Am J Transl Res 2019;11(7):4382-4396
www.ajtr.org /ISSN:1943-8141/AJTR0096450

Original Article

Integrated analysis of 10 lymphoma datasets identifies E2F8 as a key regulator in Burkitt's lymphoma and mantle cell lymphoma

An-Ping Yang[1*], Leyna G Liu[2*], Min-Min Chen[1*], Fang Liu[1], Hua You[3], Lian Liu[1], Hua Yang[1], Yang Xun[1], Jing Liu[1], Rui-Xue Wang[1], David D Brand[4], Dahai Liu[1], Song Guo Zheng[5], Wen-Xing Li[6,7]

[1]School of Stomatology and Medicine, Foshan University, Foshan 528000, Guangdong, China; [2]Portola High School, 1001 Cadence, Irvine 92618, CA, USA; [3]Affiliated Cancer Hospital and Institute of Guangzhou Medical University, Guangzhou 510095, Guangdong, China; [4]Research Service, Memphis VA Medical Center, Memphis 38104, TN, USA; [5]Department of Internal Medicine, Ohio State University College of Medicine and Wexner Medical Center, Columbus 43210, USA; [6]Key Laboratory of Animal Models and Human Disease Mechanisms, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, Yunnan, China; [7]Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming 650204, Yunnan, China. [*]Equal contributors.

Received May 4, 2019; Accepted June 18, 2019; Epub July 15, 2019; Published July 30, 2019

Abstract: Burkitt's lymphoma (BURK), diffuse large B-cell lymphoma (DLBCL) and mantle cell lymphoma (MCL) are three main types of B-cell lymphomas. This study aimed to compare the differences of affected biological functions and pathways, as well as to explore the possible regulatory mechanisms and the potential therapeutic targets in BURK, DLBCL and MCL. We performed an integrated analysis of 10 lymphoma datasets including 352 BURK patients, 880 DLBCL patients, 216 MCL patients, and 33 controls. Our results showed that signaling pathways, amino acid metabolism and several lipid metabolism pathways varies considerably among these three types of lymphoma. Furthermore, we identified several key transcription factors (TFs) and their target genes that may promote these diseases by influencing multiple carcinogenic pathways. Among these TFs, we reported first that E2F8 displayed the most significant effects in BURK and MCL. Our results demonstrate that over-expression of E2F8 activates target genes that may promote cell cycle, mitosis, immune and other cancer related functions in BURK and MCL. Therefore, we suggest that E2F8 could be used as a biomarker and potential therapeutic target for BURK and MCL. These findings would be helpful in the study of pathogenesis, and drug discovery and also in the prognosis of B cell lymphomas.

Keywords: Burkitt's lymphoma, diffuse large B-cell lymphoma, mantle cell lymphoma, transcriptional regulation network, E2F8

Introduction

Non-Hodgkin's lymphomas (NHLs) are broadly classified into two major groups: B-cell lymphomas and T-cell/NK-cell lymphomas. B-cell lymphomas develop from abnormal B lymphocytes and account for 85-90% of all NHLs [1, 2]. Burkitt's lymphoma (BURK), diffuse large B-cell lymphoma (DLBCL) and mantle cell lymphoma (MCL) are three main types of B-cell lymphomas. BURK is often associated with Epstein-Barr Virus (EBV) infection, being endemic in some parts of Africa but encountered only sporadically in the rest of the world. BURK is largely diagnosed in children. In Cameroon, the country with the highest incidence rate worldwide, the average incidence in the northwest province is 2.58 cases per 100,000 per year in children aged 15 or younger [3]. DLBCL accounts for approximately 33% of all newly diagnosed cases of NHL worldwide. It is most commonly diagnosed during the sixth decade, and is slightly more common in males [4]. Epidemiology on MCL is primarily available for the U.S. and Europe, where the average incidence is approximately 0.5 cases per 100,000 person-years. The male-to-female ratio is around 2.5:1, and the average age of onset is 65-70 years [5, 6].

Immunosuppression is the main risk factor for NHLs [1, 7]. Other putative environmental and lifestyle-related risk factors have also been proposed, including current cigarette smoking [8] and heavy alcohol intake [9].

In addition to the traditional risk factors, recent studies have identified several NHL risk genes. A previous study found that the co-expression of MYC/BCL6 may affect the survival in patients with DLBCL [10]. Another study suggested that subset-specific B-cell-associated gene signatures were better suited to identifications than the traditional classifications (biologic, morphologic, and clinical heterogeneity) in clinical diagnosis and prognosis of DLBCL [11]. In addition, the top discriminatory genes expressed in formalin-fixed, paraffin-embedded (FFPE) tissues of DLBCL patients also showed high statistical difference and contained novel biology with potential therapeutic insights [12]. Furthermore, gene expression signatures were performed well in distinguishing BURK and DLBCL [13, 14]. Still another recent study identified that the deregulation of CUL4A, ING1 and MCPH1 affects the proliferation; and DNA damage response pathways in MCL by integrating gene expression and copy number changes [15]. Palomero *et al.* showed that over-expression of SOX11 activated PDGFA and promoted tumor microvascular angiogenesis in MCL cell lines, suggesting that the SOX11-PDGFA axis may serve as a potential therapeutic target of MCL [16]. Saba *et al.* demonstrated that B-cell receptor activation and canonical NF-kB signaling may promote tumor proliferation and affect overall survival in MCL patients [17].

Reconstruction of transcriptional regulation networks (TRNs) has been a useful approach that provides a global view of gene expression profiles and their regulations, and it also provides new insights in tumorigenesis, pathways and regulations [18, 19], as well as differentiation of embryonic stem cells [20]. Our previous study has identified several tissue-specific transcription factors (TFs) that correlated with breast cancer by analyzing tissue-specific TRNs [21]. However, to our knowledge, no other studies have compared multiple types of NHLs from the perspective of TRNs. Therefore, in the present study, we integrated 10 lymphoma microarray datasets including BURK, DLBCL and MCL in order to explore common differential changes in global gene expression, pathway and TRN analyses, as well as to identify key regulators and their targets among these three types of lymphoma.

Materials and methods

Lymphoma datasets collection

Transcriptome datasets of BURK, DLBCL and MCL were downloaded from NCBI-GEO (https://www.ncbi.nlm.nih.gov/geo/) in Jun, 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. Based on the above criteria, we chose 10 datasets for our integrated analysis. The details of these datasets are illustrated in **Table 1**. In total, we collected samples from 352 BURK patients, 880 DLBCL patients, 216 MCL patients, and 33 controls.

Data preprocessing and differential expression analysis

R statistical software v3.3.3 (https://www.r-project.org/) was used to perform data preprocessing and differential gene expression gene analysies. Because these datasets contain different microarray platforms and different types of lymphoma, they were first grouped into 14 batches according to study, category (different types of lymphoma and controls), and platform. Each batch contained only one study, one type of lymphoma or control, and one platform (**Table 2**). We used the Robust Multichip Average (RMA) algorithm in oligo package [22] to normalize the raw expression data and generate normalized gene expression intensity. Gene annotation, integration and renormalization of the 14 batches was carried out using a custom designed Python code. We removed probes with no gene annotation or those that matched multiple gene symbols. Next, we calculated the average expression value of multiple probe IDs matching an official gene symbol and took this value to represent the expression intensity of the corresponding gene symbol. The renormalization method and scripts are detailed in our previous publications [23, 24]. The distributions of RMA-processed and global-renormalized gene expression values across all studies are shown in Figures S1 and S2. After global

Table 1. Details of lymphoma datasets

GEO ID	Contributor	Title	Samples
Burkitt's lymphoma			
GSE4475	Hummel M, 2006	A Biologic Definition of Burkitt's Lymphoma from Transcriptional and Genomic Profiling	221 cases
GSE69053	Sha C, 2015	Transferring genomics to the clinic: distinguishing Burkitt and Diffuse large B-cell lymphoma	131 cases
Diffuse large B-cell lymphoma			
GSE31312	Li Y, 2011	Development and application of a new immunophenotypic algorithm for molecular subtype classification of Diffuse Large B-Cell Lymphoma (DLBCL): Report from an International DLBCL Rituximab-CHOP Consortium Program Study	498 cases
GSE56315	Bødker JS, 2014	Diffuse Large B-Cell Lymphoma Classification System That Associates Normal B-Cell Subset Phenotypes With Prognosis	89 cases 33 controls
GSE64555	Linton K, 2014	Microarray gene expression analysis of FFPE tissue permits molecular classification in diffuse large B-cell lymphoma	40 cases
GSE69053	Sha C, 2015	Transferring genomics to the clinic: distinguishing Burkitt and Diffuse large B-cell lymphoma	212 cases
GSE86613	Bødker JS, 2016	Expression data from diagnostic samples of diffuse large B-cell lymphomas (DLBCL), follicular lymphoma (FL) and primary and relapsed transformed FL	41 cases
Mantle cell lymphoma			
GSE21452	Staudt LM, 2010	Integrated genomic profiling in mantle cell lymphoma	64 cases
GSE36000	Jares P, 2012	Mantle Cell Lymphoma	38 cases
GSE70910	Liu D, 2015	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	55 cases
GSE93291	Staudt LM, 2017	A new molecular assay for the proliferation signature in mantle cell lymphoma applicable to formalin-fixed paraffin-embedded biopsies	59 cases

Table 2. Grouping the lymphoma datasets for data integration and global renormalization

GEO ID	Category	Platform	Batch
GSE4475	Burkitt's lymphoma	GPL96	1
GSE21452	Mantle cell lymphoma	GPL570	2
GSE31312	Diffuse large B-cell lymphoma	GPL570	3
GSE36000	Mantle cell lymphoma	GPL570	4
GSE56315	Diffuse large B-cell lymphoma	GPL570	5
GSE64555	Diffuse large B-cell lymphoma	GPL570	6
GSE69053	Burkitt's lymphoma	GPL8432	7
GSE69053	Diffuse large B-cell lymphoma	GPL8432	8
GSE69053	Burkitt's lymphoma	GPL14951	9
GSE69053	Diffuse large B-cell lymphoma	GPL14951	10
GSE70910	Mantle cell lymphoma	GPL570	11
GSE86613	Diffuse large B-cell lymphoma	GPL570	12
GSE93291	Mantle cell lymphoma	GPL570	13
GSE56315	Control	GPL570	14

expression was renormalized, the distribution of gene expression values across all studies had a consistent range. The empirical Bayesian algorithm in the limma package [25] 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) ≥ log2(1.5) or ≤ -log2(1.5), respectively. A *P* value ≤ 0.05 was considered as significant.

Gene set enrichment analysis

We used javaGSEA Desktop Application v2.2.4 to perform gene set enrichment analysis (GSEA) of the three types of lymphoma [26]. KEGG gene sets v6.0 (including 186 gene sets) was chosen to perform KEGG pathway enrichment analysis in BURK, DLBCL and MCL. The analysis parameters were set as follows: gene sets containing fewer than 15 genes or more than 500 genes were excluded, a phenotype label was set as case vs. control, and the t-statistic mean of the genes was computed in each KEGG pathway using a permutation test with 1000 replications. The up-regulated pathways were defined by a normalized enrichment score (NES) > 0 and the down-regulated pathways were defined by an NES < 0. Pathways with a *P* value ≤ 0.05 were chosen as significantly enriched.

Transcriptional regulation network analysis

We downloaded 1,691 human transcription factors (TFs) from the AnimalTFDB 2.0 web server (http://bioinfo.life.hust.edu.cn/AnimalTFDB/) [27] and mapped 915 TFs to our lym-

Table 3. Differentially expressed genes in three types of lymphoma

Lymphoma	Cases/Comtrols	Up-regulated	Down-regulated
Burkitt's lymphoma	352/33	216	205
Diffuse large B-cell lymphoma	880/33	198	89
Mantle cell lymphoma	216/33	230	67

phoma datasets. GENIE3 (Gene Network Inference with Ensemble of trees) [28] was used to reconstruct transcriptional regulatory networks for BURK, DLBCL and MCL. We used the gene expression matrix of the three types of lymphoma and the 915 mapped transcription factors as the input data and ran GENIE3 with its default parameters. Next, we extracted the top 10,000 interactions from the original output and used Cytoscape v3.4.0 to visualize the output results. The NetworkAnalyzer tool in Cytoscape was used to analyze the three resultant networks.

Tissue-specific network analysis

GIANT web server (http://giant.princeton.edu/) [29] was implemented to construct genome-scale integrated analysis of gene networks in BURK, DLBCL and MCL. We filtered the TFs with nodes containing greater than 100 in transcriptional regulatory networks of the three types of lymphoma, and then used these TFs in each type of lymphoma as input parameters to perform gene network analysis. The tissue option chosen for the GIANT web server was "lymph node". The server generated a network of query genes and target genes, as well as performing GO Biological Process enrichment analysis of the genes in the network. GO Biological Processes with a P value ≤ 0.05 was considered as significantly enriched.

Identify key regulators and target genes

We filtered key transcription factors based on the relationship confidence in the GIANT network in BURK, DLBCL and MCL, respectively. A Venn diagram was used to screen the differentially expressed target genes in the GIANT networks in the three types of lymphoma. We used univariate linear model (the "lm" function in R) to calculate the correlation between the filtered key transcription factors and differentially expressed target genes in each type of lymphoma. The absolute value of correlation coefficient ≥ 0.5 and a P value ≤ 0.05 was considered as significant. We used the "plot" function in R to show the log2 transformed fold changes and the correlation coefficients of each paired transcription factors and target genes in the three types of lymphoma.

Results

Differentially expressed genes in three types of lymphoma

In total, we found 11,511 genes in the integrated lymphoma datasets. Differentially expressed genes in the three types of lymphoma were illustrated in **Table 3**. We obtained 216 up-regulated and 205 down-regulated genes, 198 up-regulated and 89 down-regulated genes, as well as 230 up-regulated and 67 down-regulated genes in BURK, DLBCL, and MCL, respectively. The top 20 differentially-expressed genes in three types of lymphoma were shown in Table S1. In addition, we noted that 178 common genes were deregulated in the three types of lymphoma (Figure S3A).

Commonly and differentially deregulated KEGG pathways

We screened the pathways that were significantly enriched in at least one type of lymphoma. The results were shown in **Figure 1**. There were 13 commonly enriched pathways (10 up-regulated and 3 down-regulated) in BURK, DLBCL, and MCL. These pathways include cell cycle, amino sugar and nucleotide sugar metabolism, PPAR signaling pathways, pyrimidine metabolism, Fc gamma R-mediated phagocytosis, N-glycan biosynthesis and other cancer-related pathways. Furthermore, we observed tyrosine metabolism, GnRH signaling pathway, vascular smooth muscle contraction, neuroactive ligand-receptor interaction, fatty acid and other lipid metabolism were common down-regulated in BURK and DLBCL, whereas these pathways showed no change in MCL (**Figures 1** and S4). In addition, the JAK-STAT signaling pathway, adherents junction and chemokine signaling pathways were commonly deregulated in DLBCL and MCL, whereas these pathways did not reach significant difference in BURK (**Figures 1** and S4). Furthermore, there were

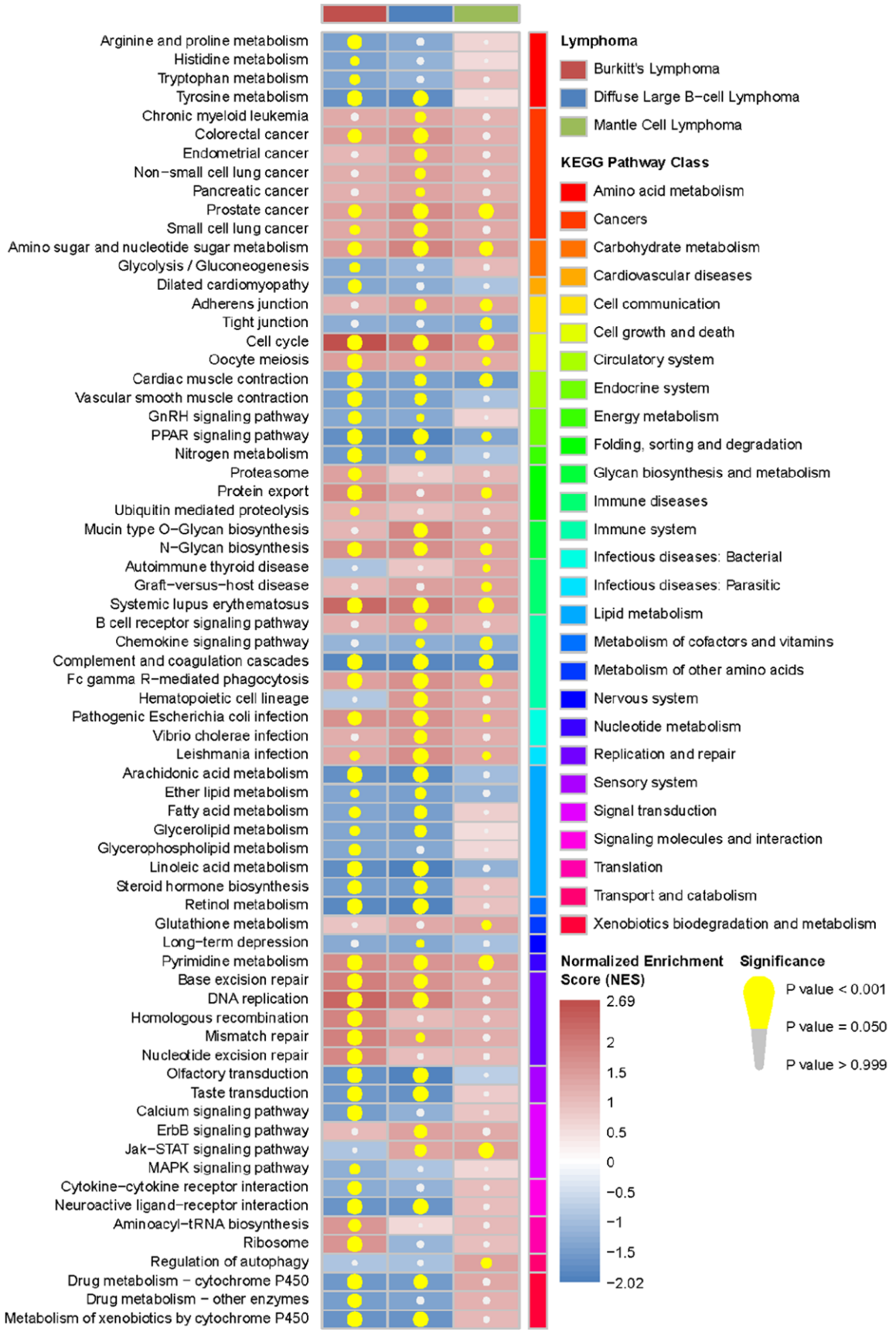

Arginine and proline metabolism
Histidine metabolism
Tryptophan metabolism
Tyrosine metabolism
Chronic myeloid leukemia
Colorectal cancer
Endometrial cancer
Non−small cell lung cancer
Pancreatic cancer
Prostate cancer
Small cell lung cancer
Amino sugar and nucleotide sugar metabolism
Glycolysis / Gluconeogenesis
Dilated cardiomyopathy
Adherens junction
Tight junction
Cell cycle
Oocyte meiosis
Cardiac muscle contraction
Vascular smooth muscle contraction
GnRH signaling pathway
PPAR signaling pathway
Nitrogen metabolism
Proteasome
Protein export
Ubiquitin mediated proteolysis
Mucin type O−Glycan biosynthesis
N−Glycan biosynthesis
Autoimmune thyroid disease
Graft−versus−host disease
Systemic lupus erythematosus
B cell receptor signaling pathway
Chemokine signaling pathway
Complement and coagulation cascades
Fc gamma R−mediated phagocytosis
Hematopoietic cell lineage
Pathogenic Escherichia coli infection
Vibrio cholerae infection
Leishmania infection
Arachidonic acid metabolism
Ether lipid metabolism
Fatty acid metabolism
Glycerolipid metabolism
Glycerophospholipid metabolism
Linoleic acid metabolism
Steroid hormone biosynthesis
Retinol metabolism
Glutathione metabolism
Long−term depression
Pyrimidine metabolism
Base excision repair
DNA replication
Homologous recombination
Mismatch repair
Nucleotide excision repair
Olfactory transduction
Taste transduction
Calcium signaling pathway
ErbB signaling pathway
Jak−STAT signaling pathway
MAPK signaling pathway
Cytokine−cytokine receptor interaction
Neuroactive ligand−receptor interaction
Aminoacyl−tRNA biosynthesis
Ribosome
Regulation of autophagy
Drug metabolism − cytochrome P450
Drug metabolism − other enzymes
Metabolism of xenobiotics by cytochrome P450
Lymphoma
Burkitt's Lymphoma
Diffuse Large B−cell Lymphoma
Mantle Cell Lymphoma
KEGG Pathway Class
Amino acid metabolism
Cancers
Carbohydrate metabolism
Cardiovascular diseases
Cell communication
Cell growth and death
Circulatory system
Endocrine system
Energy metabolism
Folding, sorting and degradation
Glycan biosynthesis and metabolism
Immune diseases
Immune system
Infectious diseases: Bacterial
Infectious diseases: Parasitic
Lipid metabolism
Metabolism of cofactors and vitamins
Metabolism of other amino acids
Nervous system
Nucleotide metabolism
Replication and repair
Sensory system
Signal transduction
Signaling molecules and interaction
Translation
Transport and catabolism
Xenobiotics biodegradation and metabolism
Normalized Enrichment Score (NES)
2.69
2
1.5
1
0.5
0
−0.5
−1
−1.5
−2.02
Significance
P value < 0.001
P value = 0.050
P value > 0.999

Figure 1. GSEA results of enriched KEGG pathways in Burkitt's lymphoma, diffuse large B-cell lymphoma and mantle cell lymphoma.

16, 10 and 5 pathways that were only significantly enriched in BURK, DLBCL, and MCL, respectively. These results suggest a relatively large difference in these three types of lymphoma.

Expression profiles of transcription factors and transcriptional regulatory networks

There were 29, 14 and 16 differentially expressed transcription factors in BURK, DLBCL, and MCL, respectively (**Figure 2A**). Among them, 11 deregulated transcription factors were shared by all the three types of lymphoma (Figure S3B). Transcriptional regulatory networks (TRNs) of three types of lymphoma were shown in Figures S5, S6, S7. Network properties of the three TRNs were listed in Table S2. We found that MCL TRN had the highest clustering coefficient of 0.120; the clustering coefficient in BURK TRN was 0.049 and in DLBCL it was 0.079. Furthermore, MCL TRN showed the highest connected component of 101, followed by DLBCL TRN (76) and BURK TRN (56). In addition, MCL TRN had the most multi-edge node pairs at 130, followed by DLBCL TRN (89) and BURK TRN (68). We sorted transcription factors for each of the TRNs according to the number of nodes, and then extracted the transcription factors with more than 100 nodes. In total, we got 8, 11 and 24 top transcription factors for BURK, DLBCL, and MCL, respectively (**Figures 2B-D**, S5, S6, S7). We revealed that no top TF was shared by all three types of lymphoma, only E2F8 was shared by BURK and MCL, and CDX2 was shared by BURK and DLBCL (**Figure 2E**). We combined these differentially expressed TFs and top-node TFs in each type of lymphoma and found the intersected TFs in BURK was E2F8, and in MCL were PRDM1, E2F8, VDR and IRF4 (Figure S8). No deregulated top-node TF was found in DLBCL.

GIANT networks and enriched biological functions

Figure 3 shows the GIANT networks of top-node TFs and their biological functions in each of the three types of lymphoma. We listed the top 20 enriched GO Biological Processes in each type of lymphoma (only 5 significantly enriched items in DLBCL). In BURK, only E2F8 showed the highest relationship confidence with target genes. In DLBCL, there were 6 genes (YY1, STAT1, HMGB1, PMS1, RNF138 and YBX1) which showed a relatively high relationship confidence with target genes. Furthermore, WHSC1, E2F8, HMGB3, E2F1 and FOXM1 had the highest relationship confidence with target genes in MCL. The enriched biological functions in BURK were mainly cell division, mitosis, microtubule cytoskeleton organization, nuclear division, chromosome segregation and other cell cycle related functions (**Figure 3A**). In DLBCL, there were 5 enriched DNA replication and transcription related biological processes. In addition, several immune-related biological functions (interleukin-13 production, interleukin-4 production, vitamin D receptor signaling pathway and other functions) and cell cycle related functions were significantly enriched in MCL. Based on the GIANT networks, we therefore speculate that these top-node TFs may play key roles in BURK and MCL.

Key transcription factors and their target genes in three types of lymphoma

We combined target genes in the GIANT network and deregulated genes in BURK, DLBCL and MCL (**Figure 4A-C**). There were 15, 3 and 8 target genes that were deregulated in BURK, DLBCL and MCL, respectively. Based on the above results, we calculated the correlations between the key TFs and deregulated target genes in each type of lymphoma. Interestingly, these target genes were all up-regulated. In BURK, there were significant correlations between E2F8 and 15 downstream target genes (**Figure 4D**). However, the key TFs in DLBCL showed no expression changes and relatively lower correlations with target genes (**Figure 4E**). Furthermore, the key TFs and target genes in MCL presented the highest correlations of all (**Figure 4F**). Among these key TFs, E2F8 was significantly up-regulated, while E2F1 and HMGB3 showed boundary significance.

Possible regulatory mechanisms in three types of lymphoma

Through integration the results of differentially expressed genes, affected pathways and biological functions, transcriptional regulation net-

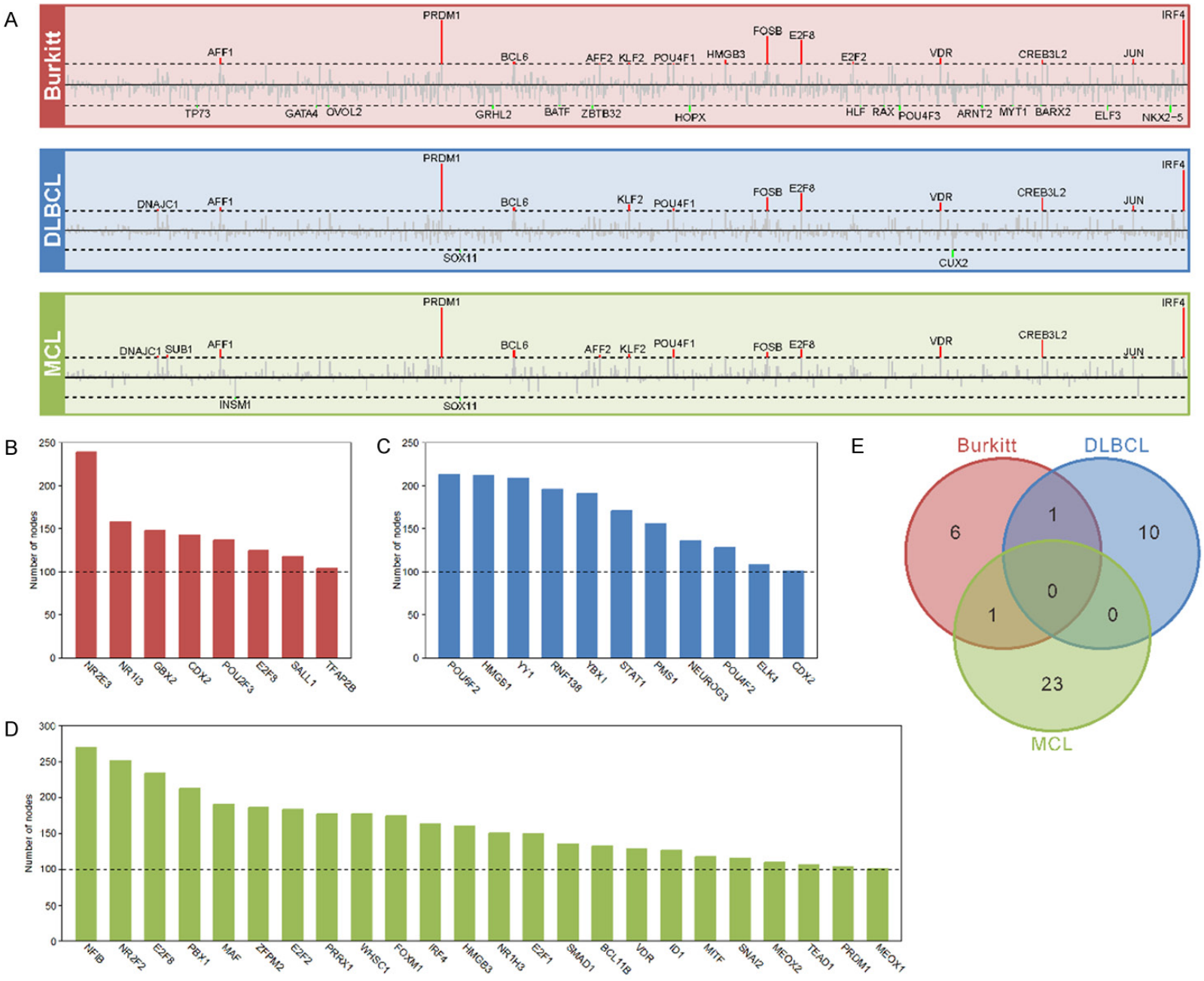

A
Burkitt
PRDM1
IRF4
FOSB
E2F8
AFF1
BCL6
AFF2
KLF2
POU4F1
HMGB3
E2F2
VDR
CREB3L2
JUN
TP73
GATA4
OVOL2
GRHL2
BATF
ZBTB32
HOPX
HLF
RAX
POU4F3
ARNT2
MYT1
BARX2
ELF3
NKX2-5
DLBCL
DNAJC1
SOX11
CUX2
MCL
SUB1
INSM1
B
C
D
Number of nodes
NR2E3
NR1I3
GBX2
CDX2
POU2F3
E2F3
SALL1
TFAP2B
POU5F2
HMGB1
YY1
RNF138
YBX1
STAT1
PMS1
NEUROG3
POU4F2
ELK1
CDX2
NFIB
NR2F2
E2F8
PBX1
MAF
ZFPM2
E2F2
PRRX1
WHSC1
FOXM1
IRF4
HMGB3
NR1H3
E2F1
SMAD1
BCL11B
VDR
ID1
MITF
SNAI2
MEOX2
TEAD1
PRDM1
MEOX1
E
Burkitt
DLBCL
MCL
6
1
10
0
1
0
23

Figure 2. Differentially expressed and multiple connected transcription factors in three types of lymphoma. A. log2 transformed fold change of all transcription factors in three different types of lymphoma. B-D. transcription factors with nodes ≥ 100 in Burkitt's lymphoma, diffuse large B-cell lymphoma and mantle cell lymphoma, respectively. E. Venn diagram of transcription factors with nodes ≥ 100 in three types of lymphoma.

works and TF-target interactions, we propose the possible regulatory mechanisms in each type of lymphoma (**Figure 5**). In BURK, the up-regulated E2F8 activated 15 downstream target genes, and these genes further affected multiple cell cycle related pathways, eventually resulting in cancer. However, in DLBCL, the effect of these transcription factors may not be carcinogenic. Furthermore, in MCL, the highly-expressed E2F8, E2F1 and HMGB3 activated 8 downstream target genes, and influenced cell cycle and immune functions, also resulting in cancer.

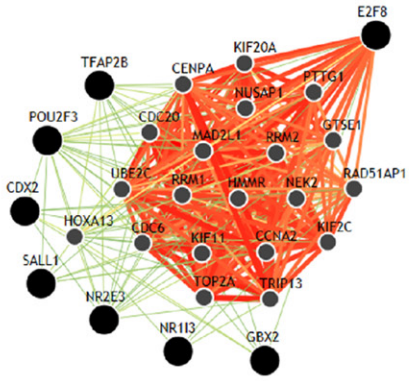
Discussion

In the present study, we observed there were 13 commonly enriched pathways and 31 specific enriched pathways in BURK, DLBCL and MCL. Furthermore, through reconstruction of TRNs and GIANT networks analysis in the three types of lymphoma, we identified several key TFs and their target genes in BURK, DLBCL and MCL, respectively. In particular, E2F8 was up-regulated in all the three types of lymphoma and was also the top TF (nodes ≥ 100) in BURK and MCL. We further found that several E2F8 target genes were over-expressed in BURK and MCL, and that these genes play crucial roles in cell cycle and immune-related functions. However, the top TFs may not cause a carcinogenic effect in DLBCL.

Previous studies have demonstrated some important pathways result in a carcinogenic effect. Multiple tumor-related metabolic pathways (energy metabolism, amino acid metabolism, fatty acid metabolism and choline phospholipid metabolism) were identified in BURK mouse models [30]. Fatty acid oxidation and up-regulated carnitine palmitoyltransferase I (CPTI) have been found to promote cancer growth, regulate gene expression and apoptosis in most of cancers, including BURK and DLBCL [31]. In addition, up-regulated fatty acid synthase (FASN) affected lipogenesis and multiple signaling pathways in several cancer cell lines and was suggested as a potential therapeutic target [32]. Our study showed several lipid metabolism pathways (fatty acid metabolism, glycerolipid metabolism, steroid hormone biosynthesis, etc.) were significantly down-regulated in BURK and DLBCL, whereas these pathways showed no change or were slightly elevated in MCL (**Figure 1**). The JAK-STAT signaling pathway is another tumor-related signaling pathway. Deregulation of JAK-STAT has been found in several lymphoma cell lines [33, 34], primary mediastinal B-cell lymphoma [35], DLBCL [36], MCL and follicular lymphoma [37]. Furthermore, the JAK-STAT signaling pathway was also found to be influenced in other hematological malignancies (leukemia, multiple myeloma, myeloproliferative neoplasms, etc.) and suggested that several JAK inhibitors have broad clinical application prospects [38]. In our study, the JAK-STAT signaling pathway was significantly up-regulated in DLBCL and MCL, whereas it exhibited no change in BURK. Furthermore, our study also showed several significantly down-regulated amino acid metabolism pathways (tyrosine metabolism etc.), sensory systems (olfactory transduction and taste transduction) and xenobiotic biodegradation and metabolism (drug metabolism - cytochrome P450 and metabolism of xenobiotics by cytochrome P450) in BURK and DLBCL. However, all these pathways showed no change or even modest increases in MCL. Based on these findings, we propose that these impairment pathways have a certain carcinogenic effect and vary widely in different types of lymphoma.

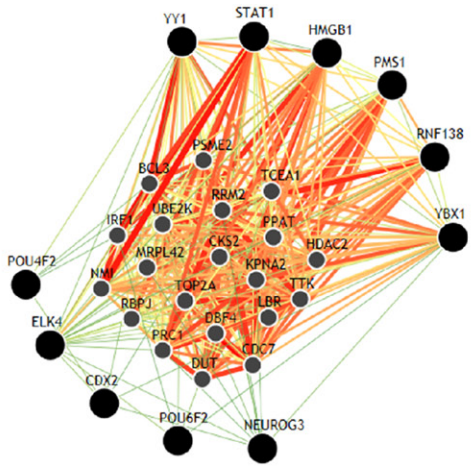
This study identified 11 up-regulated TFs in all the three types of lymphoma. Among them, PRDM1 and IRF4 were most over-expressed and also found in the top 20 list of differentially expressed genes (**Figure 2** and Table S1). Most of these TFs play dominant roles in TRNs and their target genes are related to cell cycle, mitosis, immune and other biological functions (**Figures 3** and S5, S6, S7). PRDM1 belongs to the PRDM gene family of transcription repressors containing Kruppel-type zinc fingers. The SET-related PR (PRDI-BF1-RIZ) domain is a master transcriptional regulator for terminal differentiation of B cells into plasma cells [39, 40]. Previous studies have demonstrated that

A



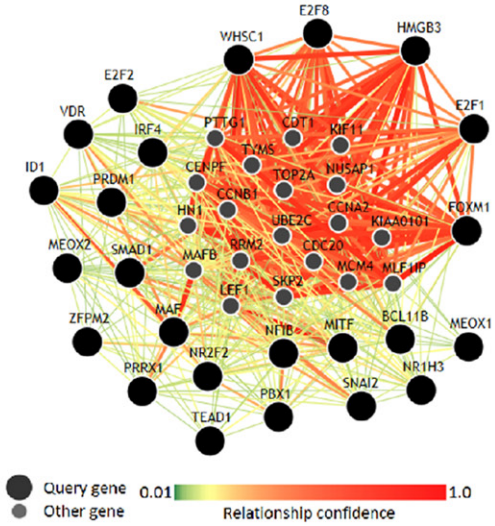
GO Biology Processes	FDR
GO0000226: microtubule cytoskeleton organization	3.54E-04
GO0051783: regulation of nuclear division	3.54E-04
GO0007067: mitosis	3.67E-04
GO0007088: regulation of mitosis	3.99E-04
GO0000280: nuclear division	4.28E-04
GO0051301: cell division	4.54E-04
GO0007017: microtubule-based process	4.55E-04
GO0048285: organelle fission	5.56E-04
GO0007059: chromosome segregation	5.58E-04
GO0031145: anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process	6.74E-04
GO0051983: regulation of chromosome segregation	7.06E-04
GO0090068: positive regulation of cell cycle process	1.51E-03
GO0051302: regulation of cell division	1.58E-03
GO0000278: mitotic cell cycle	1.69E-03
GO0040001: establishment of mitotic spindle localization	4.17E-03
GO0051293: establishment of spindle localization	8.69E-03
GO0051653: spindle localization	9.20E-03
GO0007346: regulation of mitotic cell cycle	1.35E-02
GO0030261: chromosome condensation	1.41E-02
GO0007091: metaphase/anaphase transition of mitotic cell cycle	1.56E-02

B



GO Biology Processes	FDR
GO0000122: negative regulation of transcription from RNA polymerase II promoter	3.22E-03
GO0035458: cellular response to interferon-beta	5.97E-03
GO0006266: DNA ligation	9.91E-03
GO0010225: response to UV-C	1.04E-02
GO0035456: response to interferon-beta	1.11E-02

C



GO Biology Processes	FDR
GO0010564: regulation of cell cycle process	4.85E-03
GO0007346: regulation of mitotic cell cycle	1.69E-02
GO0032616: interleukin-13 production	1.71E-02
GO0070561: vitamin D receptor signaling pathway	1.73E-02
GO0032656: regulation of interleukin-13 production	1.85E-02
GO0010839: negative regulation of keratinocyte proliferation	1.91E-02
GO0071295: cellular response to vitamin	1.91E-02
GO0032673: regulation of interleukin-4 production	2.00E-02
GO0010837: regulation of keratinocyte proliferation	2.00E-02
GO0000122: negative regulation of transcription from RNA polymerase II promoter	2.01E-02
GO0000278: mitotic cell cycle	2.15E-02
GO0032633: interleukin-4 production	2.18E-02
GO0071305: cellular response to vitamin D	2.30E-02
GO0007059: chromosome segregation	2.42E-02
GO0031145: anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process	2.55E-02
GO0051656: establishment of organelle localization	2.58E-02
GO0007067: mitosis	2.60E-02
GO0033280: response to vitamin D	2.69E-02
GO0050680: negative regulation of epithelial cell proliferation	2.71E-02
GO0000280: nuclear division	2.72E-02

Figure 3. Genome-scale integrated analysis of gene networks in Burkitt's lymphoma (A), diffuse large B-cell lymphoma (B) and mantle cell lymphoma (C). The top 20 enriched biological functions of the genes in the networks are shown in the right panel.

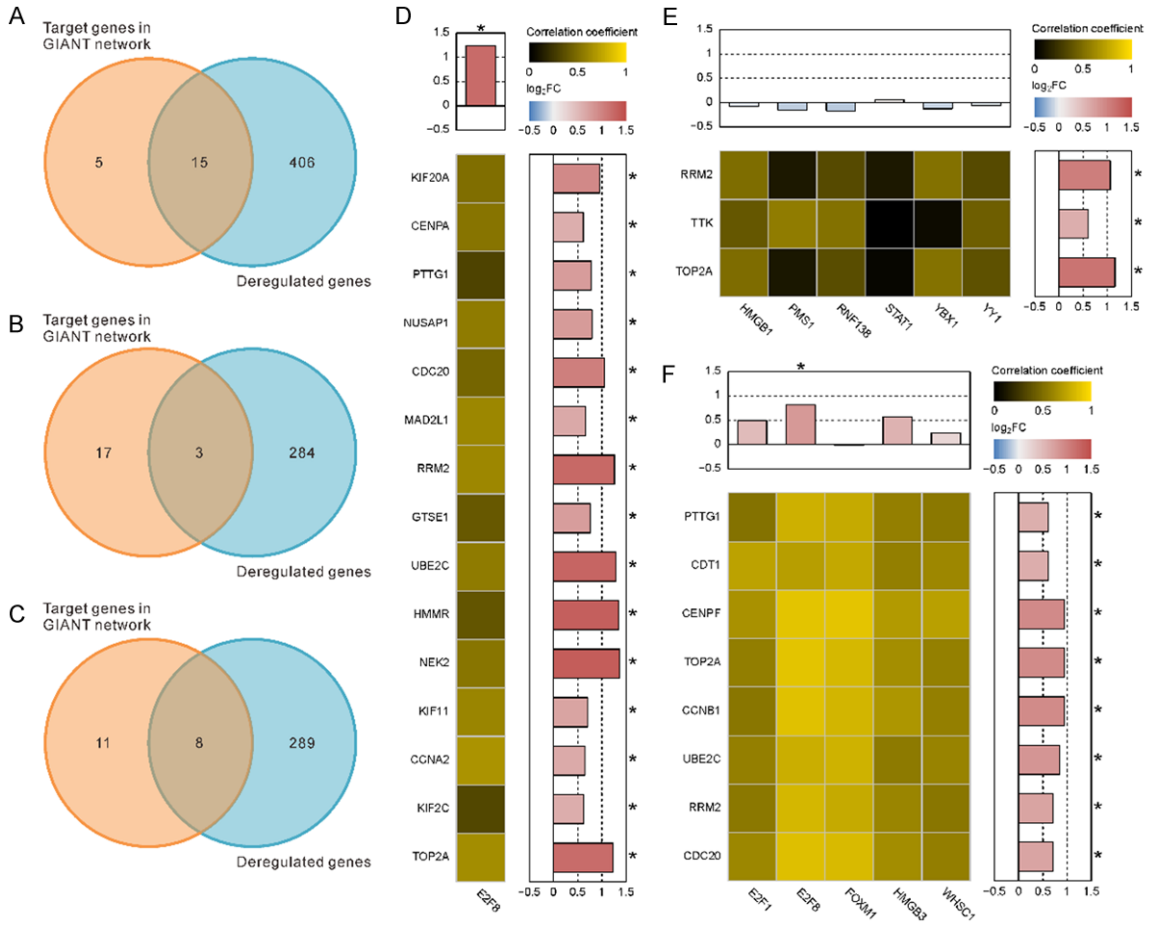


Figure 4. Correlation of key transcription factors and deregulated target genes in three types of lymphoma. A-C. Venn diagram of target genes in GIANT network and deregulated genes in Burkitt's lymphoma, diffuse large B-cell lymphoma and mantle cell lymphoma, respectively. D-F. correlation of key transcription factors and deregulated target genes in Burkitt's lymphoma, diffuse large B-cell lymphoma and mantle cell lymphoma, respectively. The black/yellow color bar indicates the correlation coefficients of key transcription factors and deregulated target genes. The blue/red color bar indicates the log2 transformed fold change of these genes. Significantly differentially expressed genes are indicated by an asterisk.

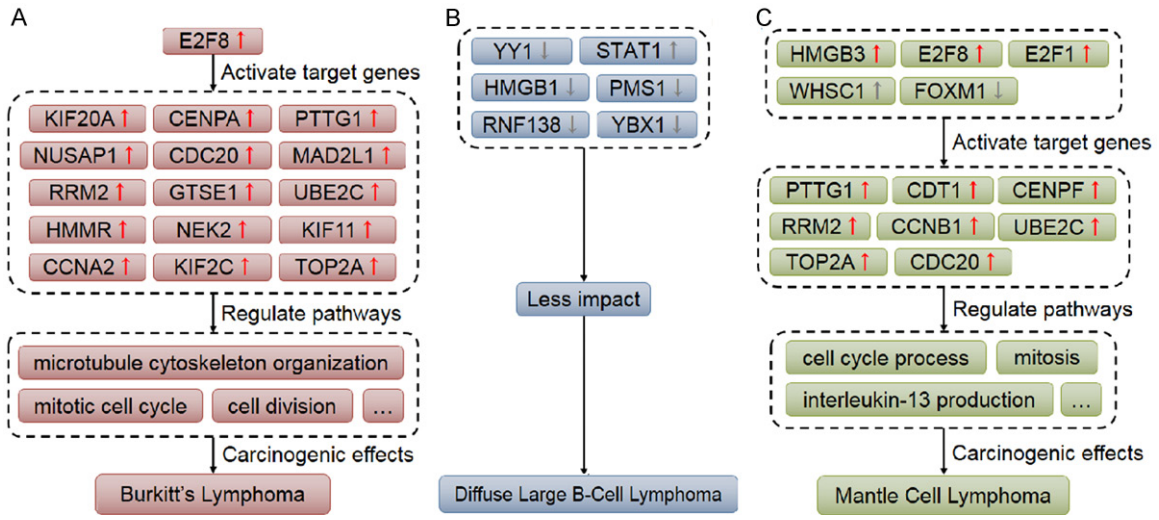


Figure 5. Possible regulatory mechanisms of key transcription factors, target genes, and pathways in three types of lymphoma. A. Up-regulated E2F8 activated several target genes and then affected microtubule cytoskeleton organization, mitotic cell cycle, cell division, and other cancer-related pathways, finally caused Burkitt's lymphoma. B. The expression of these transcription factors may have less effect on diffuse large B-cell lymphoma. C. Up-regulated HMGB3, E2F8 and E2F1 activated several target genes and then affected cell cycle process, mitosis, interleukin-13 production, and other cancer-related pathways, finally resulting in mantle cell lymphoma.

PRDM1 is over-expressed in Hodgkin lymphomas [39], T-cell lymphoma [41] and DLBCL [40, 42]. Furthermore, other studies have demonstrated that over-expression of miR-9 or let-7a reduced PRDM1 levels in multiple cell lines and thus these factors are suggested as potential therapeutic targets for lymphomas [39, 40]. IRF4 (also called MUM1) is a lymphocyte-specific member of the interferon regulatory factor (IRF) family of transcription factors and plays crucial roles in the activation of innate and adaptive immune systems [43]. IRF4 is involved in the t(6;14)(p25;q32) translocation, which causes the juxtaposition of the IRF4 gene (mapping at 6p25) to the IgH locus on 14q32 [43]. Over-expression of the IRF4 is detected in multiple myelomas, HIV-related lymphomas, lymphoplasmacytic lymphomas, peripheral T-cell lymphomas, DLBCL and MCL [43-46] and may represent a possible therapeutic target [46]. Furthermore, we also identified several top-node TFs in the TRNs of the three types of lymphoma, including E2F1, E2F8, HMGB3 etc. A previous study found E2F1 is highly expressed in BURK cell lines and sporadic BURK specimens, suggesting E2F1 as a promising therapeutic target for sporadic BURK [47]. However, recent research showed over-expression of E2F1 has no relationship with prognosis in T lymphoblastic lymphoma or leukemia [48]. The expression of E2F1 in this study was also up-regulated in BURK, DLBCL and MCL, despite all did not reach a significant difference.

E2F8 is a recently identified member of the E2F family. It controls cell cycle genes through homodimers or heterodimers from DNA-binding domains [49]. Furthermore, E2F8 plays a variety of regulatory roles including regulation of placenta architecture and embryonic development, mediation of the DNA damage-correlated cell-cycle, promotion of angiogenesis and lymphangiogenesis, and also influences polyploidization [49]. The traditional view is that E2F8 functions as a transcriptional repressor. Its mediated transcriptional repression is a critical mechanism during postnatal liver development [50] and it also represses a large set of E2F target genes in skin cancer [51]. However, new perspectives noted that E2F8 acts as a transcription activator which confers cisplatin resistance by promoting MASTL mediated mitotic progression in ER+ breast cancer cells [52]. Another breast cancer study demonstrated that E2F8 transcriptionally upregulates CCNE1 and CCNE2 *via* directly interacting with their respective gene promoters, which accelerates the transition from G1 to S phase of breast cancer cells [53]. Furthermore, a recent study suggested that E2F8 promotes papillary thyroid cancer progression by activating CCND1 and other cell cycle related genes [54]. In our study, E2F8 was up-regulated in all the three types of lymphoma, and was the most important TFs in BURK and MCL TRNs (**Figures 2**, S5 and S7). In addition, over-expressed E2F8 target genes influenced cell cycle and disrupted immune functions in BURK and MCL (**Figure 4**). However, to our knowledge, no other study reports the role of up-regulated E2F8 in lymphoma. Our study also discloses that HMGB3 is a key regulator in MCL, although the expression did not achieve significance (**Figure 4**). High mobility group-box 3 (HMGB3) is a member of the HMG superfamily and is classified with HMGB1 and HMGB2 into the HMGB subfamily. It plays important roles in DNA recombination, repair, replication, and transcription [55]. Recent studies showed that high-expression of HMGB3 promotes growth and migration in colorectal cancer [56] and reduces patient survival in esophageal squamous cell carcinoma [57]. Knockdown of HMGB3 also showed a potential therapeutic effect in gastric cancer cells [55]. However, no studies have reported the association between HMGB3 and lymphoma so far.

In summary, this study identifies a series of commonly and specifically deregulated pathways in BURK, DLBCL and MCL, respectively. Furthermore, several key TFs may activate their target genes, influencing cell cycle, mitosis, immune functions and other oncogenic pathways in BURK and MCL. However, these top TFs

may have less impact on DLBCL. It is noteworthy that we are the first to report that E2F8 plays a dominant role in BURK and MCL. Therefore, we suggest that E2F8 can be used as a biomarker and a potential therapeutic target for BURK and MCL. Future studies are required to uncover its potential mechanisms.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 815-70376, No. 81870307, No. 81570202), the Scientific Research Startup Fund of Foshan University (Grant No. CGg07025), the Featured Project of Innovation Enhancing College of Guangdong Province (No. 2015KTSCX154), the University Special Innovative Research Program of Department of Education of Guangdong Province (No. 2017KTSCX189) and Foshan Science and Technology Innovation Project (No. 2015AG10010).

Disclosure of conflict of interest

None.

Address correspondence to: Dahai Liu, School of Stomatology and Medicine, Foshan University, Foshan 528000, Guangdong, China. E-mail: seansean2014@126.com; Song Guo Zheng, Department of Internal Medicine, Ohio State University College of Medicine and Wexner Medical Center, Columbus 43210, USA. E-mail: SongGuo.Zheng@osumc.edu; Wen-Xing Li, Key Laboratory of Animal Models and Human Disease Mechanisms, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, Yunnan, China. E-mail: liwenxing2016@gmail.com

References

[1] Shankland KR, Armitage JO and Hancock BW. Non-Hodgkin lymphoma. Lancet 2012; 380: 848-857.

[2] Ansell SM. Non-Hodgkin lymphoma: diagnosis and treatment. Mayo Clin Proc 2015; 90: 1152-1163.

[3] Lewis N, Young J, Hesseling PB, McCormick P and Wright N. Epidemiology of Burkitt's lymphoma in Northwest Province, Cameroon, 2003-2010. Paediatr Int Child Health 2012; 32: 82-5.

[4] Kwak JY. Treatment of diffuse large B cell lymphoma. Korean J Intern Med 2012; 27: 369-377.

[5] Humala K and Younes A. Current and emerging new treatment strategies for mantle cell lymphoma. Leuk Lymphoma 2013; 54: 912-921.

[6] Smedby KE and Hjalgrim H. Epidemiology and etiology of mantle cell lymphoma and other non-Hodgkin lymphoma subtypes. Semin Cancer Biol 2011; 21: 293-298.

[7] Engels EA, Pfeiffer RM, Fraumeni JF, Kasiske BL, Israni AK, Snyder JJ, Wolfe RA, Goodrich NP, Bayakly AR, Clarke CA, Copeland G, Finch JL, Fleissner ML, Goodman MT, Kahn A, Koch L, Lynch CF, Madeleine MM, Pawlish K, Rao C, Williams MA, Castenson D, Curry M, Parsons R, Fant G and Lin M. Spectrum of cancer risk among US solid organ transplant recipients. JAMA 2011; 306: 1891-1901.

[8] Diver WR, Patel AV, Thun MJ, Teras LR and Gapstur SM. The association between cigarette smoking and non-Hodgkin lymphoid neoplasms in a large US cohort study. Cancer Causes Control 2012; 23: 1231-40.

[9] Gapstur SM, Diver WR, McCullough ML, Teras LR, Thun MJ and Patel AV. Alcohol intake and the incidence of non-hodgkin lymphoid neoplasms in the cancer prevention study II nutrition cohort. Am J Epidemiol 2012; 176: 60-69.

[10] Ye Q, Xu-Monette ZY, Tzankov A, Deng L, Wang X, Manyam GC, Visco C, Montes-Moreno S, Zhang L, Dybkaer K, Chiu A, Orazi A, Zu Y, Bhagat G, Richards KL, Hsi ED, Choi WW, van Krieken JH, Huh J, Ponzoni M, Ferreri AJ, Parsons BM, Moller MB, Piris MA, Winter JN, Medeiros LJ, Hu S and Young KH. Prognostic impact of concurrent MYC and BCL6 rearrangements and expression in de novo diffuse large B-cell lymphoma. Oncotarget 2016; 7: 2401-2416.

[11] Dybkaer K, Bogsted M, Falgreen S, Bodker JS, Kjeldsen MK, Schmitz A, Bilgrau AE, Xu-Monette ZY, Li L, Bergkvist KS, Laursen MB, Rodrigo-Domingo M, Marques SC, Rasmussen SB, Nyegaard M, Gaihede M, Moller MB, Samworth RJ, Shah RD, Johansen P, El-Galaly TC, Young KH and Johnsen HE. Diffuse large B-cell lymphoma classification system that associates normal B-cell subset phenotypes with prognosis. J Clin Oncol 2015; 33: 1379-1388.

[12] Linton K, Howarth C, Wappett M, Newton G, Lachel C, Iqbal J, Pepper S, Byers R, Chan WJ and Radford J. Microarray gene expression analysis of fixed archival tissue permits molecular classification and identification of potential therapeutic targets in diffuse large B-cell lymphoma. J Mol Diagn 2012; 14: 223-232.

[13] Hummel M, Bentink S, Berger H, Klapper W, Wessendorf S, Barth TF, Bernd HW, Cogliatti SB, Dierlamm J, Feller AC, Hansmann ML, Ha-

ralambieva E, Harder L, Hasenclever D, Kühn M, Lenze D, Lichter P, Martin-Subero JI, Möller P, Müller-Hermelink HK, Ott G, Parwaresch RM, Pott C, Rosenwald A, Rosolowski M, Schwaenen C, Stürzenhofecker B, Szczepanowski M, Trautmann H, Wacker HH, Spang R, Loeffler M, Trümper L, Stein H, Siebert R; Molecular Mechanisms in Malignant Lymphomas Network Project of the Deutsche Krebshilfe. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. N Engl J Med 2006; 354: 2419-2430.

[14] Sha C, Barrans S, Care MA, Cunningham D, Tooze RM, Jack A and Westhead DR. Transferring genomics to the clinic: distinguishing Burkitt and diffuse large B cell lymphomas. Genome Med 2015; 7: 64.

[15] Hartmann EM, Campo E, Wright G, Lenz G, Salaverria I, Jares P, Xiao W, Braziel RM, Rimsza LM, Chan WC, Weisenburger DD, Delabie J, Jaffe ES, Gascoyne RD, Dave SS, Mueller-Hermelink HK, Staudt LM, Ott G, Bea S and Rosenwald A. Pathway discovery in mantle cell lymphoma by integrated analysis of high-resolution gene expression and copy number profiling. Blood 2010; 116: 953-961.

[16] Palomero J, Vegliante MC, Rodriguez ML, Eguileor A, Castellano G, Planas-Rigol E, Jares P, Ribera-Cortada I, Cid MC, Campo E and Amador V. SOX11 promotes tumor angiogenesis through transcriptional regulation of PDGFA in mantle cell lymphoma. Blood 2014; 124: 2235-2247.

[17] Saba NS, Liu D, Herman SE, Underbayev C, Tian X, Behrend D, Weniger MA, Skarzynski M, Gyamfi J, Fontan L, Melnick A, Grant C, Roschewski M, Navarro A, Bea S, Pittaluga S, Dunleavy K, Wilson WH and Wiestner A. Pathogenic role of B-cell receptor signaling and canonical NF-kappaB activation in mantle cell lymphoma. Blood 2016; 128: 82-92.

[18] Pinweha P, Rattanapornsompong K, Charoensawan V and Jitrapakdee S. MicroRNAs and oncogenic transcriptional regulatory networks controlling metabolic reprogramming in cancers. Comput Struct Biotechnol J 2016; 14: 223-233.

[19] Jang SM, An JH, Kim CH, Kim JW and Choi KH. Transcription factor FOXA2-centered transcriptional regulation network in non-small cell lung cancer. Biochem Biophys Res Commun 2015; 463: 961-967.

[20] Gan L, Schwengberg S and Denecke B. Transcriptome analysis in cardiomyocyte-specific differentiation of murine embryonic stem cells reveals transcriptional regulation network. Gene Expr Patterns 2014; 16: 8-22.

[21] Li WX, He K, Tang L, Dai SX, Li GH, Lv WW, Guo YC, An SQ, Wu GY, Liu D and Huang JF. Comprehensive tissue-specific gene set enrichment analysis and transcription factor analysis of breast cancer by integrating 14 gene expression datasets. Oncotarget 2016; 8: 6775-6786.

[22] Carvalho BS and Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinformatics 2010; 26: 2363-2367.

[23] Li WX, Dai SX, Liu JQ, Wang Q, Li GH and Huang JF. Integrated analysis of Alzheimer's disease and schizophrenia dataset revealed different expression pattern in learning and memory. J Alzheimers Dis 2016; 51: 417-425.

[24] Li WX, Dai SX, Wang Q, Guo YC, Hong Y, Zheng JJ, Liu JQ, Liu D, Li GH and Huang JF. Integrated analysis of ischemic stroke datasets revealed sex and age difference in anti-stroke targets. PeerJ 2016; 4: e2470.

[25] 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 2015; 43: e47.

[26] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.

[27] Zhang HM, Liu T, Liu CJ, Song S, Zhang X, Liu W, Jia H, Xue Y and Guo AY. AnimalTFDB 2.0: a resource for expression, prediction and functional study of animal transcription factors. Nucleic Acids Res 2015; 43: D76-81.

[28] Huynh-Thu VA, Irrthum A, Wehenkel L and Geurts P. Inferring regulatory networks from expression data using tree-based methods. PLoS One 2010; 5.

[29] Greene CS, Krishnan A, Wong AK, Ricciotti E, Zelaya RA, Himmelstein DS, Zhang R, Hartmann BM, Zaslavsky E, Sealfon SC, Chasman DI, FitzGerald GA, Dolinski K, Grosser T and Troyanskaya OG. Understanding multicellular function and disease with human tissue-specific networks. Nat Genet 2015; 47: 569-576.

[30] Yang F, Du J, Zhang H, Ruan G, Xiang J, Wang L, Sun H, Guan A, Shen G, Liu Y, Guo X, Li Q and Tang Y. Serum metabolomics of burkitt lymphoma mouse models. PLoS One 2017; 12: e0170896.

[31] Qu Q, Zeng F, Liu X, Wang QJ and Deng F. Fatty acid oxidation and carnitine palmitoyltransferase I: emerging therapeutic targets in cancer. Cell Death Dis 2016; 7: e2226.

[32] Flavin R, Peluso S, Nguyen PL and Loda M. Fatty acid synthase as a potential therapeutic target in cancer. Future Oncol 2010; 6: 551-562.

[33] Melzner I, Weniger MA, Bucur AJ, Bruderlein S, Dorsch K, Hasel C, Leithauser F, Ritz O, Dyer MJ, Barth TF and Moller P. Biallelic deletion within 16p13.13 including SOCS-1 in Karpas1106P mediastinal B-cell lymphoma line is associated with delayed degradation of JAK2 protein. Int J Cancer 2006; 118: 1941-1944.

[34] Melzner I, Bucur AJ, Bruderlein S, Dorsch K, Hasel C, Barth TF, Leithauser F and Moller P. Biallelic mutation of SOCS-1 impairs JAK2 degradation and sustains phospho-JAK2 action in the MedB-1 mediastinal lymphoma line. Blood 2005; 105: 2535-2542.

[35] Ritz O, Guiter C, Castellano F, Dorsch K, Melzner J, Jais JP, Dubois G, Gaulard P, Moller P and Leroy K. Recurrent mutations of the STAT6 DNA binding domain in primary mediastinal B-cell lymphoma. Blood 2009; 114: 1236-1242.

[36] Karmali R and Gordon LI. Molecular subtyping in diffuse large B cell lymphoma: closer to an approach of precision therapy. Curr Treat Options Oncol 2017; 18: 11.

[37] Chim CS, Wong KY, Loong F and Srivastava G. SOCS1 and SHP1 hypermethylation in mantle cell lymphoma and follicular lymphoma: implications for epigenetic activation of the Jak/STAT pathway. Leukemia 2004; 18: 356-358.

[38] Furqan M, Mukhi N, Lee B and Liu D. Dysregulation of JAK-STAT pathway in hematological malignancies and JAK inhibitors for clinical application. Biomark Res 2013; 1: 5.

[39] Nie K, Gomez M, Landgraf P, Garcia JF, Liu Y, Tan LH, Chadburn A, Tuschl T, Knowles DM and Tam W. MicroRNA-mediated down-regulation of PRDM1/Blimp-1 in Hodgkin/Reed-Sternberg cells: a potential pathogenetic lesion in Hodgkin lymphomas. Am J Pathol 2008; 173: 242-252.

[40] Nie K, Zhang T, Allawi H, Gomez M, Liu Y, Chadburn A, Wang YL, Knowles DM and Tam W. Epigenetic down-regulation of the tumor suppressor gene PRDM1/Blimp-1 in diffuse large B cell lymphomas: a potential role of the microRNA let-7. Am J Pathol 2010; 177: 1470-1479.

[41] Zhao WL, Liu YY, Zhang QL, Wang L, Leboeuf C, Zhang YW, Ma J, Garcia JF, Song YP, Li JM, Shen ZX, Chen Z, Janin A and Chen SJ. PRDM1 is involved in chemoresistance of T-cell lymphoma and down-regulated by the proteasome inhibitor. Blood 2008; 111: 3867-3871.

[42] Tam W, Gomez M, Chadburn A, Lee JW, Chan WC and Knowles DM. Mutational analysis of PRDM1 indicates a tumor-suppressor role in diffuse large B-cell lymphomas. Blood 2006; 107: 4090-4100.

[43] Carbone A, Gloghini A, Larocca LM, Capello D, Pierconti F, Canzonieri V, Tirelli U, Dalla-Favera R and Gaidano G. Expression profile of MUM1/IRF4, BCL-6, and CD138/syndecan-1 defines novel histogenetic subsets of human immunodeficiency virus-related lymphomas. Blood 2001; 97: 744-751.

[44] Natkunam Y, Warnke RA, Montgomery K, Falini B and van de Rijn M. Analysis of MUM1/IRF4 protein expression using tissue microarrays and immunohistochemistry. Mod Pathol 2001; 14: 686-694.

[45] Perez-Galan P, Dreyling M and Wiestner A. Mantle cell lymphoma: biology, pathogenesis, and the molecular basis of treatment in the genomic era. Blood 2011; 117: 26-38.

[46] Feldman AL, Law M, Remstein ED, Macon WR, Erickson LA, Grogg KL, Kurtin PJ and Dogan A. Recurrent translocations involving the IRF4 oncogene locus in peripheral T-cell lymphomas. Leukemia 2009; 23: 574-580.

[47] Molina-Privado I, Rodriguez-Martinez M, Rebollo P, Martin-Perez D, Artiga MJ, Menarguez J, Flemington EK, Piris MA and Campanero MR. E2F1 expression is deregulated and plays an oncogenic role in sporadic Burkitt's lymphoma. Cancer Res 2009; 69: 4052-4058.

[48] Li J, Li P, Yu ZZ, Bu P and Xi YF. Significance of microRNA-17, microRNA-19, c-myc, and E2F1 expression in T lymphoblastic lymphoma/leukemia and their relation to prognosis. Int J Clin Exp Pathol 2017; 10: 5821-5830.

[49] Lv Y, Xiao J, Liu J and Xing F. E2F8 is a potential therapeutic target for hepatocellular carcinoma. J Cancer 2017; 8: 1205-1213.

[50] Kent LN, Rakijas JB, Pandit SK, Westendorp B, Chen HZ, Huntington JT, Tang X, Bae S, Srivastava A, Senapati S, Koivisto C, Martin CK, Cuitino MC, Perez M, Clouse JM, Chokshi V, Shinde N, Kladney R, Sun D, Perez-Castro A, Matondo RB, Nantasanti S, Mokry M, Huang K, Machiraju R, Fernandez S, Rosol TJ, Coppola V, Pohar KS, Pipas JM, Schmidt CR, de Bruin A and Leone G. E2f8 mediates tumor suppression in postnatal liver development. J Clin Invest 2016; 126: 2955-2969.

[51] Thurlings I, Martinez-Lopez LM, Westendorp B, Zijp M, Kuiper R, Tooten P, Kent LN, Leone G, Vos HJ, Burgering B and de Bruin A. Synergistic functions of E2F7 and E2F8 are critical to suppress stress-induced skin cancer. Oncogene 2017; 36: 829-839.

[52] Tian J, Lin Y and Yu J. E2F8 confers cisplatin resistance to ER+ breast cancer cells via transcriptionally activating MASTL. Biomed Pharmacother 2017; 92: 919-926.

[53] Ye L, Guo L, He Z, Wang X, Lin C, Zhang X, Wu S, Bao Y, Yang Q, Song L and Lin H. Upregulation of E2F8 promotes cell proliferation and tumorigenicity in breast cancer by modulating G1/S phase transition. Oncotarget 2016; 7: 23757-23771.

[54] Sun J, Shi R, Zhao S, Li X, Lu S, Bu H, Ma X and Su C. E2F8, a direct target of miR-144, promotes papillary thyroid cancer progression via

regulating cell cycle. J Exp Clin Cancer Res 2017; 36: 40.
[55] Guo S, Wang Y, Gao Y, Zhang Y, Chen M, Xu M, Hu L, Jing Y, Jing F, Li C, Wang Q and Zhu Z. Knockdown of High Mobility Group-Box 3 (HMGB3) expression inhibits proliferation, reduces migration, and affects chemosensitivity in gastric cancer cells. Med Sci Monit 2016; 22: 3951-3960.
[56] Zhang Z, Chang Y, Zhang J, Lu Y, Zheng L, Hu Y, Zhang F, Li X, Zhang W and Li X. HMGB3 promotes growth and migration in colorectal cancer by regulating WNT/beta-catenin pathway. PLoS One 2017; 12: e0179741.
[57] Gao J, Zou Z, Gao J, Zhang H, Lin Z, Zhang Y, Luo X, Liu C, Xie J and Cai C. Increased expression of HMGB3: a novel independent prognostic marker of worse outcome in patients with esophageal squamous cell carcinoma. Int J Clin Exp Pathol 2015; 8: 345-352.

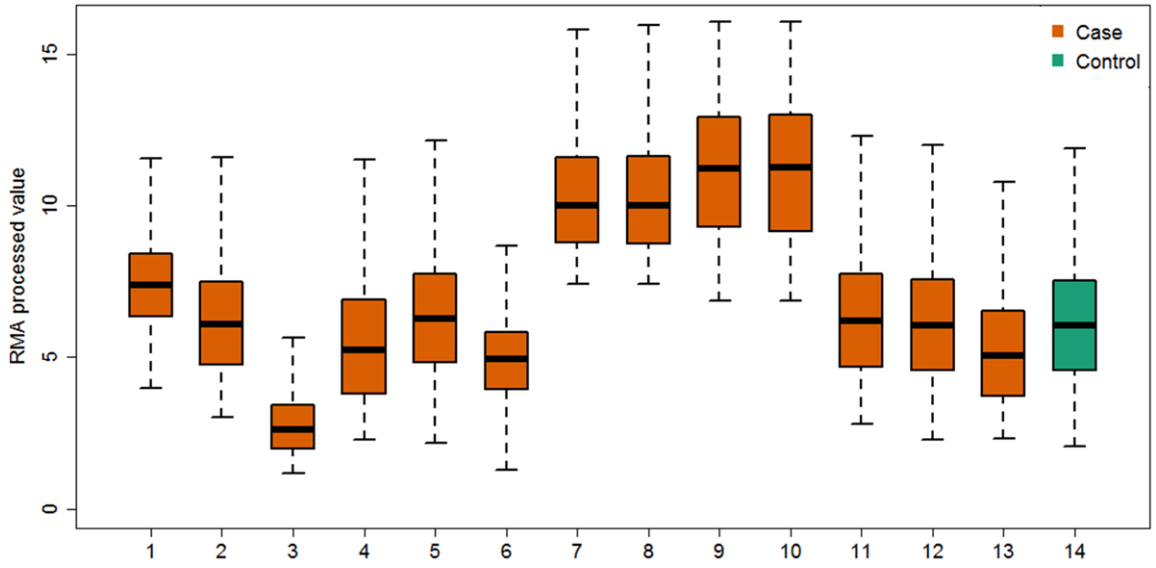


Figure S1. The distribution of RMA processed gene expression values of lymphoma datasets. The orange box shows the distribution of gene expression values in lymphoma patients and the green box shows the distribution of gene expression values in controls. There was a relatively large deviation in the distribution of gene expression values across these studies.

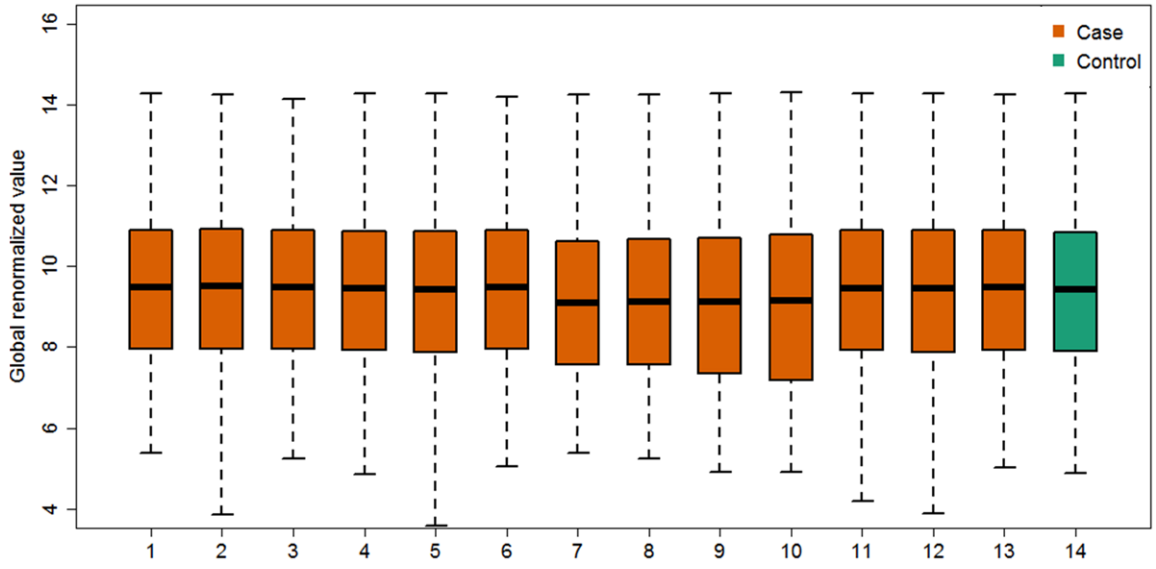


Figure S2. The distribution of global renormalized (after RMA processed) gene expression values of lymphoma datasets. The orange box shows the distribution of gene expression values in lymphoma patients and the green box shows the distribution of gene expression values in controls. The distribution of gene expression values across these studies had a consistent range.

Table S1. Top 20 differentially expressed genes in three types of lymphoma

Burkitt's lymphoma			Diffuse large B-cell lymphoma			Mantle cell lymphoma		
Symbol	logFC	*P* value	Symbol	logFC	*P* value	Symbol	logFC	*P* value
CYorf15B	2.38	4.61E-23	KRT6A	-2.88	1.27E-13	KRT13	-3.13	1.61E-14
HIST2H2BE	2.06	9.57E-25	SPRR1B	-2.36	7.19E-13	KRT6A	-2.78	9.96E-14
RPS4Y1	2.03	3.19E-07	SPRR3	-2.27	4.91E-11	KRT14	-2.40	4.06E-11
CENPE	1.92	7.45E-18	HIST2H2BE	2.14	1.68E-40	SPRR3	-2.37	3.97E-13
KRT14	-1.90	2.30E-09	PRDM1	1.98	1.57E-29	STATH	-2.10	1.71E-07
MMP1	-1.87	9.04E-11	SPRR1A	-1.97	1.96E-10	PRDM1	2.06	6.92E-22
C7orf10	1.85	3.36E-13	STATH	-1.89	1.87E-10	HIST2H2BE	2.05	8.03E-18
KIF18B	1.84	1.34E-17	RPS4Y1	1.84	8.86E-07	IRF4	2.04	8.32E-25
ORM1	-1.83	1.35E-09	CENPE	1.84	2.02E-29	SPRR1B	-1.99	6.25E-12
IRF4	1.81	1.58E-17	IRF4	1.79	9.99E-21	SPRR1A	-1.97	1.31E-11
PRDM1	1.80	2.31E-21	MMP1	-1.78	4.14E-07	C7orf10	1.95	7.34E-18
EIF1AY	1.79	6.70E-11	OLFM4	-1.77	4.07E-12	HTN3	-1.84	5.57E-07
KRT6A	-1.69	1.72E-07	KRT13	-1.75	1.93E-07	GPR15	1.79	2.92E-37
SPC25	1.68	7.87E-18	HTN3	-1.70	3.00E-10	FCGR2A	1.77	5.21E-19
FCGR2A	1.65	4.99E-20	S100A7	-1.66	5.65E-08	S100A2	-1.72	1.51E-09
KRT13	-1.64	5.16E-09	KRT14	-1.65	1.77E-05	CENPE	1.69	7.88E-11
CDKN3	1.62	9.13E-17	CYorf15B	1.62	3.78E-08	CYorf15B	1.64	1.71E-05
HIST1H2BG	1.58	2.20E-20	FCGR2A	1.62	1.92E-23	HTN1	-1.59	6.62E-06
PLAC8	1.55	5.54E-06	SPC25	1.62	3.54E-27	RPS4Y1	1.58	3.02E-03
TPX2	1.52	3.58E-15	S100A2	-1.61	5.68E-09	ZBP1	1.58	6.16E-16

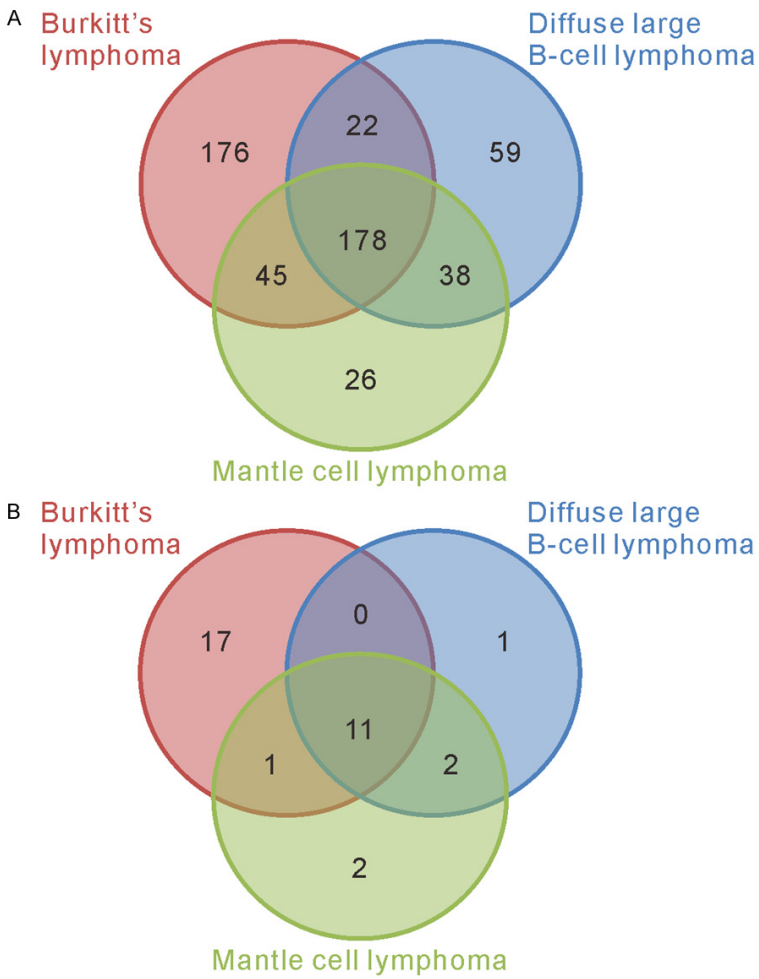


Figure S3. Venn diagram of all differentially expressed genes (A) and differentially expressed transcription factors (B) in three types of lymphoma.

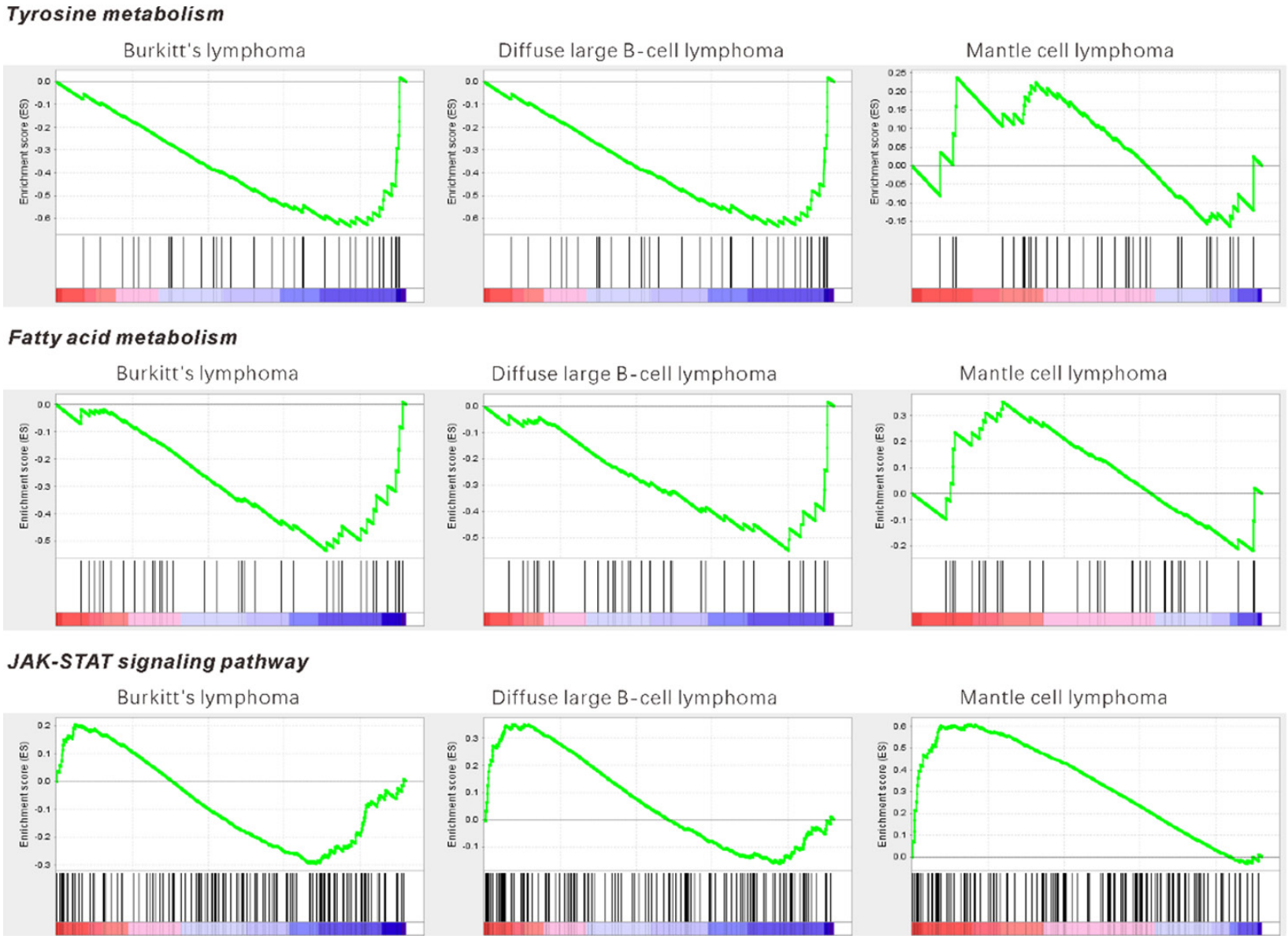


Figure S4. GSEA enrichment results of tyrosine metabolism, fatty acid metabolism and JAK-STAT signaling pathway in three types of lymphoma. A green curve below zero indicates that the pathway is down-regulated, while a green curve above zero indicates that this pathway is up-regulated.

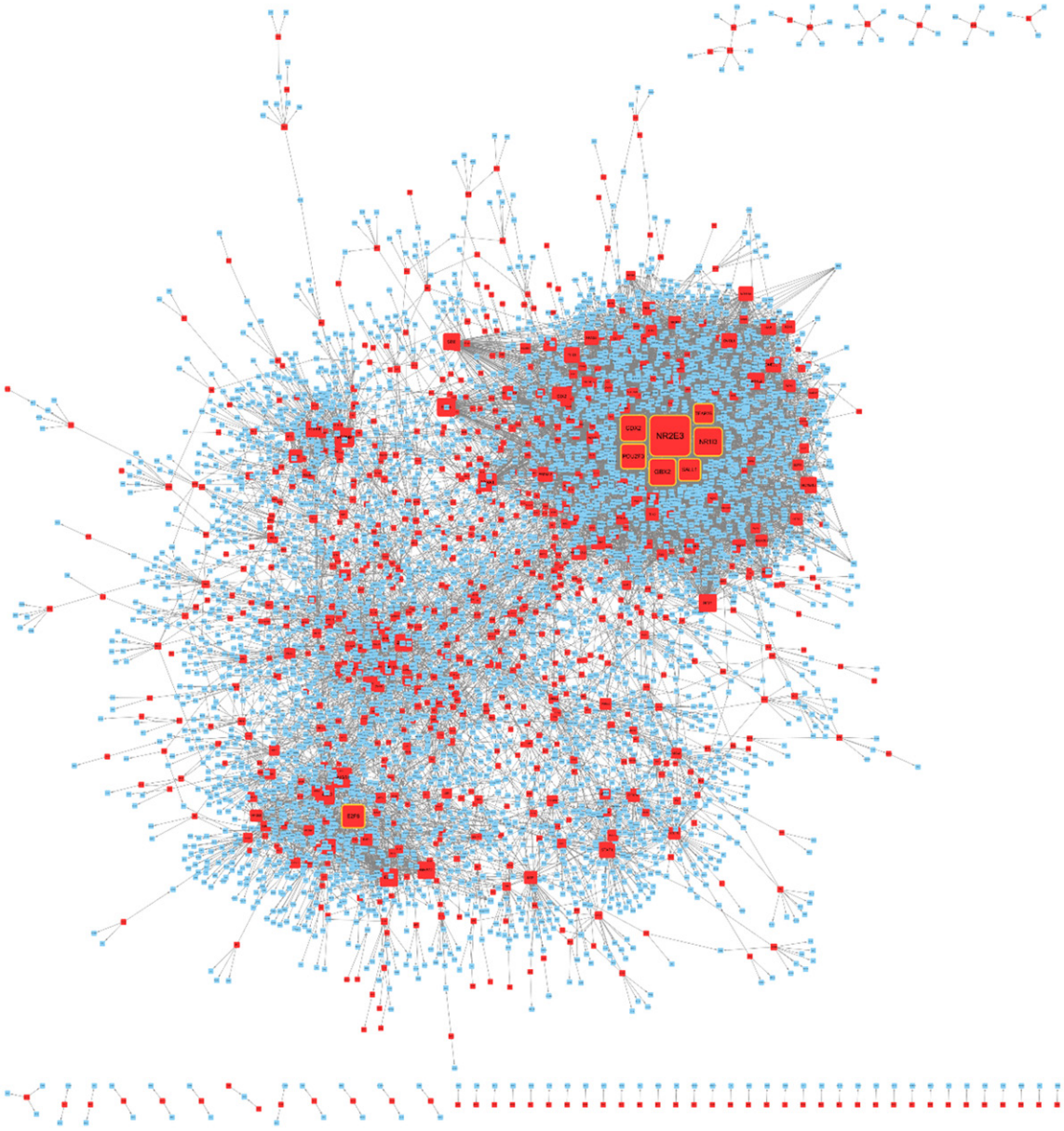


Figure S5. Transcriptional regulatory networks of Burkitt's lymphoma. Red boxes represent transcription factors and blue boxes represent target genes. The box size represents the number of nodes of a gene. Arrows indicate the direction of regulation. Transcription factors with greater than 100 nodes are marked with a yellow box.

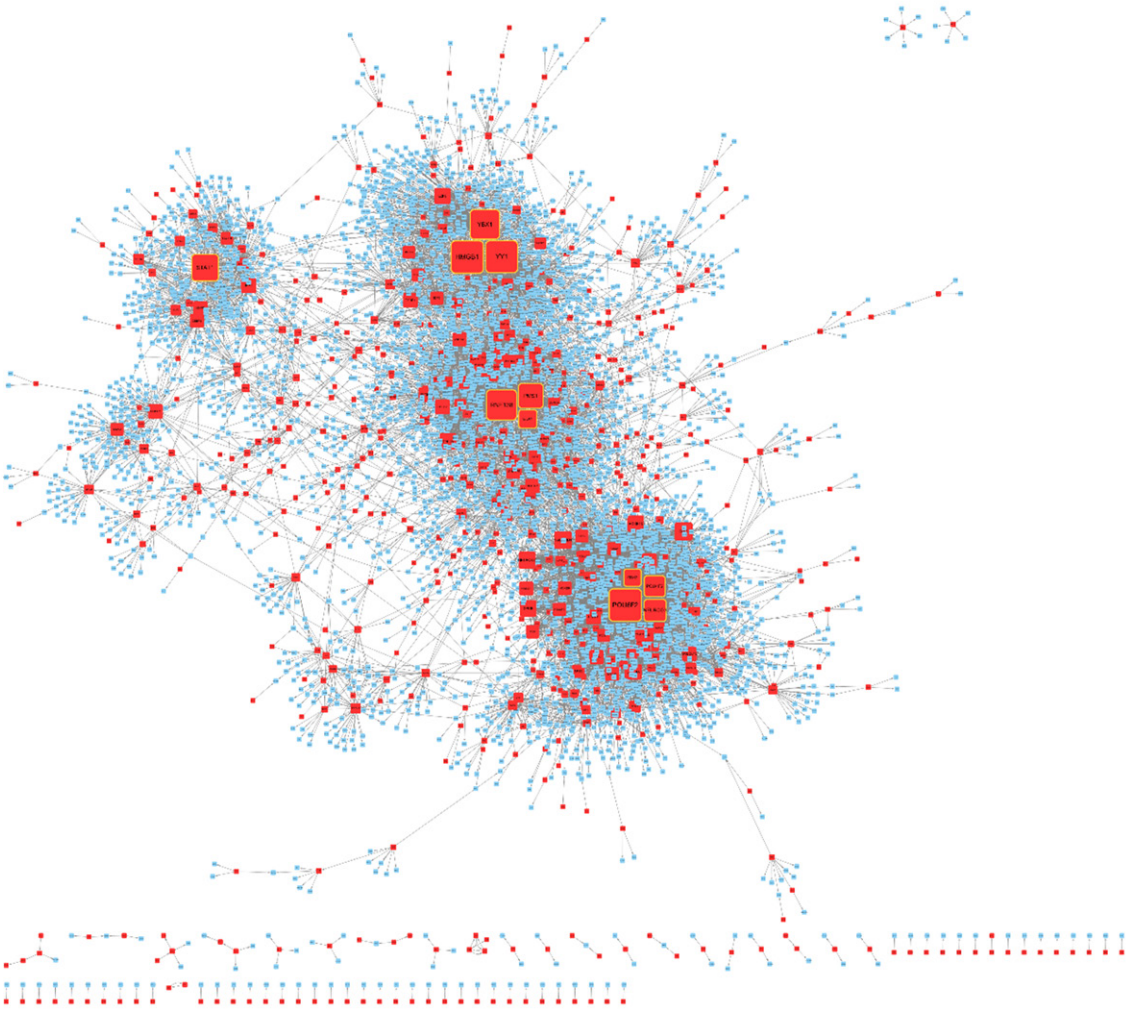


Figure S6. Transcriptional regulatory networks of diffuse large B-cell lymphoma. Red boxes represent transcription factors and blue boxes represent target genes. The box size represents the number of nodes of a gene. Arrows indicate the direction of regulation. Transcription factors with greater than 100 nodes are marked with a yellow box.

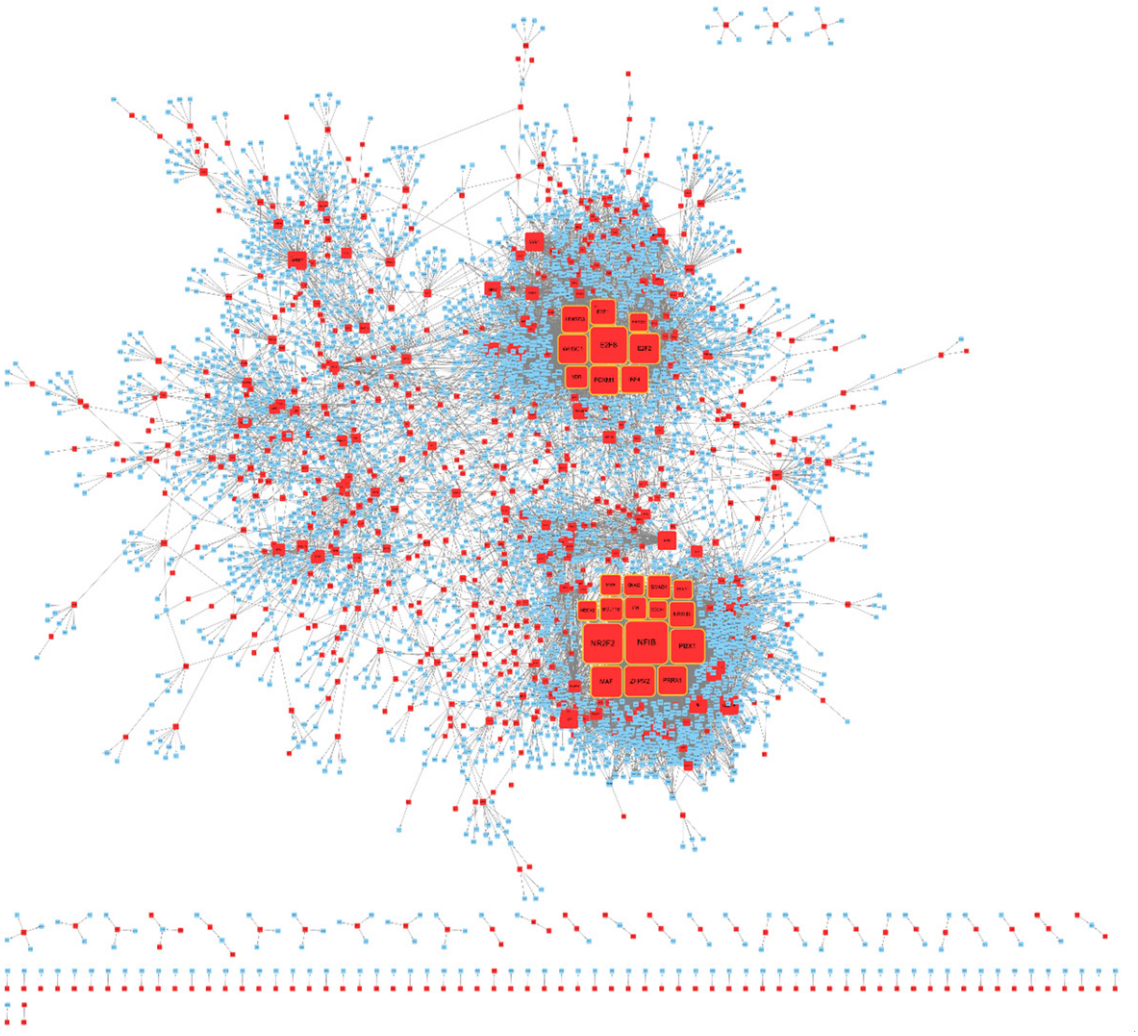


Figure S7. Transcriptional regulatory networks of mantle cell lymphoma. Red boxes represent transcription factors and blue boxes represent target genes. The box size represents the number of nodes of a gene. Arrows indicate the direction of regulation. Transcription factors with greater than 100 nodes are marked with a yellow box.

Table S2. Transcriptional regulatory network properties in three types of lymphoma

Parameters	Burkitt's lymphoma	Diffuse large B-cell lymphoma	Mantle cell lymphoma
Clustering coefficient	0.049	0.079	0.120
Connected component	56	76	101
Characteristic path length	7.027	5.624	8.316
Ave. number of neighbors	3.775	4.079	4.368
Number of nodes	5262	4859	4519
Multi-edge node pairs	68	89	130

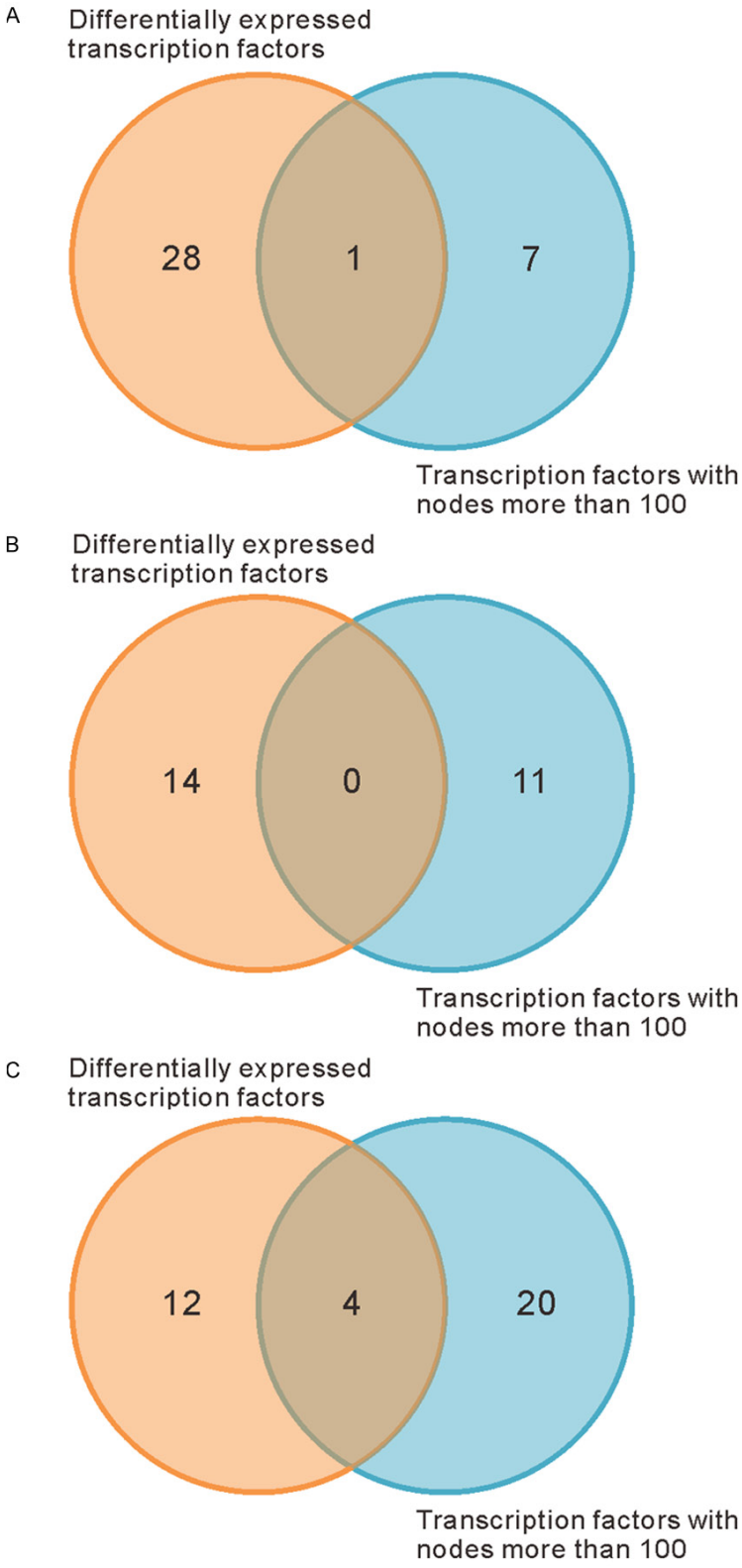


Figure S8. Venn diagram of differentially expressed transcription factors and transcription factors with greater than 100 nodes in Burkitt's lymphoma (A), diffuse large B-cell lymphoma (B) and mantle cell lymphoma (C). The common transcription factor in Burkitt's lymphoma is E2F8. The common transcription factors in mantle cell lymphoma are PRDM1, E2F8, VDR and IRF4.