A Random Walk-Based Method to Identify Candidate Genes Associated With Lymphoma

Minjie Sheng[1†], Haiying Cai[1†], Qin Yang[1], Jing Li[1], Jian Zhang[2,3,4,5,6]* and Lihua Liu[1]*

[1]Department of Ophthalmology, Yangpu Hospital, School of Medicine, Tongji University, Shanghai, China, [2]Department of Ophthalmology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [3]Shanghai Key Laboratory of Ocular Fundus Diseases, Shanghai, China, [4]Shanghai Engineering Center for Visual Science and Photomedicine, Shanghai, China, [5]National Clinical Research Center for Eye Diseases, Shanghai, China, [6]Shanghai Engineering Center for Precise Diagnosis and Treatment of Eye Diseases, Shanghai, China

Lymphoma is a serious type of cancer, especially for adolescents and elder adults, although this malignancy is quite rare compared with other types of cancer. The cause of this malignancy remains ambiguous. Genetic factor is deemed to be highly associated with the initiation and progression of lymphoma, and several genes have been related to this disease. Determining the pathogeny of lymphoma by identifying the related genes is important. In this study, we presented a random walk-based method to infer the novel lymphoma-associated genes. From the reported 1,458 lymphoma-associated genes and protein–protein interaction network, raw candidate genes were mined by using the random walk with restart algorithm. The determined raw genes were further filtered by using three screening tests (i.e., permutation, linkage, and enrichment tests). These tests could control false-positive genes and screen out essential candidate genes with strong linkages to validate the lymphoma-associated genes. A total of 108 inferred genes were obtained. Analytical results indicated that some inferred genes, such as *RAC3*, *TEC*, *IRAK2/3/4*, *PRKCE*, *SMAD3*, *BLK*, *TXK*, *PRKCQ*, were associated with the initiation and progression of lymphoma.

Keywords: lymphoma, random walk with restart algorithm, protein-protein interaction network, enrichment theory, permutation test

1 INTRODUCTION

Lymphocytes are a group of effective immune-associated cells and include two famous cell subtypes, namely, T and B lymphocytes (Mesquita Júnior et al., 2010). Lymphocytes play an irreplaceable role in humoral (B lymphocytes) and cellular (T lymphocytes) immune responses (Mesquita Júnior et al., 2010) to fight against infectious virus or bacteria and endogenous malignant cancer cells. However, even as immune cells, lymphocytes can also be malignant when transformed by exogenous stimulations, such as benzene (Guo et al., 2021) or the human immunodeficiency virus (Wang et al., 2021), and endogenous factors, such as family history (Chang et al., 2005) and aging (Parsonnet and Isaacson, 2004). Cancers that begin in the immune-associated lymphocytes are generally summarized as lymphoma (Armitage et al., 2017).

Lymphoma can be divided into two groups, namely, Hodgkin lymphoma (Mathas et al., 2016) and non-Hodgkin lymphoma (Shankland et al., 2012) according to the existence of reed-sternberg cells. Lymphoma with and without detectable reed-sternberg cells are generally regarded as Hodgkin

and non-Hodgkin lymphoma (Shankland et al., 2012; Mathas et al., 2016). Both kinds of lymphoma are quite rare compared with other cancer subtypes, such as lung and liver cancers (Siegel et al., 2021). Approximately 9,000 new cases and 1,000 deaths have been reported in 2020 by the American Cancer Society (Siegel et al., 2021). Contrary to other cancer subtypes, the risk of lymphoma is quite high for adolescents and elder adults (older than 55 years old) but relatively low for adults in their 30 and 40 s (Wilson et al., 2012). This characteristic reflects a typical age-associated disease susceptibility distribution pattern for lymphoma.

However, the cause of non-Hodgkin lymphoma remains unknown. Several reports have associated some viruses, such as *T cell leukemia lymphoma virus* (Zhang et al., 2017b), Epstein-Barr virus (Vockerodt et al., 2015), and hepatitis B virus (Ren et al., 2018), and bacteria, such as *Helicobacter pylori* (specific for gastric MALT lymphoma) (Salar, 2019), with the pathogenesis of non-Hodgkin lymphoma. For the Hodgkin lymphoma, the risk is increased in people with Human Immunodeficiency Virus and Epstein-Barr virus infections (Grewal et al., 2018). For both types of lymphoma, family history has long been considered as an important risk factor, and genetic background has also been highly associated with the initiation and progression of this cancer (Skibola et al., 2007). According to a review for the genetic susceptibility to lymphoma, seven groups of genes with the following functions are involved in the pathogenesis of lymphoma as follows: DNA repair [e.g., *NHEJ* (Lieber et al., 2010) and *DSBR* (Shen et al., 2006)]; carbon metabolism [e.g., *MTHFR* (He et al., 2014) and *MTR* (Ruiz-Cosano et al., 2013)]; immune regulation [e.g., *TNF*, *IL4*, and *IL4R* (Mottok and Steidl, 2015)]; oxidative stress [e.g., *NOS2A* (Fabisiewicz et al., 2013) and *MPO* (*Sugiyama et al., 2017*)]; energy regulation [e.g., *LEP* and *GHRL* (Argyrou et al., 2019)]; hormone production [e.g., *CYP17A1* (*Skibola et al., 2005*)]; xenobiotic [e.g., *GSTT1* (Yang et al., 2014)]; and cell cycle regulation [e.g., *CCND1* (Mohanty et al., 2019)]. The association of these genes with the pathogenesis of lymphoma has been established. Thus, the initiation and progression of lymphoma are precisely regulated by genetic background. Finding the genetic factors for lymphoma is therefore one of the most effective and straight-forward approaches to reveal the pathogenesis of such complex diseases.

Traditionally, the identification of lymphoma associated genes depends on several classical analytic approaches and methods. For familial lymphoma cases, family pedigree analyses based on Sanger sequencing (Liu et al., 2014), microarray analyses (Hedvat et al., 2002), next generation target sequencing and whole genome wide sequencing (Hung et al., 2018) on large familial samples are major traditional methods to identify potential pathogenic lymphoma associated genes or variants. As for sporadic lymphoma cases, to validate the molecular abnormalities associated with lymphoma, Southern blot analyses (Sangueza et al., 1992), *in situ* hybridization (Quintanilla-Martinez et al., 2009) and quantitative real-time PCR (Takatori et al., 2021) are also applied to explore and confirm specific distribution of genetic abnormal arrangement associated with lymphoma. There are three advantages for traditional analyses: 1) Firstly, the accuracy of traditional experimental analyses is generally higher than statistical bioinformatics analyses; 2) Secondly, independent repeat experimental analyses are easier to perform at experimental level to validate the identified potential biomarkers; 3) Thirdly, results from experimental analyses were easier to be used for further functional exploration. However, the disadvantages of experiment-based analyses are also obvious, including 1) Clinical samples are difficult to obtain, and results from experimental animals are not always consistent with human beings; 2) Low reproducibility caused by more potential unrelated variables; 3) High cost and time consuming.

Due to the high cost and time consuming of traditional experiment-based methods, we introduced a random walk-based computational method to recognize the novel candidate lymphoma-associated genes in this study. The reported lymphoma-associated genes, as summarized from the DisGeNET database (Piñero et al., 2015), and the protein–protein interaction (PPI) network collected in STRING (Szklarczyk et al., 2015), were fed into the random walk with restart (RWR) algorithm (Kohler et al., 2008; Macropol et al., 2009) to determine the raw candidate genes. Then, three screening tests (i.e., permutation, linkage, and enrichment tests) were performed to control false-positive genes and select the essential candidate genes that had strong linkages to validate the lymphoma-associated genes. The analytical results indicated that several of these genes had associations with the initiation and progression of lymphoma.

2 MATERIALS AND METHODS

2.1 Lymphoma-Associated Genes

In this study, we summarized all lymphoma-associated genes from the DisGeNET database (https://www.disgenet.org/, version 7.0, accessed in March 2021) (Piñero et al., 2015), one of the largest publicly available databases of human genes and gene associated with human diseases. A total of 1,548 genes have been associated with the pathogenesis of lymphoma in the past 5 years (**Supplementary Table S1**). Then, the related proteins of these genes were picked up and further mapped onto their Ensembl IDs. The IDs not in the PPI network as described in **Section 2.2** were excluded, resulting in 1,375 Ensembl IDs. Based on these proteins, as represented by Ensembl IDs, we set up a computational method to discover other proteins, which were highly related to these proteins. The genes encoding the identified proteins were regarded to be highly associated with the pathogenesis of lymphoma.

2.2 PPI Network

This study proposed a random walk-based method to investigate the lymphoma-associated genes. A network should be employed to execute the random walk algorithm. In recent years, the PPI network is widely used to study various problems related to proteins or genes (Ng et al., 2010; Hu et al., 2011a; Hu et al., 2011b; Zhang et al., 2016; Cai et al., 2017; Zhang et al., 2019; Zhang and Chen, 2020; Zhao et al., 2020; Gao et al., 2021). Thus,

we used the structure of one PPI network and mined new candidate genes related to lymphoma based on the validated ones.

We employed the PPI network collected in STRING (version 10, https://www.string-db.org/) (Szklarczyk et al., 2015). The file "9,606. protein.links.v10. txt.gz" was retrieved, which consisted of 4,274,001 PPIs covering 19,247 human proteins. A PPI included two proteins, encoded by Ensembl IDs. Furthermore, one confidence score with range between 1 and 999 was assigned to each PPI. Such score can comprehensively measure the associations of proteins, because it integrates several scores, including "neighborhood", "fusion", "cooccurence", "coexpression", "experimental", "database", and "textmining" scores, which assess the associations of proteins from various aspects of proteins, such as close neighborhood in (prokaryotic) genomes, gene fusion, occurrence across species, gene coexpression, scientific literature description, etc. The higher the score was, the stronger the PPI would be. Accordingly, a PPI network was constructed by taking 19,247 human proteins as nodes, and two nodes were connected by an edge if and only if their corresponding proteins could constitute a PPI with a confidence score larger than zero. In this case, each edge in the PPI network represented a PPI. To further indicate the strength of edges, a weight was assigned to each edge, which was the confidence score of the corresponding PPI.

2.3 RWR Algorithm

Based on the validated lymphoma-associated genes, we employed the RWR algorithm (Kohler et al., 2008; Macropol et al., 2009; Chen et al., 2018a; Chen et al., 2018b; Liang et al., 2020) to discover the novel genes in the PPI network. Such algorithm simulated a walker starting from one node or a set of nodes (these nodes are called seed nodes) in one network, and such walker randomly moved in the network to deliver probabilities on the seed nodes to other nodes. Given a network and m seed nodes, the RWR algorithm initialized a probability vector P_0, with the same length as the node number of the network. One node corresponded to one component. The component of one seed node was defined as $1/m$, and other components were set to 0. The RWR algorithm repeatedly updated such vector as follows:

$$P_{t+1} = (1-r)A^T P_t + rP_0, \tag{1}$$

where A denotes the column-wise normalized adjacency matrix; and r stands for the restarting probability, which was set to 0.8 as used in some previous studies (Yuan and Lu, 2017; Zhang et al., 2017a; Zhang et al., 2017c; Chen et al., 2018a). When the vectors P_{t+1} and P_t were close enough, i.e., $\|P_{t+1} - P_t\|_{L_1} < 10^{-6}$, the update procedure was stopped. P_{t+1} was selected as the outcome of the RWR algorithm. Based on such vector, the probability of each node, which was obtained from the seed nodes, was determined. Evidently, a node assigned with a high probability may have strong associations with the seed nodes.

In this study, the RWR program developed by Li and Patra (Li and Patra, 2010) was adopted. Although this program is designed for heterogeneous networks, we set the jumping probability to zero and selected seed nodes in one part of the network so that probabilities was transmitted only in one part of the network.

Here, the 1,375 Ensembl IDs were set as the seed nodes. According to the outcome of the RWR algorithm, the nodes with high probabilities were picked up. These nodes could be the novel candidate genes related to lymphoma.

2.4 Screening Tests

Some candidate genes mined by the RWR algorithm were highly related to the structure of the PPI network, and these genes could induce some extreme cases. For example, some nodes may easily receive high probabilities regardless of which nodes were seed nodes. On the other hand, the candidate genes with strong associations with validated ones had higher likelihood to be novel genes related to lymphoma. In view of this, we designed three screening tests to further filter the essential candidate genes.

Permutation test. As previously mentioned, the structure of the network may influence the outcome of the RWR algorithm. To control such influence, the permutation test was adopted. We first randomly constructed 1,000 node sets, with sizes the same as that of the seed node set. The nodes in each set were fed into the RWR algorithm as the seed nodes. Then, each candidate gene selected by the RWR algorithm was also assigned a probability. After all node sets had been tested by the RWR algorithm, all candidate genes received 1,000 probabilities, and their means and standard deviations were computed. Accordingly, the Z-score was computed for each candidate gene g as follows:

$$Z-score(g) = \frac{Pro(g) - ProM(g)}{ProSTD(g)} \tag{2}$$

where $Pro(g)$ denotes the probability of the candidate gene g obtained by using the actual seed node set; and $ProM(g)$ and $ProSTD(g)$ represent the mean and standard deviation of the probabilities, respectively, which were obtained by 1,000 randomly produced node sets. In statistics, the value of 1.96 is a widely accepted threshold of the Z-score to denote statistical significance. Thus, we could select candidate genes with Z-score>1.96. These genes were assigned much higher probability based on the actual seed node set than those based on randomly produced node sets, indicating their significant association with lymphoma.

Linkage test. The permutation test could decrease the influence of the PPI network. However, some candidate genes with weak or even without association with the validated genes may be included. Thus, we employed the linkage test. Several studies have reported that interacting proteins are more likely to have similar functions (Ng et al., 2010; Hu et al., 2011a; Hu et al., 2011b; Chen et al., 2016; Cai et al., 2017; Li et al., 2018; Zhang and Chen, 2020; Zhu et al., 2021). Considering the strength of the PPI, proteins that could comprise a PPI with a higher confidence score were more likely to exhibit similar functions. Hence, we adopted the interaction information mentioned in **Section 2.2** to design the linkage test. For two proteins p_1 and p_2, their confidence score was defined as $Q(p_1, p_2)$. The maximum linkage score (MLS) was computed for each candidate gene g as follows:

$$MLS(g) = Max\{Q(g, g'): g' \text{ is a validated lymphoma associated gene}\} \tag{3}$$

Candidate genes with high MLSs evidently had high probabilities to be novel lymphoma-associated genes and thus should be selected. Considering that 900 is the threshold of the highest confidence in STRING, we adopted such value to screen the essential candidate genes, i.e., candidate genes with MLSs no less than 900 were selected.

Enrichment test. Finally, we used the enrichment test to evaluate the importance of the candidate genes with functional terms, including gene ontology (GO) terms and KEGG pathways. The validated lymphoma-associated genes should have some similar functional terms. If a candidate gene had functional terms that were also shared by one validated lymphoma-associated gene, such gene had a high probability to be a novel lymphoma-associated gene. The enrichment score (Carmona-Saez et al., 2007) was adopted to evaluate the linkage between one gene and one GO term or KEGG pathway. The enrichment score between a gene g and one GO term/KEGG pathway F was computed as follows:

$$ES(g, F) = -\log_{10}\left(\sum_{k=m}^{n} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\right), \quad (4)$$

where N and M denote the number of human genes and genes annotated by F, respectively; n represents the number of interacting genes of g reported in STRING; and m represents the number of common genes that can be interacted with g and was annotated by F. For a gene g, enrichment scores to all GO terms and KEGG pathways were put into a vector $V(g)$. The associations of two genes g and g' could be evaluated according to their vectors as follows:

$$\Phi(g, g') = \frac{V(g) \cdot V(g')}{\|V(g)\| \cdot \|V(g')\|}, \quad (5)$$

Similar to MLS, we could further calculate the maximum enrichment score (MES) for each candidate gene g, which could be computed as follows:

$$MES(g) = Max\{\Phi(g, g'): g' \text{ is a validated lymphoma associated gene}\} \quad (6)$$

A candidate gene assigning a high MES had a high probability to be a novel lymphoma-associated gene. We set the threshold 0.98 to select important candidate genes.

2.5 Functional Enrichment Analyses on Candidate Genes
To reveal the biological meaning behind the candidate genes identified by the random walk-based method, the functional enrichment analyses were performed, which was implemented by the R package *topGO* (https://bioconductor.org/packages/topGO/, v2.42.0) (Alexa and Rahnenführer, 2009). To conduct such analyses, identified genes were regarded as gene of interest and all available human genes were termed as background. The *p*-value threshold was set as 0.001 to identify significant enrichment results.

3 RESULTS

We propose a random walk-based method to discover novel lymphoma-associated genes. The whole process is illustrated in **Figure 1**.
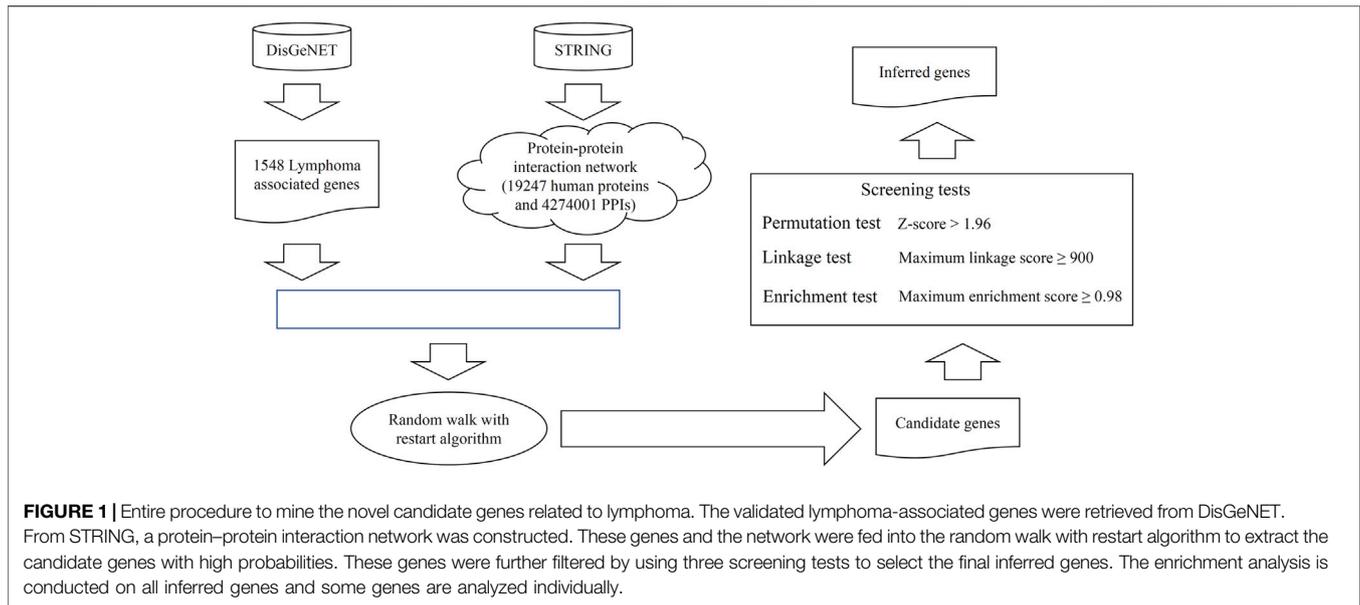
3.1 Results of the Random Walk-Based Method
The RWR algorithm was first performed on the PPI network with the proteins of lymphoma-associated genes as seed nodes. A probability was assigned to each node in the network to indicate its associations with the seed nodes. Nodes with probabilities no less than 10^{-5} were picked up, and their corresponding proteins were extracted. Thus, 4,962 proteins were obtained and are listed in **Supplementary Table S2**. The permutation test assigned a Z-score to each protein, and the scores are also listed in **Supplementary Table S2**. Proteins with Z-scores>1.96 were selected, resulting in 1,144 proteins. Afterward, these proteins were fed into the linkage test. Each protein was assigned an MLS, which is also provided in **Supplementary Table S2**. A total of 986 proteins were with MLSs no less than 900 and were selected. Finally, the enrichment test was performed to evaluate the importance of the remaining proteins. An MES was computed for each protein, and the results are listed in **Supplementary Table S2**. After setting the threshold of MES to 0.98, 108 proteins were obtained, which are the first 108 proteins in **Supplementary Table S2**. Their corresponding genes were selected and deemed to have strong associations with lymphoma. These genes are provided in **Supplementary Table S3**. In the following text, these genes were termed as inferred genes.

3.2 Associations Between Inferred Genes and Validated Genes
To indicate the reliability of the inferred genes, we conducted the following investigations. For each inferred gene, the number of its interacting lymphoma-associated genes with confidence scores no less than 900 was counted and is shown in a box plot (**Figure 2**). Some inferred genes have numerous interacting lymphoma-associated genes with confidence scores no less than 900, indicating their high relation to lymphoma. The average number of interacting lymphoma-associated genes with high confidence scores was 18.88 inferred genes, occupying 81.48%, can interact with more than five lymphoma-associated genes with high confidence score (≥900). These results implied that some hidden lymphoma-associated genes may be included in the inferred genes.

3.3 Enrichment Analysis on Inferred Genes
Of the 108 inferred genes, we conducted functional enrichment analysis on them. Thirteen GO terms were identified with significant *p*-value less than 0.001, including eight biological processes (BP) terms, four molecular function (MF) terms and one cellular component (CC) term. Detailed information of these thirteen GO terms and their *p*-values were illustrated in **Figure 3**. In **Section 4.2**. some discussions were performed.

FIGURE 1 | Entire procedure to mine the novel candidate genes related to lymphoma. The validated lymphoma-associated genes were retrieved from DisGeNET. From STRING, a protein–protein interaction network was constructed. These genes and the network were fed into the random walk with restart algorithm to extract the candidate genes with high probabilities. These genes were further filtered by using three screening tests to select the final inferred genes. The enrichment analysis is conducted on all inferred genes and some genes are analyzed individually.

4 DISCUSSION

From the random walk-based method, we identified a group of inferred genes that may be functionally associated with the initiation and progression of lymphoma. This section conducted some discussions to confirm their associations with lymphoma.
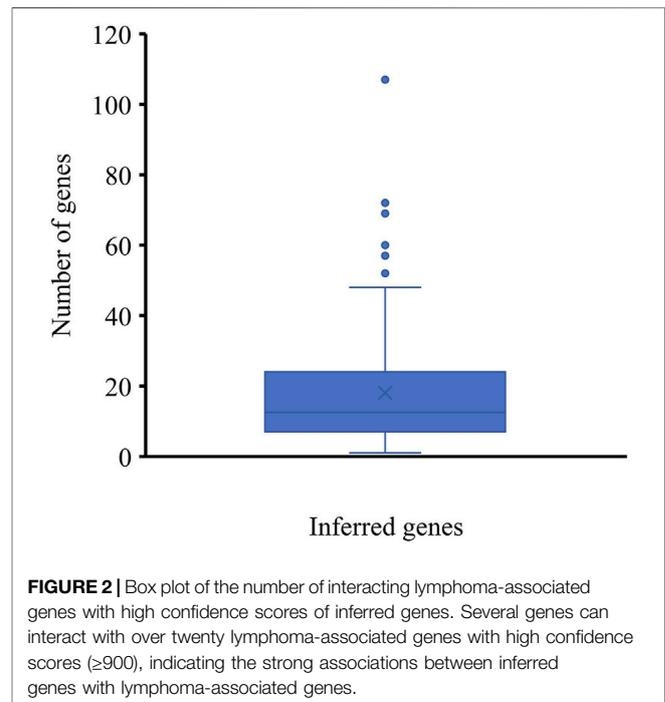
4.1 Individual Analysis on Some Inferred Genes

According to some publications, we found reliable literatures that supported the contribution of some inferred genes on lymphoma, and these genes are listed in **Table 1**.

The first gene is *RAC3* (ENSP00000304283), which had been associated with B-cell lymphoma. Early in 2006, researchers from France confirmed that the absence of *RAC3* can trigger the initiation and progression of B-cell lymphoma (Coste et al., 2006), reflecting the potential association between *RAC3* and lymphoma.

The next gene is *TEC* (ENSP00000370912). In 2004, *TEC* has been shown to mediate the abnormal proliferation and apoptosis of lymphoma cells (Zilberman et al., 2004). In 2015, another member of the *TEC* family, *BTK* has been shown to be an effective biomarker for Hodgkin and B cell non-Hodgkin lymphoma (Watson et al., 1970).

IRAK2 (ENSP00000256458), as the next predicted gene, has been reported to contain multiple significant variants associated with lymphoma through interactions with Toll-like receptors (Wang et al., 2014). In 2020, researchers from the University of North Carolina have validated that *IRAK2*-associated signaling pathway participates in the initiation and progression of lymphoma primarily triggered by the herpes virus (Seltzer et al., 2020). *IRAK4* (ENSP00000390651) is also a participant in the IRAK signaling pathway, which is essential for the pathogenesis of lymphoma. Therefore, predicting such gene

FIGURE 2 | Box plot of the number of interacting lymphoma-associated genes with high confidence scores of inferred genes. Several genes can interact with over twenty lymphoma-associated genes with high confidence scores (≥900), indicating the strong associations between inferred genes with lymphoma-associated genes.

(*IRAK4*) as another lymphoma biomarker is quite reasonable. Similarly, another component of the IRAK signaling pathway, *IRAK3* (ENSP00000261233), has also been identified, validating the reliability of our results.

PRKCE (ENSP00000306124) is the next predicted gene. According to recent publications, such gene is associated with lymphoma at different omic levels. In 2006, a methylation analyses on the small B-cell lymphoma showed that *PRKCE* is a specific methylation biomarker for different clinical outcomes and prognosis of small B-cell lymphoma (Rahmatpanah et al., 2006). Further studies on transcriptomics profiling also
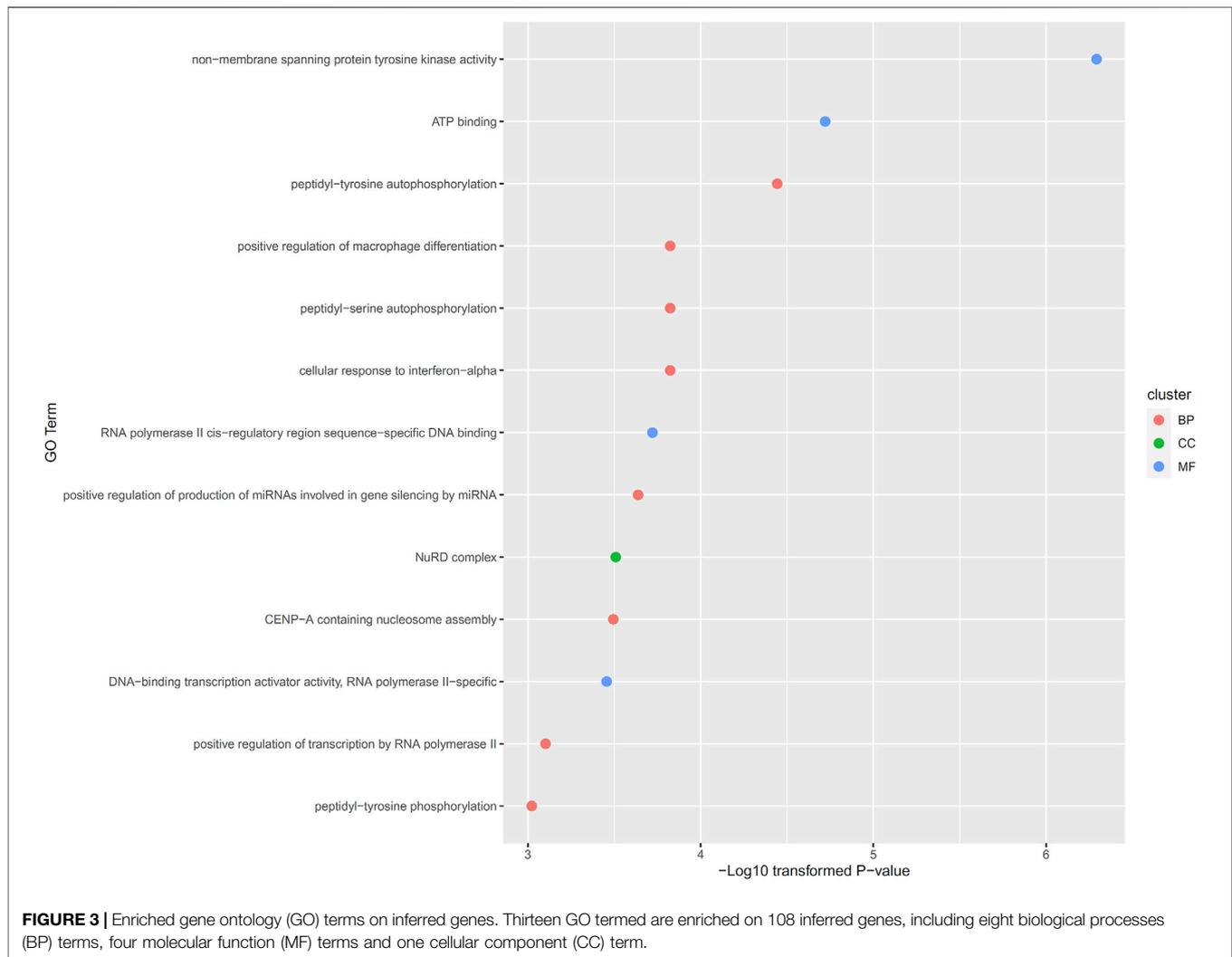
FIGURE 3 | Enriched gene ontology (GO) terms on inferred genes. Thirteen GO termed are enriched on 108 inferred genes, including eight biological processes (BP) terms, four molecular function (MF) terms and one cellular component (CC) term.

TABLE 1 | Some important inferred lymphoma-associated genes.

Ensembl id	Gene symbol	Description	Probability	Z-score	Maximum linkage score	Maximum enrichment score	Reference
ENSP00000304283	RAC3	Rac Family Small GTPase 3	9.700E−05	5.0457	998	0.9984	Coste et al. (2006)
ENSP00000370912	TEC	Tec Protein Tyrosine kinase	2.900E−05	3.2291	990	0.9979	Watson et al. (1970; Zilberman et al. (2004)
ENSP00000256458	IRAK2	Interleukin 1 Receptor Associated kinase 2	3.530E−05	3.3867	999	0.9976	Wang et al. (2014), Seltzer et al. (2020)
ENSP00000390651	IRAK4	Interleukin 1 Receptor Associated kinase 4	4.190E−05	4.3253	999	0.9967	Wang et al. (2014), Seltzer et al. (2020)
ENSP00000306124	PRKCE	Protein kinase C Epsilon	4.260E−05	3.9254	984	0.9967	Rahmatpanah et al. (2006), Wang et al. (2012)
ENSP00000261233	IRAK3	Interleukin 1 Receptor Associated kinase 3	3.540E−05	3.4101	999	0.9965	Wang et al. (2014), Seltzer et al. (2020)
ENSP00000332973	SMAD3	SMAD Family Member 3	8.280E−05	7.1221	999	0.9965	Park et al. (2001), Nakahata et al. (2010)
ENSP00000259089	BLK	BLK Proto-Oncogene, Src Family Tyrosine kinase	4.630E−05	6.2082	983	0.9964	Petersen et al. (2014)
ENSP00000264316	TXK	TXK Tyrosine kinase	2.890E−05	2.6307	915	0.9963	Liu et al. (2020)
ENSP00000263125	PRKCQ	Protein kinase C Theta	4.750E−05	4.8540	999	0.9963	Rosenwald et al. (2003)

confirmed that *PRKCE* is a specific biomarker to identify follicular lymphoma, one of the major subtypes of non-Hodgkin lymphoma (Wang et al., 2012), reflecting the specific association between *PRKCE* and lymphoma.

SMAD3 (ENSP00000332973), as the next predicted biomarker, has been associated with lymphoma by multiple independent publications. In 2001, *SMAD3* and its homolog, *SMAD4*, have been shown to mediate the expression of autoimmune antibodies during B-cell lymphoma (Park et al., 2001). In 2010, associations between T-cell lymphoma and *SMAD4* have also been revealed (Nakahata et al., 2010). Both T-linkage and B-linkage lymphoma have been associated with *SMAD4* or related pathways, implying the specific role of *SMAD4* during the initiation and progression of lymphoma. Other inferred genes, such as *BLK* (ENSP00000259089) (Petersen et al., 2014), *TXK* (ENSP00000264316) (Liu et al., 2020), and *PRKCQ* (ENSP00000263125) (Rosenwald et al., 2003), have also been associated with lymphoma.

Thus, some inferred genes can be validated to be associated with lymphoma-related biological processes, confirming that the inferred genes discovered in this study were quite reliable.

4.2 Analysis of Enrichment Results on Inferred Genes

As described in **Section 3.3**, thirteen GO terms were identified, which were enriched by 108 inferred genes. Generally, these GO terms should be associated with the pathogenesis of lymphoma. The enriched GO terms can be further divided into two groups: transcription regulation associated GO terms and immune associated GO terms. There are multiple enriched terms associated with RNA polymerase II (RNA polymerase II *cis*-regulatory region sequence-specific DNA binding, DNA binding transcription activator activity, RNA polymerase II-specific and positive regulation of transcription by RNA polymerase II). RNA polymerase II has been widely reported to be associated with the pathogenesis of lymphoma (Kawahata et al., 1983; Devaiah et al., 2012). As for another group of GO terms, there are multiple immune responses associated GO terms, including positive regulation of macrophage differentiation and cellular response

to interferon-alpha. According to recent publications, macrophage differentiation (Kant et al., 2013; Arlt et al., 2020) and interferon-alpha (Hermine et al., 2002) associated immune responses have both been reported to be associated with the pathogenesis of lymphoma.

5 CONCLUSION

In this study, a random walk-based computational method was proposed to determine the novel lymphoma-associated genes. Based on the powerful RWR algorithm and three screening tests, 108 inferred genes were obtained. The analytical results showed that some of these genes (*RAC3*, *TEC*, *IRAK2/3/4*, *PRKCE*, *SMAD3*, *BLK*, *TXK*, *PRKCQ*) could be novel lymphoma-associated genes. These findings may give new insights to investigate lymphoma and improve the understanding on the pathogeny of lymphoma.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.disgenet.org/.

AUTHOR CONTRIBUTIONS

JZ and LL designed the study. MS and QY performed the experiments. HC and JL analyzed the results. MS and HC wrote the manuscript. All authors contributed to the research and reviewed the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.792754/full#supplementary-material

REFERENCES

Alexa, M., and Adamson, A. (2009). Interpolatory point Set Surfaces-Convexity and Hermite Data. *ACM Trans. Graph.* 28, 1–10. doi:10.1145/1516522.1516531

Argyrou, C., Hatziagapiou, K., Theodorakidou, M., Nikola, O. A., Vlahopoulos, S., and Lambrou, G. I. (2019). The Role of Adiponectin, LEPTIN, and Ghrelin in the Progress and Prognosis of Childhood Acute Lymphoblastic Leukemia. *Leuk. Lymphoma* 60, 2158–2169. doi:10.1080/10428194.2019.1569230

Arlt, A., Bonin, F., Rehberg, T., Perez-Rubio, P., Engelmann, J. C., Limm, K., et al. (2020). High CD206 Levels in Hodgkin Lymphoma-educated Macrophages Are Linked to Matrix-remodeling and Lymphoma Dissemination. *Mol. Oncol.* 14, 571–589. doi:10.1002/1878-0261.12616

Armitage, J. O., Gascoyne, R. D., Lunning, M. A., and Cavalli, F. (2017). Non-Hodgkin Lymphoma. *The Lancet* 390, 298–310. doi:10.1016/s0140-6736(16)32407-2

Cai, Y.-D., Zhang, Q., Zhang, Y.-H., Chen, L., and Huang, T. (2017). Identification of Genes Associated with Breast Cancer Metastasis to Bone on a Protein-Protein Interaction Network with a Shortest Path Algorithm. *J. Proteome Res.* 16, 1027–1038. doi:10.1021/acs.jproteome.6b00950

Carmona-Saez, P., Chagoyen, M., Tirado, F., Carazo, J. M., and Pascual-Montano, A. (2007). GENECODIS: a Web-Based Tool for Finding Significant Concurrent Annotations in Gene Lists. *Genome Biol.* 8, R3. doi:10.1186/gb-2007-8-1-r3

Chang, E. T., Smedby, K. E., Hjalgrim, H., Porwit-Macdonald, A., Roos, G., Glimelius, B., et al. (2005). Family History of Hematopoietic Malignancy and Risk of Lymphoma. *J. Natl. Cancer Inst.* 97, 1466–1474. doi:10.1093/jnci/dji293

Chen, L., Liu, T., and Zhao, X. (2018a). Inferring Anatomical Therapeutic Chemical (ATC) Class of Drugs Using Shortest Path and Random Walk with Restart Algorithms. *Biochim. Biophys. Acta (Bba) - Mol. Basis Dis.* 1864, 2228–2240. doi:10.1016/j.bbadis.2017.12.019

Chen, L., Xing, Z., Huang, T., Shu, Y., Huang, G., and Li, H.-P. (2016). Application of the Shortest Path Algorithm for the Discovery of Breast Cancer-Related Genes. *Cbio* 11, 51–58. doi:10.2174/1574893611666151119220024

Chen, L., Zhang, Y.-H., Zhang, Z., Huang, T., and Cai, Y.-D. (2018b). Inferring Novel Tumor Suppressor Genes with a Protein-Protein Interaction Network

and Network Diffusion Algorithms. *Mol. Ther. - Methods Clin. Dev.* 10, 57–67. doi:10.1016/j.omtm.2018.06.007

Coste, A., Antal, M. C., Chan, S., Kastner, P., Mark, M., O'malley, B. W., et al. (2006). Absence of the Steroid Receptor Coactivator-3 Induces B-Cell Lymphoma. *EMBO J.* 25, 2453–2464. doi:10.1038/sj.emboj.7601106

Devaiah, B. N., Lewis, B. A., Cherman, N., Hewitt, M. C., Albrecht, B. K., Robey, P. G., et al. (2012). BRD4 Is an Atypical Kinase that Phosphorylates Serine2 of the RNA Polymerase II Carboxy-Terminal Domain. *Proc. Natl. Acad. Sci.* 109, 6927–6932. doi:10.1073/pnas.1120422109

Fabisiewicz, A., Pacholewicz, K., Paszkiewicz-Kozik, E., Walewski, J., and Siedlecki, J. A. (2013). Polymorphisms of DNA Repair and Oxidative Stress Genes in B-Cell Lymphoma Patients. *Biomed. Rep.* 1, 151–155. doi:10.3892/br.2012.31

Gao, J., Hu, B., and Chen, L. (2021). A Path-Based Method for Identification of Protein Phenotypic Annotations. *Cbio* 16, 1214–1222. doi:10.2174/1574893616666210531100035

Grewal, R., Irimie, A., Naidoo, N., Mohamed, N., Petrushev, B., Chetty, M., et al. (2018). Hodgkin's Lymphoma and its Association with EBV and HIV Infection. *Crit. Rev. Clin. Lab. Sci.* 55, 102–114. doi:10.1080/10408363.2017.1422692

Guo, H., Ahn, S., and Zhang, L. (2021). Benzene-associated Immunosuppression and Chronic Inflammation in Humans: a Systematic Review. *Occup. Environ. Med.* 78, 377–384. doi:10.1136/oemed-2020-106517

He, J., Liao, X.-Y., Zhu, J.-H., Xue, W.-Q., Shen, G.-P., Huang, S.-Y., et al. (2014). Association of MTHFR C677T and A1298C Polymorphisms with Non-hodgkin Lymphoma Susceptibility: Evidence from a Meta-Analysis. *Sci. Rep.* 4, 6159. doi:10.1038/srep06159

Hedvat, C. V., Hegde, A., Chaganti, R. S. K., Chen, B., Qin, J., Filippa, D. A., et al. (2002). Application of Tissue Microarray Technology to the Study of Non-hodgkin's and Hodgkin's Lymphoma. *Hum. Pathol.* 33, 968–974. doi:10.1053/hupa.2002.127438

Hermine, O., Allard, I., Lévy, V., Arnulf, B., Gessain, A., and Bazarbachi, A. (2002). A Prospective Phase II Clinical Trial with the Use of Zidovudine and Interferon-Alpha in the Acute and Lymphoma Forms of Adult T-Cell Leukemia/lymphoma. *Hematol. J.* 3, 276–282. doi:10.1038/sj.thj.6200195

Hu, L., Huang, T., Liu, X.-J., and Cai, Y.-D. (2011a). Predicting Protein Phenotypes Based on Protein-Protein Interaction Network. *PLoS One* 6, e17668. doi:10.1371/journal.pone.0017668

Hu, L., Huang, T., Shi, X., Lu, W.-C., Cai, Y.-D., and Chou, K.-C. (2011b). Predicting Functions of Proteins in Mouse Based on Weighted Protein-Protein Interaction Network and Protein Hybrid Properties. *PLoS One* 6, e14556. doi:10.1371/journal.pone.0014556

Hung, S. S., Meissner, B., Chavez, E. A., Ben-Neriah, S., Ennishi, D., Jones, M. R., et al. (2018). Assessment of Capture and Amplicon-Based Approaches for the Development of a Targeted Next-Generation Sequencing Pipeline to Personalize Lymphoma Management. *J. Mol. Diagn.* 20, 203–214. doi:10.1016/j.jmoldx.2017.11.010

Kant, S., Kumar, A., and Singh, S. M. (2013). Myelopoietic Efficacy of Orlistat in Murine Hosts Bearing T Cell Lymphoma: Implication in Macrophage Differentiation and Activation. *PLoS One* 8, e82396. doi:10.1371/journal.pone.0082396

Kawahata, R. T., Chuang, L. F., Holmberg, C. A., Osburn, B. I., and Chuang, R. Y. (1983). Inhibition of Human Lymphoma DNA-dependent RNA Polymerase Activity by 6-mercaptopurine Ribonucleoside Triphosphate. *Cancer Res.* 43, 3655–3659.

Köhler, S., Bauer, S., Horn, D., and Robinson, P. N. (2008). Walking the Interactome for Prioritization of Candidate Disease Genes. *Am. J. Hum. Genet.* 82, 949–958. doi:10.1016/j.ajhg.2008.02.013

Li, J., Chen, L., Wang, S., Zhang, Y., Kong, X., Huang, T., et al. (2018). A Computational Method Using the Random Walk with Restart Algorithm for Identifying Novel Epigenetic Factors. *Mol. Genet. Genomics* 293, 293–301. doi:10.1007/s00438-017-1374-5

Li, Y., and Patra, J. C. (2010). Genome-wide Inferring Gene-Phenotype Relationship by Walking on the Heterogeneous Network. *Bioinformatics* 26, 1219–1224. doi:10.1093/bioinformatics/btq108

Liang, H., Chen, L., Zhao, X., and Zhang, X. (2020). Prediction of Drug Side Effects with a Refined Negative Sample Selection Strategy. *Comput. Math. Methods Med.* 2020, 1573543. doi:10.1155/2020/1573543

Lieber, M. R., Gu, J., Lu, H., Shimazaki, N., and Tsai, A. G. (2010). Nonhomologous DNA End Joining (NHEJ) and Chromosomal Translocations in Humans. *Subcell Biochem.* 50, 279–296. doi:10.1007/978-90-481-3471-7_14

Liu, J., Liang, Q., Wang, A., Zou, F., Qi, Z., Yu, K., et al. (2020). Discovery of a Highly Potent and Selective Bruton's Tyrosine Kinase Inhibitor Avoiding Impairment of ADCC Effects for B-Cell Non-hodgkin Lymphoma. *Sig Transduct Target. Ther.* 5, 200. doi:10.1038/s41392-020-00309-1

Liu, Y., Abdul Razak, F. R., Terpstra, M., Chan, F. C., Saber, A., Nijland, M., et al. (2014). The Mutational Landscape of Hodgkin Lymphoma Cell Lines Determined by Whole-Exome Sequencing. *Leukemia* 28, 2248–2251. doi:10.1038/leu.2014.201

Macropol, K., Can, T., and Singh, A. K. (2009). RRW: Repeated Random Walks on Genome-Scale Protein Networks for Local Cluster Discovery. *BMC bioinformatics* 10, 283. doi:10.1186/1471-2105-10-283

Mathas, S., Hartmann, S., and Küppers, R. (2016). Hodgkin Lymphoma: Pathology and Biology. *Semin. Hematol.* 53, 139–147. doi:10.1053/j.seminhematol.2016.05.007

Mesquita Júnior, D., Araújo, J. A. P., Catelan, T. T. T., Souza, A. W. S. d., Cruvinel, W. d. M., Andrade, L. E. C., et al. (2010). Sistema imunitário - parte II: fundamentos da resposta imunológica mediada por linfócitos T e B. *Rev. Bras. Reumatol.* 50, 552–580. doi:10.1590/s0482-50042010000500004

Mohanty, A., Sandoval, N., Phan, A., Nguyen, T. V., Chen, R. W., Budde, E., et al. (2019). Regulation of SOX11 Expression through CCND1 and STAT3 in Mantle Cell Lymphoma. *Blood* 133, 306–318. doi:10.1182/blood-2018-05-851667

Mottok, A., and Steidl, C. (2015). Genomic Alterations Underlying Immune Privilege in Malignant Lymphomas. *Curr. Opin. Hematol.* 22, 343–354. doi:10.1097/moh.0000000000000155

Nakahata, S., Yamazaki, S., Nakauchi, H., and Morishita, K. (2010). Downregulation of ZEB1 and Overexpression of Smad7 Contribute to Resistance to TGF-B1-Mediated Growth Suppression in Adult T-Cell Leukemia/lymphoma. *Oncogene* 29, 4157–4169. doi:10.1038/onc.2010.172

Ng, K.-L., Ciou, J.-S., and Huang, C.-H. (2010). Prediction of Protein Functions Based on Function-Function Correlation Relations. *Comput. Biol. Med.* 40, 300–305. doi:10.1016/j.compbiomed.2010.01.001

Park, S.-R., Lee, J.-H., and Kim, P.-H. (2001). Smad3 and Smad4 Mediate Transforming Growth Factor-B1-Induced IgA Expression in Murine B Lymphocytes. *Eur. J. Immunol.* 31, 1706–1715. doi:10.1002/1521-4141(200106)31:6<1706:aid-immu1706>3.0.co;2-z

Parsonnet, J., and Isaacson, P. G. (2004). Bacterial Infection and MALT Lymphoma. *N. Engl. J. Med.* 350, 213–215. doi:10.1056/nejmp038200

Petersen, D. L., Krejsgaard, T., Berthelsen, J., Fredholm, S., Willerslev-Olsen, A., Sibbesen, N. A., et al. (2014). B-lymphoid Tyrosine Kinase (Blk) Is an Oncogene and a Potential Target for Therapy with Dasatinib in Cutaneous T-Cell Lymphoma (CTCL). *Leukemia* 28, 2109–2112. doi:10.1038/leu.2014.192

Piñero, J., Queralt-Rosinach, N., Bravo, A., Deu-Pons, J., Bauer-Mehren, A., Baron, M., et al. (2015). DisGeNET: a Discovery Platform for the Dynamical Exploration of Human Diseases and Their Genes. *Database* 2015, bav028. doi:10.1093/database/bav028

Quintanilla-Martinez, L., Slotta-Huspenina, J., Koch, I., Klier, M., Hsi, E. D., De Leval, L., et al. (2009). Differential Diagnosis of Cyclin D2+ Mantle Cell Lymphoma Based on Fluorescence *In Situ* Hybridization and Quantitative Real-Time-PCR. *haematologica* 94, 1595–1598. doi:10.3324/haematol.2009.010173

Rahmatpanah, F. B., Carstens, S., Guo, J., Sjahputera, O., Taylor, K. H., Duff, D., et al. (2006). Differential DNA Methylation Patterns of Small B-Cell Lymphoma Subclasses with Different Clinical Behavior. *Leukemia* 20, 1855–1862. doi:10.1038/sj.leu.2404345

Ren, W., Ye, X., Su, H., Li, W., Liu, D., Pirmoradian, M., et al. (2018). Genetic Landscape of Hepatitis B Virus-Associated Diffuse Large B-Cell Lymphoma. *Blood* 131, 2670–2681. doi:10.1182/blood-2017-11-817601

Rosenwald, A., Wright, G., Leroy, K., Yu, X., Gaulard, P., Gascoyne, R. D., et al. (2003). Molecular Diagnosis of Primary Mediastinal B Cell Lymphoma Identifies a Clinically Favorable Subgroup of Diffuse Large B Cell Lymphoma Related to Hodgkin Lymphoma. *J. Exp. Med.* 198, 851–862. doi:10.1084/jem.20031074

Ruiz-Cosano, J., Torres-Moreno, D., and Conesa-Zamora, P. (2013). Influence of Polymorphisms in ERCC5, XPA and MTR DNA Repair and Synthesis Genes in

B-Cell Lymphoma Risk. A Case-Control Study in Spanish Population. *J. BUON* 18, 486–490.

Salar, A. (2019). Gastric MALT Lymphoma and *Helicobacter pylori*. *Medicina Clínica (English Edition)* 152, 65–71. doi:10.1016/j.medcle.2018.09.009

Sangueza, O. P., Yadav, S., White, C. R., Jr, and Braziel, R. M. (1992). Evolution of B-Cell Lymphoma from Pseudolymphoma A Multidisciplinary Approach Using Histology, Immunohistochemistry, and Southern Blot Analysis. *The Am. J. dermatopathology* 14, 408–413. doi:10.1097/00000372-199210000-00006

Seltzer, J., Moorad, R., Schifano, J. M., Landis, J. T., and Dittmer, D. P. (2020). Interleukin-1 Receptor-Associated Kinase (IRAK) Signaling in Kaposi Sarcoma-Associated Herpesvirus-Induced Primary Effusion Lymphoma. *J. Virol.* 94, e02123–02119. doi:10.1128/JVI.02123-19

Shankland, K. R., Armitage, J. O., and Hancock, B. W. (2012). Non-Hodgkin Lymphoma. *The Lancet* 380, 848–857. doi:10.1016/s0140-6736(12)60605-9

Shen, M., Zheng, T., Lan, Q., Zhang, Y., Zahm, S. H., Wang, S. S., et al. (2006). Polymorphisms in DNA Repair Genes and Risk of Non-hodgkin Lymphoma Among Women in Connecticut. *Hum. Genet.* 119, 659–668. doi:10.1007/s00439-006-0177-2

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer Statistics, 2021. *CA A. Cancer J. Clin.* 71, 7–33. doi:10.3322/caac.21654

Skibola, C. F., Bracci, P. M., Paynter, R. A., Forrest, W. A., Agana, L., Woodage, T., et al. (2005). Polymorphisms and Haplotypes in the Cytochrome P450 17A1, Prolactin, and Catechol-O-Methyltransferase Genes and Non-hodgkin Lymphoma Risk. *Cancer Epidemiol. Biomarkers Prev.* 14, 2391–2401. doi:10.1158/1055-9965.epi-05-0343

Skibola, C. F., Curry, J. D., and Nieters, A. (2007). Genetic Susceptibility to Lymphoma. *Haematologica* 92, 960–969. doi:10.3324/haematol.11011

Sugiyama, A., Kobayashi, M., Daizo, A., Suzuki, M., Kawashima, H., Kagami, S.-i., et al. (2017). Diffuse Cerebral Vasoconstriction in a Intravascular Lymphoma Patient with a High Serum MPO-ANCA Level. *Intern. Med.* 56, 1715–1718. doi:10.2169/internalmedicine.56.8051

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING V10: Protein-Protein Interaction Networks, Integrated over the Tree of Life. *Nucleic Acids Res.* 43, D447–D452. doi:10.1093/nar/gku1003

Takatori, M., Sakihama, S., Miyara, M., Imaizumi, N., Miyagi, T., Ohshiro, K., et al. (2021). A New Diagnostic Algorithm Using Biopsy Specimens in Adult T-Cell Leukemia/lymphoma: Combination of RNA *In Situ* Hybridization and Quantitative PCR for HTLV-1. *Mod. Pathol.* 34, 51–58. doi:10.1038/s41379-020-0635-8

Vockerodt, M., Yap, L.-F., Shannon-Lowe, C., Curley, H., Wei, W., Vrzalikova, K., et al. (2015). The Epstein-Barr Virus and the Pathogenesis of Lymphoma. *J. Pathol.* 235, 312–322. doi:10.1002/path.4459

Wang, H., Flannery, S. M., Dickhöfer, S., Huhn, S., George, J., Kubarenko, A. V., et al. (2014). A Coding IRAK2 Protein Variant Compromises Toll-like Receptor (TLR) Signaling and Is Associated with Colorectal Cancer Survival. *J. Biol. Chem.* 289, 23123–23131. doi:10.1074/jbc.m113.492934

Wang, W., Corrigan-Cummins, M., Hudson, J., Maric, I., Simakova, O., Neelapu, S. S., et al. (2012). MicroRNA Profiling of Follicular Lymphoma Identifies microRNAs Related to Cell Proliferation and Tumor Response. *Haematologica* 97, 586–594. doi:10.3324/haematol.2011.048132

Wang, Z., Zhang, R., Liu, L., Shen, Y., Chen, J., Qi, T., et al. (2021). Incidence and Spectrum of Infections Among HIV/AIDS Patients with Lymphoma during Chemotherapy. *J. Infect. Chemother.* 27, 1459–1464. doi:10.1016/j.jiac.2021.06.012

Watson, R. L., Pasewark, R. A., and Fitzgerald, B. J. (1970). Use of the Edwards Personal Preference Schedule with Delinquents. *Psychol. Rep.* 26, 963–965. doi:10.2466/pr0.1970.26.3.963

Wilson, L. D., Hinds, G. A., and Yu, J. B. (2012). Age, Race, Sex, Stage, and Incidence of Cutaneous Lymphoma. *Clin. Lymphoma Myeloma Leuk.* 12, 291–296. doi:10.1016/j.clml.2012.06.010

Yang, F., Xiong, J., Jia, X.-E., Gu, Z.-H., Shi, J.-Y., Zhao, Y., et al. (2014). GSTT1 Deletion Is Related to Polycyclic Aromatic Hydrocarbons-Induced DNA Damage and Lymphoma Progression. *PLoS One* 9, e89302. doi:10.1371/journal.pone.0089302

Yuan, F., and Lu, W. (2017). Prediction of Potential Drivers Connecting Different Dysfunctional Levels in Lung Adenocarcinoma via a Protein-Protein Interaction Network. *Biochim. Biophys. Acta Mol. Basis Dis.* 1864, 2284–2293. doi:10.1016/j.bbadis.2017.11.018

Zhang, J., Suo, Y., Liu, M., and Xu, X. (2017a). Identification of Genes Related to Proliferative Diabetic Retinopathy through RWR Algorithm Based on Protein-Protein Interaction Network. *Biochim. Biophys. Acta Mol. Basis Dis.* 1864, 2369–2375. doi:10.1016/j.bbadis.2017.11.017

Zhang, J., Yang, J., Huang, T., Shu, Y., and Chen, L. (2016). Identification of Novel Proliferative Diabetic Retinopathy Related Genes on Protein-Protein Interaction Network. *Neurocomputing* 217, 63–72. doi:10.1016/j.neucom.2015.09.136

Zhang, L.-l., Wei, J.-y., Wang, L., Huang, S.-l., and Chen, J.-l. (2017b). Human T-Cell Lymphotropic Virus Type 1 and its Oncogenesis. *Acta Pharmacol. Sin* 38, 1093–1103. doi:10.1038/aps.2017.17

Zhang, X., Chen, L., Guo, Z.-H., and Liang, H. (2019). Identification of Human Membrane Protein Types by Incorporating Network Embedding Methods. *IEEE Access* 7, 140794–140805. doi:10.1109/access.2019.2944177

Zhang, X., and Chen, L. (2020). Prediction of Membrane Protein Types by Fusing Protein-Protein Interaction and Protein Sequence Information. *Biochim. Biophys. Acta (Bba) - Proteins Proteomics* 1868, 140524. doi:10.1016/j.bbapap.2020.140524

Zhang, Y., Dai, L., Liu, Y., Zhang, Y., and Wang, S. (2017c). Identifying Novel Fruit-Related Genes in *Arabidopsis thaliana* Based on the Random Walk with Restart Algorithm. *PLoS One* 12, e0177017. doi:10.1371/journal.pone.0177017

Zhao, R., Hu, B., Chen, L., and Zhou, B. (2020). Identification of Latent Oncogenes with a Network Embedding Method and Random Forest. *Biomed. Res. Int.* 2020, 5160396. doi:10.1155/2020/5160396

Zhu, Y., Hu, B., Chen, L., and Dai, Q. (2021). iMPTCE-Hnetwork: A Multilabel Classifier for Identifying Metabolic Pathway Types of Chemicals and Enzymes with a Heterogeneous Network. *Comput. Math. Methods Med.* 2021, 6683051. doi:10.1155/2021/6683051

Zilberman, Y., Zafrir, E., Ovadia, H., Yefenof, E., Guy, R., and Sionov, R. V. (2004). The Glucocorticoid Receptor Mediates the Thymic Epithelial Cell-Induced Apoptosis of CD4+8+ Thymic Lymphoma Cells. *Cell Immunol.* 227, 12–23. doi:10.1016/j.cellimm.2004.01.005