



Systematic Analysis of Competing Endogenous RNA Networks in Diffuse Large B-Cell Lymphoma and Hodgkin's Lymphoma

Juanjuan Kang^{1,2}, Pengcheng Yao¹, Qiang Tang³, Ying Wang¹, Yuwei Zhou¹ and Jian Huang^{1*}

¹ Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, China, ² Affiliated Foshan Maternity & Child Healthcare Hospital, Southern Medical University, Foshan, China, ³ Innovative Institute of Chinese Medicine and Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu, China

OPEN ACCESS

Edited by:

Hongwei Wang,
Sun Yat-sen University, China

Reviewed by:

Xiaopei Shen,
Fujian Medical University, China
Yongsheng Kevin Li,
Harbin Medical University, China

*Correspondence:

Jian Huang
hj@uestc.edu.cn

Specialty section:

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

Received: 23 July 2020

Accepted: 28 August 2020

Published: 30 September 2020

Citation:

Kang J, Yao P, Tang Q, Wang Y,
Zhou Y and Huang J (2020)
Systematic Analysis of Competing
Endogenous RNA Networks in Diffuse
Large B-Cell Lymphoma
and Hodgkin's Lymphoma.
Front. Genet. 11:586688.
doi: 10.3389/fgene.2020.586688

Lymphoma is a systemic malignancy, originating from the lymphatic system, which accounts for 3 to 4% of all tumors. There are two major subtypes of lymphoma, namely, diffuse large B-cell lymphoma (DLBCL) and Hodgkin's lymphoma (HL). Elucidation of the pathogenesis of these two lymphoma types is crucial for the identification of potential therapeutic targets. Compared with the corresponding knowledge of other diseases, the understanding of the regulatory networks involved in DLBCL and HL is relatively deficient. To address this, we comprehensively analyzed the mRNAs, lncRNAs, and miRNAs that were differentially expressed between normal and tumor samples of DLBCL and HL. In addition, functional enrichment analysis of the differentially expressed mRNAs was performed. We constructed two specific ceRNA networks of DLBCL and HL. The pathways enriched by dysregulated mRNAs in DLBCL and HL were mainly involved in immune responses, transcription process, and metabolism process. The ceRNA network analysis revealed that 45 ceRNAs were shared between the two ceRNA networks, including five pivotal lncRNAs (*MALAT1*, *CTBP1-AS*, *THUMPD3-AS*, *PSMA3-AS1*, and *NUTM2A-AS1*). In addition, we proposed a DLBCL survival risk model based on a DLBCL-specific network constructed by Lasso regression analysis. The model, which is based on eight mRNAs, exhibited excellent performance in regard to predicting outcomes in DLBCL patients, with a *p* value of 0.0017 and AUC of 0.9783. In summary, although the molecular mechanisms underlying tumorigenesis in DLBCL and HL were quite different, the same pivotal lncRNAs acted as key regulators. Our findings identify novel potential prognostic and therapeutic targets for DLBCL and HL.

Keywords: lymphoma, diffuse large B-cell lymphoma, Hodgkin's lymphoma, competing endogenous RNA, regulatory network

INTRODUCTION

Lymphoma is a systemic malignancy, arising in the lymphatic system, that accounts for about 3–4% of all tumors. It can be traditionally divided into two major pathological forms: Hodgkin's lymphoma (HL) and non-Hodgkin's lymphoma (NHL) (Mugnaini and Ghosh, 2016). Worldwide, nearly 510,000 new cases of NHL were estimated in 2018, with approximately 250,000 deaths, representing 2.8% and 2.6% of all tumors. At the same time, nearly 80,000 new cases (0.4% of all

tumors) and 26,000 deaths (0.3% of all tumors) were reported for HL (Bray et al., 2018). Diffuse large B-cell lymphoma (DLBCL), which originates in the germinal center, is the most prevalent subtype of NHL and accounts for 24% of new cases of this form (Liu and Barta, 2019). Even though investigations of DLBCL and HL have resulted in encouraging progress in many important areas, particularly in the identification of driver genes of DLBCL (Reddy et al., 2017) and an improved understanding of HL pathogenesis (Mathas et al., 2016), studies on the regulatory networks involved in the tumorigenesis of DLBCL and HL are lacking. Moreover, there are significant variations between DLBCL patients in regard to survival. Therefore, it is necessary to identify new diagnostic and prognostic biomarkers.

Advances in next-generation sequencing technologies have resulted in the discovery of non-coding RNAs, including micro RNAs (miRNAs, with a length of ~22 nucleotides) and long non-coding RNAs (lncRNAs, longer than 200 nt in length), as important regulatory molecules. Functional studies have shown that miRNAs and lncRNAs participate in regulating a broad range of biological processes, such as cell differentiation and development (Di Lisio et al., 2012; Fatica and Bozzoni, 2014). The changes in their expression are associated with many human pathologies including, but not limited to, lymphoma (Sole et al., 2017; Zhou et al., 2017). For example, the aberrant downregulation of miR-150 activates the PI3K-AKT pathway, inducing telomerase activation and immortalization of malignant lymphoma cells (Watanabe et al., 2011). Yan et al. found that the lncRNA *HOTAIR* acts as an oncogene whose knockdown leads to cell cycle arrest and the induction of apoptosis in DLBCL; further, its higher expression was found to be associated with poorer prognosis (Yan et al., 2016). Although the role of lncRNAs in DLBCL has been relatively well described (Zhou et al., 2017), their biological functions in HL have not been fully elucidated.

The competing endogenous RNA (ceRNA) hypothesis proposed by Salmena et al. (Salmena et al., 2011) provides a new angle on the potential functions of lncRNAs. According to this hypothesis, mRNA and lncRNA can cross talk with each other through the competitive binding of miRNA. On the basis of this hypothesis, ceRNA networks have been constructed to elucidate pathological mechanisms and identify biomarkers in various types of tumors such as squamous cell carcinoma (Zhou R.S. et al., 2019), myeloid leukemia (Sen et al., 2018), gastric cancer (Zhu et al., 2019), and so on (Xu et al., 2015, 2016; Li et al., 2018). Thus, a delineation of malignant lymphoma from the aspect of ceRNA networks is practicable and necessary.

In this study, we identified differentially expressed lncRNAs, miRNAs, and mRNAs involved in the pathogenesis of DLBCL and HL. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed to investigate the potential biological functions of dysregulated genes. Then, two ceRNA networks for the two lymphoma types were constructed. The shared ceRNA pairs across the two types of lymphoma were analyzed to derive new lncRNA biomarkers for DLBCL and HL. In addition, we constructed a survival risk model to stratify patients with DLBCL according to survival.

MATERIALS AND METHODS

Microarray Data

Two publicly available microarray expression profiling datasets consisting of gene expression data for DLBCL and HL were downloaded from the Gene Expression Omnibus (GEO) database. The DLBCL GSE56315 dataset (Dybkaer et al., 2015) and HL dataset of GSE25986 (Steidl et al., 2011) were obtained using the HG-U133 Plus 2.0 expression array from Affymetrix. The samples in the GSE56315 dataset were derived from 55 patients diagnosed with DLBCL and 33 B cells from normal human tonsil samples. The GSE25986 dataset contained data for five samples of both commonly used HL cell lines and reactive germinal center (GC) cell lines, respectively.

We searched the GEO database for all available miRNA expression profiling datasets related to DLBCL and HL and identified two relevant datasets: one dataset consisted of 45 DLBCL and 10 normal lymph node samples in frozen biopsies (accession number GSE42906) (Caramuta et al., 2013); the second dataset (accession number, GSE45264) contained data for eight normal lymph node samples and 14 classical HL lymph node samples in formalin-fixed paraffin-embedded tissue. The microarray expression profiling datasets are summarized in **Table 1**.

Data Processing

The “affy” package in R was used to process the raw CEL files of these two gene datasets. Probe names were annotated with gene IDs according to platform annotation files. After excluding probes without gene IDs or with more than one gene ID, the average values for multiple probes mapped to the same gene IDs were calculated. For miRNA datasets, we downloaded the express matrix directly for further analysis.

We identified differentially expressed RNAs by using the “samr” package in R. Genes with $|\log_2FC| \geq 1$ and $q < 0.01$ were considered significantly differentially expressed genes; $|\log_2FC| \geq 1$ and $q < 0.05$ for GSE45264 and only $q < 0.05$ for GSE42906 were the cutoff values for differentially expressed miRNAs.

lncRNA Identification Pipeline

In order to evaluate the expression pattern of lncRNAs in microarray data, we identified the probe sets mapped to lncRNAs from gene expression profiling using the following criteria: (1) We selected 44135 probes with Ensemble gene IDs. (2) Among these, probes annotated with “non-protein coding,” “non-coding,” or “antisense RNA” were saved. (3) To improve the accuracy of recognition, probes that were annotated with microRNA, miRNA, rRNA, tRNA, snRNA, or snoRNA were excluded. Finally, 1867 probes mapping to lncRNAs were identified. The genes mapped by these probes were lncRNA genes, and the remaining were mRNA genes.

Functional Enrichment Analysis

Gene Ontology and KEGG enrichment analyses were employed to determine the biological functions of the differentially expressed mRNAs in DLBCL and HL. We used Metascape

TABLE 1 | Summary of the microarray expression profiling datasets.

Date type	Accession number	Lymphoma type	Tumor samples	Normal samples
Gene expression profiling	GSE56315	DLBCL	55	33
	GSE25986	HL	5	5
miRNA expression profiling	GSE42906	DLBCL	45	10
	GSE45264	HL	14	8

(Zhou Y. et al., 2019)¹ to conduct functional enrichment analysis. Terms with p value < 0.01 , minimum count of 3, and enrichment factor > 1.5 were collected and grouped into clusters based on their membership similarities. The most statistically significant term within a cluster was chosen to represent the cluster. Heatmaps were constructed with the top 20 enriched GO biological processes and KEGG clusters to analyze the functional differences between differentially expressed mRNAs in DLBCL and HL.

Identification of miRNA Targets

Based on the ceRNA hypothesis, ceRNA pairs should have shared miRNA targets. The mRNA targets of dysregulated miRNAs were identified using ENCORI, a new version of starBase (Li et al., 2014). The targets should be predicted by at least four programs out of seven (PITA, RNA22, miRmap, DIANA-microT, miRanda, PicTar, and TargetScan). StarBase was used to identify lncRNA targets.

Construction of the ceRNA Network

There are two necessary conditions that potential lncRNA and mRNA pairs must meet for the construction of a ceRNA network.

Firstly, they should regulate the same miRNA, to a statistically significant extent. We used hypergeometric cumulative distribution function tests to identify candidate lncRNA and mRNA pairs. A p value was calculated for each pair using the formula,

$$p = 1 - F(k|A, M, N) = 1 - \sum_{i=0}^{k-1} \frac{\binom{M}{i} \binom{A-M}{N-i}}{\binom{A}{N}} \quad (k > 3) \quad (1)$$

where k is the number of target miRNAs shared with RNA₁ and RNA₂; M and N are the number of target miRNAs of RNA₁ and RNA₂, respectively; and A is the total number of dysregulated miRNAs. The Benjamini-Hochberg adjusted p value < 0.05 was used to filter out the potential lncRNA and mRNA pairs.

Secondly, the lncRNA and mRNA pairs should exhibit a co-expression pattern. For potential lncRNA and mRNA pairs, we calculated the Pearson correlation coefficient (R) based on their expression profiling; $R > 0.7$ were applied to select the candidate lncRNA and mRNA pairs.

Finally, Cytoscape 3.7.2 was used to visualize the DLBCL- and HL-specific ceRNA networks.

¹<http://metascape.org>

Collection of mRNAs, lncRNAs, and miRNAs Associated With DLBCL and HL

A list of published mRNAs, lncRNAs, and miRNAs related to lymphoma was collated from the state-of-the-art databases. A list of mRNAs associated with DLBCL and HL was collected from DisGeNET v2019 (Pinero et al., 2020). A total of 574 DLBCL-related and 645 HL-related mRNAs was identified by searching the database. A list of lncRNAs associated with the two types of lymphoma was derived from the MNDR v2.0 (Huang et al., 2019), LncRNADisease 2.0 (Bao et al., 2019), and Lnc2Cancer 2.0 (Gao et al., 2019) databases. In the three databases, 15 lncRNAs were related to DLBCL, while only one lncRNA, *PVT1*, was related to HL and also included in the list of lncRNAs associated with DLBCL (**Supplementary Table S1**). In addition, a list of miRNAs associated with the two types of lymphoma was obtained from HMDD v3.0 (Cui et al., 2018). A list of miRNAs containing 53 DLBCL- and 29 HL-related miRNA precursors was constructed.

Survival Analysis

To assess the prognostic characteristics of RNAs in the DLBCL-specific ceRNA network, we obtained survival and expression data for the DLBCL samples from the TCGA database and used the “survival” package of R to perform Kaplan-Meier curve analysis. Samples were divided into higher-expression and lower-expression groups based on the median expression level of each RNA. Then, Lasso regression analysis for the significant RNAs was performed to build the prognostic risk model. The performance of the model was evaluated by the area under the receiver operating characteristic curve (AUC) based on 5-year survival outcome.

RESULTS

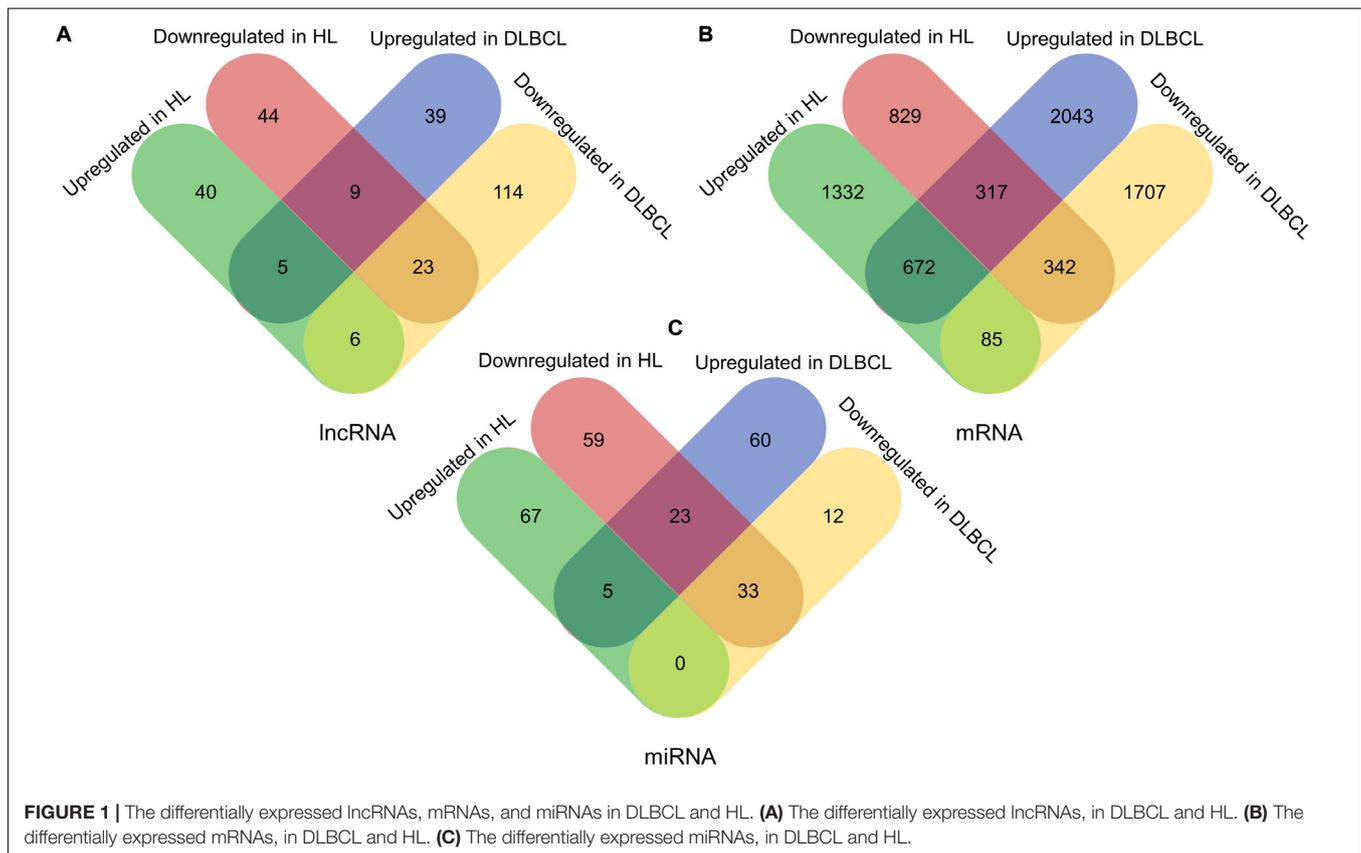
Identification of Differentially Expressed lncRNAs, mRNAs, and miRNAs

In total, 196 lncRNAs (143 downregulated and 53 upregulated lncRNAs) and 127 lncRNAs (76 downregulated and 51 upregulated lncRNAs) were significantly differentially expressed in DLBCL and HL, respectively. In addition, 3032 upregulated and 2134 downregulated mRNAs in DLBCL, and 2089 upregulated and 1488 downregulated mRNAs in HL, were found. Additionally, we identified 133 dysregulated miRNAs (45 downregulated and 88 upregulated) in DLBCL and 187 dysregulated (115 downregulated and 72 upregulated) miRNAs in HL. The dysregulated RNAs are summarized in **Table 2**.

TABLE 2 | Summary of dysregulated mRNA, lncRNA, and miRNA.

Lymphoma type	Date type	RNA type	Up*	Down*	Total
DLBCL	Gene expression profiling	mRNA	3032	2134	5166
		lncRNA	53	143	196
	miRNA expression profiling	miRNA	88	45	133
HL	Gene expression profiling	mRNA	2089	1488	3577
		lncRNA	51	76	127
	miRNA expression profiling	miRNA	72	115	187

*Up stands for the upregulated RNAs. Down represents the downregulated RNAs.



Among them, 43 lncRNAs, 1416 mRNAs, and 61 miRNAs were shared between the two types of lymphoma (**Figure 1**).

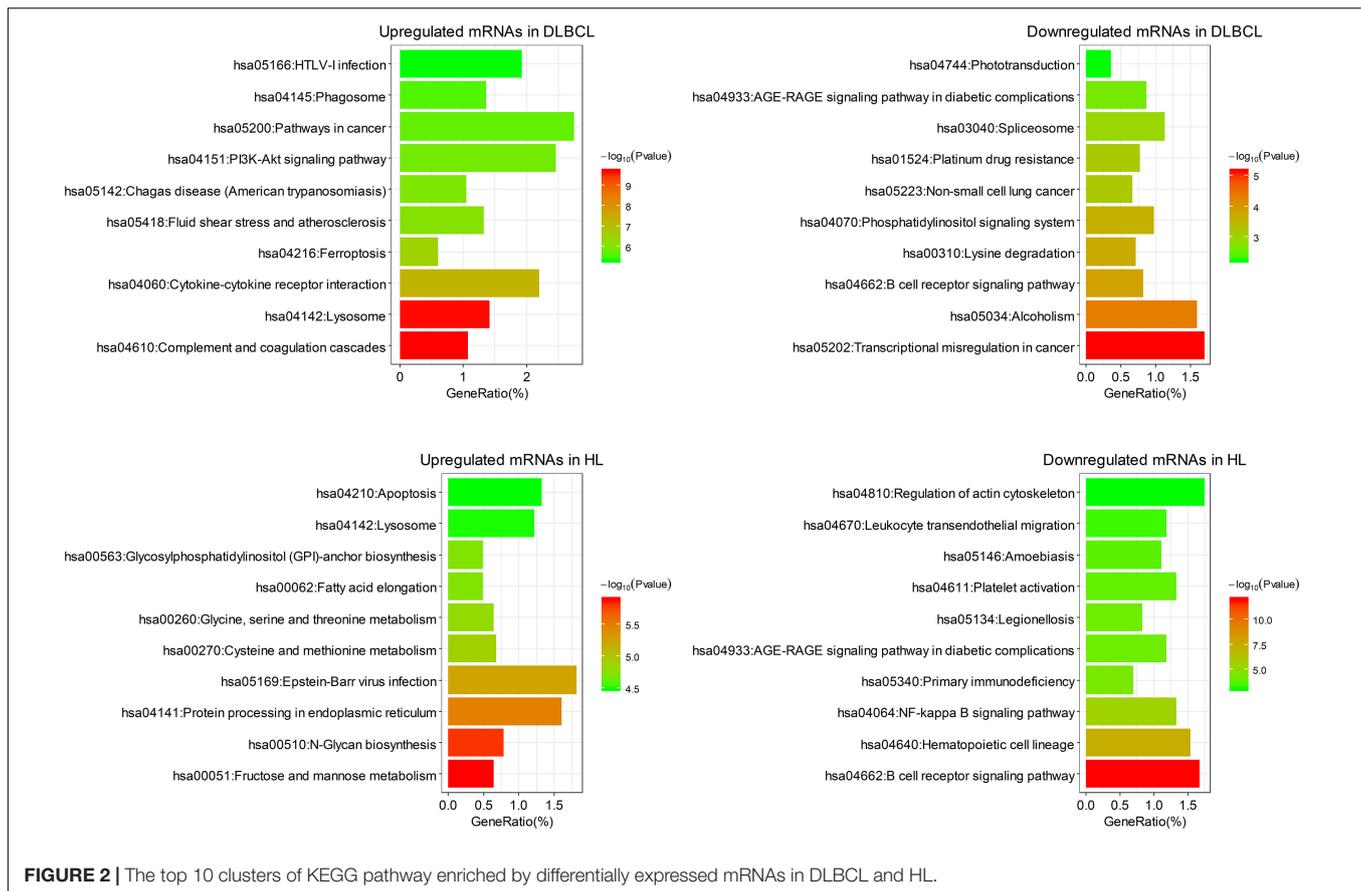
Compared with the list of published mRNAs, lncRNAs, and miRNAs related to lymphoma, 190 mRNAs and 24 miRNA precursors were dysregulated in DLBCL, and 164 mRNAs and 14 miRNA precursors were dysregulated in HL. Some typical biomarkers of lymphoma were included, such as the five miRNA biomarkers (miR-150, miR-21, miR-155, miR-17, and miR-146a) for distinguishing lymphoma from reactive lymphoid hyperplasia (Kang et al., 2020) and the T-cell lineage protein GATA3 (Atayar et al., 2005).

GO and KEGG Enrichment Analyses of Differentially Expressed mRNAs

Functional enrichment analysis of the differentially expressed mRNAs was performed. These mRNAs were enriched in

pathways previously reported to be associated with B cell lymphoma, such as the PI3K-Akt signaling pathway (Uddin et al., 2006; Yuan et al., 2019) and the NF- κ B signaling pathway (Weniger and Kuppers, 2016; **Figure 2**). In addition, infection with the Epstein-Barr virus has been shown to be highly correlated with pathogenesis of HL, and it affects the survival of subgroups of this form of lymphoma (Myriam et al., 2017). Consistent with this finding, the Epstein-Barr virus infection pathway was enriched in the upregulated mRNAs in HL. Interestingly, downregulated mRNAs in HL were enriched in the B-cell receptor signaling pathway, reflecting the defective B-cell program in classical HL (Kuppers, 2009).

Among the enriched pathways, the upregulated mRNAs in DLBCL were mainly related to parasitic or viral infection, such as cytokine-cytokine receptor interaction and parasitic infection, and the downregulated mRNAs in DLBCL were enriched in



transcription pathways. The upregulated mRNAs in HL were mainly involved in pathways associated with metabolism and biosynthesis, and the downregulated mRNAs in HL were enriched in signal transduction (Figure 2).

The enrichment results for GO biological process were similar to that for the KEGG pathway analysis. The upregulated mRNAs in DLBCL were mainly involved in myeloid leukocyte activation, cytokine production, response to wounding, etc. The downregulated mRNAs in DLBCL were mainly involved in transcriptional processes such as mRNA processing and RNA splicing and localization. The mRNAs upregulated in HL mainly participated in a large number of metabolic and biosynthetic processes, including carbohydrate derivative biosynthetic process, cellular amino acid metabolic process, and cofactor metabolic process, whereas the mRNAs that were downregulated in HL were mainly involved in biological processes related to the lymphocyte immune response (Figure 3).

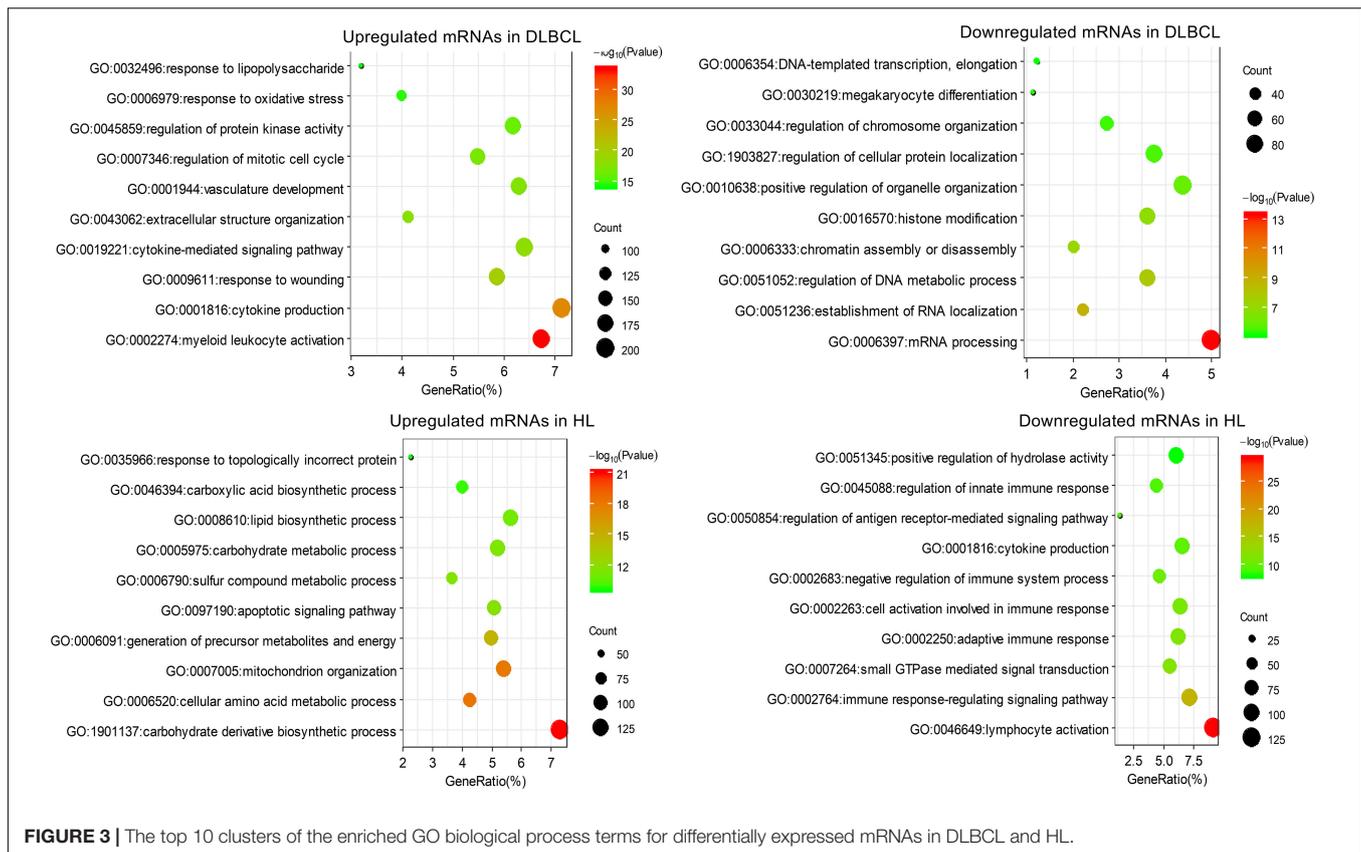
The results of functional enrichment analyses show that there was a significant difference between DLBCL and HL. However, hierarchical cluster analysis of the top 20 enriched pathways and biological processes revealed that the GO terms and KEGG pathways enriched in the mRNAs with the same direction of dysregulation were much more similar than those enriched by the miRNAs of opposite direction of dysregulation in DLBCL and HL (Supplementary Figure S1).

At the same time, we also analyzed the enrichment pathways and biological processes of commonly dysregulated

mRNAs in DLBCL and HL (Supplementary Figure S2). The mRNAs that were downregulated in both lymphoma subtypes mainly participated in the biological processes of B cell activation and immunity. The upregulated mRNAs were significantly enriched in metabolism and biosynthesis process, such as cofactor metabolic process, cellular acid metabolic process, and small-molecule catabolic process. The result indicated that the biological processes of metabolism and biosynthesis are simultaneously activated in DLBCL and HL, providing nutrition and energy for tumor cells to promote the continuous proliferation of tumor cells.

Topological Properties of the ceRNA Networks

After the filter process based on shared targeted miRNA and co-expression pattern as mentioned in Methods, we constructed DLBCL- and HL-specific ceRNA networks separately. The DLBCL-specific ceRNA network comprised 24 lncRNA nodes, 358 mRNA nodes, and 1237 edges (Figure 4A). The average node degree of mRNA was 3.46 (95% CI: 3.14–3.77): that is, one mRNA had an indirect regulatory relationship with approximately three lncRNAs. The HL-specific ceRNA network contained 516 lncRNA nodes, 33 mRNA nodes, and 1050 edges (Figure 4B). Similar to the DLBCL-specific ceRNA network, one mRNA in the HL-specific ceRNA network may interact with two lncRNAs,



and the average node degree of mRNAs was 1.97 (95% CI: 1.87–2.07).

True biological networks possess the important characteristic of being scale-free, which means that nodes with a higher degree are far fewer than those with a lower degree. We evaluated the node degree distributions of the two ceRNA networks. As shown in **Figure 4C**, the node degree distributions in both DLBCL- and HL-specific networks had long tails and fit the power law distribution, with $R^2 > 0.95$, indicating that the two ceRNA networks were scale-free and similar to true biological networks.

Functional Analysis of HL and DLBCL ceRNA Networks

In the DLBCL-specific ceRNA network, 18 out of 358 mRNAs were covered by the list of DLBCL-related genes, and 30 out of 533 mRNAs in the HL ceRNA network were reported to be associated with HL (**Supplementary Table S2**). In the DLBCL-specific ceRNA network, the mRNAs recorded in the DLBCL-related gene list had a higher degree, although a Mann–Whitney U test indicated that this tendency was not significant ($p > 0.05$). A similar tendency of miRNAs associated with HL was not found in HL-specific ceRNA network (**Supplementary Figure S3**).

Among the lncRNAs related to DLBCL and HL, only *MALAT1* and *GAS5* were involved in the DLBCL ceRNA network ($p = 0.027$), with a degree of 147 and 28, respectively. Interestingly, *HOTAIR*, which was associated with DLBCL, was involved in the HL ceRNA network, with a degree of

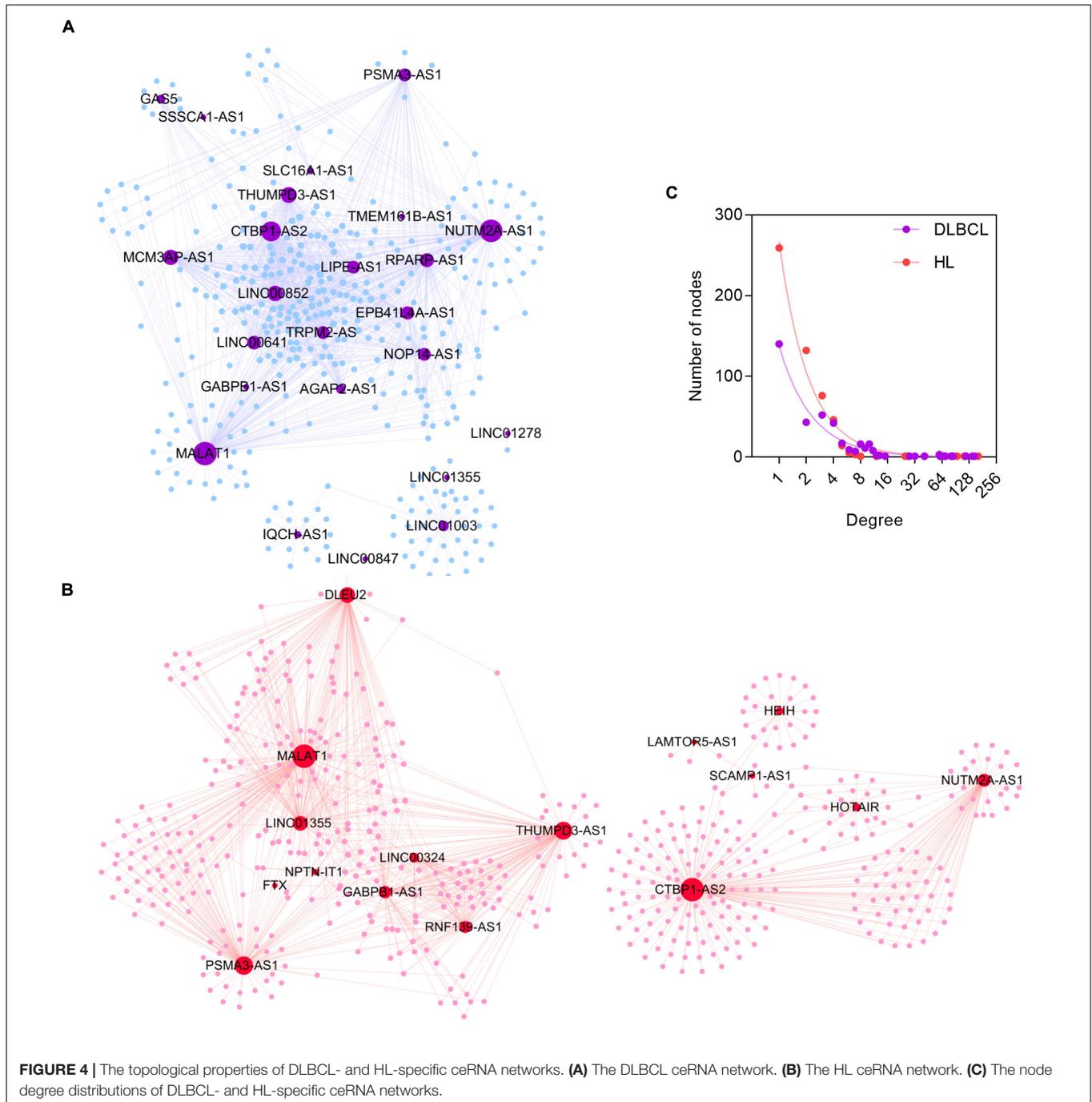
27. By contrast, *PVT1* was the only lncRNA recorded to be associated with HL, however, it was not included in the two ceRNA networks.

The pivot lncRNA, *MALAT1*, with the largest degree in both ceRNA networks, regulates the proliferation, invasion, and migration potential of metastatic tumor cells by sponging miR-195 in DLBCL (Isin et al., 2014; Wang et al., 2019). However, its regulatory function in HL has not yet been elucidated. lncRNAs associated with lymphoma, especially with HL, have been relatively less explored. The lncRNAs in the ceRNA networks may present promising candidate regulatory molecules.

Network Analysis of ceRNA Subnetwork

As shown in **Figure 5**, five lncRNAs (*MALAT1*, *CTBPI-AS*, *THUMPD3-AS*, *PSMA3-AS1*, and *NUTM2A-AS1*) were shared by DLBCL and HL ceRNA networks, which accounted for 20.83 and 31.25% of all lncRNAs corresponding to DLBCL and HL. By contrast, only 45 ceRNA pairs shared and occupied a minority of ceRNA networks (3.64% in DLBCL, 4.29% in HL). These results suggest that DLBCL and HL share a large proportion of lncRNAs involved in gene regulation, but that these lncRNAs have differing mRNA targets in their respective networks. The expression heatmaps of the shared ceRNAs are shown in **Supplementary Figure S4**.

Except for *MALAT1*, there was no evidence that the other four lncRNAs were associated with DLBCL or HL. However, it has been reported in previous studies that some of these lncRNAs



are closely related to the development of other tumor types (Cui and Geng, 2019; Xu et al., 2019). For example, the overexpressed *CTBPI1-AS1* inhibits cell apoptosis and promotes tumorigenesis by sponging miRNA-940 in breast cancer (Cui and Geng, 2019).

Moreover, we searched the 35 mRNAs included in the ceRNA subnetwork in DisGeNET database and found that the majority of those mRNAs (31/35) were associated with the neoplastic process except for *TMX3*, *SNX29*, *FNBP4*, and *RAB30*. Among them, *FBXO11*, *HDAC9*, and *GNA13* were biomarkers for DLBCL. Furthermore, *HDAC9* was

a biomarker for HL. These findings indicated that the lncRNAs and mRNAs in the ceRNA subnetwork represent common key molecules for the pathogenic processes involved in DLBCL and HL.

Effect of mRNAs in the DLBCL-Specific ceRNA Network on Survival

Because only the clinical information and gene expression data of DLBCL patients are available from the TCGA database, we

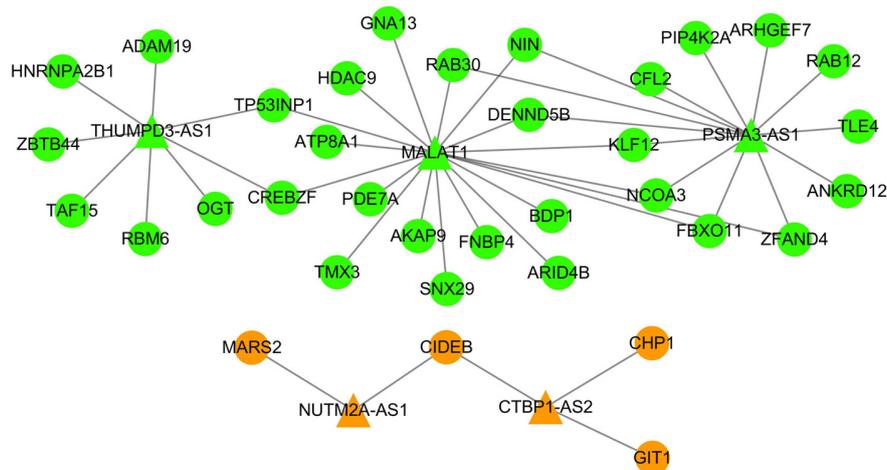


FIGURE 5 | The ceRNA subnetwork between DLBCL and HL. The triangle nodes represent lncRNAs, and the circle nodes indicate mRNAs. The green nodes have the same direction of dysregulation in DLBCL and HL, while the orange nodes have the opposite regulation direction in these two types of lymphoma.

identified candidate biomarkers predictive of survival outcomes based on the DLBCL-specific ceRNA network. Kaplan–Meier curve analysis showed that 10 mRNAs (ZNF3, PPP1R2, RABEP1, PEX11B, SIK1, KXD1, TSG101, TLE4, CDV3, and STX16) were significantly differentially expressed ($p < 0.05$) (Figure 6). The groups with higher expression of ZNF3, PPP1R2, PEX11B, KXD1, TSG101, and CDV3 had better overall survival, while elevated expression of RABEP1, SIK1, TLE4, and STX16 was associated with poor prognosis.

We constructed a DLBCL survival risk model based on Lasso regression analysis. For each sample, a risk score analysis was conducted based on the model: risk score = $-0.3905 \times \text{ZNF3} + 0.1006 \times \text{RABEP1} + 0.0961 \times \text{SIK1} - 0.259 \times \text{KXD1} - 0.0835 \times \text{TSG101} + 0.2224 \times \text{TLE4} - 0.0524 \times \text{CDV3} + 0.0840 \times \text{STX16}$. The distribution of survival risk score and the relationship between survival time and risk score are shown in Figures 7A,B. The higher the risk score, the shorter the overall survival duration and the higher the death rate. Then, the DLBCL samples were divided into higher- and lower-risk score groups based on the median risk score (value = -8.9363). Results of Kaplan–Meier tests revealed that the low-risk group was associated with better overall survival ($p = 0.0017$) (Figure 7C). ROC analysis demonstrated that the AUC was 0.9783 (95% CI: 0.9300–1.0000) (Figure 7D). These results indicate that the prognostic model showed potentially good performance in terms of predicting the survival outcomes of DLBCL patients.

DISCUSSION

The ceRNA language is not dependent on a single protein-encoding mRNA, but instead focuses on the communication of mRNAs with other RNA categories, such as lncRNAs, in the transcriptome. Thus, the hypothesis offers a new layer of understanding into the complexity of tumorigenesis and

progression. In this study, we systematically constructed DLBCL- and HL-specific ceRNA regulatory networks, which provide a theoretical basis for analyzing the regulatory functions of ceRNAs in DLBCL and HL.

We identified 10 mRNAs (ZNF3, PPP1R2, RABEP1, PEX11B, SIK1, KXD1, TSG101, TLE4, CDV3, and STX16) that were significantly correlated with overall survival in patients with DLBCL. Interestingly, although none of the 10 prognostic mRNAs in this work have been reported to be associated with lymphoma, some of them act as prognostic biomarkers in other malignant tumors (Yau et al., 2010; Andres et al., 2013; Liu et al., 2014; Wang et al., 2016; Xiao et al., 2018). For example, Yau et al. (Yau et al., 2010) have demonstrated that ZNF3 is a prognostic candidate gene in breast cancer. It has been reported that the overexpression of RABEP1, another biomarker of breast cancer, is predictive of reduced survival (Andres et al., 2013). Liu et al. (Liu et al., 2014) showed that TSG101 is an independent factor of poor prognosis in gallbladder cancer. Wang et al. (Wang et al., 2016) suggested that TLE4, which plays an important role in the development and progression of colorectal cancer, is a potential prognostic biomarker in this cancer. In addition, higher CDV3 expression is associated with poor survival rate in hepatocellular carcinoma (Xiao et al., 2018). Therefore, the 10 identified genes might represent promising prognostic biomarkers for DLBCL.

We were only able to identify a few DLBCL-related lncRNAs. The reasons for this may be as follows: (1) Some probes mapping to lncRNAs were filtered because they also mapped to other genes: for instance, the 225155_at probe mapped to SNHG5, SNORD50A, and SNORD50B; therefore, we removed this probe. (2) Because the array platform was designed to detect coding RNAs, some key lncRNAs, such as lincRNA-p21 and PANDAR, were not included. (3) Some lncRNAs were not significantly differentially expressed between lymphoma and control samples, such as TUG1 and PEG10. Despite these limitations, the study provided a general landscape of the ceRNA regulation models in DLBCL and HL.

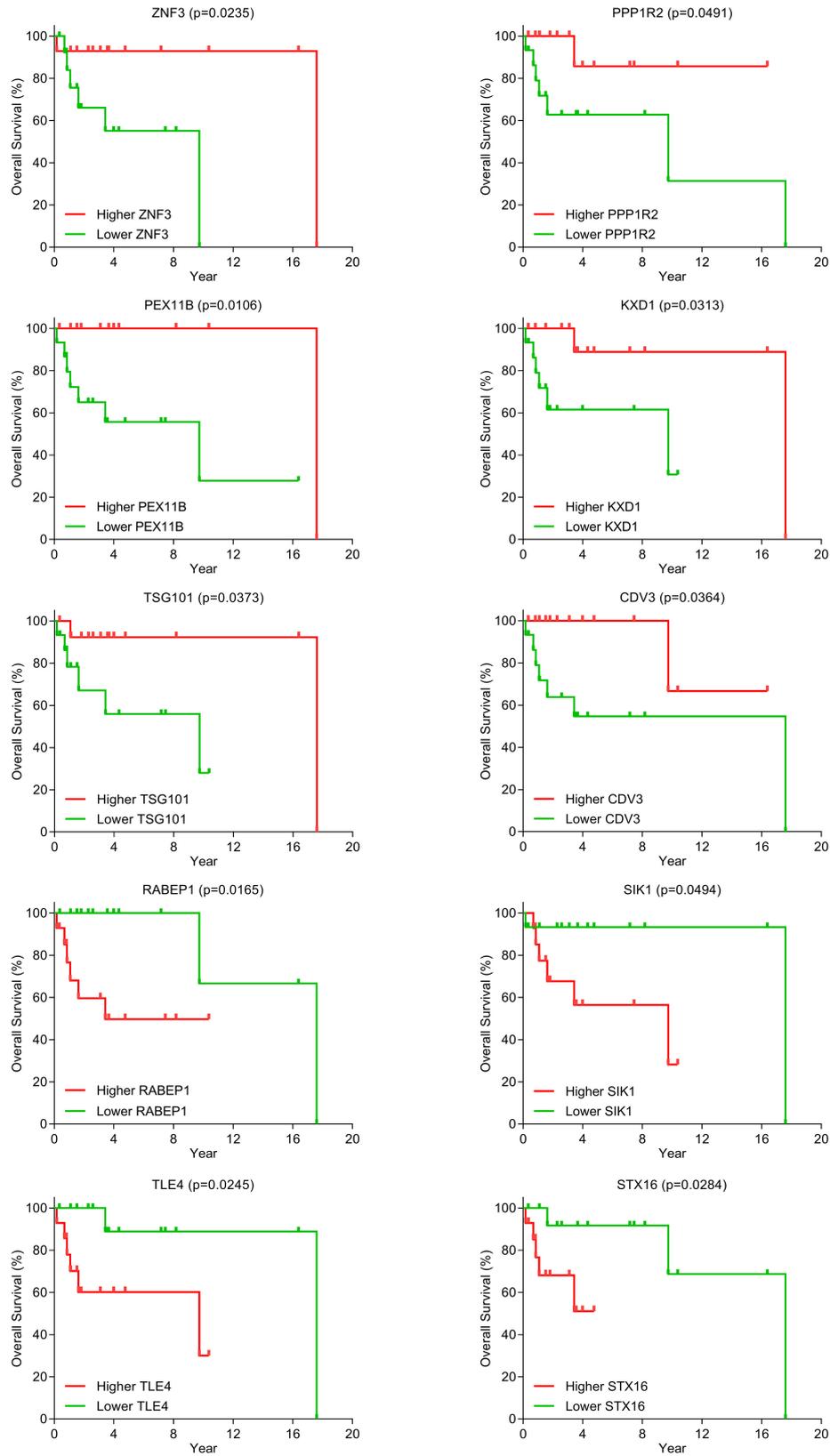
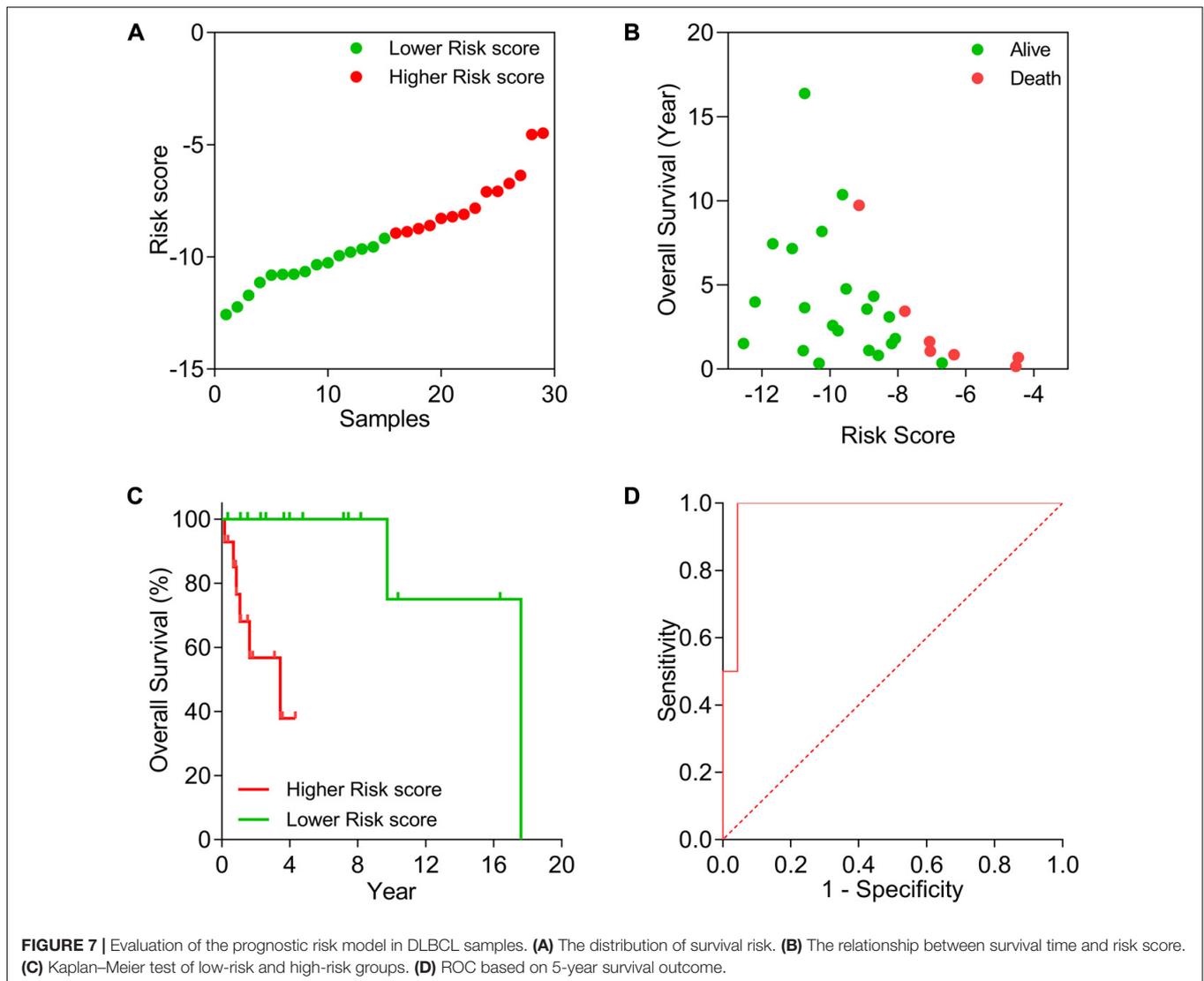


FIGURE 6 | Kaplan–Meier curve analysis for mRNAs in the DLBCL ceRNA network.



In summary, in the present study, we constructed DLBCL- and HL-specific ceRNA networks and identified indirect interactions between lncRNAs and mRNAs. The results provide a method to identify lncRNA and mRNA biomarkers and further elucidate the mechanisms underlying the pathogenesis of DLBCL and HL.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JH conceived and designed this study. JK and QT wrote the manuscript. PY and YW analyzed the data. YZ drew the

figures. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (61571095, 62071099, and 61901130) and the China Postdoctoral Science Foundation Grant (2019M653369).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.586688/full#supplementary-material>

REFERENCES

- Andres, S. A., Brock, G. N., and Wittliff, J. L. (2013). Interrogating differences in expression of targeted gene sets to predict breast cancer outcome. *BMC Cancer* 13:326. doi: 10.1186/1471-2407-13-326
- Atayar, C., Poppema, S., Blokzijl, T., Harms, G., Boot, M., and Van Den Berg, A. (2005). Expression of the T-cell transcription factors, GATA-3 and T-bet, in the neoplastic cells of Hodgkin lymphomas. *Am. J. Pathol.* 166, 127–134.
- Bao, Z., Yang, Z., Huang, Z., Zhou, Y., Cui, Q., and Dong, D. (2019). LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases. *Nucleic Acids Res.* 47, D1034–D1037.
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424.
- Caramuta, S., Lee, L., Ozata, D. M., Akcakaya, P., Georgii-Hemming, P., Xie, H., et al. (2013). Role of microRNAs and microRNA machinery in the pathogenesis of diffuse large B-cell lymphoma. *Blood Cancer J.* 3:e152.
- Cui, H. B., and Geng, C. Z. (2019). Molecular mechanisms of long chain non-coding RNA CTBP1-AS in regulation of invasion and migration of breast cancer cells. *J. Biol. Regul. Homeost. Agents* 33, 773–785.
- Cui, T., Zhang, L., Huang, Y., Yi, Y., Tan, P., Zhao, Y., et al. (2018). MNDR v2.0: an updated resource of ncRNA-disease associations in mammals. *Nucleic Acids Res.* 46, D371–D374.
- Di Liso, L., Martinez, N., Montes-Moreno, S., Piris-Villaespesa, M., Sanchez-Beato, M., and Piris, M. A. (2012). The role of miRNAs in the pathogenesis and diagnosis of B-cell lymphomas. *Blood* 120, 1782–1790. doi: 10.1182/blood-2012-05-402784
- Dybkaer, K., Bogsted, M., Falgreen, S., Bodker, J. S., Kjeldsen, M. K., Schmitz, A., et al. (2015). Diffuse large B-cell lymphoma classification system that associates normal B-cell subset phenotypes with prognosis. *J. Clin. Oncol.* 33, 1379–1388. doi: 10.1200/jco.2014.57.7080
- Fatica, A., and Bozzoni, I. (2014). Long non-coding RNAs: new players in cell differentiation and development. *Nat. Rev. Genet.* 15, 7–21. doi: 10.1038/nrg3606
- Gao, Y., Wang, P., Wang, Y., Ma, X., Zhi, H., Zhou, D., et al. (2019). Lnc2Cancer v2.0: updated database of experimentally supported long non-coding RNAs in human cancers. *Nucleic Acids Res.* 47, D1028–D1033.
- Huang, Z., Shi, J., Gao, Y., Cui, C., Zhang, S., Li, J., et al. (2019). HMDD v3.0: a database for experimentally supported human microRNA-disease associations. *Nucleic Acids Res.* 47, D1013–D1017.
- Isin, M., Ozgur, E., Cetin, G., Erten, N., Aktan, M., Gezer, U., et al. (2014). Investigation of circulating lncRNAs in B-cell neoplasms. *Clin. Chim. Acta* 431, 255–259. doi: 10.1016/j.cca.2014.02.010
- Kang, J., Yu, S., Lu, S., Xu, G., Zhu, J., Yan, N., et al. (2020). Use of a 6-miRNA panel to distinguish lymphoma from reactive lymphoid hyperplasia. *Signal Transduct. Target Ther.* 5:2.
- Kuppers, R. (2009). The biology of Hodgkin's lymphoma. *Nat. Rev. Cancer* 9, 15–27.
- Li, J. H., Liu, S., Zhou, H., Qu, L. H., and Yang, J. H. (2014). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 42, D92–D97.
- Li, Y., Huo, C., Lin, X., and Xu, J. (2018). Computational identification of cross-talking ceRNAs. *Adv. Exp. Med. Biol.* 1094, 97–108. doi: 10.1007/978-981-13-0719-5_10
- Liu, Y., and Barta, S. K. (2019). Diffuse large B-cell lymphoma: 2019 update on diagnosis, risk stratification, and treatment. *Am. J. Hematol.* 94, 604–616.
- Liu, Z., Yang, Z., Liu, D., Li, D., Zou, Q., Yuan, Y., et al. (2014). TSG101 and PEG10 are prognostic markers in squamous cell/adenosquamous carcinomas and adenocarcinoma of the gallbladder. *Oncol. Lett.* 7, 1128–1138. doi: 10.3892/ol.2014.1886
- Mathas, S., Hartmann, S., and Kuppers, R. (2016). Hodgkin lymphoma: pathology and biology. *Semin. Hematol.* 53, 139–147.
- Mugnaini, E. N., and Ghosh, N. (2016). Lymphoma. *Prim. Care* 43, 661–675.
- Myriam, B. D., Sonia, Z., Hanene, S., Teheni, L., and Mounir, T. (2017). Prognostic significance of Epstein-Barr virus (EBV) infection in Hodgkin lymphoma patients. *J. Infect. Chemother.* 23, 121–130. doi: 10.1016/j.jiac.2016.09.004
- Pinero, J., Ramirez-Anguita, J. M., Sauch-Pitarch, J., Ronzano, F., Centeno, E., Sanz, F., et al. (2020). The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.* 48, D845–D855.
- Reddy, A., Zhang, J., Davis, N. S., Moffitt, A. B., Love, C. L., Waldrop, A., et al. (2017). Genetic and functional drivers of diffuse large B cell lymphoma. *Cell* 171, 481–494.e415.
- Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011). A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 146, 353–358. doi: 10.1016/j.cell.2011.07.014
- Sen, K., Bhattacharyya, D., Sarkar, A., Das, J., Maji, N., Basu, M., et al. (2018). Exploring the major cross-talking edges of competitive endogenous RNA networks in human Chronic and acute myeloid leukemia. *Biochim. Biophys. Acta Gen. Subj.* 1862, 1883–1892. doi: 10.1016/j.bbagen.2018.06.002
- Sole, C., Larrea, E., Di Pinto, G., Tellaetxe, M., and Lawrie, C. H. (2017). miRNAs in B-cell lymphoma: Molecular mechanisms and biomarker potential. *Cancer Lett.* 405, 79–89. doi: 10.1016/j.canlet.2017.07.020
- Steidl, C., Shah, S. P., Woolcock, B. W., Rui, L., Kawahara, M., Farinha, P., et al. (2011). MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. *Nature* 471, 377–381. doi: 10.1038/nature09754
- Uddin, S., Hussain, A. R., Siraj, A. K., Manogaran, P. S., Al-Jomah, N. A., Moorji, A., et al. (2006). Role of phosphatidylinositol 3'-kinase/AKT pathway in diffuse large B-cell lymphoma survival. *Blood* 108, 4178–4186. doi: 10.1182/blood-2006-04-016907
- Wang, Q. M., Lian, G. Y., Song, Y., Huang, Y. F., and Gong, Y. (2019). LncRNA MALAT1 promotes tumorigenesis and immune escape of diffuse large B cell lymphoma by sponging miR-195. *Life Sci.* 231:116335. doi: 10.1016/j.lfs.2019.03.040
- Wang, S. Y., Gao, K., Deng, D. L., Cai, J. J., Xiao, Z. Y., He, L. Q., et al. (2016). TLE4 promotes colorectal cancer progression through activation of JNK/c-Jun signaling pathway. *Oncotarget* 7, 2878–2888. doi: 10.18632/oncotarget.6694
- Watanabe, A., Tagawa, H., Yamashita, J., Teshima, K., Nara, M., Iwamoto, K., et al. (2011). The role of microRNA-150 as a tumor suppressor in malignant lymphoma. *Leukemia* 25, 1324–1334. doi: 10.1038/leu.2011.81
- Weniger, M. A., and Kuppers, R. (2016). NF-kappaB deregulation in Hodgkin lymphoma. *Semin. Cancer Biol.* 39, 32–39. doi: 10.1016/j.semcancer.2016.05.001
- Xiao, H., Zhou, B., Jiang, N., Cai, Y., Liu, X., Shi, Z., et al. (2018). The potential value of CDV3 in the prognosis evaluation in Hepatocellular carcinoma. *Genes Dis.* 5, 167–171. doi: 10.1016/j.gendis.2018.01.003
- Xu, H., Han, H., Song, S., Yi, N., Qian, C., Qiu, Y., et al. (2019). Exosome-transmitted PSMA3 and PSMA3-AS1 promote proteasome inhibitor resistance in multiple myeloma. *Clin. Cancer Res.* 25, 1923–1935. doi: 10.1158/1078-0432.ccr-18-2363
- Xu, J., Feng, L., Han, Z., Li, Y., Wu, A., Shao, T., et al. (2016). Extensive ceRNA-ceRNA interaction networks mediated by miRNAs regulate development in multiple rhesus tissues. *Nucleic Acids Res.* 44, 9438–9451.
- Xu, J., Li, Y., Lu, J., Pan, T., Ding, N., Wang, Z., et al. (2015). The mRNA related ceRNA-ceRNA landscape and significance across 20 major cancer types. *Nucleic Acids Res.* 43, 8169–8182. doi: 10.1093/nar/gkv853
- Yan, Y., Han, J., Li, Z., Yang, H., Sui, Y., and Wang, M. (2016). Elevated RNA expression of long noncoding HOTAIR promotes cell proliferation and predicts a poor prognosis in patients with diffuse large B cell lymphoma. *Mol. Med. Rep.* 13, 5125–5131. doi: 10.3892/mmr.2016.5190
- Yau, C., Esserman, L., Moore, D. H., Waldman, F., Sninsky, J., and Benz, C. C. (2010). A multigene predictor of metastatic outcome in early stage hormone receptor-negative and triple-negative breast cancer. *Breast Cancer Res.* 12:R85.
- Yuan, T., Zhang, F., Zhou, X., Li, Y., Zhang, Y., Xu, Y., et al. (2019). Inhibition of the PI3K/AKT signaling pathway sensitizes diffuse large B-cell lymphoma cells to treatment with proteasome inhibitors via suppression of BAG3. *Oncol. Lett.* 17, 3719–3726.
- Zhou, M., Zhao, H., Xu, W., Bao, S., Cheng, L., and Sun, J. (2017). Discovery and validation of immune-associated long non-coding RNA biomarkers associated with clinically molecular subtype and prognosis in diffuse large B cell lymphoma. *Mol. Cancer* 16:16.

- Zhou, R. S., Zhang, E. X., Sun, Q. F., Ye, Z. J., Liu, J. W., Zhou, D. H., et al. (2019). Integrated analysis of lncRNA-miRNA-mRNA ceRNA network in squamous cell carcinoma of tongue. *BMC Cancer* 19:779. doi: 10.1186/1471-2407-13-779
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10:1523.
- Zhu, Y., Sun, X., Lin, J., Zhang, T., Liu, X., and Shen, X. (2019). Investigating potential molecular mechanisms of carcinogenesis and genes as biomarkers for prognosis of gastric cancer based on integrated bioinformatics analysis. *Pathol. Oncol. Res.* 25, 1125–1133.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Kang, Yao, Tang, Wang, Zhou and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.