Integrated analysis of 14 lymphoma datasets revealed high expression of CXCL14 promotes cell migration in mantle cell lymphoma

Dahai Liu  
Foshan University

Fei Qi  
Peking University First Hospital

Wei Liu  
Anhui Medical College

Fang Liu  
Foshan University

Yang Xun  
Foshan University

Min-Min Chen  
Foshan University

Xin Chen  
Foshan University

Shu-Ting Yang  
Foshan University

Wen-Qiao Jiao  
Foshan University

Zong-Ye Li  
Foshan University

Hua Yang  
Foshan University

Wen-Xing Li  
E liwenxin@mail.kiz.ac.cn  
Kunming Institute of Zoology Chinese Academy of Sciences  
https://orcid.org/0000-0001-9984-8439

Research article

Keywords: mantle cell lymphoma, cytokines, gene expression, CXCL14, cell migration

Posted Date: October 4th, 2019

DOI: https://doi.org/10.21203/rs.2.15629/v1
Abstract

Background Lymphoma is accompanied by multiple immune functions impairment. Cytokines play an important role in a variety of immune-related functions and affect the tumor microenvironment. However, the exact regulatory mechanisms between them remains unclear. This study aimed to explore the cytokines expression and function in Hodgkin's lymphoma (HL), diffuse large B-cell lymphoma (DLBCL) and mantle cell lymphoma (MCL). Results We performed a transcriptome integration analysis of 14 lymphoma datasets including 240 Hodgkin's lymphoma, 891 diffuse large B-cell lymphoma, 216 mantle cell lymphoma, and 64 health samples. The results showed that multiple immune function and signal pathway damage were shared by all the three types of lymphoma, and these functions were related to cytokines. Furthermore, through co-expression network and functional interaction network analysis, we identified CXCL14 as a key regulator and it affects cell chemotaxis and migration functions. Functional experiment showed that CXCL14 knockdown inhibited cell migration in MCL cell lines. Conclusions High expression of CXCL14 may aggravate MCL via promote cell migration. Our findings provide novel insights into the biology of MCL and would be helpful for the pathogenesis study and drug discovery of lymphomas.

Introduction

Hodgkin's lymphoma (HL), a B cell-derived cancer originating in lymphocytes and involving the lymphatic system, is one of the most commonly diagnosed forms of lymphoma in the Western world. Patients are commonly diagnosed with HL in their 20s and 30s, and they present with supradiaphragmatic lymphadenopathy, often with systemic B symptoms [1, 2]. Hodgkin's lymphoma also stands out as one of the most highly curable forms of cancer, with the vast majority (80% or more) of patients achieving clinical cure using modern chemo- and radiotherapy combinations [3]. Non-Hodgkin's lymphoma (NHL) is an umbrella term referring to various closely related lymphoproliferative malignancies. According to the World Health Organization (WHO), there are more than 60 different types of cancer classified under the broader heading of NHL. Diffuse large B-cell lymphoma (DLBCL), of which more than a dozen subtypes exist, is the most common form of NHL in all countries and age groups [4, 5]. B-cell lymphomas can be classified according to their rate of growth as low grade (indolent) or high grade (aggressive). High-grade forms include DLBCL, transformed follicular lymphoma, Burkitt's lymphoma and lymphoblastic lymphoma. Mantle cell lymphoma (MCL) is unique in that it has features of both indolent and aggressive disease [6, 7].

Despite advances in treatments for HL and NHL, more than 20% patients still died of disease progression. Immunosuppression is the most important risk factor for non-Hodgkin's lymphoma. Various diseases and conditions can induce a state of compromised immune function which may place the patient at increased risk of NHL [7]. In addition, several immunotherapeutics have been successfully applied to HL treatment, such as Brentuximab vedotin, an antibody-drug conjugate that targets CD30 [8]. Recently published results have shown that chimeric antigen receptor (CAR) T-cell therapy is active and safe in patients with refractory large B-cell lymphoma [9]. Transcriptomic studies have also shown that
impairment of immune functions can affect the prognosis of lymphoma. A previous study identified disturbed of cell-mediated immune response, cell-to-cell signaling and interaction, and up-regulation of pathway genes involved in interleukin–12 signaling and production in macrophages and apoptosis were associated with poor prognosis of classic Hodgkin's lymphoma (cHL) [10]. Another study showed a distinct tolerogenic host immune response between T-cell/histiocyte-rich large B-cell lymphoma (THRLBCL) and nodular lymphocyte-predominant Hodgkin's lymphoma (NLPHL), and further identified some key regulatory genes, such as CCL8 [11]. The distinct immune status may also be the main difference between THRLBCL and NLPHL [12].

Cytokines, a broad and loose category of small proteins, plays an important role in innate immune response and adaptive immune response. Cytokines can serve as metabolic hormones to provide adaptations to nutrient fluctuations, affect key metabolic processes such as glucose metabolism and lipid metabolism, eventually leading to multiple inflammation, metaflammation and immunometabolic disorders [13]. A previous report demonstrated that EBV+ and EBV- cHL tissues can be clearly separated from each other by a series of gene markers, such as CXCL9, CXCL10, CCL20 and other genes that involving innate immunity and antiviral responses in EBV+ tumors [14]. Other studies have also reported that cytokines (including CCL3, CCL4 and CCL18) may affect lymphoma microenvironment and patient’s survival both in DLBCL and MCL [15, 16]. However, there is a lack of systematic studies on these cytokines in different lymphomas. Therefore, this study aim to explore the expression and function of multiple cytokines, as well as explore key cytokines and their regulations in HL, DLBCL and MCL through transcriptome integration analysis.

**Materials And Methods**

**Lymphoma datasets and cytokines collection**

Transcriptome datasets of HL, DLBCL and MCL were downloaded from NCBI-GEO (https://www.ncbi.nlm.nih.gov/geo/) in Dec, 2017. The data selection criteria were as follows: (1) all datasets were genome-wide; (2) the samples in each dataset must be ≥ 10; (3) all samples were non-cell-line samples; and (4) complete microarray raw or normalized data were available. If the data contains the following items: (1) the number of samples is less than 3 in cases or controls; (2) samples were treatment with drugs or other agents; (3) RNA degradation serious or the number of detected genes is too small, we will exclude these datasets. Based on the above criteria, we chose 14 datasets for our integrated analysis (GSE12453, GSE13996, GSE17920, GSE21452, GSE31312, GSE36000, GSE47044, GSE56315, GSE64555, GSE69053, GSE70910, GSE7788, GSE86613 and GSE93291). All samples are human lymphocytes and the details of these datasets are illustrated in Table 1. In total, we collected 240 HL samples, 891 DLBCL samples, 216 MCL samples, and 64 health samples.

Human cytokine genes were collected from Cytokines & Cell Online Pathfinder Encyclopedia (COPE) database (http://www.cells-talk.com/index.php/page/about). According to the records of the database, cytokines are mainly divided into 5 categories: interleukins, chemokines, interferons, colony stimulating...
factors and tumor necrosis factors. In total, we got 69 interleukins, 59 chemokines, 18 interferons, 3 colony stimulating factors and 18 tumor necrosis factors from COPE (Table S1). Then we mapped these cytokines to our datasets for subsequent analysis. The receptors of collected cytokines were obtained from previous literatures [17, 18].

Data preprocessing and differential expression analysis

R statistical software v3.4.1 (https://www.r-project.org/) was used to perform data preprocessing and differential gene expression gene analysis. Because these datasets contain different microarray platforms, they were grouped into 15 batches according to study and platform. Each batch contained only one study and one platform (Table S2). We used the Robust Multichip Average (RMA) algorithm in oligo package [19] to normalize the raw expression data and generate normalized gene expression intensity. Gene annotation, integration and renormalization of the 15 batches was carried out using a custom designed Python code. The renormalization method and scripts are detailed in our previous publications [20, 21]. Due to there are missing values of genes in some samples, we used the mean expression of these genes in whole samples to replace the missing data. Fortunately, the missing values have little effect on the data (Figure S1). The distributions of RMA-processed and global-renormalized gene expression values across all studies are shown in Figures S2 and S3. After global expression was renormalized, the distribution of gene expression values across all studies had a consistent range. The empirical Bayesian algorithm in the “limma” package [22] was used to detect differentially expressed genes between lymphoma patients and controls. Up- and down-regulated genes were defined as a log2 transformed fold-change (logFC) ≥ 1 or ≤ −1, respectively. A false discovery rate (FDR) corrected P value ≤ 0.05 was considered as significant.

KEGG and GO enrichment analysis

Differentially expressed genes in three types of lymphoma were used to performed KEGG pathway enrichment analysis. The corresponding relationships between human genes and pathways were downloaded from Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.kegg.jp/). The formula for enrichment analysis see the previous report [23]. The enrichment percentage in each subsystem was calculated as the number of differentially expressed genes divided by the number of all genes. GO biological process enrichment analysis of filtered gene list was using BiNGO plugin in Cytoscape [24]. BiNGO is a tool to determine which Gene Ontology (GO) categories are statistically overrepresented in a set of genes or a subgraph of a biological network. We use the default parameters to perform enrichment analysis, an FDR P value ≤ 0.05 was considered significantly enriched.

Gene networks analysis
Protein-protein interaction (PPI) networks were used to explore the interactions of common deregulated cytokines in three types of lymphomas. We used STRING web server (https://string-db.org/cgi/input.pl) to construct PPI networks of cytokines and other related genes. The parameter settings were: (1) the meaning of network edges was set as molecular action (line shape indicates the predicted mode of action); (2) the active interaction sources were chose all types (including text-mining, experiments, databases, co-expression, neighborhood, gene fusion and co-occurrence); (3) the minimum required interaction score was set as medium confidence of 0.4; and (4) the max number of interactors in first shell was query proteins only, the max number of interactors in second shell was set none or no more than 10 according to the number of query proteins. Gene co-expression networks were used to explore key cytokines. Firstly, we calculate the correlation coefficients of significantly differentially expressed cytokines and other genes in HL, DLBCL and MCL. Then chose the cytokines-genes pairs with correlation coefficients ≥ 0.7 and FDR P values ≤ 0.05 to construct the gene co-expression networks in each lymphoma. Cytokines with the most nodes in the network were defined as key cytokines. GeneMANIA plugin in Cytoscape was used to explore the functional interaction of key cytokines and related genes [25]. Functional association networks of queried key cytokines and related genes were generated based on their relationships, such as co-expression, co-localization, pathway, physical interactions, genetic interactions, shared protein domains and predicted. Biological functions of these genes were automatic generated and an FDR P value ≤ 0.05 was considered significantly enriched.

Cell culture

The MCL cell lines Z138 and G519 and human normal B cell line GM12878 were purchased from the Shanghai cell bank of Chinese Academy of Sciences. The cells were maintained in RPMI 1640 medium (Gibco, USA) supplemented with 10% fetal bovine serum (Gibco, USA), 100 units/mL penicillin and 100 µg/mL streptomycin (Gibco, USA), in a humidified atmosphere with 5% CO2 at 37 °C.

Transfections

The experimental group was divided into control group (untreated Z138 and G519 cells), scramble group (mock treatment), and siCXCL14 group (CXCL14 knockdown). Z138 and G519 cells in logarithmic growth phase were inoculated in 6-well plate at 5×105 cells per well. The medium was replaced with the new medium without double antibiotics. 100-pmol siRNA and 5-µL Lipofectamine 3000 was incubated at room temperature and then mixed. Consequently, the mixture was incubated for another 20 min and added into 6-well plate. The medium was replaced after 6 h of regular cultivation. Protein expression level was detected 48 h after transfection. All siRNA duplexes targeting CXCL14 and negative control siRNAs were purchased from Ribobio Co. (Guangzhou, China), with sequences of Si CXCL14: sense, 5′- GGG UCC AAA UGC U –3′ and antisense, 5′- AGC ACU UGC AUU UGG ACC C –3′. Besides, RFect siRNA Transfection Reagent (BIO-TRAN) was used for transfection of siRNA duplexes.
RT-qPCR

Total RNA was extracted from individual types of Z138 and G519 cells using the RNeasy kit according to the manufacturers’ protocol (Omega biotek, USA) and reversely transcribed into cDNA using the SuperScript II reverse transcriptase (Fisher, Pittsburgh, USA). The relative levels of CXCL14 mRNA transcripts in individual samples were determined using the CFX Connect (Bio-Rad, USA). The sequences of specific primers were forward 5’- GTT CTC TGA GGA ACT CAA GTT TGG –3’ and reverse 5’- CTT TAA GGA TCA TTT GTC TCG CC –3’ for CXCL14. All samples were normalized to internal controls, and fold changes were calculated through relative quantification $2^{-\Delta\Delta Ct}$. The experiments were performed in triplicate.

Western blot

Western blot was performed as previously described [26]. The primary antibodies CXCL14 is purchased from Peprotech (0.2μg/ml, Peprotech, USA, Cat:500-P237). Band intensities were quantified by densitometry using ImageJ software [27].

Cell proliferation assay

Cell proliferation assay was measured using the cell counting kit–8 (CCK8) method. 2 × 103 tumor cells/well were seeded into 96-well plates and grown for 24 h, then treated as follows: control group, scramble group, and siCXCL14 group. After incubation for 24 h, 48 h, 72 h and 96 h), 10 μl of CCK–8 solution (Dojindo Laboratories, Japan) was added to each well of the plate. The plate was incubated for an additional 4 h, and the absorbance was measured at 450 nm using a microplate reader (BioTek, Winooski, VT, USA). The percentage of cell viability (relative proliferation rate) was calculated by comparison with the control group. The experiments were performed in triplicate.

Cell migration assay

MCL cells treated as follows: control group, scramble group, and siCXCL14 group. After transfection of 48h, 2 × 106 cells/mL of each cell type was starved in serum-free 1640 for 12 hour at 37°C in 5% CO2. Migration assays were subsequently performed using Transwell chambers with 8-μm pore filters (Corning, USA). Cell suspensions (2 × 105 in 100 μL) were added to the upper chambers and 600 μL of medium either containing 10% FBS was added to each of the lower chambers. After transwells were incubated for 24 hours at 37°C in 5% CO2, the cells in each lower chamber were recovered and counted using CCK8 assay; the entire assay was repeated 3 times.
Statistical analysis

R statistical software v3.4.1 was used for statistical analysis. The difference of cell proliferation and cell migration in three groups were compared using the one-way analysis of variance (ANOVA). Tukey's honestly significant difference (HSD) test was used to compared the difference between two groups. P<0.05 was considered to indicate a statistically significant difference.

Results

Deregulated genes overview in three types of lymphoma

In total, we got 18,116 unique genes in the integrated lymphoma datasets. Deregulated genes in each lymphoma dataset was illustrated in Table 2. We used different fold-change (FC) thresholds to show the trend of deregulated genes in three types of lymphoma. Within the FC cutoff of 2.0 (the threshold value used in this study), we obtained 2590 up-regulated and 541 down-regulated genes, 2441 up-regulated and 903 down-regulated genes, as well as 2560 up-regulated and 515 down-regulated genes in HL, DLBCL, and MCL, respectively. Within the FC cutoff of 1.5, there were more than 4000 up-regulated genes and more than 2000 down-regulated genes in each lymphoma dataset. Within the FC cutoff of 4.0, there were also more than 600 up-regulated genes in each lymphoma dataset, however, only a handful down-regulated genes in these lymphoma datasets. These results suggested that up-regulated genes may play a more important role in the progression of lymphoma. The top 20 differentially-expressed genes in three types of lymphoma are shown in Table S3, all these top genes were over-expressed. Furthermore, 5 over-expressed cytokines (CCL18, CCL19, CXCL9, CXCL10 and CXCL13) were in these top 20 genes.

Damaged immune functions and signaling pathways in lymphomas

KEGG pathway enrichment results of HL, DLBCL and MCL were showed in Tables S4-S6. The results showed that several immune functions and signaling pathways were significantly enriched in all the three lymphomas (Figure 1). Among these immune functions, we found that chemokine signaling pathway, complement and coagulation cascades, systemic lupus erythematosus, rheumatoid arthritis and primary immunodeficiency were enriched in all the three lymphomas. Furthermore, hematopoietic cell lineage, Th1 and Th2 cell differentiation, allograft rejection and graft-versus-host disease were enriched in HL and MCL. These results indicated that the immune functions were seriously damaged in HL, DLBCL and MCL. Among signaling pathways, we found cytokine-cytokine receptor interaction, NF-κB signaling pathway and TNF signaling pathway were enriched in all the three lymphomas. In view of the large number of
cytokines involved in these pathways, it necessary to further analyze the expression and function of cytokines in different lymphomas.

**Differentially expressed cytokines and their functions**

We used the collected cytokines to match the lymphoma database and got 44 interleukins, 43 chemokines, 17 interferons, 3 colony stimulating factors and 18 tumor necrosis factors. Expression profiles of these cytokines in HL, DLBCL and MCL were showed in Figure 2. There were 34 up-regulated cytokines and 1 down-regulated cytokine in all the three lymphomas. In addition, we found IL8, IL9, IL12B, IL17B, IL26, CCL20, CCL24, LIF and TNFSF10 expressed varies in three types of lymphoma. Protein-protein interaction network of the 35 commonly deregulated cytokines was showed in Figure 3. By k-means clustering, we divided these cytokines into three categories. We found there were strong interactions between multiple C-C motif chemokine ligands and C-X-C motif chemokine ligands. Only IL17C has no interaction with other cytokines. There were 32 significantly enriched KEGG pathways of these commonly deregulated cytokines. Most of these pathways are associated with immune function and signaling pathway, such as cytokine-cytokine receptor interaction, chemokine signaling pathway, TNF signaling pathway, NF-κB signaling pathway, NOD-like receptor signaling pathway and Jak-STAT signaling pathway.

**Identify key cytokines and their functions through network analysis**

Co-expression networks of deregulated cytokines and related genes in HL, DLBCL and MCL were showed in Figure 4. According to our screening criteria, there were 19 cytokines and 174 co-expressed genes in HL network (Figure 4A), 4 cytokines and 14 co-expressed genes in DLBCL network (Figure 4B), and 22 cytokines and 422 co-expressed genes in MCL network (Figure 4C). The number of nodes in the DLBCL network is too small and lacks statistical power. The MCL network had the highest clustering coefficient and network centralization (Table S7). There was no more than 100 nodes of cytokines in HL or DLBCL network. However, there were 11 cytokines with nodes more than 100 in MCL network (CXCL14, CCL19, IL33, CXCL12, CXCL13, CCL21, IL32, CCL2, CXCL9, CXCL10 and IL18). Therefore, we identified these genes as key cytokines in MCL. Among these genes, CXCL14 has the largest number of nodes and its function on lymphoma has not been reported yet. Functional interaction network of showed there was physical interaction between CXCL14 and CXCR4 (Figure 4D). GO biological process enrichment analysis of these CXCL14 related genes suggested that CCL13, CCL23, CXCL12, CXCR4 and PROS1 were involved in the following 5 biological processes: leukocyte migration, cell chemotaxis, leukocyte chemotaxis, chemokine activity, and chemokine receptor binding.
Functional experiment of CXCL14 on mantle cell lymphoma

We examined the mRNA and protein expression of CXCL14 in normal B cell line (GM12878) and MCL cell lines (Z138 and G519). CXCL14 mRNA and protein level in Z138 were higher than in GM12878 whereas in G519 showed a lower expression (Figure S4). To explore the function of CXCL14 on MCL, we constructed CXCL14 knockdown MCL cell lines (Z138 and G519) and performed cell proliferation and migration assay. The results showed a lower CXCL14 protein level in siCXCL14 group compared with control and scramble groups in Z138 and G519 cell lines (Figure 5A–5C). There was no influence of CXCL14 knockdown on cell proliferation (Figure 5D and 5E). Interestingly, siCXCL14 group showed significantly decreased relative migration rate compared with control and scramble groups in both two MCL cell lines (Figure 5F and 5G). The above analysis showed that high expressed CXCL14 affected cell chemotaxis and migration functions in MCL. Combined with the experimental results, we speculated that high expressed CXCL14 may promote cell migration and aggravate MCL.

Discussion

Lymphomas are solid tumors of the immune system. Physiological immune checkpoint pathways are important to regulate self-tolerance, limit immune reactions, and moderate autoimmunity. Several immune checkpoint inhibitors and other cellular immunotherapies have already shown great success in HL patients [3, 28]. The most well established risk factor for the development of NHL is immunosuppression. Patients with HIV have an increased risk of developing high-grade NHL [7]. Recently, adoptive T-cell therapy with chimeric antigen receptor T cells (CAR-Ts) has achieved clinically successful application in patients with B-cell malignancies, new data from clinical trials have demonstrated the benefits of CAR-T therapy in the NHL setting [29]. In this study, we revealed multiple damaged immune-related pathways and deregulated immune genes shared by HL, DLBCL and MCL. Most of these pathways and genes are related to cytokine functions.

Cytokines play important roles in B-cell activation, proliferation, and apoptosis. A previous study showed dysregulated circulating cytokines (such as IL5, IL13, TNF, etc.) were correlated with B-cell non-Hodgkin lymphoma [30]. A recent study found that IL6, IL8, TNF and other cytokines expression varies in lymphoma patients (both HL and NHL) and health population [31]. Our results showed that 35 cytokines were consistent differentially expressed in HL, DLBCL and MCL. Among them, over-expressed CCL18, CCL19, CXCL9, CXCL10 and CXCL13 were listed in the top 20 deregulated genes in all lymphoma datasets. CCL18 and CCL19 are belong to C-C motif chemokine ligand family. The expression of CCL18 in DLBCL was higher than control group and patients with higher level of CCL18 had a shorter overall survival than those with lower level [32]. CCL19 and CXCL12 were previous found as chemoattractants for MCL B cells, and suggested that MCL B cells were induced to migrate by CXCL12 and CCL19 [33]. CXCL9, CXCL10 and CXCL13 are belong to C-X-C motif chemokine ligand family. CXCL9 and CXCL10 are ligands for CXCR3 and attract CXCR3-expressing natural killer (NK) cells and both are up-regulated in cHL.
tissues. Studies showed that CXCL9 and CXCL10 may cause functional NK cell deficiencies and lead to deterioration of the tumor microenvironment of cHL [34]. In addition, high-expressed CXCL9 and CXCL10 were also found in lymphoma-associated hemophagocytic syndrome (LAHS) [35]. CXCL13 was over-production within the central nervous system (CNS) of CNS lymphoma patients, studies suggested that CXCL13 may as potential biomarker of CNS lymphoma [36, 37].

Our data indicated that CXCL14 was critical for the immune function and high expression CXCL14 may promote cell migration and aggravate MCL. CXCL14 is a non-ELR (glutamic acid-leucine-arginine) chemokine with a broad spectrum of biological activities and expressed by a variety of immune and nonimmune cells. CXCL14 mainly contributes to the regulation of immune cell migration, also executes antimicrobial immunity [38]. The functional interaction results showed that CXCL14 interacts with CXCL12 and CXCR4. Activated CXCL12-CXCR4 axis promote cell migration has been widely reported [39, 40]. Although CXCR4 as a receptor for CXCL14 is controversial in previous reports [17, 18, 41]. Recent evidences verified that CXCL14 binds to CXCR4 and shows synergistic effect with CXCL12 in cancers [42, 43]. Upregulated CXCL14 expression was also found in breast implant-associated anaplastic large cell lymphoma [44], and the over-expressed CXCL14 was associated with poor survival in non-small cell lung cancer (NSCLC) patients after curative resection [45]. In this study, CXCL14 was highly expressed in all lymphomas. Therefore, we speculate that CXCL14 may play an important role in a variety of lymphomas.

Conclusions

In conclusion, this study identified a series of deregulated cytokines among HL, DLBCL and MCL. These cytokines are mainly involved in immune or inflammation-related functions and signaling pathways. This study firstly reported that high expression of CXCL14 may aggravate MCL via promote cell migration. Therefore, we suggest that CXCL14 can be used as a biomarker and a potential therapeutic target for MCL. Future studies are required to uncover its potential mechanisms.

Abbreviations

HL: Hodgkin’s lymphoma
DLBCL: diffuse large B-cell lymphoma
MCL: mantle cell lymphoma
COPE: Cytokines & Cell Online Pathfinder Encyclopedia
RMA: robust multichip average
FC: fold-change
KEGG: Kyoto Encyclopedia of Genes and Genomes
GO: Gene Ontology
PPI: protein-protein interaction

**FDR: false discovery rate**

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Availability of data and materials**

Transcriptome datasets of HL, DLBCL and MCL were downloaded from NCBI-GEO (https://www.ncbi.nlm.nih.gov/geo/). Human cytokine genes were collected from Cytokines & Cell Online Pathfinder Encyclopedia (COPE) database (http://www.cells-talk.com/index.php/page/about). Original experimental data are available from the corresponding author on reasonable request.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

This work was supported by the National Natural Science Foundation of China (No. 81570376, No. 81870307, No. 81570202), the Scientific Research Startup Fund of Foshan University (Grant No. CGg07025), the University Special Innovative Research Program of Department of Education of Guangdong Province (No.2017KTSCX189), the Featured Project of Innovation Enhancing College of Guangdong Province (No. 2015KTSCX154) and Foshan Science and Technology Innovation Project (No. 2015AG10010).

**Authors’ contributions**
WXL, HY and DL designed the study. WXL, HY, DL, FQ, and WL performed data collection. WXL, DL, FQ, WL, FL, and YX performed bioinformatics analysis. HY, MMC, XC, STY, WQJ and ZYL performed experimental verification. WXL, HY, DL and FL wrote and revised the manuscript. All authors read and approved the final manuscript.

**Acknowledgements**

Not applicable.

**References**


**Tables**

**Table 1. Information of three types of lymphoma datasets.**
<table>
<thead>
<tr>
<th>GEO ID</th>
<th>Contributor</th>
<th>Title</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE77881</td>
<td>Van Loo P, 2007</td>
<td>Nodular lymphocyte predominant Hodgkin's lymphoma vs T cell/histiocyte rich B cell lymphoma</td>
<td>10 cases 1 control</td>
</tr>
<tr>
<td>GSE124532</td>
<td>Brune V, 2008</td>
<td>Origin and pathogenesis of lymphocyte-predominant Hodgkin lymphoma as revealed by global gene expression analysis</td>
<td>17 cases 25 controls</td>
</tr>
<tr>
<td>GSE13996</td>
<td>Chetaille B, 2008</td>
<td>Molecular profiling of classical Hodgkin's lymphoma tissues</td>
<td>64 cases</td>
</tr>
<tr>
<td>GSE17920</td>
<td>Steidl C, 2009</td>
<td>Expression data of diagnostic biopsy samples from Hodgkin lymphoma patients</td>
<td>130 cases</td>
</tr>
<tr>
<td>GSE47044</td>
<td>Hartmann S, 2013</td>
<td>Nodular lymphocyte predominant Hodgkin lymphoma and T cell/histiocyte rich large B cell lymphoma - endpoints of a spectrum of one disease?</td>
<td>19 cases 5 controls</td>
</tr>
</tbody>
</table>

**Diffuse large B-cell lymphoma (DLBCL)**

<table>
<thead>
<tr>
<th>GEO ID</th>
<th>Contributor</th>
<th>Title</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE124532</td>
<td>Brune V, 2008</td>
<td>Origin and pathogenesis of lymphocyte-predominant Hodgkin lymphoma as revealed by global gene expression analysis</td>
<td>11 cases</td>
</tr>
<tr>
<td>GSE56315</td>
<td>Bødker JS, 2014</td>
<td>Diffuse Large B-Cell Lymphoma Classification System That Associates Normal B-Cell Subset Phenotypes With Prognosis</td>
<td>89 cases 33 controls</td>
</tr>
<tr>
<td>GSE64555</td>
<td>Linton K, 2014</td>
<td>Microarray gene expression analysis of FFPE tissue permits molecular classification in diffuse large B-cell lymphoma</td>
<td>40 cases</td>
</tr>
<tr>
<td>GSE69053</td>
<td>Sha C, 2015</td>
<td>Transferring genomics to the clinic: distinguishing Burkitt and Diffuse large B-cell lymphoma</td>
<td>212 cases</td>
</tr>
<tr>
<td>GSE86613</td>
<td>Bødker JS, 2016</td>
<td>Expression data from diagnostic samples of diffuse large B-cell lymphomas (DLBCL), follicular lymphoma (FL) and primary and relapsed transformed FL</td>
<td>41 cases</td>
</tr>
</tbody>
</table>

**Mantle cell lymphoma (MCL)**

<table>
<thead>
<tr>
<th>GEO ID</th>
<th>Contributor</th>
<th>Title</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE21452</td>
<td>Staudt LM, 2010</td>
<td>Integrated genomic profiling in mantle cell lymphoma</td>
<td>64 cases</td>
</tr>
<tr>
<td>GSE36000</td>
<td>Jares P, 2012</td>
<td>Mantle Cell Lymphoma</td>
<td>38 cases</td>
</tr>
<tr>
<td>GSE70910</td>
<td>Liu D, 2015</td>
<td>Direct in vivo evidence for B-cell receptor and NF-KB activation in mantle cell lymphoma: role of the lymph node microenvironment and activating mutations</td>
<td>55 cases</td>
</tr>
<tr>
<td>GSE93291</td>
<td>Staudt LM, 2017</td>
<td>A new molecular assay for the proliferation signature in mantle cell lymphoma applicable to formalin-fixed paraffin-embedded biopsies</td>
<td>59 cases</td>
</tr>
</tbody>
</table>

1 The one control sample in this study was the mixed five control samples.
2 This dataset including 17 HL samples, 11 DLBCL samples, and 25 controls. The number of control samples was showed in HL group and was not repeated in DLBCL group.
Table 2. Number of DEGs in merged datasets of the three types of lymphoma defined by different fold change cutoff values.

<table>
<thead>
<tr>
<th>Fold-change (log2)</th>
<th>Total number of DEGs (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HL</td>
</tr>
<tr>
<td></td>
<td>Up</td>
</tr>
<tr>
<td>1.5 (0.58)</td>
<td>4583</td>
</tr>
<tr>
<td>2.0 (1.0)</td>
<td>2590</td>
</tr>
<tr>
<td>4.0 (2.0)</td>
<td>668</td>
</tr>
</tbody>
</table>

Abbreviations: DEGs, differentially expressed genes; HL, Hodgkin’s lymphoma; DLBCL, diffuse large B-cell lymphoma; MCL, mantle cell lymphoma.

Up, up-regulated genes in lymphoma patients compared with controls, with log2 fold change high than the cutoffs and FDR smaller than 0.05.

Down, down-regulated genes in lymphoma patients compared with controls, with log2 fold change less than the negative cutoffs and FDR smaller than 0.05.

The percentage (%) = number of DEGs in the merged datasets / total number of genes in the merged datasets.

Figures
Figure 1

Immune function and signaling pathway enrichment results in three types of lymphoma. (A) Immune-related pathways enrichment results. (B) Signal transduction related pathways enrichment results. The orange, blue and green lines represent the HL, DLBCL and MCL, respectively. The line width indicates the enrichment percentage. The dotted line in the box indicates the significance threshold (FDR P value = 0.05). HL, Hodgkin's lymphoma; DLBCL, diffuse large B-cell lymphoma; MCL, mantle cell lymphoma.
Figure 2

Expression profiles of cytokines in three types of lymphoma. (A) Interleukins. (B) Chemokines. (C) Interferons. (D) Colony stimulating factors. (E) Tumor necrosis factors. The length of the bar indicates the log2(fold-change) between lymphoma samples and controls. The red, green and gray color represent the up-regulated genes, down-regulated genes and no change genes, respectively. The dotted line in the box indicates the log2FC = 1. HL, Hodgkin's lymphoma; DLBCL, diffuse large B-cell lymphoma; MCL, mantle cell lymphoma.
**Figure 3**

Interactions and functions of commonly deregulated cytokines. (A) Protein-protein interaction network of 35 commonly deregulated cytokines. The network using k-means clustering method and is clustered to 3 specified groups. (B) Significantly enriched KEGG pathways of genes in the network. A pathway with a FDR P value less than 0.05 was considered significant.

**Figure 4**

Network analysis identified CXCL14 as a key cytokine gene. Panel (A), (B), and (C) showed the gene co-expression networks of differentially expressed cytokines and other genes in Hodgkin's lymphoma, diffuse large B-cell lymphoma, and mantle cell lymphoma. Cytokine-gene pairs with correlation coefficients higher than 0.7 were chose to build the network. The red circles represent the cytokines. The orange, blue and green circles represent the genes in each type of lymphoma. The size of the circle indicates the number of nodes. (D) Functional interaction network analysis of CXCL14. The network shows the CXCL14 related genes and their functions. Each color line represents a different interaction; color line width indicates the weight of the interactions. There were 5 significantly enriched biological functions and the green square indicates the gene is involved in the biological function.

**Figure 5**

Effect of CXCL14 on cell proliferation and migration. (A) Western blot of CXCL14 in three treatment group. Z138 and G519 cells were grown and transiently transfected with CXCL14 or negative control for 48 h and then subjected to western blotting. (B) Quantified band intensities in Z138 cell lines. (C) Quantified band intensities in G519 cell lines. The dotted line indicates the reference integrated density (control group). (D) Knockdown of CXCL14 on cell proliferation in Z138 cells. (E) Knockdown of CXCL14 on cell proliferation in G519 cells. The reference proliferation rate was defined as in control group in each day. (F) Knockdown of CXCL14 on cell migration in Z138 cells. (G) Knockdown of CXCL14 on cell migration in G519 cells. The reference migration rate was defined as in control group. Significance: *** P < 0.001, ** P < 0.01, * P < 0.05. (H) High expressed CXCL14 promote cell migration and aggravate mantle cell lymphoma.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigures.docx](#)
- [SupplementaryTables.xlsx](#)