Cancer Immunology, Immunotherapy (2025) 74:131
https://doi.org/10.1007/s00262-025-03972-x

RESEARCH

Identification of intratumoral microbiome-driven immune modulation and therapeutic implications in diffuse large B-cell lymphoma

Zheng Yijia[1] · Xiaoyu Li[1] · Lina Ma[1] · Siying Wang[1] · Hong Du[1] · Yun Wu[2] · Jing Yu[1] · Yunxia Xiang[3,4] · Daiqin Xiong[3,4] · Huiting Shan[3,4] · Yubo Wang[3,4] · Zhi Wang[3,4] · Jianping Hao[5] · Jie Wang[3,4]

Received: 28 October 2024 / Accepted: 7 February 2025 / Published online: 3 March 2025

Abstract

Objective Diffuse large B-cell lymphoma (DLBCL) is the most common subtype of non-Hodgkin lymphoma, with significant clinical heterogeneity. Recent studies suggest that the intratumoral microbiome may influence the tumor microenvironment, affecting patient prognosis and therapeutic responses. This study aims to identify microbiome-related subtypes in DLBCL and assess their impact on prognosis, immune infiltration, and therapeutic sensitivity.

Methods Transcriptomic and microbiome data from 48 DLBCL patients were obtained from public databases. Consensus clustering was used to classify patients into distinct microbiome-related subtypes. Functional enrichment analysis, immune infiltration assessments, and single-cell RNA sequencing were performed to explore the biological characteristics of these subtypes. Drug sensitivity predictions were made using the OncoPredict tool. Hub genes' expression and biological function were validated and inferred in cell lines and independent cohorts of DLBCL.

Results Two distinct microbiome-related subtypes were identified. Patients in Cluster 1 exhibited significantly better overall survival ($P < 0.05$), with higher immune infiltration of regulatory T cells and M0 macrophages compared to Cluster 2, which was associated with poorer outcomes. Functional enrichment analysis revealed that genes in Cluster 1 were involved in immune regulatory pathways, including cytokine–cytokine receptor interactions and chemokine signaling, suggesting enhanced anti-tumor immune responses. In contrast, genes in Cluster 2 were enriched in immunosuppressive pathways, contributing to a less favorable prognosis. Single-cell RNA sequencing analysis revealed significant heterogeneity in immune cell populations within the tumor microenvironment. B cells exhibited the most notable heterogeneity, as indicated by stemness and differentiation potential scoring. Intercellular communication analysis demonstrated that B cells played a key role in immune cell interactions, with significant differences observed in MIF signaling between B-cell subgroups. Pseudo-time analysis further revealed distinct differentiation trajectories of B cells, highlighting their potential heterogeneity across different immune environments. Metabolic pathway analysis showed significant differences in the average expression levels of metabolic pathways among B-cell subgroups, suggesting functional specialization. Furthermore, interaction analysis between core genes involved in B-cell differentiation and microbiome-driven differentially expressed genes identified nine common genes (GSTM5, LURAP1, LINC02802, MAB21L3, C2CD4D, MMEL1, TSPAN2, and CITED4), which were found to play critical roles in B-cell differentiation and were influenced by the intratumoral microbiome. DLBCL cell lines and clinical cohorts validated that MMEL1 and CITED4 with important biologically function in DLBCL cell survival and subtype classification.

Conclusions This study demonstrates the prognostic significance of the intratumoral microbiome in DLBCL, identifying distinct microbiome-related subtypes that impact immune infiltration, metabolic activity, and therapeutic responses. The findings provide insights into the immune heterogeneity within the tumor microenvironment, focusing on B cells and their differentiation dynamics. These results lay the foundation for microbiome-based prognostic biomarkers and personalized treatment approaches, ultimately aiming to enhance patient outcomes in DLBCL.

Zheng Yijia, Xiaoyu Li, and Lina Ma have contributed equally to this research.

Extended author information available on the last page of the article

Keywords Diffuse large B-cell lymphoma · Intratumoral microbiome · Immune infiltration · Tumor microenvironment · Personalized medicine

Introduction

Diffuse large B-cell lymphoma (DLBCL) is the most prevalent subtype of non-Hodgkin lymphoma, accounting for approximately 30% of cases worldwide [1]. Despite advances in therapeutic regimens such as R-CHOP (rituximab combined with cyclophosphamide, doxorubicin, vincristine, and prednisone), a significant proportion of patients experience relapse or develop resistance to therapy, resulting in poor prognosis [2]. The clinical heterogeneity of DLBCL underscores the necessity for a deeper understanding of the biological mechanisms driving disease progression and therapeutic outcomes [3].

Recent research has increasingly highlighted the critical role of the tumor microenvironment (TME) in cancer progression and response to therapy [4]. Components of the TME, including immune cells, stromal cells, and signaling molecules, can significantly influence tumor behavior [5]. Among these components, the intratumoral microbiome has emerged as a potential modulator of the immune landscape in various cancers [6]. Studies in gastrointestinal malignancies such as colorectal and pancreatic cancers have demonstrated that specific microbiomes within tumors can affect immune responses, potentially promoting or inhibiting tumor growth [7, 8]. These findings suggest that the intratumoral microbiome may influence tumor progression, immune evasion, and sensitivity to therapy.

In contrast to solid tumors of the gastrointestinal tract, the role of the intratumoral microbiome in hematological malignancies like DLBCL remains unexplored mainly [9]. While some studies have investigated the gut microbiome's impact on lymphoma development and treatment response, few have focused on the microbiome within the tumor microenvironment [10]. Emerging evidence suggests that microbial diversity within tumors may affect prognosis and therapeutic responses in DLBCL by influencing immune cell infiltration and function [11]. Understanding the interactions between the intratumoral microbiome and immune cells, such as regulatory T cells and macrophages, could reveal novel mechanisms underlying disease heterogeneity and treatment resistance [12].

This study aims to fill this knowledge gap by systematically investigating the relationship between the intratumoral microbiome and prognosis in DLBCL patients. Utilizing advanced bioinformatics approaches—including consensus clustering of microbiome profiles, single-cell RNA sequencing data analysis, and immune infiltration assessments—we seek to identify distinct microbiome-associated subtypes of DLBCL and evaluate their prognostic significance, exploring how variations in the intratumoral microbiome may modulate the TME by affecting immune cell composition and function, potentially influencing therapeutic responses. Furthermore, based on cell line expression analysis, knockout inference, and clinical cohort validation—we validated the expression of the identified hub genes and explored their impact on DLBCL classification. The overall study design and key findings are summarized in the graphical abstract (Fig. 1). The findings of this research have substantial clinical implications. Identifying microbiome-related subtypes could enable personalized treatment strategies for DLBCL patients, where microbial profiles inform therapeutic decisions. Patients with specific microbiome signatures might benefit from tailored chemotherapy regimens or novel interventions targeting the microbiome, such as probiotics or antibiotics, to modulate the TME and enhance treatment efficacy. Additionally, microbiome-associated biomarkers could contribute to early diagnosis and risk stratification, ultimately improving patient outcomes. In summary, this study provides novel insights into the complex interactions between the intratumoral microbiome and DLBCL, advancing our understanding of lymphoma biology and offering potential avenues for personalized therapy. By elucidating the role of the intratumoral microbiome in influencing immune responses and therapeutic sensitivity, we aim to contribute to developing more effective strategies for managing DLBCL.

Materials and methods

Data sources and microbiomes identification

Transcriptional and clinical data for 48 DLBCL cases were obtained through The Cancer Genome Atlas (TCGA) at https://portal.gdc.cancer.gov/ [13]. DLBCL microbiome data were sourced from the cBioPortal platform at https://www.cbioportal.org/ [14]. Microbiomes associated with DLBCL prognosis were identified using univariate Cox regression analysis and the Kaplan–Meier (KM) method [15].

Analysis of clustering based on microbiome abundance in DLBCL

We performed consensus clustering on DLBCL samples based on the relative abundance of microbiomes associated with prognosis using the 'ConsensusClusterPlus' R package [16]. The optimal number of clusters was determined by examining the cumulative distribution function (CDF) curve and its area under the curve. We used

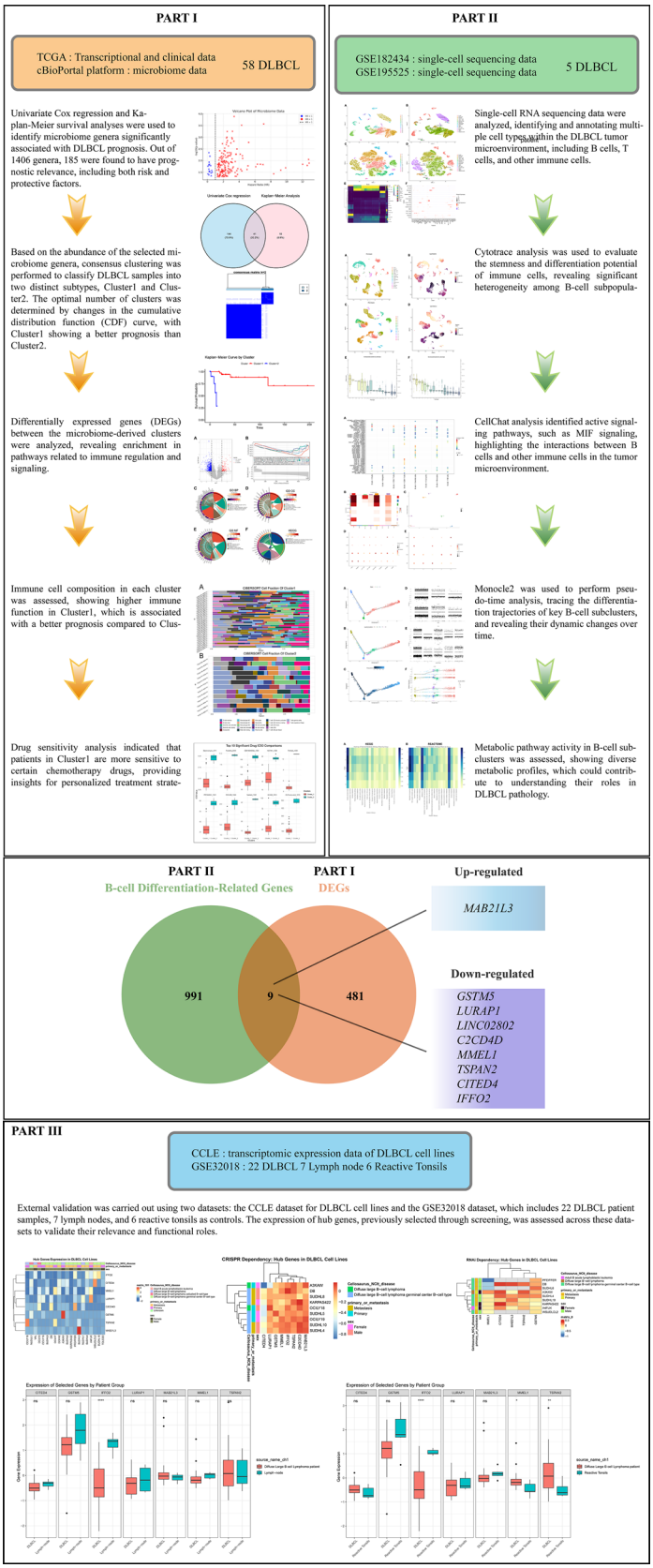


Fig. 1 Graphical Abstract. This study investigates the role of the tumor microbiome in diffuse large B-cell lymphoma (DLBCL) through three distinct parts. The first part analyzes microbiome-related data, using consensus clustering to classify DLBCL patients into two microbiome subtypes and assess their prognostic significance. The second part involves single-cell RNA sequencing data, revealing immune cell heterogeneity in the tumor microenvironment, particularly B cells' differentiation trajectories and immune functions. The third part utilizes external validation with the CCLE cell line data and the GSE32018 GEO dataset, including DLBCL patient samples, lymph nodes, and reactive tonsil controls. These datasets confirm the differential expression and functional dependency of key genes identified in the previous parts, providing additional support for the relevance of these findings in DLBCL pathology

principal component analysis (PCA) and t-SNE to validate the accuracy of the clustering results. Subsequently, the KM method was employed further to compare the prognostic characteristics of different clustering groups.

Analysis of differentially expressed genes between microbiome clusters

The clustering results categorized TCGA patients into different subtypes. To further analyze differences between microbial subtypes within tumors, we analyzed differential gene expression using the 'limma' package to identify genes associated with intratumoral microbiomes between different subtypes [17]. The selection criteria were |log2FC|> 1 and false discovery rate (FDR) < 0.05.

Analysis of functional enrichment on DEGs

We performed functional enrichment analysis on selected DEGs to identify gene ontology (GO) annotations using the org.Hs.eg.db R package [18]. Subsequently, gene sets were mapped to this background dataset and enriched using clusterProfiler [19] (version 4.11.0) to identify significant functional enrichments. Our analysis parameters included a minimum of 5 genes and a maximum of 5000 per set, with significance levels set at $P < 0.05$ and FDR < 0.25. Additionally, we downloaded gene set files related to KEGG and Reactome pathways from the Molecular Signatures Database (MSigDB) [20] and analyzed them for enrichment using the clusterProfiler R package.

Analysis of immune infiltration on microbiome clusters

Immune infiltration analysis on each microbiome cluster was conducted using the IOBR (Immune-Oncology Biological Research) R package [21] to assess the components of immune cells within the tumor microenvironment. The IOBR is a dedicated tool for tumor immunology research, providing a wide range of analytical and visualization features. The CIBERSORT algorithm from the IOBR package was used in this analysis.

Prediction of drug sensitivity

After identifying differences in the intratumoral microbiome and single-cell characteristics, we used the OncoPredict tool to assess variations in drug sensitivity between patients with different microbiome subtypes [22]. OncoPredict is an R package developed to predict drug responses and biomarkers based on cell line data and cancer patient profiles.

Single-cell sequencing analysis

Two datasets (5 DLBCL samples, 4 from the GSE182434 and 1 from the GSE195525 dataset) were integrated for single-cell sequencing analysis (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE195525) [23]. Low-quality cells with fewer than 300 genes or high mitochondrial gene expression were filtered out. We corrected batch effects between datasets using the Harmony package [24]. Single-cell sequencing data were then analyzed using the Seurat package [25]. Gene expression was normalized to identify variable genes, and the top 3000 genes with the highest variability were selected. The data were analyzed using principal component analysis (PCA) and UMAP for clustering, with optimal resolution determined using the clustertreeR package [26]. Cell types were annotated using the SingleR package [27].

Predicting the differentiation status of immune cells in the DLBCL microenvironment

To further explore the impact of the identified DLBCL intratumoral microbes on the immune microenvironment, the stemness and differentiation potential of different cell types were accessed using the CytoTRACE V2 tool [28]. CytoTRACE V2 uses AI algorithms to predict cell differentiation from single-cell RNA data, providing a score from 0 (fully differentiated) to 1 (totipotent). Cells are classified into categories based on their differentiation potential. The results were visualized using UMAP clustering and differentiation potential plots.

Intercellular communication and signaling strength analysis

To explore cell interactions in the DLBCL immune microenvironment, we used the CellChat package to analyze single-cell sequencing data. This tool helped identify key cell subpopulations and their communication pathways in the tumor immune environment [29]. We visualized intercellular communication patterns based on network analysis, using a two-dimensional map to highlight functional and structural similarities in signaling pathways.

Pseudo-time analysis of key immune cell clusters

To further investigate the developmental trajectories, differentiation status, and transcriptional regulatory features of critical immune cells, pseudo-time analyses of annotated different cell types were performed using the Monocle2

package to construct trajectories of cell-to-cell changes and thus remodel the process of cellular changes over time [30]. The list of crucial cell marker genes obtained from the analysis in Seurat was used as a feature for cell differentiation trajectory inference, and cells were sorted according to the proposed temporal order to determine the differentiation STATE and subpopulation of different vital cells. The significance of each gene was returned in the temporal order direction to determine the gene expression profile of the critical signaling pathways with different states and clusters. In addition, the differentiation trajectories and time-series characteristics of critical cell subpopulation development were further inferred by analyzing the correlation between gene expression trends of crucial signaling pathways and immune cell status.

Metabolic pathway analysis of key immune cell clusters

Analyzing quantitative metabolic activities at single-cell resolution makes it possible to predict different expression patterns between cells and cell clusters, thus revealing heterogeneity between cells/cell clusters and demonstrating complex metabolic processes. Therefore, we further calculated the cellular metabolic intensities of key cell clusters using the scMetabolism R package [31]. scMetabolism R package was pre-populated with a set of human metabolic genes containing 85 KEGG pathways and 82 Reactome entries. We obtained activity scores for key immune cell subclusters in each metabolic pathway using the vision algorithm. Finally, we selected the 20 metabolic pathways with the highest variability, and the R program was used for heatmap presentation and analysis.

Identification of the hub genes in microbe-driven B-cell differentiation

B cells undertake different stages of development and differentiation, and with their key role in DLBCL progression, it is critical to understand the genetic alterations that influence their differentiation and function. To explore how different microbiome-driven genetic alterations in DLBCL affect B-cell development, we focused on the interaction of genes involved in B-cell differentiation with microbiome-driven differentially expressed genes. By focusing on these genes, we can investigate how microbiome-driven changes may affect B-cell maturation and lead to immune evasion or treatment resistance in DLBCL. We used interaction analysis and Venn plots to visualize genes common to both datasets and defined them as hub genes for microbe-driven B-cell differentiation.

Validation of hub genes expression and biological function in cell lines and independent cohort of DLBCL

This study utilized R and related packages to analyze hub genes based on the RNAseq gene expression data and metadata (e.g., disease subtype, sex, metastasis status, etc.) of 20 DLBCL cell lines from the Cancer Cell Line Encyclopedia (CCLE) database [32]. Additionally, independent cohort data from the GSE32018 dataset [33], which includes DLBCL patient samples, lymph nodes, and reactive tonsil controls, were used for independent cohort validation. RNA inference knockout and CRISPR-Cas9 data were obtained via the DepMap package [34]. Data processing involved filtering, cleaning, and wide-format transformation using the dplyr and tidyr packages, with missing values filled and standardized. The expression and inference of hub genes' clustering features were visualized using a pheatmap and a complex heatmap to generate heatmaps. Boxplots were created using ggplot2 to visualize the distribution of CRISPR knockout and RNAi knockdown dependencies and assess group differences. The dependency and expression data were then integrated to identify key genes with functional relevance in DLBCL, with external validation conducted using the GEO dataset to confirm the reproducibility of the findings across different sources.

Results

Identification of intratumoral microbiome subtypes in DLBCL

From the Cancer Genome Atlas (TCGA) database, this research gathered transcriptomic, microbiomes, and clinical data for 47 DLBCL samples. We acquired 1406 genus from the DLBCL microenvironment through the cBioPortal platform (Supplementary Table S1). Univariate Cox regression analysis selected 185 genus that significantly impacted prognosis. Of these, 167 were risk factors (162 bacterial genus and 5 viral genus), and 18 were protective factors (12 bacterial genus and 6 viral genus) (Fig. 2A & 1B, Supplementary Table S2). Kaplan–Meier analysis further identified 59 genus associated with DLBCL patient prognosis (Supplementary Table S3). The intersection of microbe genus identified through univariate Cox regression and Kaplan–Meier analysis included 41 genus (Fig. 2C): Carlavirus, Muromegalovirus, Mamastrovirus, Thermoanaerobacterium, Labilithrix, Trueperella, Sorangium, Aggregatibacter, Mahella, Actinosynnema, Anaerotruncus, Anaerostipes, Endomicrobium, Elizabethkingia, Mycetocola, Fictibacillus, Erythrobacter, Thalassospira, Caedibacter, Micavibrio, Gemmata, Salinivibrio, Belnapia, Candidatus_Endolissoclinum, Luminiphilus,

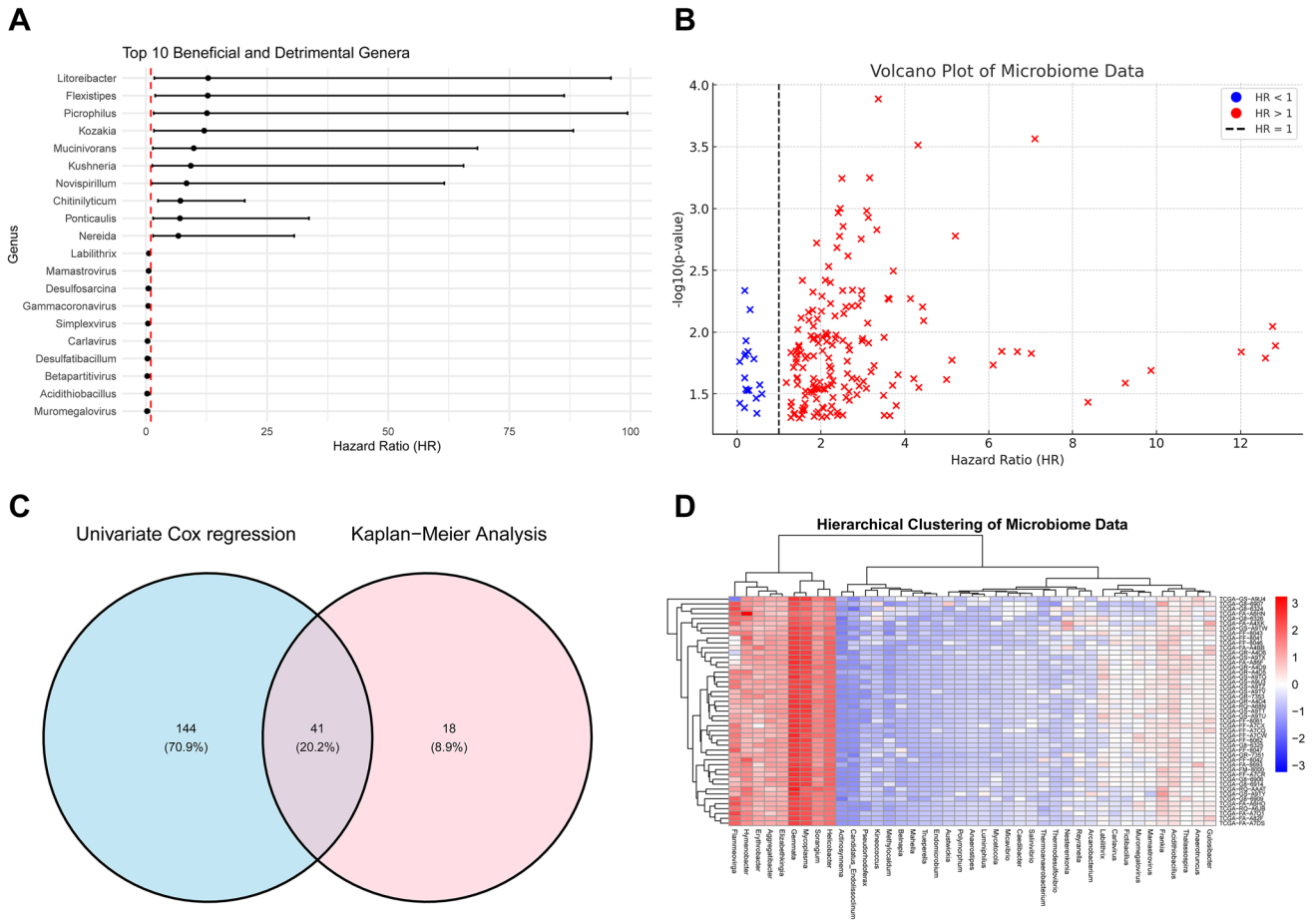


Fig. 2 Integrated analysis of microbial genus as prognostic factors in DLBCL. **A** represents a forest plot of the top 10 microbial genus with notable risk and protective impacts. The forest plot, derived from univariate Cox regression analysis, shows microbial genus related to patient outcomes. Each point denotes a genus's hazard ratio (HR), with error lines marking the 95% confidence intervals. Genus to the left of the dashed line (HR < 1) is beneficial, suggesting a lower risk of poor outcomes, whereas those to the right (HR > 1) are harmful, suggesting a higher risk of poor outcomes. **B** represents a volcano plot of microbiome data. The volcano plot displays HR values on the x-axis and -log10(p-values) on the y-axis, with each point representing a microbial genus. Blue points indicate protective factors (HR < 1), and red points indicate risk factors (HR > 1). **C** Venn diagram shows the significant findings from univariate Cox regression and Kaplan–Meier analysis. This Venn diagram illustrates the relationship between significant results from univariate Cox regression analysis (blue circle) and Kaplan–Meier survival analysis (pink circle). Numbers within the blue circle indicate microbial genus significant only in the Cox regression analysis; numbers within the pink circle denote genus significant only in the Kaplan–Meier analysis; numbers in the gray intersection area represent genus significant in both analyses. **D** represents a hierarchical clustering heatmap of expression data for 41 microbial genus. The heatmap is based on a hierarchical clustering analysis of expression data for 41 microbial genus, which is significant in both analysis methods. The x-axis represents the 41 microbial genus, the y-axis represents patient IDs, and the color intensity indicates the relative expression levels, with red indicating high expression and blue indicating low expression. All significant results are based on a *P*-value < 0.05

Polymorphum, Pseudorhodoferax, Reyranella, Frankia, Gulosibacter, Hymenobacter, Mycoplasma, Thermodesulfovibrio, Acidithiobacillus, Austwickia, Arcanobacterium, Kineococcus, Helicobacter, Methylocaldum, Flammeovirga, Nesterenkonia. Subsequently, consensus clustering was performed using the abundance of 41 tumor taxonomies (Fig. 2D). The consensus cumulative distribution function (CDF) graph, along with the area under the curve and its changes, is depicted in (Fig. S3A, C). As illustrated, the optimal k-value is 2, and the consensus matrix for K = 2 is displayed in (Fig. S3B, Supplementary Fig. S1). DLBCL was divided into two intratumoral microbiome subtypes, Cluster 1 and Cluster 2, which were distinctly differentiated by PCA and t-SNE (Supplementary Fig. S2). Kaplan–Meier survival analysis indicated that Cluster 1 has a significantly better prognosis compared to Cluster 2 (Fig. S3D).

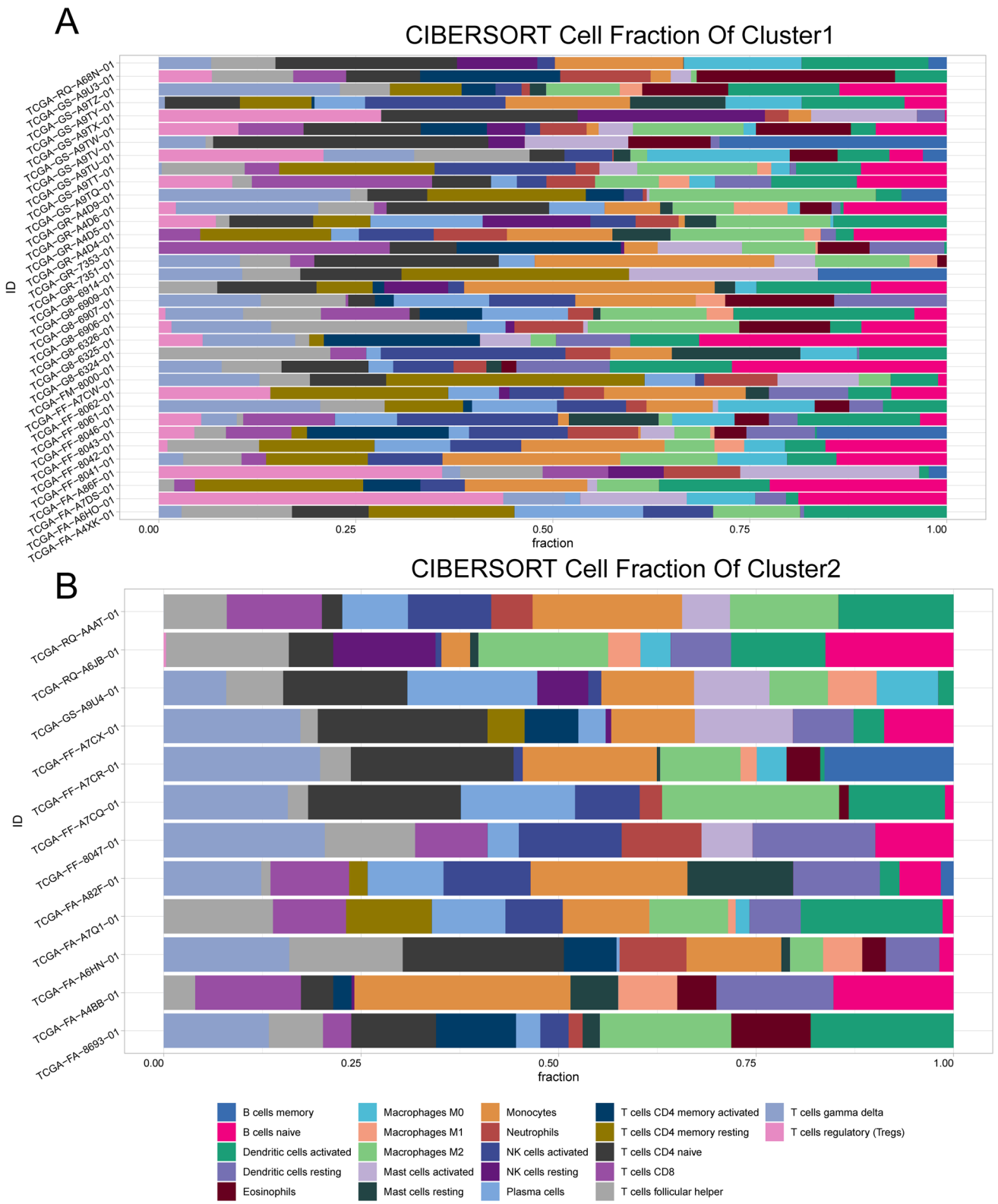


Fig. 3 CIBERSORT immune cell infiltration analysis. **A** Proportion of immune cells in Cluster 1. Each horizontal band represents one patient, and the color's width indicates the proportion of different immune cell types in that patient. **B** Proportion of immune cells in Cluster 2. The colors are the same as in the A plot and represent the same cell types

Functional enrichment analysis of differentially expressed genes

Consensus clustering divided patients into two subtypes: Cluster 1 and Cluster 2. Further differential gene expression (DEG) analysis revealed 658 significantly differentially expressed genes between these subtypes, with 106 genes up-regulated and 552 genes down-regulated (Fig. S4A). To explore the underlying molecular mechanisms between microbiome-derived subtypes within the tumor microenvironment, we conducted extensive functional enrichment analyses of differentially expressed genes between Cluster 1 and Cluster 2, and the results revealed that the differentially expressed genes were significantly enriched in several immunomodulatory pathways, such as cytokine–cytokine receptor interactions and chemokine signaling pathways (Fig. S4B). Furthermore, GO analysis showed that the biological functions of the differentially expressed genes were closely related to inflammatory response and immune regulation (Fig. S4C), were mainly enriched in membrane transporter receptors and various ion channel complexes (Fig. S4D), and the molecular functions were mainly involved in receptor and ligand regulatory activity, chemokine and receptor activity (Fig. S4E). In addition, KEGG pathway analysis showed that key biological pathways such as cytokine–cytokine receptor action, chemokine signaling pathway, and hedgehog signaling pathway were significantly enriched (Fig. S4F). Notably, cytokine–cytokine receptor interactions play a key role in the immune response by regulating the proliferation, differentiation, and activation of immune cells, which can enhance the anti-tumor immune response, and the enrichment of these genes suggests that patients with different subtypes may have a more robust immune response and be able to efficiently identify and attack tumor cells, leading to improved survival. The chemokine signaling pathway also plays an essential role in regulating the recruitment of immune cells to the tumor microenvironment. Enhanced chemokine signaling in patients with different subtypes may contribute to infiltrating more immune effector cells into the tumor site, thereby enhancing anti-tumor immunity. This enhanced immune cell infiltration and activation may be one of the critical reasons for the better prognosis of Cluster 1 patients. In contrast, the enrichment of genes in Cluster 2 in immunosuppression-related pathways may lead to an enhanced immunosuppressive state in the tumor microenvironment, thereby suppressing anti-tumor immune responses. Activation of these immunosuppressive pathways may make tumors in Cluster 2 patients more immune evasive, leading to poorer survival.

Immune infiltration analysis on microbiome-derived clusters

We assessed changes in immune infiltration between subtypes using CIBERSORT's algorithm to explore the relationship between microbiome-derived subtypes within the tumor and the tumor microenvironment. Compared to Cluster 2, higher levels of infiltration of regulatory T cells (Tregs) and M0-type macrophages have been observed in patients with Cluster 1, and the role of these immune cells in the tumor microenvironment substantially impacts patient prognosis (Fig. 3). Tregs are cells capable of suppressing the immune response and are often thought to play a role in tumor immune escape. However, Tregs can also maintain immune homeostasis by modulating excessive inflammatory responses, which has been associated with better tumor control and patient prognosis in some cases. Higher infiltration of Tregs in Cluster 1 may help inhibit excessive inflammation, thereby enhancing the effectiveness of the anti-tumor immune response. In addition, M0 macrophages are considered to be the more primitive, unpolarized state of the tumor microenvironment, with plasticity and the ability to polarize in either an anti-tumor (M1) or pro-tumor (M2) direction when stimulated by appropriate microenvironmental signals. The higher levels of M0 macrophages in Cluster 1 may indicate that in this microenvironment, the tumor can still maintain some degree of immune homeostasis without being entirely biased toward a pro-tumor immunosuppressive state of the tumor. Therefore, M0 macrophages may provide a potentially plastic basis for subsequent anti-tumor signaling to promote a more favorable immune response.

Prediction of drug sensitivity in different clusters

Drug adjuvant therapy is a crucial means to improve the prognosis of DLBCL patients. Selecting susceptible drugs for different patients is crucial to optimize therapeutic outcomes. Predictive results for drug sensitivity show that Cluster 1 patients exhibit higher sensitivity to chemotherapy drugs or compounds like Dactinomycin (1811), Foretinib (2040), GSK1904529A (1093), NU7441 (1038), Pictilisib (1058), PRT062607 (1631), RVX.208 (1625), Taselisib (1561), VE.822 (1613), and 5-Fluorouracil (1073) (Fig. 4). This finding suggests that utilizing intratumoral microbiome characteristics to subtype patients can provide a compelling new strategy for personalized medicine in DLBCL.

Single-cell sequencing analysis with cell clustering and annotation

Following drug sensitivity prediction, we found that different microbial subtypes exhibited significantly different

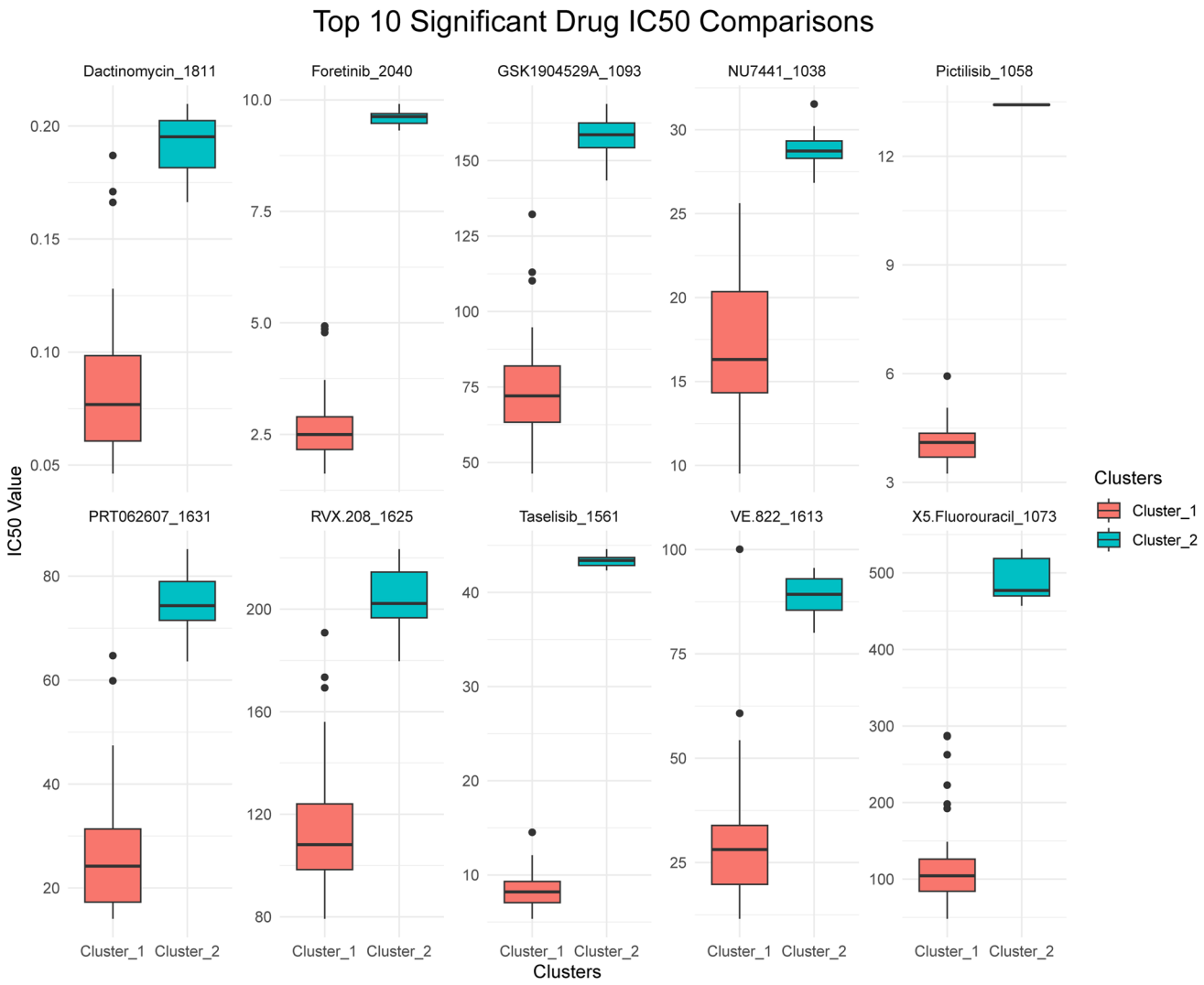
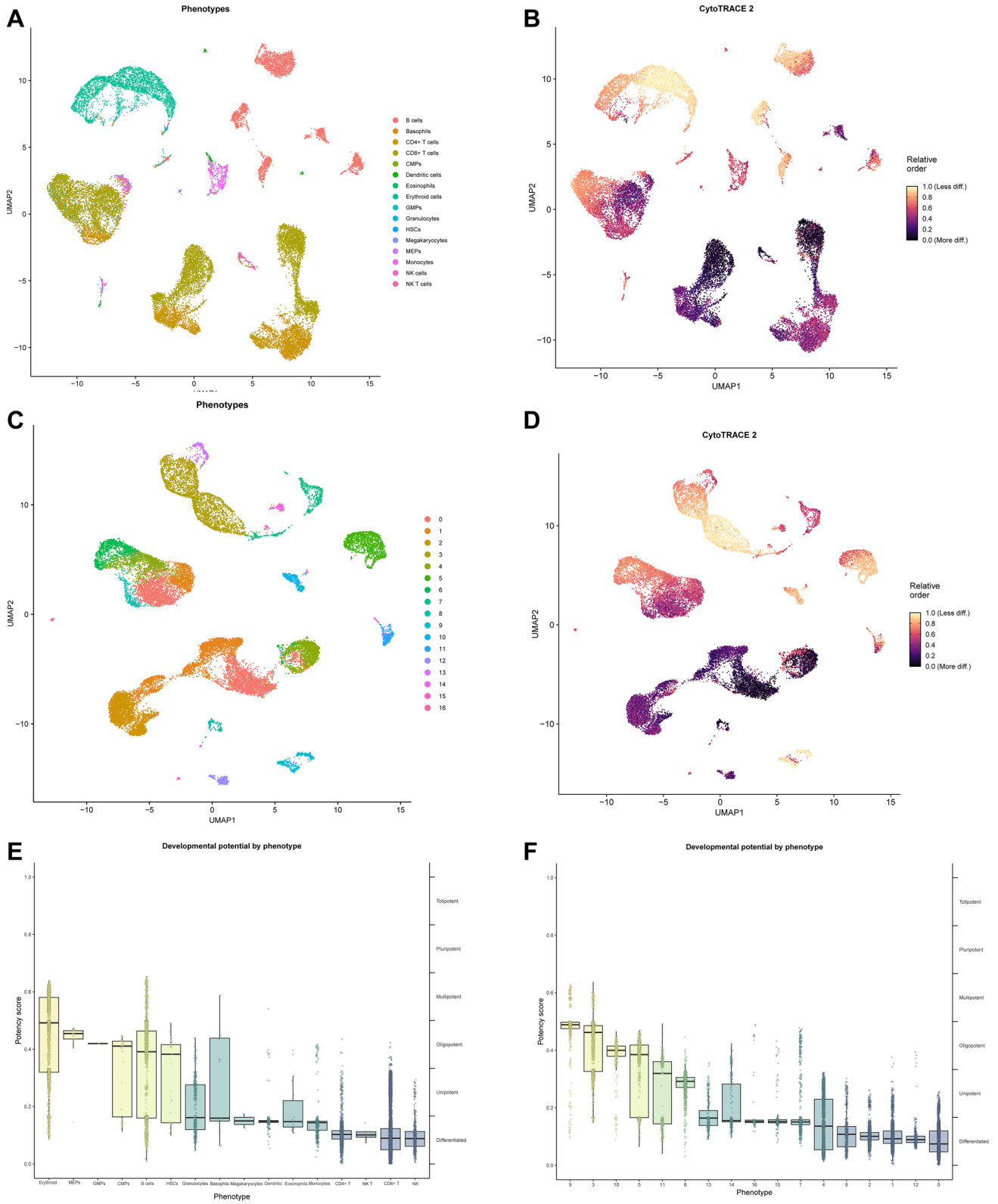


Fig. 4 The comparison results of drug sensitivity between the two subgroups. The present boxplots of drug sensitivity for Cluster 1 and 2 illustrate the IC50 values for ten drugs among different patient groups. Red boxplots indicate Cluster 1, while blue boxplots represent Cluster 2. The median, quartiles, and extremes of the IC50 values are visually depicted in the boxplots, with dot plots showing the outliers

sensitivities to multiple chemotherapeutic agents. We used single-cell RNA sequencing analyses to elucidate the mechanisms behind these differences and further explore immune cell heterogeneity in the tumor microenvironment. By characterizing the immune cells in more detail, we hope to understand the underlying biological mechanisms in terms of drug response in different subtypes of patients, in particular, how the microbiome affects the functional state of the immune cells and whether changes in these immune cells are associated with differences in drug sensitivity. To validate the specific impact of intratumoral subtypes on cell behavior in the tumor immune microenvironment, we performed single-cell RNA sequencing (scRNA-seq) analysis on 23,332 DLBCL patient-derived cells after quality control and filtering (QC map). Using UMAP dimensionality reduction, we classified these cells into 17 distinct cell clusters (Fig. S5A). These clusters were annotated to 16 major immune cell types, including B cells, monocytes, NK cells, CD4 + T cells, CD8 + T cells, dendritic cells, basophils, granulocytes, eosinophils, hematopoietic stem cells (HSCs), common myeloid progenitor cells (CMPs), NK T cells, erythrocytes, megakaryocytic progenitors (MEPs), granulocyte-monocyte progenitors (GMPs), and megakaryocytes (Fig. S5B). To validate the clustering results further, we applied t-SNE dimensionality reduction, which showed consistent cell population distributions compared to UMAP. As shown in Fig. S5C and D, distinct cell subclusters were identified, and the annotations confirmed the stable classification of immune cell types. Cell types were further annotated



based on reference data using the SingleR package, and a heatmap of similarity scores was generated (Fig. S5E). This analysis confirmed the identification of key immune

◂**Fig. 5** Single-cell transcriptome data for cell developmental potential assessment. **A** UMAP map of cell phenotypes. Cells were classified into phenotypes by known markers based on single-cell RNA sequencing data. Each point represents a cell, and different colors indicate different cell types. **B** UMAP plot of CytoTRACE scores indicated the relative developmental order of cells, with colors ranging from light to dark indicating changes in developmental potential, with darker colors indicating higher developmental potential. **C** UMAP plot of cell clusters indicated the cell population based on cluster analysis, with different colors and numbers representing different cell clusters. **D** UMAP plot of CytoTRACE score represented the relative developmental order of cells, with color coding indicating changes in developmental potential. **E** The box line plot of developmental potential by phenotype represented the distribution of CytoTRACE scores for different cell phenotypes, reflecting the developmental potential of each cell type. **F** Boxplots of developmental potential by cell cluster showed CytoTRACE scores for different cell clusters

cell populations in DLBCL. The heatmap highlights the immune landscape and demonstrates the functional diversity among the identified clusters. Additionally, we analyzed the top three differentially expressed genes for each cell type, visualized in Fig. S5F. This analysis revealed cell type-specific gene expression patterns, providing insights into the functional roles of these immune cell populations within the tumor microenvironment. For example, certain genes were highly expressed in specific B-cell subsets, suggesting that these clusters may play distinct roles in tumor progression and immune modulation.

Assessing the stemness and differentiation potential of cells exploring intercellular heterogeneity

We assessed the stemness and differentiation potential of different cell subpopulations and types using the CytoTRACE2 tool to explore the heterogeneity among these cell types further. The results showed that B cells had the highest stemness score and higher differentiation potential among all immune cells (Fig. 5A, B, and E). However, there was also significant heterogeneity in different B-cell subclusters (5, 9, 10, 11, 12) regarding differentiation potential (Fig. 5C, D, and F).

Cell–cell communication and signaling pathway analysis

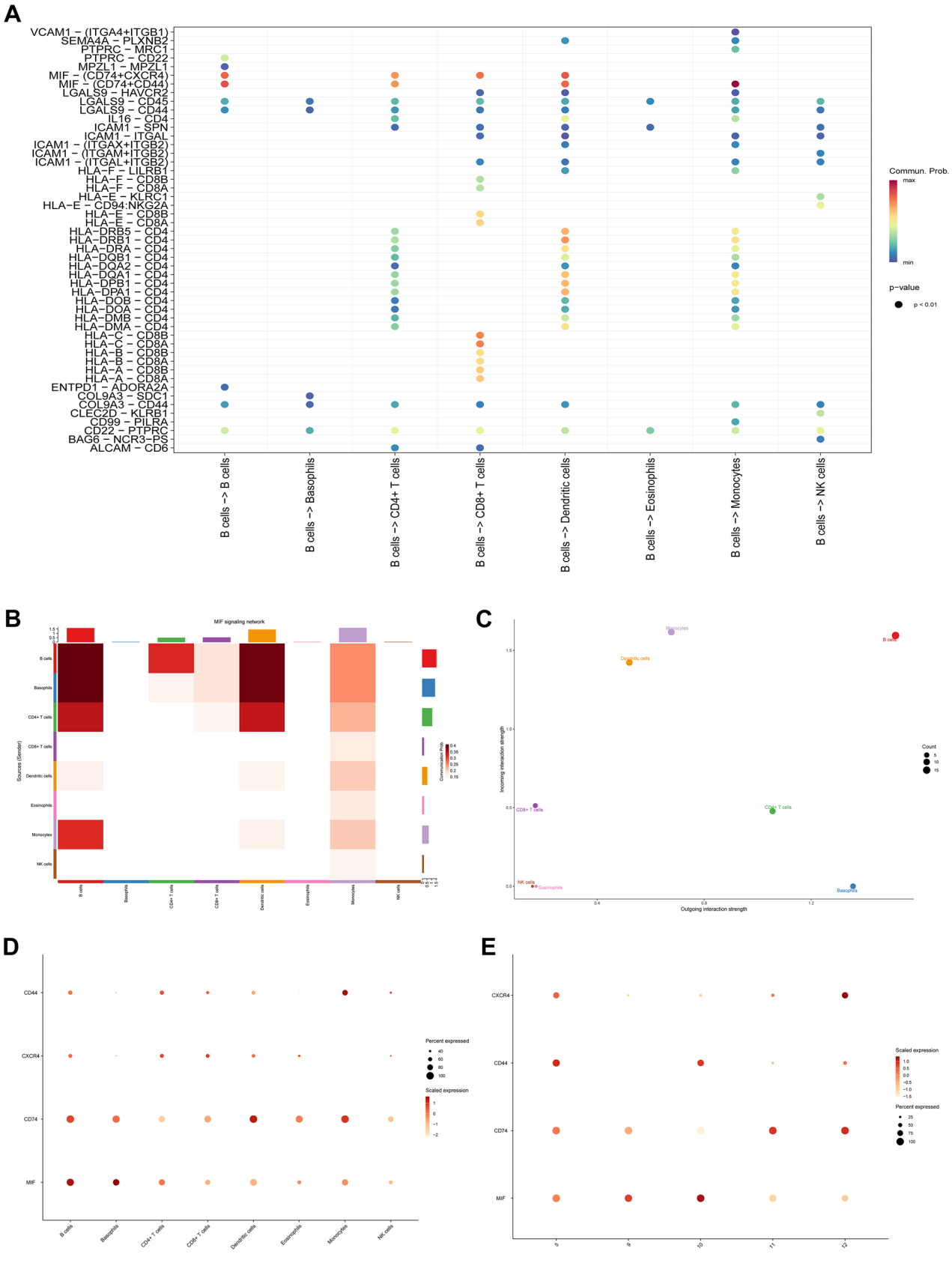
To further explore the potential causes of the heterogeneity generated among cell types, CellChat analysis revealed significant communication exercises among all immune cells in DLBCL, and in particular, B cells played a key role in immune cell communication interactions (Fig. 6). Furthermore, it is noteworthy that MIF signaling had the most significant regulatory role in different immune cell types in DLBCL. Similarly, B cells, as an essential source and receiver of MIF signals, had the most robust interactions with a wide range of other immune cells (e.g., CD4 + T cells, CD8 + T cells, dendritic cells, and monocytes) (Supplementary Table S4, Fig. S7A, B and C). Further molecular analysis of the interactions revealed that key MIF-interacting molecules included MIF-(CD74 + CD44) and MIF-(CD74 + CXCR4) ligand–receptor pairs. In addition, the MIF signaling pathway involves multiple ligand–receptor pairs that transmit information between different types of immune cells (Fig. S7D), and not only this, the ligand–receptor expression of the MIF signaling pathway also varies significantly between different B-cell subpopulations (Fig. S7E).

Pseudo-time analysis of key immune cell subclusters

In the single-cell dimension, we identified the potential heterogeneity of B cells across different immune microenvironments. To gain a deeper understanding of the differentiation trajectories of these B-cell subpopulations in the temporal dimension, we performed a proposed time-series analysis using Monocle2. The result revealed that the B-cell subpopulations formed three main states during differentiation, and there were significant differences in the distribution of each subpopulation in different states. Subpopulation 5 was mainly distributed in state 1, subpopulation 12 was distributed primarily in state 2, while subpopulation 9 was mainly distributed in state 3. Subclusters 10 and 11 were more uniformly distributed in each state but had relatively low expression levels. Notably, subgroup 9 was located at the root of differentiation, while subgroup 5 was at the end position of differentiation (Fig. 7A, B and C). In addition, the expression levels of MIF, CD74, CD44, and CXCR4 showed a high degree of heterogeneity between different stages and subgroups (Fig. 7D, E). Figure 7F demonstrates the relative expression of these proteins at different stages of the proposed chronology, further supporting the heterogeneity of B-cell subpopulations in the MIF signaling pathway.

Metabolic pathway analysis of key immune cells

From the perspective of cellular metabolism, we calculated the average expression levels of each B-cell subcluster (5, 9, 10, 11, and 12) on each metabolic pathway. Then, the top 20 metabolic pathways with the most significant variance were selected for heatmap presentation to highlight the heterogeneity among B-cell subpopulations. The results showed that the expression levels of different B-cell subpopulations on metabolic pathways were significantly different. Specifically, subcluster 5 showed high activity on several metabolic


A
VCAM1 – (ITGA4+ITGB1)
SEMA4A – PLXNB2
PTPRC – MRC1
PTPRC – CD22
MPZL1 – MPZL1
MIF – (CD74+CXCR4)
MIF – (CD74+CD44)
LGALS9 – HAVCR2
LGALS9 – CD45
LGALS9 – CD44
IL16 – CD4
ICAM1 – SPN
ICAM1 – ITGAL
ICAM1 – (ITGAX+ITGB2)
ICAM1 – (ITGAM+ITGB2)
ICAM1 – (ITGAL+ITGB2)
HLA–F – LILRB1
HLA–F – CD8B
HLA–F – CD8A
HLA–E – KLRC1
HLA–E – CD94:NKG2A
HLA–E – CD8B
HLA–E – CD8A
HLA–DRB5 – CD4
HLA–DRB1 – CD4
HLA–DRA – CD4
HLA–DQB1 – CD4
HLA–DQA2 – CD4
HLA–DQA1 – CD4
HLA–DPB1 – CD4
HLA–DPA1 – CD4
HLA–DOB – CD4
HLA–DOA – CD4
HLA–DMB – CD4
HLA–DMA – CD4
HLA–C – CD8B
HLA–C – CD8A
HLA–B – CD8B
HLA–B – CD8A
HLA–A – CD8B
HLA–A – CD8A
ENTPD1 – ADORA2A
COL9A3 – SDC1
COL9A3 – CD44
CLEC2D – KLRB1
CD99 – PILRA
CD22 – PTPRC
BAG6 – NCR3–PS
ALCAM – CD6
B cells –> B cells
B cells –> Basophils
B cells –> CD4+ T cells
B cells –> CD8+ T cells
B cells –> Dendritic cells
B cells –> Eosinophils
B cells –> Monocytes
B cells –> NK cells
Commun. Prob.
max
min
p–value
p < 0.01
B
MIF signaling network
Sources (Sender)
Communication Prob.
C
Incoming interaction strength
Outgoing interaction strength
Count
D
CD44
CXCR4
CD74
MIF
Percent expressed
Scaled expression
E
CXCR4
CD44
CD74
MIF
Scaled expression
Percent expressed

◂**Fig. 6** Analysis of communication networks between immune cells. **A** The intercellular communications showed the number of communications between different types of immune cells. Each node represents an immune cell type, the thickness of the connecting lines between the nodes indicates the number of communications, and the colors represent different communication pathways. **B** Strength of communication between immune cells. The color coding of the nodes and connecting lines indicate the strength of the communication, with thicker connecting lines indicating higher communication strength. The color coding of the nodes and connecting lines indicate the strength of the communication, with thicker connecting lines indicating higher communication strength. **C** Specific pathways of intercellular communication. Each subgraph shows the communication pathways of one cell type with other immune cells. The direction and thickness of the arrows indicate the direction and strength of communication, and the colors indicate different communication pathways

pathways, including polyamine metabolism, vitamin C and ascorbic acid metabolism, and RNA metabolism, suggesting that it may have an essential role in cell proliferation, antioxidant response, and RNA processing. Subcluster 9 showed high activity on TP53-regulated metabolic genes and RNA metabolism, suggesting that it may be essential in regulating stress response and gene expression. Subcluster 10 mainly showed high activity in RNA metabolism, which might be related to its vital role in protein synthesis and gene expression regulation. Subcluster 11 was more evenly distributed across metabolic pathways, but the overall activity level was relatively low, suggesting its metabolic activity was more moderate. Subcluster 12 showed high activity in oxidative phosphorylation and metabolism of fructose and mannose, indicating that it had a significant role in energy metabolism and sugar metabolism. These metabolic differences reflect the diversity of metabolic requirements and functions of different B-cell subpopulations, providing new insights into the understanding of the metabolic characteristics of B cells in DLBCL and their pathological mechanisms and offering potential targets for the development of individualized therapeutic strategies (Fig. S8A, B, Supplementary Tables S5, S6).

Identification of the hub genes in intratumoral microbe-driven B-cell differentiation

Given the critical role of B cells in the progression of DLBCL, it is crucial to understand the genetic alterations that affect their differentiation and function. Under the single-cell dimension, our findings have confirmed significant differences in the B-cell differentiation and development processes in DLBCL. Therefore, understanding whether intratumoral microbe-driven DLBCL differential gene genes are involved in B-cell differentiation and developmental processes remains worthy of further exploration. We then extracted 1000 genes that play key roles in B-cell differentiation (Supplementary Table S7) and performed an interaction analysis with 490 differentially expressed genes driven by intratumoral microbes. Interestingly, this analysis revealed 9 common genes, such as GSTM5, LURAP1, LINC02802, MAB21L3, C2CD4D, MMEL1, TSPAN2, and CITED4 (Fig. S9), and these genes were demonstrated play pivotal roles in the differentiation process of B cells while also being influenced by the intratumoral microbiome, making them potential biomarkers for both DLBCL progression and therapeutic response.

Validation of hub genes in

The expression validation of hub genes in 20 DLBCL cell lines revealed distinct expression and dependency profiles. The heatmap (Fig. S10) shows the expression of selected hub genes across various DLBCL cell lines. GSTM5, C2CD4D, MAB21L3, and TSPAN2 exhibited relatively high expression across most cell lines, with GSTM5 showing higher expression. In contrast, MMEL1, IFFO2, and LURAP1 displayed lower expression levels across most cell lines. These results suggest that GSTM5, C2CD4D, MAB21L3, and TSPAN2 may play more significant roles in DLBCL, while MMEL1, IFFO2, and LURAP1 may have a lesser impact on DLBCL progression.

Furthermore, we performed functional inference analysis of hub genes in DLBCL cell lines using RNAi knockdown and CRISPR knockout data from the DepMap database. The results showed significant differences in the dependence score of different genes in cell survival, revealing their potential roles in DLBCL cell growth and disease progression. RNAi knockdown-low inference analyses revealed a gene-dependent distribution of MMEL1 in different DLBCL cell lines (Fig. 8A) and showed a strong survival dependence (Fig. 8B). Further analysis showed a significant dependence of MMEL1 in the invasion and metastasis of the germinal center B-cell subtype, suggesting that it may have an important function in this subtype and play a key role in DLBCL cell survival (Fig. 8C). Besides, the results of CRISPR knockdown inference further validated the findings of the RNAi knockdown assay and revealed additional information on gene function. The results showed that CITED4 and MMEL1 were gene-dependent in several DLBCL cell lines (Fig. 8D), and comparative analyses indicated that DLBCL was more sensitive to CITED4 knockdown (Fig. 8E), further subtype analysis indicated that CITED4 was higher dependence in DLBCL primary samples (Fig. 8F), suggesting that it may play an important role in the primary of DLBCL.

Moreover, we validate the differential expression of hub genes in the clinical samples to confirm the gene-dependent inference obtained from RNAi and CRISPR knockdown. Specifically, compared to normal lymph node tissues, the expression level of IFFO2 was significantly higher in DLBCL patient samples ($P < 0.0001$). However,

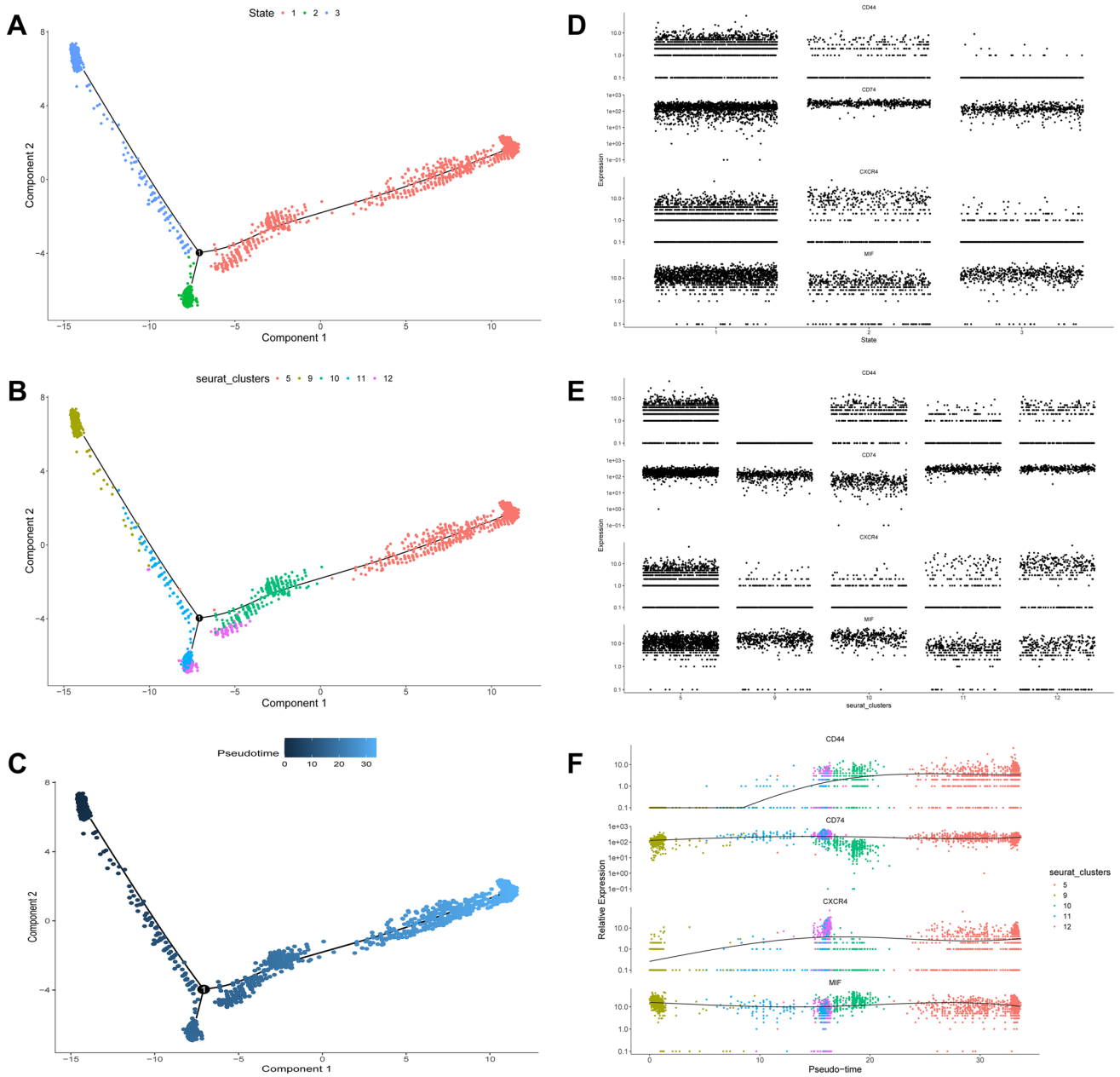


Fig. 7 Pseudo-time analysis of B-cell subclusters in the single-cell transcriptome. **A** The plot of the proposed chronological trajectories of cell states showed the trajectories of different cell states (State 1, 2, 3) in the proposed chronological analysis. Each point represents a cell, and the colors indicate the different states of the cell. **B** Pseudo-time trajectory plot of cell clusters. Trajectories of different B-cell clusters (Cluster 5, 9, 10, 11, 12) in the pseudo-time analysis are shown based on cluster analysis results. Different colors indicate different cell clusters. **C** The plot of cell distribution in the pseudo-time shows the distribution of cells in the proposed chronology, with color shades indicating the change in the cells' pseudo-time from early to late. **D** A scatter plot of gene expression versus cell state showed the relative expression of MIF pathway-related genes (CD44, CXCR4, CD74, MIF) in different cell states. **E** A scatter plot of gene expression versus cell cluster showed the relative expression of MIF pathway-related genes (e.g., CD44, CXCR4, CD74, MIF) in different cell clusters. **F** A scatter plot of gene expression versus pseudo-time showed the expression trend of MIF-related genes on pseudo-time, with colors indicating different cell clusters

the expression of MMEL1 was not significantly different between the two groups, suggesting that its expression was

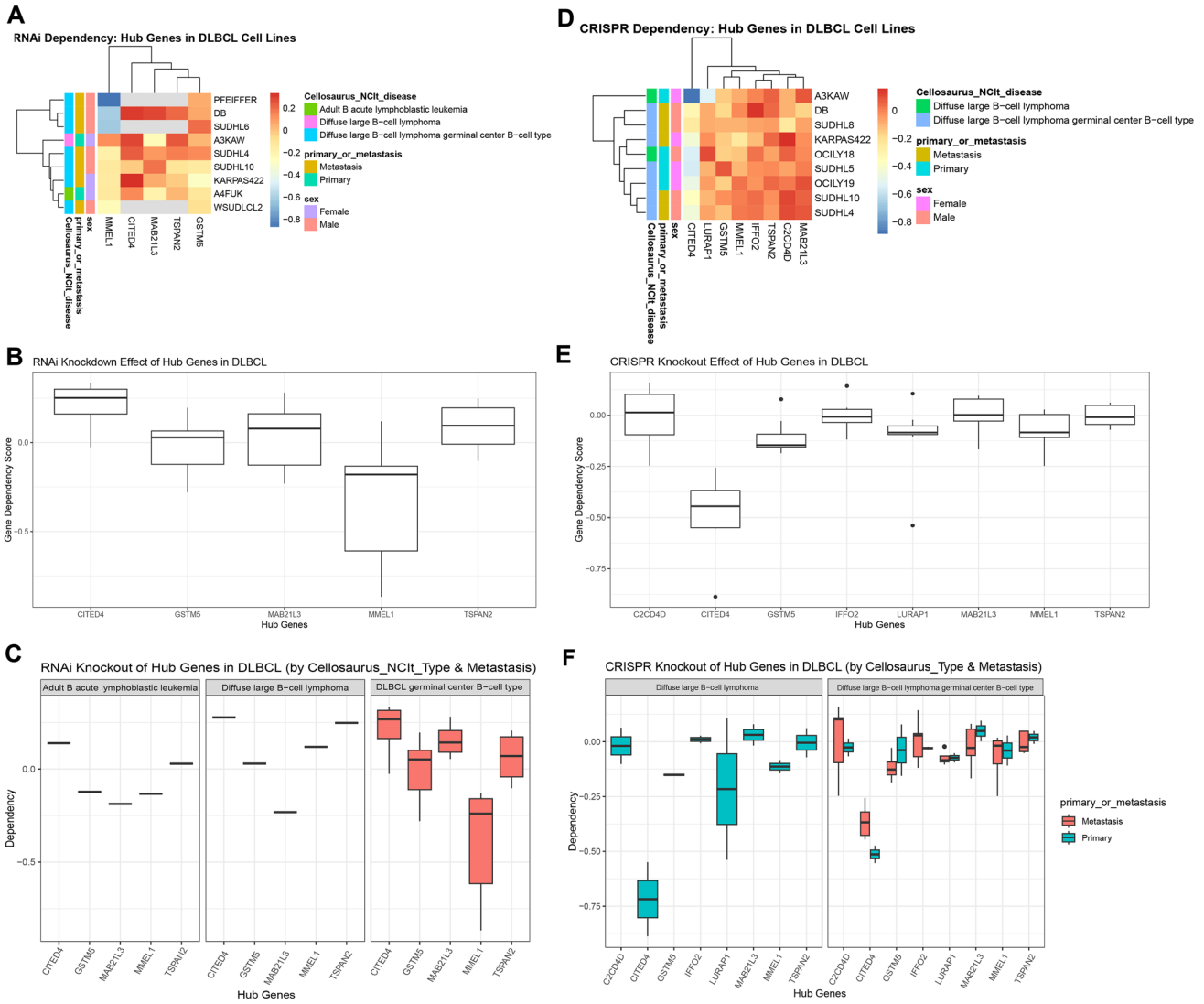


Fig. 8 Functional dependency analysis of hub genes in DLBCL cell lines. **A** Heatmap illustrating RNAi dependency scores of hub genes across DLBCL cell lines, with lower scores (blue) indicating higher dependency. **B** Boxplot summarizing RNAi dependency scores, showing the impact of gene knockdown on cell survival. **C** Subtype-specific RNAi dependency analysis stratified by disease subtype (diffuse large B-cell lymphoma vs. germinal center B-cell type) and metastasis status (primary vs. metastatic), highlighting MMEL1's strong dependency in the germinal center B-cell subtype. **D** The heatmap depicts CRISPR dependency scores for the same hub genes, with blue shades indicating higher dependency. **E** Boxplot summarizes CRISPR dependency scores, validates RNAi findings, and reveals additional functional insights. **F** Subtype and metastasis-stratified CRISPR dependency analysis showing higher CITED4 dependency in primary DLBCL samples, suggesting its potential role in disease initiation

relatively stable in DLBCL versus normal lymphoid tissues and may have a limited role in the early stages of tumorigenesis (Fig. S11A). Similarly, we found a significant upregulation of IFFO2 expression when comparing DLBCL patient samples with reactive tonsil tissue ($P < 0.0001$), further demonstrating this gene's high enrichment in tumor tissues. Notably, the expression levels of MMEL1 and TSPAN2 were also significantly up-regulated in DLBCL samples ($P < 0.05$; $P < 0.01$), suggesting that these two genes may have a more important role in DLBCL tumor progression or specific molecular subtypes (Fig. S11B).

Discussion

DLBCL is a clinically diverse malignant lymphoma that, although sensitive to initial chemotherapy, still maintains a high mortality rate due to rapid disease progression and a high relapse rate. Despite continuous improvements in lymphoma treatment strategies, the prognosis for DLBCL patients remains poor [35]. Consequently, developing biomarkers for early diagnosis and treatment response assessment in DLBCL holds significant clinical importance. Recent studies have increasingly revealed the role of the

tumor microbiome in the development, progression, and treatment responses of lymphoma, affecting tumor behavior and potentially influencing therapeutic efficacy through modulation of the host's immune response [36]. Our research analyzes the tumor microbiome in DLBCL patients to explore its association with patient prognosis and to identify potential microbial biomarkers, hoping to furnish new approaches for the early diagnosis and personalized treatment of DLBCL.

The microbiota contributes to cancer development by influencing the host's immune system and metabolic pathway. Studies indicate that gut microbiota modulates immune responses, affecting tumor resistance and susceptibility [37]. For instance, metabolic products from certain bacteria can regulate the host's anti-tumor immunity, thereby influencing tumor growth and treatment responses [38, 39]. The present study obtained transcriptomic and clinical data for 48 cases of DLBCL from the TCGA database, conducting an in-depth analysis of intratumoral microbiome subtypes and their relationship with the tumor microenvironment. Initially, we used univariate Cox regression analysis and the Kaplan–Meier method to identify 41 microbial genera from 1406 that significantly impact DLBCL prognosis, revealing the critical role of the tumor microbiome in the development of DLBCL. Through consensus clustering, patients were classified into two intratumoral microbiome subtypes (Cluster 1 and Cluster 2), with patients in Cluster 1 showing significantly better prognosis than those in Cluster 2. This finding indicates that intratumoral microbiome diversity is closely related to the prognosis of DLBCL and confirms that different microbial compositions can serve as potential biomarkers. Furthermore, the immune infiltration analysis indicated that patients in Cluster 1 have more substantial immune functions. In contrast, those in Cluster 2 have a significant reduction in key immune cells such as regulatory T cells, eosinophils, and M0 macrophages, which may be a critical factor leading to the poorer prognosis in Cluster 2. Also, the GSEA, GO, and KEGG analysis results revealed that the DEGs, which were compared in the two intratumoral clusters, are significantly enriched in membrane transport receptors and ion channels, reflecting their vital role in intercellular material exchange and signal transduction. This discovery suggests that intratumoral microbiota modulates the immune environment and may indirectly contribute to tumor growth and progression by affecting chemokine activity and receptor-ligand interactions [40]. Moreover, limited studies have proved that the varied composition of the intratumoral microbiome may affect patient responses to specific treatments, such as R-CHOP chemotherapy [41]. Our study demonstrated that different intratumoral microbiome subtypes exhibited significant differences in drug sensitivity, with patients with the Cluster 1 subtype showing higher sensitivity to multiple chemotherapeutic agents. This finding not only strengthens the position of the intratumoral microbiota as a potential biomarker for DLBCL but also provides new ideas for personalized medicine. By subtyping the intratumoral microbiome of patients, clinicians can select chemotherapeutic agents with greater precision, thereby improving treatment efficacy. This microbiome-based typing strategy may lead to groundbreaking clinical applications, especially when dealing with patients with limited effectiveness of conventional therapies.

Immune cells play a crucial role in the pathogenesis and treatment of DLBCL, a heterogeneous malignancy with different molecular subtypes. The tumor microenvironment of DLBCL consists of a wide range of immune cells, such as B and T cells, natural killer (NK) cells, dendritic cells (DC), tumor-associated macrophages (TAM), and myeloid-derived suppressor cells (MDSC), which interact with tumor cells and influence disease progression [42]. Notably, the MIF-(CD74 + CD44) and MIF-(CD74 + CXCR4) signaling pathways play critical roles in the immune modulation of the tumor microenvironment. These pathways are involved in immune cell migration, immune evasion, and tumor progression, which are key features in DLBCL. The interaction of MIF with CD74 and CD44 has been linked to immune suppression, while MIF-(CD74 + CXCR4) signaling facilitates cell migration and metastasis. We propose that these pathways may be influenced by the diversity of the intratumoral microbiome, potentially contributing to immune modulation and therapeutic resistance. Future studies should investigate how microbiome-induced alterations in these pathways could affect tumor progression and response to treatment, providing new targets for therapeutic intervention. Recently, studies have confirmed that the immune cells in the DLBCL TME influence treatment efficacy through complex interactions, which can significantly influence patient survival [43]. Besides, interactions between immune cells can also promote or inhibit anti-tumor responses, such as CD8 cytotoxic T cells targeting cancer cells and regulatory T cells suppressing immunoreactivity [44]. For instance, an inflammatory T-cell microenvironment with high T-cell content and low macrophage content is associated with better overall survival (OS) and progression-free survival (PFS) [45]. Regulatory T cells (Tregs) in the tumor microenvironment are associated with better prognosis, as these cells help tumors evade immune surveillance by suppressing immune responses [46]. Increasing FOXP3 + lymphocytes

are associated with favorable prognosis in DLBCL patients [47]. Thus, to better understand the specific cellular features in the DLBCL immune microenvironment, we have further analyzed tumor samples from DLBCL patients at a cellular level with high resolution using single-cell sequencing technology. The results of single-cell sequencing results revealed the complex heterogeneity of different cell types in the DLBCL tumor microenvironment, and in particular, we identified the distinct heterogeneity of B-cell subpopulations in DLBCL after scoring all immune cells for stemness and differentiation potential. Similarly, we identified important signaling pathways between B cells and other immune cells by intercellular communication analysis to determine their important role in regulating the immune response and cell differentiation process in DLBCL. Notably, as an important source and re-receiver of MIF signals, B cells had significant interactions with various other immune cells (e.g., CD4 + T cells, CD8 + T cells, dendritic cells, and monocytes). Consequently, molecular analyses of the interactions showed that the expression of key MIF-interacting molecules varied considerably between different B-cell subpopulations. In addition, to gain insights into the differentiation trajectories of these B-cell subpopulations from the temporal dimension, we further revealed the dynamics of B-cell subpopulations on differentiation trajectories based on pseudo-time analyses. The results suggested that B-cell subpopulations formed three main states during differentiation, and there were significant differences in the distribution of each subpopulation in different states. Among them, MIF, CD74, CD44, and CXCR4 expression levels showed a high degree of heterogeneity among different stages and subgroups, further confirming the heterogeneity of B-cell subpopulations in the MIF signaling pathway. Also, analysis of cellular metabolism has revealed a diversity of options for different B-cell subpopulations in terms of cellular metabolic demands and biological functions. These results are consistent with the report that activating MIF signaling reflects the importance of B cells in tumor immunity and suggests the different functions of different B-cell subpopulations in the pathomechanism of DLBCL [48]. Additionally, DLBCL arising from immunodominant loci exhibits unique immunological features, including immune evasion mechanisms and altered immune recognition [49]. Furthermore, it has been reported that immune cells (such as regulatory T cells and macrophages) are closely linked to prognosis after treatment, as these cells interact with tumor cells through various mechanisms, affecting patient survival rates [50], increased Tregs correlate with enhanced infiltration of cytotoxic T lymphocytes, which also play a crucial role in controlling tumors [51].

Following an interaction analysis of 1000 B-cell differentiation genes and 490 DEGs driven by intratumoral microbes, we identified 9 common genes, including GSTM5, LURAP1, LINC02802, MAB21L3, C2CD4D, MMEL1, TSPAN2, and CITED4, which further validates their role as important value as potential biomarkers in the assessment of DLBCL progression and treatment response. Besides, these common genes can be grouped into three major functional categories: metabolism and oxidative stress, cellular differentiation and migration, and transcriptional regulation and signaling. Similarly, several common genes are involved in metabolic processes and the response to oxidative stress, both essential for B-cell survival within the tumor microenvironment. For instance, GSTM5 (Glutathione S-transferase Mu 5) plays a key role in detoxification, protecting cells from reactive oxygen species (ROS). This function is critical in the oxidative stress-rich environment of DLBCL, where B cells must adapt to survive [52]. CITED4 is another gene that responds to hypoxic conditions, helping cells adapt to low-oxygen environments commonly found in tumors and potentially aiding B-cell survival under metabolic stress [39]. LURAP1, TSPAN2, and MMEL1 are associated with cellular motility and differentiation processes. LURAP1 and TSPAN2 are involved in immune cell migration and interaction with the extracellular matrix, suggesting they may regulate the positioning and movement of B cells within lymphoid tissues or the tumor microenvironment [53]. It suggested that these genes could influence how B cells interact with other immune cells or respond to tumor-derived signals. MMEL1, on the other hand, contributes to the remodeling of the extracellular space through proteolytic activity, which may facilitate tumor invasion or immune cell infiltration. LINC02802, MAB21L3, and C2CD4D are involved in gene expression regulation and signaling pathways critical for B-cell differentiation and activation. Long non-coding RNA LINC02802 may regulate gene expression in B cells, while MAB21L3 is linked to developmental processes that could influence B-cell progenitor maturation. C2CD4D is a key player in calcium-dependent signaling, essential for B-cell activation and function, affecting how B cells respond to external stimuli within the tumor microenvironment [54–56]. Therefore, by grouping these genes according to their biological functions, we can better understand how they might contribute to B-cell heterogeneity in DLBCL and suggest that the intratumoral microbiota may be involved in the immune regulation and tumor progression of

DLBCL by regulating the differentiation status and function of B cells [57].

Although this study provides strong evidence to support the critical role of the intratumoral microbiota in DLBCL, several limitations remain. First, the sample size was small, limiting the results' statistical validity and generalizability. While our findings are promising, the relatively small cohort of 48 patients may not fully capture the heterogeneity of DLBCL across diverse patient populations. To strengthen the robustness and consistency of our findings, future studies should aim to expand the sample size and include multicenter cohorts. These efforts will help enhance statistical power and ensure conclusions are more broadly applicable. Additionally, incorporating publicly available datasets such as those from TCGA and GEO will allow for validation of our findings in larger, more diverse cohorts, which will further support the generalizability of our conclusions. In addition, while bioinformatics methods have helped us identify associations between the microbiota and host immune cells, these results remain correlative, and further experimental validation is needed to elucidate the specific mechanistic pathways. Specifically, future experiments could involve manipulating the composition of particular microbiota in DLBCL models, such as through microbial gene knockdown or overexpression assays, to observe their impact on immune responses and tumor progression directly. Our study was mainly based on transcriptomic data; however, future research should incorporate additional data modalities such as proteomics, metabolomics, and single-cell RNA sequencing to provide a more comprehensive view of the intratumoral microbiota's function and mechanisms of action. Furthermore, experimental validation of the interaction between microbiota and immune regulation remains essential. The laboratory research could include exploring the role of microbial species in modulating immune cell infiltration and cytokine profiles within the tumor microenvironment. In addition, studies should explore how interventions with the microbiota can improve the therapeutic response and prognosis of DLBCL. Examples include modulating the microbiota composition through probiotics, antibiotics, or fecal microbiota transplantation, which could enhance immune responses and therapeutic efficacy. These interventions might help reverse immune suppression or increase the sensitivity of DLBCL tumors to chemotherapy, providing new avenues for personalized treatment [12]. Through a multidimensional analysis, this study revealed the diversity of the intratumoral microbiota in DLBCL and its complex relationship with the tumor microenvironment. Our findings suggest that the intratumoral microbiota may not only serve as a potential prognostic biomarker for DLBCL but also provide new directions for personalized therapeutic strategies. In particular, manipulating the microbiome with probiotics or antibiotics could impact tumor progression, modulate immune responses, and enhance the efficacy of chemotherapy or immunotherapy. Future studies, including phenotypic and molecular functional studies of DLBCL cell lines and model animals, 16S-RNA sequencing, and in vitro and in vivo experiments with microbial supplements, are needed to assess the therapeutic potential of microbiota modulation and its impact on patient prognosis. Targeting the interaction mechanisms between intratumoral microorganisms and the immune microenvironment could offer new insights into the pathomechanisms of DLBCL and pave the way for the development of targeted therapies, ultimately improving patient prognosis.

In summary, the intratumoral microbiota plays a key role in the pathogenesis and prognosis of DLBCL by influencing the function of immune regulatory cells and their interactions with tumor cells. By thoroughly studying these interactions, future therapies could be developed targeting specific intratumoral microbial environments, thus improving treatment outcomes and survival rates for DLBCL patients. Furthermore, based on our findings, genes involved in metabolism and oxidative stress may help B cells adapt to the challenging conditions of the tumor microenvironment. At the same time, those governing differentiation and migration could influence B cells' spatial organization and immune function. Finally, genes that regulate transcription and signaling are likely critical in directing B-cell maturation and response to external cues, including microbial and tumor-derived signals. Future studies could explore these genes as potential drug targets, particularly for those involved in metabolic pathways, such as IFFO2, MMEL1, and CITED4, using bioinformatics tools like DGIdb. Functional studies could also be conducted to validate their roles in B-cell function and therapeutic response, which may lead to the development of novel targeted therapies for DLBCL.

Conclusions

This study reveals the potential role of microbiota in the pathogenesis and therapeutic prognosis of DLBCL. It also novelly redefined the influence of DLBCL intratumoral microbes and the tumor immune microenvironment on pathogenesis from the transcriptomic and single-cell sequencing perspectives of DLBCL samples. It was demonstrated that the diversity of the intratumoral microbiota is strongly associated with the prognosis of DLBCL, and the heterogeneity of the B-cell developmental process further highlights how the microbiota influences the therapeutic response and patient prognosis in DLBCL by affecting the tumor microenvironment and the immune response of the host. Our study provides new perspectives for understanding

the complex pathogenic mechanisms of DLBCL. It guides the development of future diagnostic and therapeutic strategies with the hope that targeted microbiota modulation could improve the prognosis and survival of DLBC patients.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00262-025-03972-x.

Acknowledgements We would like to thank Jianhua Wang and Jianhua Yang, Director of the Department of Pharmacy at the First Affiliated Hospital of Xinjiang Medical University in Urumqi, China, who provided significant assistance with fundraising and research projects.

Author contributions The idea, methodology, and interpretation of the research data are the work of all authors. Zheng-yi Jia Xiaoyu Li, Lina Ma, and Jie Wang were in charge of ideation, methodology, validation, formal analysis, software, and first draft. Siying Wang, Hong Du, Yun Wu, and Jing Yu were responsible for data curation, visualization, and investigation. Yunxia Xiang, Daiqin Xiong, Huiting Shan, Yubo Wang, Zhi Wang, and Jianping Hao were handled the resources of this project through analysis, validation, and investigation. Jie Wang was responsible for all aspects of the integrity of this study, including resources, writing, review, editing, supervision, project management, and funding acquisition. All authors examined and approved the final draft before it was submitted.

Funding This study was supported by the Xinjiang Uygur Autonomous Region Distinguished Young Scientists Fund Project (foundation No.: 2022D01E72) and the Xinjiang Uygur Autonomous Region Youth Science and Technology Top Talent Project-Youth Science and Technology Innovation Talent Training Program(foundation NO.: 2022TSYCCX0027).

Data availability The data used to identify the DLBCL intratumoral microbial genus were derived from the cBioPortal platform (https://www.cbioportal.org/). DLBCL transcriptomic and clinical phenotypic data were obtained from The Cancer Genome Atlas database (TCGA, https://portal.gdc.cancer.gov/). Single-cell sequencing and independent cohort data, including the GSE182434, GSE195525, and GSE32018 datasets, were obtained from the Gene Expression Omnibus database (GEO, https://www.ncbi.nlm.nih.gov/geo/).

Declarations

Conflict of interest The authors declare no competing interests.

Ethical approval Not applicable.

References

1. Shi Y, Xu Y, Shen H, Jin J, Tong H, Xie W (2024) Advances in biology, diagnosis and treatment of DLBCL. Ann Hematol 103:3315–3334. https://doi.org/10.1007/s00277-024-05880-z
2. Sehn LH, Salles G (2021) Diffuse large B-cell lymphoma. N Engl J Med 384:842–858. https://doi.org/10.1056/NEJMra2027612
3. Dabrowska-Iwanicka A, Nowakowski GS (2023) DLBCL: who is high risk and how should treatment be optimized? Blood. https://doi.org/10.1182/blood.2023020779
4. Peled JU, Gomes ALC, Devlin SM, Littmann ER, Taur Y, Sung AD, Weber D, Hashimoto D, Slingerland AE, Slingerland JB et al (2020) Microbiota as predictor of mortality in allogeneic hematopoietic-cell transplantation. N Engl J Med 382:822–834. https://doi.org/10.1056/NEJMoa1900623
5. Uribe-Herranz M, Klein-González N, Rodríguez-Lobato LG, Juan M, Fernández de Larrea C, (2021) Gut microbiota influence in hematological malignancies: from genesis to cure. Int J Mol Sci, 22, 1026, https://doi.org/10.3390/ijms22031026
6. Galeano Niño JL, Wu H, LaCourse KD, Kempchinsky AG, Baryiames A, Barber B, Futran N, Houlton J, Sather C, Sicinska E et al (2022) Effect of the intratumoral microbiota on spatial and cellular heterogeneity in cancer. Nature 611:810–817. https://doi.org/10.1038/s41586-022-05435-0
7. Pushalkar S, Hundeyin M, Daley D, Zambirinis CP, Kurz E, Mishra A, Mohan N, Aykut B, Usyk M, Torres LE et al (2018) The pancreatic cancer microbiome promotes oncogenesis by induction of innate and adaptive immune suppression. Cancer Discov 8:403–416. https://doi.org/10.1158/2159-8290.CD-17-1134
8. Qin Y, Tong X, Mei W-J, Cheng Y, Zou Y, Han K, Yu J, Jie Z, Zhang T, Zhu S et al (2024) Consistent signatures in the human gut microbiome of old- and young-onset colorectal cancer. Nat Commun 15:3396. https://doi.org/10.1038/s41467-024-47523-x
9. Guevara-Ramírez P, Cadena-Ullauri S, Paz-Cruz E, Tamayo-Trujillo R, Ruiz-Pozo VA, Zambrano AK (2023) Role of the gut microbiota in hematologic cancer. Front Microbiol 14:1185787. https://doi.org/10.3389/fmicb.2023.1185787
10. Jiang P, Yu F, Zhou X, Shi H, He Q, Song X (2024) Dissecting causal links between gut microbiota, inflammatory cytokines, and DLBCL: a mendelian randomization study. Blood Adv 8:2268–2278. https://doi.org/10.1182/bloodadvances.2023012246
11. Yuan L, Wang W, Zhang W, Zhang Y, Wei C, Li J, Zhou D (2021) Gut microbiota in untreated diffuse large B cell lymphoma patients. Front Microbiol 12:646361. https://doi.org/10.3389/fmicb.2021.646361
12. Yoon SE, Kang W, Choi S, Park Y, Chalita M, Kim H, Lee JH, Hyun D-W, Ryu KJ, Sung H et al (2023) The influence of microbial dysbiosis on immunochemotherapy-related efficacy and safety in diffuse large B-cell lymphoma. Blood 141:2224–2238. https://doi.org/10.1182/blood.2022018831
13. Carrot-Zhang J, Yao X, Devarakonda S, Deshpande A, Damrauer JS, Silva TC, Wong CK, Choi HY, Felau I, Robertson AG et al (2021) Whole-genome characterization of lung adenocarcinomas lacking alterations in the RTK/RAS/RAF pathway. Cell Rep 34:108784. https://doi.org/10.1016/j.celrep.2021.108784
14. de Bruijn I, Kundra R, Mastrogiacomo B, Tran TN, Sikina L, Mazor T, Li X, Ochoa A, Zhao G, Lai B et al (2023) Analysis and visualization of longitudinal genomic and clinical data from the AACR project GENIE biopharma collaborative in cBioPortal. Cancer Res 83:3861–3867. https://doi.org/10.1158/0008-5472.CAN-23-0816
15. Zhang B, Liu J, Li H, Huang B, Zhang B, Song B, Bao C, Liu Y, Wang Z (2023) Integrated multi-omics identified the novel intratumor microbiome-derived subtypes and signature to predict the outcome, tumor microenvironment heterogeneity, and

immunotherapy response for pancreatic cancer patients. Front Pharmacol. https://doi.org/10.3389/fphar.2023.1244752
16. Wilkerson MD, Hayes DN (2010) ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics (Oxford, England) 26:1572–1573. https://doi.org/10.1093/bioinformatics/btq170
17. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) Limma powers differential expression analyses for rna-sequencing and microarray studies. Nucleic Acids Res 43:e47. https://doi.org/10.1093/nar/gkv007
18. Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, Ebert D, Feuermann M, Gaudet P, Harris NL, Hill DP (2023) The gene ontology knowledgebase in 2023. Genetics 224(1):31. https://doi.org/10.1093/genetics/iyad031
19. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan LI, Fu X (2021) clusterprofiler 4.0: a universal enrichment tool for interpreting omics data. The innovation 2(3):100141
20. Castanza AS, Recla JM, Eby D, Thorvaldsdóttir H, Bult CJ, Mesirov JP (2023) Extending support for mouse data in the molecular signatures database (MSigDB). Nat Methods 20:1619–1620. https://doi.org/10.1038/s41592-023-02014-7
21. Zeng D, Ye Z, Shen R, Yu G, Wu J, Xiong Y, Zhou R, Qiu W, Huang N, Sun L et al (2021) IOBR: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. Front Immunol 12:687975. https://doi.org/10.3389/fimmu.2021.687975
22. Maeser D, Gruener RF, Huang RS (2021) oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. Brief Bioinform 22:260. https://doi.org/10.1093/bib/bbab260
23. Steen CB, Luca BA, Esfahani MS, Azizi A, Sworder BJ, Nabet BY, Kurtz DM, Liu CL, Khameneh F, Advani RH et al (2021) The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma. Cancer Cell 39:1422-1437.e10. https://doi.org/10.1016/j.ccell.2021.08.011
24. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh P, Raychaudhuri S (2019) Fast, sensitive and accurate integration of single-cell data with harmony. Nat Methods 16:1289–1296. https://doi.org/10.1038/s41592-019-0619-0
25. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, Srivastava A, Molla G, Madad S, Fernandez-Granda C et al (2024) dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol 42:293–304. https://doi.org/10.1038/s41587-023-01767-y
26. Zappia L, Oshlack A (2018) Clustering trees: a visualization for evaluating clusterings at multiple resolutions. GigaScience 7:83. https://doi.org/10.1093/gigascience/giy083
27. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR et al (2019) Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol 20:163–172. https://doi.org/10.1038/s41590-018-0276-y
28. Vahid MR, Brown EL, Steen CB, Zhang W, Jeon HS, Kang M, Gentles AJ, Newman AM (2023) High-resolution alignment of single-cell and spatial transcriptomes with CytoSPACE. Nat Biotechnol 41:1543–1548. https://doi.org/10.1038/s41587-023-01697-9
29. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan C-H, Myung P, Plikus MV, Nie Q (2021) Inference and analysis of cell-cell communication using cell chat. Nat Commun 12:1088. https://doi.org/10.1038/s41467-021-21246-9
30. Qiu X, Hill A, Packer J, Lin D, Ma Y-A, Trapnell C (2017) Single-cell mRNA quantification and differential analysis with census. Nat Methods 14:309–315. https://doi.org/10.1038/nmeth.4150
31. Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, Cheng Y, Huang S, Liu Y, Jiang S et al (2022) Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. Cancer Discov 12:134–153. https://doi.org/10.1158/2159-8290.CD-21-0316
32. DepMap, Broad (2024). DepMap 24Q4 Public. Figshare+. Dataset. https://doi.org/10.25452/figshare.plus.27993248.v1.
33. Gómez-Abad C, Pisonero H, Blanco-Aparicio C, Roncador G, González-Menchén A, Martinez-Climent JA, Mata E, Rodríguez ME, Muñoz-González G, Sánchez-Beato M et al (2011) PIM2 inhibition as a rational therapeutic approach in B-cell lymphoma. Blood 118:5517–5527. https://doi.org/10.1182/blood-2011-03-344374
34. Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, Gill S, Harrington WF, Pantel S, Krill-Burger JM et al (2017) Defining a cancer dependency map. Cell 170:564-576.e16. https://doi.org/10.1016/j.cell.2017.06.010
35. Garg M, Takyar J, Dhawan A, Saggu G, Agrawal N, Hall A, Raut M, Ryland KE (2022) Diffuse large B-cell lymphoma (DLBCL): a structured literature review of the epidemiology, treatment guidelines, and real-world treatment patterns. Blood 140:12106–12107. https://doi.org/10.1182/blood-2022-169045
36. Simpson RC, Shanahan ER, Scolyer RA, Long GV (2023) Towards modulating the gut microbiota to enhance the efficacy of immune-checkpoint inhibitors. Nat Rev Clin Oncol 20:697–715. https://doi.org/10.1038/s41571-023-00803-9
37. Zheng D, Liwinski T, Elinav E (2020) Interaction between microbiota and immunity in health and disease. Cell Res 30:492–506. https://doi.org/10.1038/s41422-020-0332-7
38. Savage TM, Vincent RL, Rae SS, Huang LH, Ahn A, Pu K, Li F, de Los S-A, Coker C, Danino T, Arpaia N (2023) Chemokines expressed by engineered bacteria recruit and orchestrate antitumor immunity. Sci Adv 9(10):9436
39. Shi Z, Li Z, Zhang M (2024) Emerging roles of intratumor microbiota in cancer: tumorigenesis and management strategies. J Transl Med 22:837. https://doi.org/10.1186/s12967-024-05640-7
40. Gunjur A, Shao Y, Rozday T, Klein O, Mu A, Haak BW, Markman B, Kee D, Carlino MS, Underhill C et al (2024) A gut microbial signature for combination immune checkpoint blockade across cancer types. Nat Med 30:797–809. https://doi.org/10.1038/s41591-024-02823-z
41. Stefoni V, Argnani L, Casadei B, Musuraca G, Nanni L, Pellegrini C, Broccoli A, Carella M, Caldarulo E, Mammoli F et al (2023) Gut microbiome in DLBCL patients undergoing first-line R-CHOP regimen—the oncopassport study. Hematol Oncol 41:283–285. https://doi.org/10.1002/hon.3164_195
42. Takahara T, Nakamura S, Tsuzuki T, Satou A (2023) The immunology of DLBCL. Cancers 15:835. https://doi.org/10.3390/cancers15030835
43. Autio M, Leivonen S-K, Brück O, Mustjoki S, Jørgensen JM, Karjalainen-Lindsberg M-L, Beiske K, Holte H, Pellinen T, Leppä S (2021) Immune cell constitution in the tumor microenvironment predicts the outcome in diffuse large B-cell lymphoma. Haematologica 106:718–729. https://doi.org/10.3324/haematol.2019.243626
44. Visweshwar N, Rico JF, Killeen R, Manoharan A (2023) Harnessing the immune system: an effective way to manage diffuse large B-cell lymphoma. Journal of Hematology 12:145–160
45. Tamma R, Ranieri G, Ingravallo G, Annese T, Oranger A, Gaudio F, Musto P, Specchia G, Ribatti D (2020) Inflammatory cells in diffuse large B cell lymphoma. J Clin Med 9:2418. https://doi.org/10.3390/jcm9082418
46. Maharaj K, Uriepero A, Sahakian E, Pinilla-Ibarz J (2022) Regulatory T cells (Tregs) in lymphoid malignancies and the impact of novel therapies. Front Immunol. https://doi.org/10.3389/fimmu.2022.943354

47. Bai Y, He T, Zhang L, Liu Q, Yang J, Zhao Z, Yang K, Zhang M (2022) Prognostic value of FOXP3 $^{+}$ regulatory T cells in patients with diffuse large B-cell lymphoma: a systematic review and meta-analysis. BMJ Open 12:e060659. https://doi.org/10.1136/bmjopen-2021-060659
48. Dai L, Fan G, Xie T, Li L, Tang L, Chen H, Shi Y, Han X (2024) Single-cell and spatial transcriptomics reveal a high glycolysis B cell and tumor-associated macrophages cluster correlated with poor prognosis and exhausted immune microenvironment in diffuse large B-cell lymphoma. Biomarker Res 12:58. https://doi.org/10.1186/s40364-024-00605-w
49. Leppä S (2023) Data from clinical impact of immune cells and their spatial interactions in diffuse large B-cell lymphoma microenvironment. https://doi.org/10.1158/1078-0432.c.6532694.v1
50. Xu Z-F, Yuan L, Zhang Y, Zhang W, Wei C, Wang W, Zhao D, Zhou D, Li J (2024) The gut microbiome correlated to chemotherapy efficacy in diffuse large B-cell lymphoma patients. Hematol Rep 16:63–75. https://doi.org/10.3390/hematolrep16010007
51. Li C, Jiang P, Wei S, Xu X, Wang J (2020) Regulatory T cells in tumor microenvironment: new mechanisms, potential therapeutic strategies and future prospects. Mol Cancer 19:116. https://doi.org/10.1186/s12943-020-01234-1
52. Sfanos KS (2023) Intratumoral bacteria as mediators of cancer immunotherapy response. Cancer Res 83:2985–2986. https://doi.org/10.1158/0008-5472.CAN-23-1857
53. Qiu J, Jiang Y, Ye N, Jin G, Shi H, Qian D (2024) Leveraging the intratumoral microbiota to treat human cancer: are engineered exosomes an effective strategy? J Transl Med 22:728. https://doi.org/10.1186/s12967-024-05531-x
54. Fruman DA, Limon JJ (2012) Akt and mTOR in B cell activation and differentiation. Front Immunol. https://doi.org/10.3389/fimmu.2012.00228
55. Tsai D-Y, Hung K-H, Chang C-W, Lin K-I (2019) Regulatory mechanisms of B cell responses and the implication in B cell-related diseases. J Biomed Sci 26:64. https://doi.org/10.1186/s12929-019-0558-1
56. Xu Z, Boothby MR, Chaudhuri J (2021) Editorial: B cell activation and differentiation: new perspectives on an enduring topic. Front Immunol. https://doi.org/10.3389/fimmu.2021.797548
57. Liu M, Bertolazzi G, Sridhar S, Lee RX, Jaynes P, Mulder K, Syn N, Hoppe MM, Fan S, Peng Y et al (2024) Spatially-resolved transcriptomics reveal macrophage heterogeneity and prognostic significance in diffuse large B-cell lymphoma. Nat Commun 15:2113. https://doi.org/10.1038/s41467-024-46220-z

Authors and Affiliations

Zheng Yijia[1] · Xiaoyu Li[1] · Lina Ma[1] · Siying Wang[1] · Hong Du[1] · Yun Wu[2] · Jing Yu[1] · Yunxia Xiang[3,4] · Daiqin Xiong[3,4] · Huiting Shan[3,4] · Yubo Wang[3,4] · Zhi Wang[3,4] · Jianping Hao[5] · Jie Wang[3,4]

✉ Jie Wang
JieW629@163.com

[1] School of Pharmacy, Xinjiang Medical University, Urumqi 830011, China

[2] Department of General Medicine, The First Affiliated Hospital of the Xinjiang Medical University, Urumqi 830011, China

[3] Department of Pharmacy, The First Affiliated Hospital of Xinjiang Medical University, Urumqi 830011, China

[4] Xinjiang Key Laboratory of Clinical Drug Research, Urumqi 830011, China

[5] Department of Haematology, The First Affiliated Hospital of the Xinjiang Medical University, Urumqi 830011, China