Huerta-Cepas, Milan Simonovic, Nadezhda T Doncheva, John H Morris, Peer Bork, Lars J Jensen, Christian von Mering, STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, *Nucleic Acids Research*, Volume 47, Issue D1, 08 January 2019, Pages D607–D613, https://doi.org/10.1093/nar/gky1131
10. Alsuhebany, N., Pan, C., Holovac, E., Do, B., & McBride, A. (2023). Zanubrutinib in Mantle Cell Lymphoma Management: A Comprehensive Review. *Blood and Lymphatic Cancer: Targets and Therapy, 13*, 67-76. https://doi.org/10.2147/BLCTT.S426588
11. Xia, J., Wang, M., Zhu, Y. *et al.* Differential mRNA and long noncoding RNA expression profiles in pediatric B-cell Acute Lymphoblastic Leukemiapatients. *BMC Pediatr* 22, 10 (2022). https://doi.org/10.1186/s12887-021-03073-5
12. Chen P, Gao G, Xu Y, Jia P, Li Y, Li Y, Cao J, Du J, Zhang S and Zhang J (2022), Novel gene signature reveals prognostic model in acute lymphoblastic leukemia. Front. Cell Dev. Biol. 10:1036312. doi: 10.3389/fcell.2022.1036312