

# Identification of key genes and pathways associated with classical Hodgkin lymphoma by bioinformatics analysis

ZHIXING KUANG, LI GUO and XUN LI

Department of Lymphoma and Breast Cancer, The Affiliated Tumor Hospital of Xinjiang Medical University, Urumqi, Xinjiang Uyghur Autonomous Region 830011, P.R. China

Received September 11, 2016; Accepted May 22, 2017

DOI: 10.3892/mmr.2017.7158

**Abstract.** The current study aimed to explore the mechanisms associated with classic Hodgkin lymphoma (cHL) to identify novel diagnostic and therapeutic targets. The GSE12453 microarray dataset was downloaded from the Gene Expression Omnibus database; the differentially expressed genes (DEGs) between cHL samples and normal B cell samples by were identified using the limma package. Gene ontology (GO) and pathway enrichment analysis of DEGs gene were performed. Furthermore, construction and analysis of protein-protein interaction (PPI) network was performed, and co-expression modules of DEGs were produced. A total of 450 DEGs were identified, comprising 216 upregulated and 234 downregulated genes in cHL compared with normal B cell samples. The DEGs were enriched in biological processes associated with immune response. The upregulated genes were mainly associated with the pathway of transcriptional misregulation in cancer, while downregulated genes were associated with B cell receptor signaling. PPI network analysis demonstrated that IL6 had the highest connectivity degree. Interleukin-6 (IL6) and signal transducer and activator of transcription 1 (STAT1) were demonstrated to be involved with the response to cytokine GO term in co-expression module 1. Spleen tyrosine kinase (SYK), B-cell linker protein (BLNK), CD79B, phospholipase C  $\gamma$ 2 (PLCG2) were enriched in the B cell receptor signaling pathway in module 2. Matrix metalloproteinase 9 (MMP9), protein tyrosine phosphatase receptor type C had the highest connectivity degrees in module 3 and module 4, respectively. The results suggested that DEGs, including IL6, STAT1, MMP9, SYK, BLNK, PLCG2 and CD79B, and the pathways of B cell receptor signaling, Epstein-Barr virus infection and

transcriptional misregulation in cancer have strong potential to be useful as targets for diagnosis or treatment of cHL.

## Introduction

Classic Hodgkin lymphoma (cHL) is a germinal center-derived B-cell neoplasm characterized by a small number of tumor cells in a wide range of reactive infiltrating cells. cHL accounts for ~95% of HL cases (1). The incidence of the HL is 2.8 per 100,000 individuals (2). It was estimated that 8,500 people in the United States would be diagnosed with HL and 1,120 mortalities would be caused by HL in 2016 (3). Currently, cHL treatments include cytotoxic chemotherapy and involved-field radiation therapy. These treatments significantly reduce the mortality rate and produce a 5-year overall survival rate of 70-80%; however, ~20% of patients still develop progressive disease following these treatments (4). However, the pathogenesis of cHL is multifactorial and complex, therefore, improving the understanding of the molecular mechanisms involved in the pathogenesis of cHL is important for identification of novel therapeutic targets. In recent years, the pathogenic mechanisms of cHL have been extensively investigated; for example, CD30 has been demonstrated to be expressed on Hodgkin and Reed/Sternberg (HRS) cells in cHL and involved in development of cHL. The monoclonal antibody against CD30, termed Brentuximab vedotin, has been demonstrated to be a highly effective treatment for relapsed/refractory cHL (5), which supports that it is important to identify more potential targets for the treatment of cHL.

Microarray analysis has become an extensively used tool for the generation of gene expression data on a genomic scale. In the current study, GSE12453 microarray data from Brune *et al* (6) was downloaded. Brune *et al* (6) investigated differentially expressed genes (DEGs) among cHL, B-cell non-Hodgkin lymphoma and normal B cell samples (6,7). In the current study, DEGs were identified by comparing cHL and normal B cells samples, and gene function enrichment analysis, protein-protein interaction (PPI) network construction and co-expression sub-network construction were performed to identify the molecular markers of cHL. The results of the current study present candidate genes for future research into the pathogenic mechanisms of cHL and may help to establish novel diagnostic markers and therapeutic targets.

---

*Correspondence to:* Dr Xun Li, Department of Lymphoma and Breast Cancer, The Affiliated Tumor Hospital of Xinjiang Medical University, 789 Suzhou Road, Urumqi, Xinjiang Uyghur Autonomous Region 830011, P.R. China  
E-mail: smxlixun@sina.com

**Key words:** classic Hodgkin lymphoma, differentially expressed genes, pathway, protein-protein interaction network, module analysis

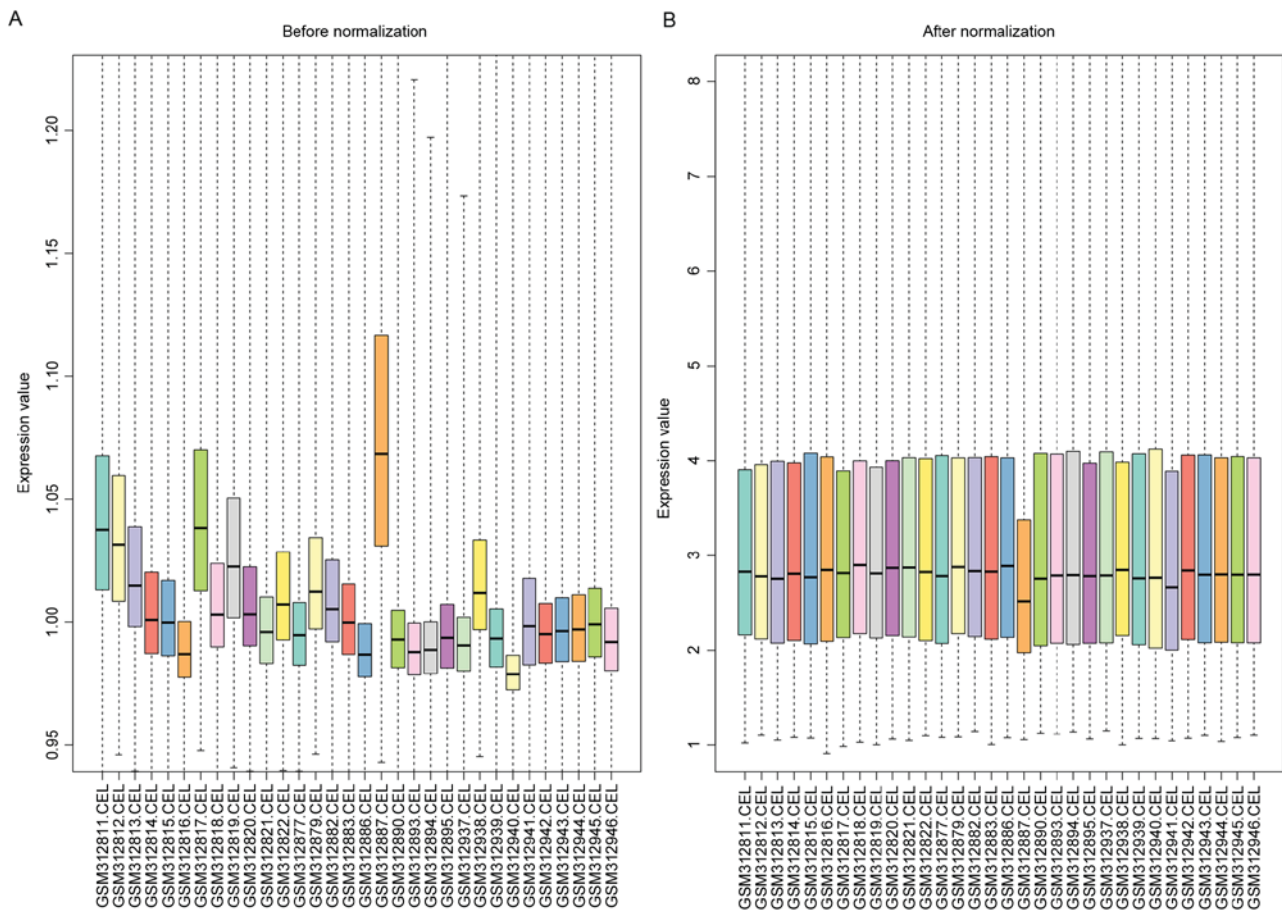


Figure 1. Box plot of expression data before and after normalization. The x-axis presents the different samples and the y-axis presents the expression value. The black line in each box represents the expression median for each sample. (A) Data before normalization and (B) data after normalization. After preprocessing, the expression medians were similar across samples, except for GSM312887.

## Materials and methods

**Microarray data.** The raw data of GSE12453 based on the GPL570 platform (HG-U133\_Plus\_2) contributed by Brune *et al* (6) were downloaded from the Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo/>). The GSE12453 dataset comprises 67 samples, including 12 cHL samples, 5 nodular lymphocyte-predominant Hodgkin lymphoma samples, 25 non-Hodgkin lymphoma samples and 25 normal B cells samples. The normal B cell samples are further divided into five subtypes: Normal naive B cells, normal memory B cells, normal centrocytes, normal centroblasts and normal plasma cells. Research has previously identified that HRS cells originate from germinal center (GC) B cells (8), which contain centrocytes and centroblasts (9). Under normal circumstances, GCB cells differentiate into memory B cells or plasma cells; however, GCB cells can transform into malignant B cells. Therefore, in the current study, the cHL samples and the normal B cell samples, which included GCB cells, memory B cells and plasma cells, were used for analysis.

**Data preprocessing and differential expression analysis.** Several R packages were used for raw microarray data preprocessing, including concluding quality assessment (QA), quality control (QC), background correction and normalization. The R packages using for QA/QC were simpleaffy (10),

affyPLM and arrayQualityMetrics (11). Through QC checks, the sample GSM312887 within the dataset was disqualified (Fig. 1) and excluded from subsequent analysis. Robust multi-array average (RMA) (12) was used to perform background correction, normalization and probe summarization. The empirical Bayesian method in limma package (13) was used to perform the differential analysis,  $|\log_2$  fold change (FC)|  $> 2.5$  and adjusted  $P < 0.001$  were considered as statistically significant.

**Gene ontology (GO) and pathway enrichment analyses.** GO is a popular approach to categorize the representation of genes and the attributes of gene expression. The Kyoto Encyclopedia of Genes and Genomes (KEGG; <http://www.genome.jp/kegg/>) (14) is an online database with collected information of genomes, biological chemicals and enzymatic pathways. Database for Annotation, Visualization and Integrated Discovery (DAVID; <https://david-d.ncifcrf.gov/>), is an exploratory visualization tool that help investigators determine the biological function a series of genes. To understand the function of the DEGs, the GO and pathway enrichment analyses were performed by using DAVID 6.8 (Beta) online tool, the threshold was set as  $P < 0.01$ .

**PPI network and co-expression module construction.** The Search Tool for the Retrieval of Interacting Genes

Table I. GO enrichment analysis for differentially expressed genes.

A, Upregulated genes		
Term	Count	P-value
GOTERM_BP_FAT		
GO:0006955 immune response	73	8.10x10 <sup>-29</sup>
GO:0006952 defense response	64	1.81x10 <sup>-19</sup>
GO:0001775 cell activation	46	2.13x10 <sup>-18</sup>
GOTERM_CC_FAT		
GO:0005615 extracellular space	49	1.56x10 <sup>-13</sup>
GO:0009897 external side of plasma membrane	19	5.93x10 <sup>-10</sup>
GO:0005576 extracellular region	86	4.44x10 <sup>-9</sup>
GOTERM_MF_FAT		
GO:0008009 chemokine activity	8	4.51x10 <sup>-7</sup>
GO:0042379 chemokine receptor binding	8	1.74x10 <sup>-6</sup>
GO:0005518 collagen binding	8	5.25x10 <sup>-6</sup>
B, Downregulated genes		
Term	Count	P-value
GOTERM_BP_FAT		
GO:0050853 B cell receptor signaling pathway	14	2.61x10 <sup>-15</sup>
GO:0042113 B cell activation	18	1.10x10 <sup>-11</sup>
GO:0050851 antigen receptor-mediated signaling pathway	16	5.08x10 <sup>-10</sup>
GOTERM_CC_FAT		
GO:0019814 immunoglobulin complex	7	6.34x10 <sup>-8</sup>
GO:0098552 side of membrane	16	5.16x10 <sup>-6</sup>
GO:0042571 immunoglobulin complex, circulating	5	3.76x10 <sup>-5</sup>
GOTERM_MF_FAT		
GO:0034987 immunoglobulin receptor binding	5	6.35x10 <sup>-5</sup>
GO:0003823 antigen binding	7	3.36x10 <sup>-4</sup>
GO:0005070 SH3/SH2 adaptor activity	5	8.38x10 <sup>-4</sup>

GO, gene ontology; BP, biological process; CC, Cell Component; MF, molecular function.

(STRING 10.0) is a database that was designed to predict the PPI between DEGs (15) and combination score >0.4 was considered to be significant. Cytoscape software was used to visualize the results of the PPI network (16), the sub-networks of co-expression were constructed by ClusterONE (17) at the base of PPI network. ClusterONE is the plug-in component of Cytoscape designed to identify closely connected and potentially overlapping regions in the PPI network (ClusterONE parameters: Minimum size, 3; minimum density, 0; overlap threshold, 0.8). GO and KEGG pathway enrichment analysis of modular genes was implemented using the DAVID online tool.

## Results

*DEGs between Hodgkin lymphoma and normal B cells.* A total of 450 DEGs were obtained from analysis of the dataset

GSE12453, of which 214 were upregulated and 234 were downregulated genes. The distribution of DEGs is presented in Fig. 2.

*GO and pathway enrichment analyses.* Analysis of GO and pathway enrichment for upregulated and downregulated genes was performed. The top three GO terms of functional classification are presented in Table I. The results indicated that upregulated DEGs primarily were involved in the GO terms relevant to immune response ( $P=8.10 \times 10^{-29}$ ), defense response ( $P=1.81 \times 10^{-19}$ ) and cell activation ( $P=2.13 \times 10^{-18}$ ). The GO terms predominantly associated with downregulated DEGs were the B cell receptor signaling pathway ( $P=2.61 \times 10^{-15}$ ), B cell activation ( $P=1.10 \times 10^{-11}$ ) and antigen receptor-mediated signaling pathway ( $P=5.08 \times 10^{-10}$ ).

The KEGG pathways enrichment of DEGs are presented in Table II. The upregulated DEGs were mainly enriched

Table II. Pathway enrichment analysis for DEGs (top three pathways with lower P-value).

A, Upregulated DEGs				
KEGG-ID	Pathway	Count	P-value	Genes
hsa05202	Transcriptional misregulation in cancer	13	6.64x10 <sup>-6</sup>	IL6, MMP9, GZMB, CD14, CCR7, CCND2, CDKN1A, ID2, IL1R2, ILRB, PLAU, MAF, CEBPB
hsa04620	Toll-like receptor signaling pathway	10	2.59x10 <sup>-5</sup>	IL6, STAT1, CCL3L3, CCL5, CXCL10, CXCL9, CD14, TLR8, CTSK, NFKBIA
hsa04060	Cytokine-cytokine receptor interaction	14	3.23x10 <sup>-5</sup>	IL6, TNFRSF8, CCL3L3, IL2RB, CXCL10, CXCL13, TNFSF13B, CXCL9, HGF, IL7R, IL2RA, CCR7, ILR2, CCL5,
B, Downregulated DEGs				
KEGG-ID	Pathway	Count	P-value	Genes
hsa04662	B cell receptor signaling pathway	8	1.90x10 <sup>-6</sup>	BLNK, PLCG2, PIK3CG, SYK, CD22, CR2, CD79B, RASGRP3
hsa04666	Fc $\gamma$ receptor-mediated phagocytosis	5	0.006327	PLCG2, SYK, PTPRC, PIK3CG, PRKCB
hsa04064	Nuclear factor- $\kappa$ B signaling pathway	5	0.007157	BLNK, PLCG2, SYK, ATM, TRAF5

KEGG, Kyoto encyclopedia of genes and genomes.

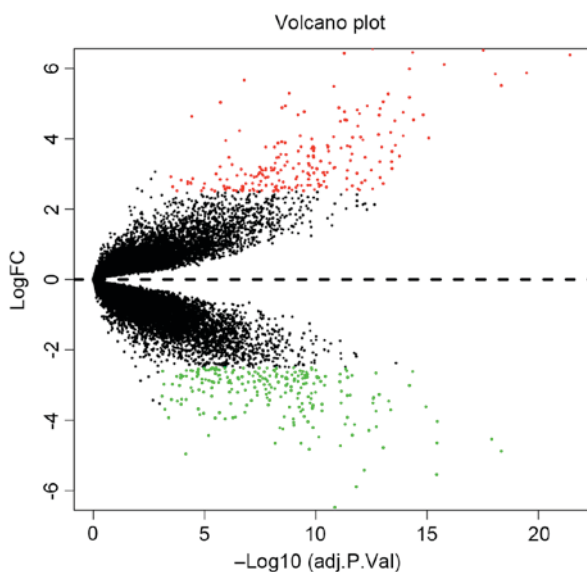


Figure 2. Volcano plot of differentially expressed genes in classical Hodgkin lymphoma. Red, significantly upregulated genes; green, significantly downregulated genes. FC, fold change.

in transcriptional misregulation in cancer ( $P=6.64 \times 10^{-6}$ ), Toll-like receptor signaling pathway ( $P=2.59 \times 10^{-5}$ ), cytokine-cytokine receptor interaction ( $P=3.23 \times 10^{-5}$ ), the downregulated DEGs were enriched in B cell receptor signaling pathway ( $P=1.90 \times 10^{-6}$ ), Fc  $\gamma$  receptor-mediated

phagocytosis ( $P=0.006327$ ) and nuclear factor- $\kappa$ B signaling pathway ( $P=0.007157$ ).

**PPI network and module analysis.** The PPI network of DEGs was constructed (Fig. 3). There were 230 nodes and 694 interactions, the top 10 genes with the highest degrees in the PPI network were interleukin 6 (IL6; degree, 43), protein tyrosine phosphatase, receptor type C (PTPRC; degree, 34), signal transducer and activator of transcription 1 (STAT1; degree, 34), matrix metalloproteinase 9 (MMP9; degree, 30), spleen tyrosine kinase (SYK; degree, 27), C-C motif chemokine receptor 7 (degree, 25), fibronectin 1 (degree, 23), TIMP metalloproteinase inhibitor 1 (degree, 22), integrin subunit  $\beta$ 1 (degree, 21) and clusterin (degree, 19; Table III). These genes were selected as the hub nodes. Four modules were obtained from the PPI network of DEGs (Fig. 4). The significant GO-biological process terms and KEGG pathway enrichment analysis for co-expression modules genes is presented in Table IV.

In module 1, all the DEGs were upregulated genes. STAT1 (degree, 7), C-X-C motif chemokine ligand (CXCL) 9 (degree, 5), CXCL10 (degree, 4) and IL6 (degree, 4) had high degrees. The analysis of GO function and pathway enrichment for DEGs demonstrated that all DEGs in module 1 were most enriched in the GO terms response to cytokine ( $P=3.21 \times 10^{-12}$ ) and the chemokine signaling pathway ( $P=1.65 \times 10^{-5}$ ). All DEGs in module 2 were downregulated genes, and the enriched GO terms were B cell activation ( $P=1.55 \times 10^{-6}$ ) and B cell receptor signaling pathway ( $P=1.86 \times 10^{-6}$ ). The

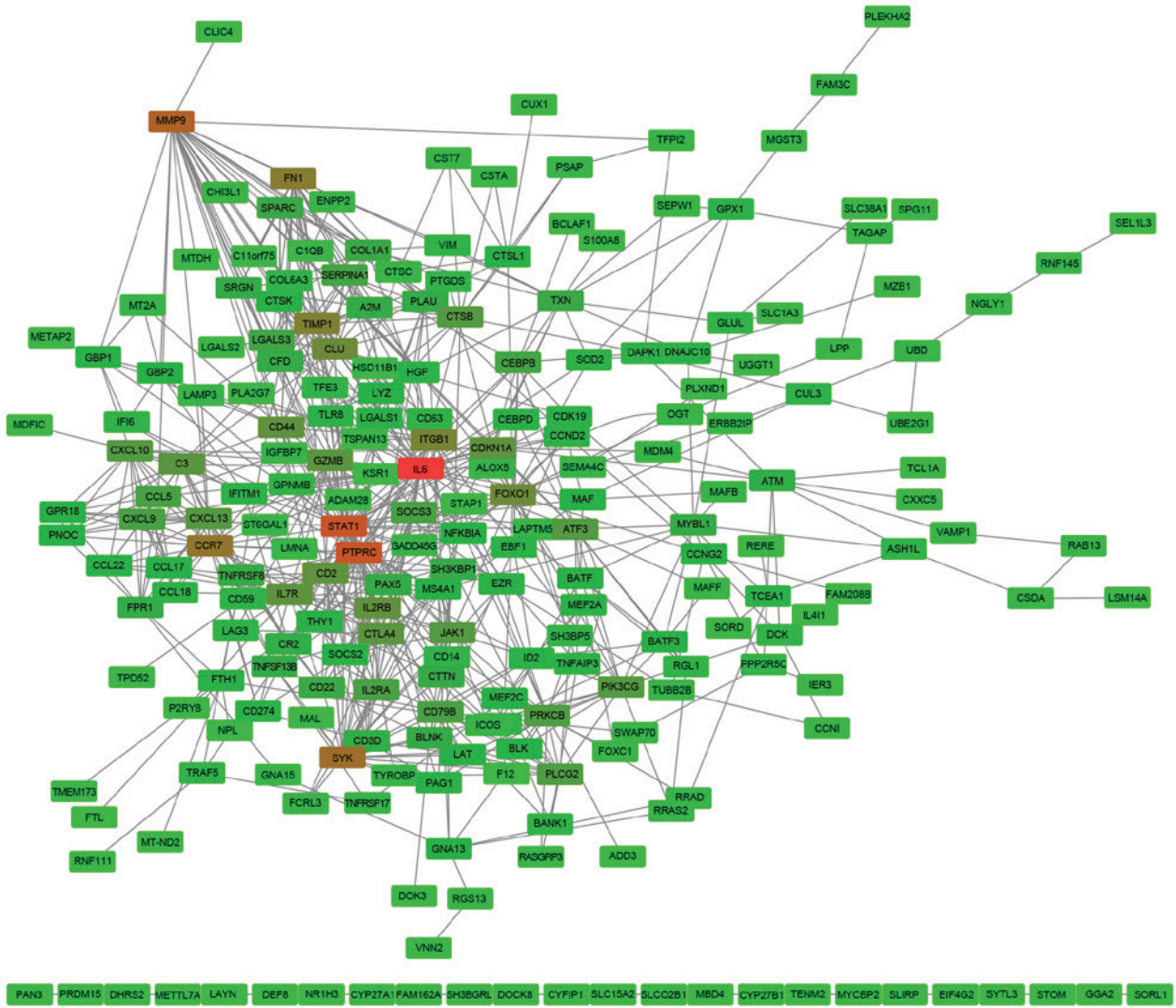


Figure 3. Protein-protein interaction network. The square nodes represent proteins and the grey lines represent interactions. The green color gradually changes to red as the degree increases and the depth of the red color is positively correlated with the degree of genes.

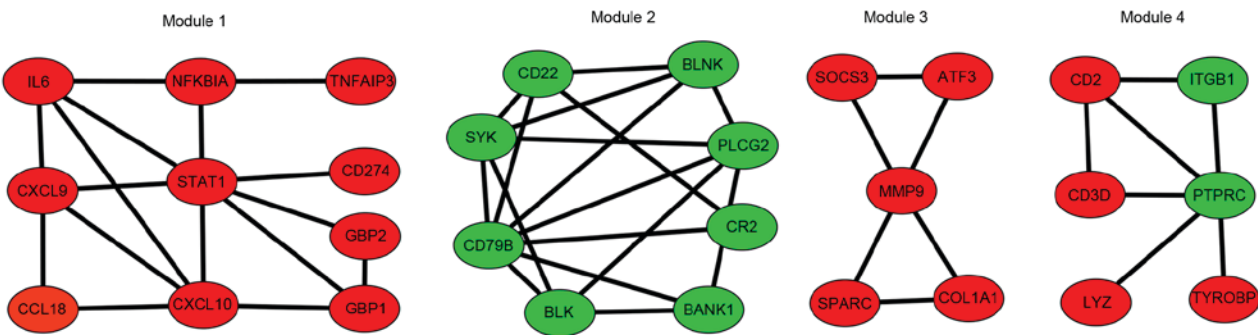


Figure 4. Co-expression modules of the DEGs. Red nodes represent upregulated genes and green nodes represent downregulated genes.

DEGs in module 2 were most enriched in the B cell receptor signaling pathway ( $P=8.06 \times 10^{-11}$ ). Furthermore, in module 2, B-cell linker protein (BLNK) and phospholipase C  $\gamma 2$  (PLCG2) were the two genes which were listed in the top 10 of all downregulated genes. The DEGs with the high degree in module 2, were CD79B (degree, 7), PLCG2 (degree, 5) and SYK (degree, 5).

In module 3, the most enriched KEGG pathway for the DEGs was the tumor necrosis factor signaling pathway ( $P=0.045025$ ), which was associated with suppressor of cytokine signaling 3 and MMP9 genes. MMP9 was one of the most significant upregulated genes. In module 4, the most enriched GO term was leukocyte differentiation ( $P=3.27 \times 10^{-6}$ ), and the enriched KEGG pathways included cell adhesion molecules

Table III. Top 30 genes with high node degrees in protein-protein interaction network.

Node	Degree	Log <sub>2</sub> FC	P-value
IL6	43	6.104794	1.70x10 <sup>-16</sup>
PTPRC	34	-2.90851	5.81x10 <sup>-8</sup>
STAT1	34	3.312368	2.32x10 <sup>-11</sup>
MMP9	30	6.104794	1.70x10 <sup>-16</sup>
SYK	27	-3.26862	1.42x10 <sup>-13</sup>
CCR7	25	3.468633	4.71x10 <sup>-9</sup>
FN1	23	2.724853	4.07x10 <sup>-5</sup>
TIMP1	22	3.760919	1.10x10 <sup>-13</sup>
ITGB1	21	-2.51558	1.79x10 <sup>-7</sup>
CLU	19	3.648265	6.08x10 <sup>-10</sup>
FOXO1	19	-2.75234	3.69x10 <sup>-11</sup>
CD2	17	3.81545	3.05x10 <sup>-11</sup>
CD44	17	2.723706	1.01x10 <sup>-5</sup>
CDKN1A	17	2.675301	5.78x10 <sup>-11</sup>
IL2RB	17	2.508374	7.95x10 <sup>-9</sup>
IL7R	17	2.992393	2.53x10 <sup>-8</sup>
C3	16	2.959825	3.96x10 <sup>-11</sup>
CTLA4	16	2.526753	6.14x10 <sup>-8</sup>
GZMB	16	3.199213	1.98x10 <sup>-6</sup>
ATF3	15	3.309377	6.60x10 <sup>-9</sup>
CTSB	15	4.85476	1.31x10 <sup>-13</sup>
CXCL10	15	4.932454	2.34x10 <sup>-9</sup>
CXCL13	15	3.913023	3.35x10 <sup>-9</sup>
JAK1	15	-2.71728	3.14x10 <sup>-10</sup>
PIK3CG	15	-3.42779	1.20x10 <sup>-8</sup>
PLCG2	14	-3.00981	6.24x10 <sup>-15</sup>
PRKCB	14	-3.14768	1.48x10 <sup>-7</sup>
CCL5	13	5.047877	9.47x10 <sup>-14</sup>
CD79B	13	-3.73922	1.95x10 <sup>-13</sup>
CXCL9	13	5.293566	1.59x10 <sup>-9</sup>

Negative log<sub>2</sub>FC values correspond to downregulated genes and positive log<sub>2</sub>FC values correspond to upregulated genes. FC, fold change.

and primary immunodeficiency. PTPRC had the highest degree (degree, 5) in module 4.

## Discussion

cHL is a type of malignancy that originates from B cells. The pathogenic mechanisms of cHL remain unclear. In the present study, the molecular mechanisms of cHL were investigated using bioinformatics with aim to provide data that could lead to the development of useful therapeutics. A total of 450 DEGs were identified by comparing expression in cHL and normal B cells, among which 216 were upregulated and 234 were downregulated. The GO-BP that was most enriched among all DEGs was immune response. The immune response has a vital role in the pathogenesis of HL, and HRS cells apply various mechanisms to escape innate immune responses by secreting various factors and inducing an immunosuppressive phenotype

in T cells and monocytes (18). The upregulated DEGs were predominantly enriched in the KEGG pathways, transcriptional misregulation in cancer, Toll-like receptor signaling and cytokine-cytokine receptor interaction. Downregulated DEGs were most enriched in the KEGG pathway B cell receptor signaling.

B cell receptor signaling has been previously demonstrated to be a major driver of the pathological mechanisms of B-cell malignancies (19). The B cell receptor is a transmembrane protein located on the external surface of B cells and consists of heavy (IgH) and light (IgL) chains linked to a CD79A (Ig $\alpha$ ) and CD79B (Ig $\beta$ ) molecule (20). BLNK and SYK act as crucial downstream components to transduce B cell receptor signals (21). In the current study, SYK, BLNK, CD79B and PLCG2 were identified to be differentially expressed and enriched in the B cell receptor signaling pathway. SYK, BLNK and CD79B were downregulated the present study, which is in line with previously published data (22,23); however, an association between PLCG2 and cHL not been previously reported. Another report suggested that CD79B and BLNK downregulation in HRS cells was mediated by the effect of latent membrane protein 1 (LMP1) (24); however, the production of LMP1 is caused by Epstein-Barr virus infection. SYK was identified as a hub gene in PPI network in the current study, BLNK was the most significantly downregulated genes, and CD79B and PLCG2 had the highest degrees in module 2. Thus, the downregulation of SYK, BLNK, CD79B and PLCG2 potential has a crucial role in the pathogenesis of cHL, through B cell receptor signaling pathway and Epstein-Barr virus infection.

In cHL, defective transcription regulation can lead to a loss of immunoglobulin expression and B cell receptor downregulation, which facilitates evasion of apoptosis by HRS cells (24-26). IL6 and MMP9 were enriched in the transcriptional misregulation in cancer KEGG pathway, therefore we hypothesize that IL6 and MMP9 may have an important role in cHL pathogenesis. HRS cells can create a tumorigenic microenvironment by secreting certain cytokines and chemokines (27); IL6 is considered to be an important cytokines associated with the occurrence of HL (28) and may predict patient prognosis (29). Furthermore, IL6 participates in the pathway of cytokine-cytokine receptor interaction. Notably, the results of the current study indicated that CD30 (also termed TNF receptor superfamily member 8) is also involved in the same pathway. As it has been previously established that expression of CD30 is characteristic in cHL (30), there may be an interactive effect between IL6 and CD30. IL-6 had the highest degree of connectivity in PPI network. Taken together, this suggests that IL-6 participates in the pathogenesis of HL. The another molecule, MMP9, enriched in the transcriptional misregulation in cancer pathway and was a hub protein in the PPI network in the present study, and has also been previously reported to be associated with reduced overall survival in patients with cHL (31) through effects on immune system regulation (32). A previous study has suggested that the patients with aggressive non-Hodgkin lymphoma and MMP9 expression had poor overall survival (33). Therefore, we hypothesize that MMP9 may be a key molecule involved in the development of classic Hodgkin lymphoma.

Table IV. Significant GO-BP terms and KEGG pathway enrichment analysis of genes in co-expression modules.

A, module 1				
ID	Description	Count	P-value	Genes
GO:0034097	Response to cytokine	10	3.21x10 <sup>-12</sup>	IL6, STAT1, TNFAIP3, CCL18, NFKBIA, GBP2, CXCL9, GBP1, CXCL10, CD274
GO:0071345	Cellular response to cytokine stimulus	9	1.88x10 <sup>-10</sup>	IL6, STAT1, CCL18, NFKBIA, GBP1, GBP2, CXCL9, CXCL10, TNFAIP3
GO:0006955	Immune response	10	1.43x10 <sup>-9</sup>	IL6, STAT1, TNFAIP, CCL18, NFKBIA, GBP2, CXCL9, GBP1, CXCL10, CD274
hsa04062	Chemokine signaling pathway	5	1.65x10 <sup>-5</sup>	STAT1, CXCL9, CXCL10, GBP2, CCL18
hsa04668	TNF signaling pathway	4	1.15x10 <sup>-4</sup>	IL6, CXCL10, TNFAIP3, GBP2
B, module 2				
ID	Description	Count	P-value	Genes
GO:0042113	B cell activation	5	1.55x10 <sup>-6</sup>	SYK, BLNK, PLCG2, BANK1, CR2
GO:0050853	B cell receptor signaling pathway	4	1.86x10 <sup>-6</sup>	SYK, CD79B, PLCG2, BLK
GO:0002429	Immune response-activating cell surface receptor signaling pathway	5	6.63x10 <sup>-6</sup>	SYK, BLK, PLCG2, CD79B, CR2
hsa04662	B cell receptor signaling pathway	6	8.06x10 <sup>-11</sup>	SYK, BLNK, PLCG2, CD22, CD79B, CR2
hsa05169	Epstein-Barr virus infection	3	0.007086	SYK, PLCG2, CR2
C, module 3				
ID	Description	Count	P-value	Genes
GO:0031667	Response to nutrient levels	4	3.93x10 <sup>-4</sup>	SOCS3, ATF3, COL1A1, SPARC
hsa04668	TNF signaling pathway	2	0.045025	SOCS3, MMP9
D, module 4				
ID	Description	Count	P-value	Genes
GO:0002521	Leukocyte differentiation	5	3.27x10 <sup>-6</sup>	PTPRC, ITGB1, CD2, CD3D, TYROBP
GO:0030097	Hemopoiesis	5	2.00x10 <sup>-5</sup>	PTPRC, ITGB1, CD2, CD3D, TYROBP
hsa04514	Cell adhesion molecules	3	0.004048	PTPRC, ITGB1, CD2,

GO-BP, gene ontology-biological process; KEGG, Kyoto encyclopedia of genes and genomes.

STAT1 has interconnectivity with IL6 and CD274/programmed death-ligand 1 (PD-L1) in the GO term of response to cytokine and immune response. It has been well

documented that tumor cells expressing CD274/PD-L1 on their surface and using the programmed cell death 1 (PD-1) pathway to evade antitumor immune response, so CD274/PD-L1

has been regarded as an independent prognostic factor for cHL (34-36). Furthermore, it was reported that PD-1 blockade with nivolumab has been successful for cHL treatment (37). module 1 in the current study suggests that STAT1 has a cross communication with CD274/PD-L1 and IL6 in pathogenesis of HL. Additionally, Butterbach *et al* (38) demonstrated that STAT1 was significantly associated with lymphoma risk (39), indicating that STAT1 may have a key role in pathogenesis of cHL. PTPRC is also known as CD45, which was a down-regulated DEGs that represent the highest degree in module 4 in the current study. PTPRC/CD45 has become a cell marker criteria for the diagnosis of cHL (39,40), this is consistent with the findings of the current study and further confirmed the reliability the results.

In conclusion, DEGs and pathways that may be involved in pathogenesis of cHL were identified by a comprehensive bioinformatics analysis in the present study. The findings may be useful for understanding the underlying molecular mechanisms of cHL. DEGS including IL6, STAT1, MMP9, SYK, BLNK, PLCG2 and CD79B, and the pathways of B cell receptor signaling, Epstein-Barr virus infection and transcriptional misregulation in cancer may potentially to be used as diagnostic and therapeutic targets in cHL. However, the present study had limited of sample size, which may give rise to a high false-positive rate. Additionally, this study lacked experimental verification, therefore, the findings require additional genetic and experimental studies to validate the results.

### Acknowledgements

This work was supported by Technological Innovation and Talents Cultivation Project of Xinjiang Province for Outstanding Youth (grant no. 2014721044).

### References

- Fend F: Classical Hodgkin lymphoma and its differential diagnoses. *Diagn Histopathol* 21: 400-407, 2015.
- Salati M, Cesaretti M, Macchia M, Mistiri ME and Federico M: Epidemiological overview of Hodgkin lymphoma across the mediterranean basin. *Mediterr J Hematol Infect Dis* 6: e2014048, 2014.
- Siegel RL, Miller KD and Jemal A: Cancer statistics, 2016. *CA Cancer J Clin* 66: 7-30, 2016.
- Steidl C, Connors JM and Gascoyne RD: Molecular pathogenesis of Hodgkin's lymphoma: Increasing evidence of the importance of the microenvironment. *J Clin Oncol* 29: 1812-1826, 2011.
- Makita S, Maeshima AM, Taniguchi H, Kitahara H, Fukuhara S, Munakata W, Suzuki T, Maruyama D, Kobayashi Y and Tobinai K: Classical Hodgkin lymphoma primary refractory to brentuximab vedotin, with transformation to CD30-positive diffuse large B-cell lymphoma. *Int J Hematol* 104: 396-399, 2016.
- Brune V, Tiacci E, Pfeil I, Döring C, Eckerle S, van Noesel CJ, Klapper W, Falini B, von Heydebreck A, Metzler D, *et al*: Origin and pathogenesis of nodular lymphocyte-predominant Hodgkin lymphoma as revealed by global gene expression analysis. *J Exp Med* 205: 2251-2268, 2008.
- Tiacci E, Döring C, Brune V, van Noesel CJ, Klapper W, Mechttersheimer G, Falini B, Küppers R and Hansmann ML: Analyzing primary Hodgkin and reed-sternberg cells to capture the molecular and cellular pathogenesis of classical Hodgkin lymphoma. *Blood* 120: 4609-4620, 2012.
- Matsuki E and Younes A: Lymphomagenesis in Hodgkin lymphoma. *Semin Cancer Biol* 34: 14-21, 2015.
- De Silva NS and Klein U: Dynamics of B cells in germinal centres. *Nat Rev Immunol* 15: 137-148, 2015.
- Wilson CL and Miller CJ: Simpleaffy: A BioConductor package for Affymetrix quality control and data analysis. *Bioinformatics* 21: 3683-3685, 2005.
- Kauffmann A, Gentleman R and Huber W: arrayQualityMetrics - a bioconductor package for quality assessment of microarray data. *Bioinformatics* 25: 415-416, 2009.
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U and Speed TP: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4: 249-264, 2003.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK: limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43: e47, 2015.
- Kanehisa M and Goto S: KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28: 27-30, 2000.
- von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P and Snel B: STRING: A database of predicted functional associations between proteins. *Nucleic Acids Res* 31: 258-261, 2003.
- Doerks T, Copley RR, Schultz J, Ponting CP and Bork P: Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res* 12: 47-56, 2002.
- Nepusz T, Yu H and Paccanaro A: Detecting overlapping protein complexes in protein-protein interaction networks. *Nat Methods* 9: 471-472, 2012.
- Aldinucci D, Celegato M and Casagrande N: Microenvironmental interactions in classical Hodgkin lymphoma and their role in promoting tumor growth, immune escape and drug resistance. *Cancer Lett* 380: 243-252, 2016.
- Bojarczuk K, Bobrowicz M, Dwojak M, Miazek N, Zapala P, Bunes A, Siernicka M, Rozanska M and Winiarska M: B-cell receptor signaling in the pathogenesis of lymphoid malignancies. *Blood Cells Mol Dis* 55: 255-265, 2015.
- Young RM, Shaffer AL, III, Phelan JD and Staudt LM: B-cell receptor signaling in diffuse large B-cell lymphoma. *Semin Hematol* 52: 77-85, 2015.
- Seda V and Mraz M: B-cell receptor signalling and its crosstalk with other pathways in normal and malignant cells. *Eur J Haematol* 94: 193-205, 2015.
- Marafioti T, Pozzobon M, Hansmann ML, Delsol G, Pileri SA and Mason DY: Expression of intracellular signaling molecules in classical and lymphocyte predominance Hodgkin disease. *Blood* 103: 188-193, 2004.
- Ushmorov A, Leithäuser F, Sakk O, Weinhäusel A, Popov SW, Möller P and Wirth T: Epigenetic processes play a major role in B-cell-specific gene silencing in classical Hodgkin lymphoma. *Blood* 107: 2493-2500, 2006.
- Vockerodt M, Morgan SL, Kuo M, Wei W, Chukwuma MB, Arrand JR, Kube D, Gordon J, Young LS, Woodman CB and Murray PG: The Epstein-Barr virus oncoprotein, latent membrane protein-1, reprograms germinal centre B cells towards a Hodgkin's reed-sternberg-like phenotype. *J Pathol* 216: 83-92, 2008.
- Marafioti T, Hummel M, Foss HD, Laumen H, Korbjuhn P, Anagnostopoulos I, Lammert H, Demel G, Theil J, Wirth T and Stein H: Hodgkin and reed-sternberg cells represent an expansion of a single clone originating from a germinal center B-cell with functional immunoglobulin gene rearrangements but defective immunoglobulin transcription. *Blood* 95: 1443-1450, 2000.
- Küppers R: Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer* 5: 251-262, 2005.
- Maggio E, van den Berg A, Diepstra A, Kluiver J, Visser L and Poppema S: Chemokines, cytokines and their receptors in Hodgkin's lymphoma cell lines and tissues. *Ann Oncol* 13 (Suppl 1): S52-S56, 2002.
- de Winde CM, Veenbergen S, Young KH, Xu-Monette ZY, Wang XX, Xia Y, Jabbar KJ, van den Brand M, van der Schaaf A, Elfrink S, *et al*: Tetraspanin CD37 protects against the development of B cell lymphoma. *J Clin Invest* 126: 653-666, 2016.
- Ghesquière H, Maurer MJ, Casasnovas O, Ansell SM, Larrabee BR, Lech-Maranda E, Novak AJ, Borrel AL, Slager SL, Brice P, *et al*: Cytokine gene polymorphisms and progression-free survival in classical Hodgkin lymphoma by EBV status: Results from two independent cohorts. *Cytokine* 64: 523-531, 2013.
- Alperovich A and Younes A: Targeting CD30 using brentuximab vedotin in the treatment of Hodgkin lymphoma. *Cancer J* 22: 23-26, 2016.
- Campos AH, Vassallo J and Soares FA: Matrix metalloproteinase-9 expression by Hodgkin-reed-sternberg cells is associated with reduced overall survival in young adult patients with classical Hodgkin lymphoma. *PLoS One* 8: e74793, 2013.



32. Wang BQ, Zhang CM, Gao W, Wang XF, Zhang HL and Yang PC: Cancer-derived matrix metalloproteinase-9 contributes to tumor tolerance. *J Cancer Res Clin Oncol* 137: 1525-1533, 2011.
33. Sakata K, Satoh M, Someya M, Asanuma H, Nagakura H, Oouchi A, Nakata K, Kogawa K, Koito K, Hareyama M and Himi T: Expression of matrix metalloproteinase 9 is a prognostic factor in patients with non-Hodgkin lymphoma. *Cancer* 100: 356-365, 2004.
34. Chen BJ, Chapuy B, Ouyang J, Sun HH, Roemer MG, Xu ML, Yu H, Fletcher CD, Freeman GJ, Shipp MA and Rodig SJ: PD-L1 expression is characteristic of a subset of aggressive B-cell lymphomas and virus-associated malignancies. *Clin Cancer Res* 19: 3462-3473, 2013.
35. Koh YW, Jeon YK, Yoon DH, Suh C and Huh J: Programmed death 1 expression in the peritumoral microenvironment is associated with a poorer prognosis in classical Hodgkin lymphoma. *Tumour Biol* 37: 7507-7514, 2016.
36. Paydas S, Bağır E, Seydaoglu G, Ercolak V and Ergin M: Programmed death-1 (PD-1), programmed death-ligand 1 (PD-L1), and EBV-encoded RNA (EBER) expression in Hodgkin lymphoma. *Ann Hematol* 94: 1545-1552, 2015.
37. Ansell SM, Lesokhin AM, Borrello I, Halwani A, Scott EC, Gutierrez M, Schuster SJ, Millenson MM, Cattray D, Freeman GJ, *et al*: PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N Engl J Med* 372: 311-319, 2015.
38. Butterbach K, Beckmann L, de Sanjosé S, Benavente Y, Becker N, Foretova L, Maynadie M, Cocco P, Staines A, Boffetta P, *et al*: Association of JAK-STAT pathway related genes with lymphoma risk: Results of a European case-control study (EpiLymph). *Br J Haematol* 153: 318-333, 2011.
39. O'Malley DP, Fedoriw Y and Weiss LM: Distinguishing classical Hodgkin lymphoma, gray zone lymphoma, and large B-cell lymphoma: A proposed scoring system. *Appl Immunohistochem Mol Morphol* 24: 535-540, 2016.
40. Weiss LM: Atypical Phenotypes in classical Hodgkin lymphoma. *Surg Pathol Clin* 6: 729-742, 2013.