Identification of a 14-Gene Prognostic Signature for Diffuse Large B Cell Lymphoma (DLBCL)

Pengcheng Feng1, Hongxia Li2, Jinhong Pei1, Yan Huang1 and Guixia Li1*

1 Department of Basic Medicine, Changzhi Medical College, Changzhi, China, 2 Affiliated Hospital of Changzhi Institute of Traditional Chinese Medicine, Changzhi, China

Although immunotherapy is a potential strategy to resist cancers, due to the inadequate acknowledge, this treatment is not always effective for diffuse large B cell lymphoma (DLBCL) patients. Based on the current situation, it is critical to systematically investigate the immune pattern. According to the result of univariate and multivariate cox proportional hazards, LASSO regression and Kaplan-Meier survival analysis on immune-related genes (IRGs), a prognostic signature, containing 14 IRGs (AQP9, LMBR1L, FGF20, TANK, CRP, ORM1, JAK1, BACH2, MTCP1, IFITM1, TNFSF10, FGF12, RFX5, and LAP3), was built. This model was validated by external data, and performed well. DLBCL patients were divided into low- and high-risk groups, according to risk scores from risk formula. The results of CIBERSORT showed that different immune status and infiltration pattern were observed in these two groups. Gene set enrichment analysis (GSEA) indicated 12 signaling pathways were significantly enriched in the high-risk group, such as natural killer cell-mediated cytotoxicity, toll-like receptor signaling pathway, and so on. In summary, 14 clinically significant IRGs were screened to build a risk score formula. This formula was an accurate tool to provide a certain basis for the treatment of DLBCL patients.

Keywords: diffuse large B cell lymphoma, immune-related gene, immune prognostic model, risk score formula, immune infiltration

INTRODUCTION

Diffuse large B cell lymphoma (DLBCL) is the most common subtype of non-hodgkin lymphoma (NHL), it can be divided into three molecular subtypes [germinal center B cell (GCB) subtype, activated B cell (ABC)-like subtype, and the unclassified subtypes] according to the unique genetic signatures (Calado et al., 2010; Zhang et al., 2016). It has been thought as an aggressive disease caused by rapidly dividing malignant B cells. With further research of deeper genome sequencing and transcriptomic profiling, it has been proven that the complexity of DLBCL biology was seriously underestimated (Scott and Gascoyne, 2014; Opinto et al., 2020).

Abbreviations: DLBCL, diffuse large B cell lymphoma; NHL, non-hodgkin lymphoma; TCGA, the cancer genome atlas; GEO, gene expression omnibus; IRGs, immune-related genes; ImmPort, immunology database and analysis portal; OS, overall survival; LASSO, least absolute shrinkage and selection operator; K-M, Kaplan-meier; GSEA, gene set enrichment analysis; PD-L1, programmed cell death ligand 1; PD-1, programmed cell death protein 1; CTLA-4, cytotoxic T lymphocyte antigen-4; LAG-3, lymphocyte activation gene 3; TIM-3, T cell immunoglobulin-3.
The majority of DLBCL patients could be relieved after a standard regimen of rituximab in combination with chemotherapy, however, 40% of DLBCL patients had a poor prognosis without suitable curative therapies (Kim et al., 2012; Coiffier and Sarkozy, 2016; Carpio et al., 2020; Zhou et al., 2020). Based on this situation, the researches of the treatment strategies on DLBCL remain important.

One of the features in the occurrence and development of carcinoma is the change of immune status. Tumor immune evading mechanisms were increasingly recognized crucial in the formation and development of multiple cancer (Motzer et al., 2014; Velchetti et al., 2014; Borghaei et al., 2015; Lin et al., 2016; Wallin et al., 2016). The fact decreased immunity stimulated the growth of cancer cells could be reversed, with the emergence of immunotherapy (Silver et al., 2015). Hence, cancer immunotherapy has become one of the major strategies to treat cancer and the researches about the relationship between immune cell and tumor have become a hot topic (Schumacher and Schreiber, 2015; Liu et al., 2017; Popovic et al., 2018; Sebastian et al., 2018). It is generally believed that a single immune marker is too farfetched to illustrate the complex immune environment. Therefore, it is necessary to find a multi-immune relevant-gene-based signature to help the physician predict patients' prognosis and characteristic of tumor microenvironment.

The therapy of immune checkpoint blockades had achieved unprecedented success in helping many cancer patients to extend overall survival (OS) (Gettinger et al., 2016; Reck et al., 2016; Rittmeyer et al., 2016). So, in the process of curing cancer, immunotherapy is always an important consideration. However, the benefited population was limited due to high heterogeneity in biological and clinical appearances (Georg et al., 2010; Dobashi and Akito, 2016; Gentzler et al., 2016). Several immune checkpoint inhibitors could enhance cytotoxicity by targeting programmed cell death protein 1 (PD-1) (CD279), programmed cell death ligand 1 (PD-L1) (CD274), cytotoxic T lymphocyte antigen-4 (CTLA-4), lymphocyte activation gene 3 (LAG-3), and T cell immunoglobulin-3 (TIM-3) (HAVCR2). PD-1/PD-L1 could cause the host immune evasion and promotion of metastasis (Velchetti et al., 2014). CTLA-4 belonged to immunoglobulin-related receptors family and could respond to T-cell immune negative regulation (Rowshanravan et al., 2017; Hosseini et al., 2020). Blocking the expression of PD-1 and CTLA-4 improved the outcomes of patients in different cancers, but immune-related adverse events were observed. LAG-3, an immune inhibitory receptor, was regarded as the foremost target next to PD-1.

In this work, we combined clinical information with immune-related genes (IRGs) expression profiles from 216 DLBCL patients to evaluate the OS. The risk score formula was constructed to predict the individual survival time. Furthermore, the prognosis significance of multiple immune biomarkers was confirmed by the cancer genome atlas (TCGA)-DLBC and GSE32918. This result provided a model for immune-related work and was the critical step toward developing personalized strategies for DLBCL.

**MATERIALS AND METHODS**

**Data Collection**

The level-3 RNA-seq data and clinical data of DLBCL were downloaded from the TCGA and normalized by TCGAbiolinks R package. The raw datasets (GSE136971 and GSE32918) of DLBCL were downloaded from the gene expression omnibus (GEO) database. Limma package was used to screen the differential expression genes. Perl was used to transform ensemble IDs and probe names to symbols, separately. IRGs were obtained from the Immunology Database and Analysis Portal (ImmPort)\(^1\). Univariate cox proportional hazard regression was used to associate the IRGs with DLBCL patients’ OS. Only IRGs with \(P\) value less than 0.05 were selected as putative genes for further analysis. Least absolute shrinkage and selection operator (LASSO) regression was performed to prevent the model overfitting, using ten-fold cross-validation to exam penalty parameter. Multivariate cox regression analysis was performed to assess the risk value of each IRGs signature, then a risk score was established as following:

\[
\text{Risk score} = \beta_{\text{gene1}} \times \text{gene1 expression value} + \beta_{\text{gene2}} \times \text{gene2 expression value} + \beta_{\text{gene3}} \times \text{gene3 expression value} + \cdots + \beta_{\text{gn}} \times \text{gene n expression value}.
\]

\(n\) is the number of relative IRGs, \(\beta\) is the coefficient generated by the multivariate cox regression.

All data were downloaded from public databases and did not apply for approval of the local ethics committees. A methodological flowchart of this research was shown in Supplementary Figure 1.

**IRGs Signature Construction and Confirmation**

Low- and high-risk groups were generated based on the median risk scores of DLBCL patients. Kaplan-Meier (K-M) was performed to estimate survival distribution. “TimeROC” and “survival” packages were used to examine the suitability of survival prediction among risk models.

**Tumor-Infiltrating Immune Cells**

CIBERSORT from sangerbox\(^2\) was used to explore the abundance of tumor-infiltrating immune cells. Ninety-eight IRGs were submitted to CIBERSORT, to predict the roles of immune infiltration in DLBCL. The correlations between IRGs and four immune checkpoints were analyzed using TCGA-DLBC tumor data by GEPIA\(^3\). The significant level was less than 0.05.

**GSEA-Enrichment Analysis**

To explore the potential biological function of IRGs, gene set enrichment analysis (GSEA) (v 4.1.0) was carried out, based on the gene expression data from low- and high-risk groups. C2.cp.KEGG.v7.2. symbols. gmt was selected as reference gene set database. Enrichment pathways were filtered under the condition of \(P\) value less than 0.05 and FDR \(P\) value less than 0.25.

---

1. https://immport.nih.gov
2. http://sangerbox.com/AllTools?tool_id=9703341
3. http://geopia.cancer-pku.cn/
CMap Analysis
Connectivity Map (CMap) (version 02)\(^4\), was adopted to screen putative drugs targeting 14 IRGs. CMap is a website used to search connections among genes, diseases and drugs. All probe IDs, corresponding to 14 IRGs on HG-U133A, were obtained according to GPL96. The genes that hazard ratios were greater than 1 were marked “up” and less than were marked “down.” The probe ID was input into files with “up” and “down” tags saved as “.grp” format. Small molecular drugs that were negatively correlated with the 14 IRGs signature might have the potential to treat DLBCL.

The Analysis of 14 IRGs Expression Level
The expression matrices were searched in GEO database using lymph as a keyword to explore the expression level of 14 IRGs. The samples (GSM217767, GSM217768, GSM217769, GSM217770, GSM217771, GSM217772, GSM217773, GSM217774, and GSM217775) in GSE8762 were used as control. GSE64555 and GSE159472 were used as disease data sets. These three data were annotated with GPL570. In order to reduce the differences caused by different standardization methods in GEO data, two R packages, Affy and affPLM, were used to re-standardize the original data.

\(^4\)https://portals.broadinstitute.org/cmap

RESULTS
The Preparation and Description of Clinical Data and Expression Data
Only individuals with complete clinical information could be used as experimental samples. In order to reduce errors as much as possible and make our model more reliable, the subjects which the survival time were less than 100 days and no survival information were abandoned. GSE136971, containing 216 samples, were used as a training group. No survival status information was recorded for GSM2329007 and GSM2329133. The survival time of GSM2329094, GSM2329071, GSM2329069, GSM2329022, and GSM2329976 was less than 100 days. These seven individual samples were deleted. GSE32918 (189 samples) and TCGA-DLBC (44 samples) were used as a validating group.

A Risk Formula Was Constructed Using Fourteen-Four IRGs
All symbol and synonyms of IRGs from ImmPort were downloaded, to avoid omissions. 7887 IRGs were obtained and summarized in Supplementary Table 1, 1328 IRGs were screened by merging the expression data of 7887 IRGs and GSE136971. Ninety-eight IRGs were related to OS and screened by univariate Cox proportional hazard regression.
TABLE 1 | The risk coefficient of 14 IRGs.

| ID    | exp(coef) | exp(-coef) | Low 95% CI | Low 95% CI | P value |
|-------|-----------|------------|------------|------------|---------|
| AQP9  | 1.80452   | 0.55416    | 0.867163   | 3.7551     | 0.11439 |
| BACH2 | 2.49894   | 0.40017    | 1.191857   | 5.2395     | 0.015325|
| CRP   | 0.2329    | 4.29369    | 0.077665   | 0.6894     | 0.009307|
| FGF12 | 0.07525   | 13.28981   | 0.012446   | 0.4549     | 0.004833|
| FGF20 | 0.46076   | 2.17031    | 0.18434    | 1.1517     | 0.097356|
| IFITM1| 4.874     | 0.20517    | 0.195216   | 121.6904   | 0.334628|
| JAK1  | 0.03042   | 32.86862   | 0.004527   | 0.2045     | 0.000327|
| LAP3  | 0.2187    | 4.57247    | 0.002935   | 16.2981    | 0.489525|
| LMBR1L| 0.09529   | 10.49381   | 0.016692   | 0.544      | 0.008172|
| MTCP1 | 0.40574   | 2.46462    | 0.075827   | 2.1711     | 0.29186 |
| ORM1  | 1.94444   | 0.51429    | 1.189246   | 3.7192     | 0.008028|
| RFX5  | 42.58626  | 0.02348    | 1.080917   | 1677.8247  | 0.045341|
| TANK  | 12.11217  | 0.08256    | 0.910807   | 161.0709   | 0.058865|
| TNFSF10| 6.24265   | 0.16019    | 0.551469   | 70.667     | 0.139076|

The detailed results of univariate Cox proportional hazard regression were shown in Supplementary Table 1. Fourteen IRGs (AQP9, LMBR1L, FGF20, TANK, CRP, ORM1, JAK1, BACH1, MTCP1, IFITM1, TNFSF10, FGF12, RFX5, and LAP3) were identified by LASSO regression (Figures 1A,B). These IRGs were used to predict risk score by multivariate cox regression (Figure 1C). According to the risk coefficient of 14 IRGs from multivariate cox regression, a risk score formula was constructed as follows.

Risk score = (expression of AQP9 × 1.80452) + (expression of BACH2 × 2.49894) + (expression of CRP × 0.2329) + (expression of FGF12 × 0.07525) + (expression of FGF20 × 0.46076) + (expression of IFITM1 × 4.874) + (expression of JAK1 × 0.03042) + (expression of LAP3 × 0.2187) + (expression of LMBR1L × 0.09529) + (expression of MTCP1 × 0.40574) + (expression of ORM1 × 1.94444) + (expression of RFX5 × 42.58626) + (expression of TANK × 12.11217) + (expression of TNFSF10 × 6.24265). The result of multivariate Cox regression was shown in Table 1.

In this part, the relationship between gender and OS of DLBCL patients were explored, but there was no significant correlation between gender and survival.
Using 14 IRGs Construct the Prognostic Risk Signature for DLBCL

Diffuse large B cell lymphoma patients were divided into low- and high-risk groups, according to the risk score calculated by formula, the median of the risk value was served as the cutoff value (cutoff = 1682). Survival curve and ROC curve were performed to test the suitability of the module. As shown in Figures 2A–H, the high expression of JAK1, CRP, and FGF12, may increase the risk of death, while, the high expression of AQP9, LAP3, ORM1, TANK, and TNFSF10, may increase the chance of survival. The K-M curve also indicated worse prognosis in the high-risk groups (Figure 2I). The areas under the curve (AUC) of 3- and 5-year ROC curve for the prognosis model were 0.813 and 0.884 (Figure 3).

Functional Annotation of the IRGs

Based on the GSE136971 expression data, we explored the difference between low- and high-risk groups using GSEA. Several significant enrichment signaling pathways were detected. Twelve significant pathways were differentially enriched in the low and high-risk groups, including chemokine signaling pathway, allograft rejection, viral myocarditis, leishmania infection, natural killer cell-mediated cytotoxicity, type 1 diabetes mellitus, graft versus host disease, amyotrophic lateral sclerosis (ALS), nod like receptor signaling pathway, apoptosis, Alzheimers’ disease and toll-like receptor signaling pathway (Figure 5 and Table 2).

Correlation Analysis Between IRGs and Immune Checkpoints

The four important immune checkpoints (LAG3, TIM3, CTLA-4, and PD-1/PD-L1) were widely used in cancer immunotherapy. To investigate the possible role of fourteen IRGs in ICB (immune checkpoint Blockade) therapy, the association of fourteen IRGs and four immune checkpoints were analyzed by Pearson’s correlation analysis. Only the gene pairs which the P-value was less than 0.05 were shown in Figure 6. MTCP1 was negatively related to HAVCR2 (−0.42, P = 0.0033), TIM3 (R = −0.31, P = 0.033), and CTLA4 (R = −0.3, 0.038). The other 21 gene pairs were positive which R values were from 0.38 to 0.9, P values were less than 0.05.
Immunocyte Infiltration in the Microenvironment

CIBERSORT was performed to understand the connection between IRGs and immune cell infiltration. The proportion of 22 immune cells was estimated according to the expression data of GSE136971. The immune score, stromal score and ESTIMATE score were calculated by ESTIMATE algorithm (Figure 7). Low-risk groups had higher level of immune infiltration. Furthermore, a significant difference was observed for the immune score ($P = 1.9e-5$) and ESTIMATE score ($P = 4.4e-6$). However, the difference of stromal score between low- and high-risk groups was not significant ($P \approx 0.056$). Six immune cells were observed between low- and high-risk groups (Figure 8). The higher expression level of CD8 T cells, CD4 memory T cells activated and M1 macrophages were shown in low-risk samples. Naïve B cells, regulatory T cells (Tregs) and monocytes were higher in high-risk individuals. In other 16 immune-related cells, expression differences were not statistically significant (Supplementary Figure 2).

Small Molecular Drugs Predicted by CMap

More than one probe in HG-U133A array was found to correspond to 14 IRGs, 15 probes were input “up” file and nine were input “down” file. Fourteen IRGs were uploaded to CMap to identify compounds that cured DLBCL, and ranked based on enrichment score (from $-0.976$ to 0.979) to screen the top 79 small molecular compounds ($P \leq 0.05$) (Supplementary Table 2). The drugs without $P$ values were excluded. Therefore, these drugs might be the most promising novel candidates for DLBCL treatment.

Detect the Expression Level of 14 IRGs

The expression data of 14 IRGs were extracted from GSE8762, GSE64555, and GSE159472. The results of the differential expression of the 14 IRGs in the three data sets were same, except for JAK1. The differential expressions of JAK1 were significant higher expressed in disease samples both GSE136971 and GSE159472, however, in GSE64555, the expression of
the normal samples was higher and in GSE159472 disease samples higher (Figure 9). It is possible that in the process of selecting test populations, differences in human bodies in different regions, or differences in some test populations, resulted in completely opposite results in the same disease sequencing process.

**DISCUSSION**

Although the combination treatment of Rituximab and standard CHOP chemotherapy had achieved unprecedented success in the prognosis and cure of DLBCL patients. However, the treatment of DLBCL is still tricky. Recently, immunotherapy is considered as a most potential treatment strategy and has shown strong strength in the treatment of cancers (Pitt et al., 2016; Llovet et al., 2018). As the present single biomarkers were not reliable enough to predict benefit from ICB therapy, the beneficiary group in DLBCL patients was few. It is essential to construct a multi-immune relevant-gene-based signature and analyze the correlation between the IRGs genes and immune checkpoints (Mushtaq et al., 2018).

Numerous previous studies had shown multi-immune related genes (IRGs) could be used as diagnostic tools and provided advice for the physician in multiple cancers (She et al., 2020; Zhang et al., 2020). However, the potential role of IRGs was not clear in DLBCL. Thus, a prognostic model was developed, which fourteen genes were included (AQP9, LMBR1L, FGF20, TANK, CRP, ORM1, JAK1, BACH1, MTCP1, IFITM1, TNFSF10, FGF12, RFX5, and LAP3), at the same time, its value of prognostic and prediction were analyzed.

The same gene may have different functions in different diseases. Poor prognostic factors had similar roles in different cancers, such as stimulating the proliferation and metastasis of tumor cell. High expression levels of AQP9 in renal cell carcinoma individual had the trend of bad prognosis (Yamada et al., 2019). However, high expression levels of AQP9 in gastric cancer and colorectal cancer patients were correlated with better OS (Huang et al., 2017; Thapa et al., 2018). The detailed information of the role of 13 IRGs play in different disease was shown in Table 3. All in all, our analysis results were in agreement with previous researches about these 14 genes. Therefore, all the present genes in this paper could be predicted as candidates for prognostic markers of DLBCL.

The impact of the sample size and external validation on the model were taken into account. GSE136971, containing 216 available individual samples, was used as a training group. Two external data were used to validate the validity of the model. DLBCL patients were divided into low- and high-risk score groups, risk score value was used as a parameter. The gene expression evaluating strategies of the different datasets might be different, especially, data was from different sources platforms. Different cutoff was used to eliminate the potential difference in training groups and validation groups. The high-risk groups tended to have worse OS; the prognostic value of this model

![FIGURE 5](image_url)

**TABLE 2** Detailed information of KEGG from GSEA.

| KEGG NAMES                               | ES   | NES   | NOM p | FDR q |
|------------------------------------------|------|-------|-------|-------|
| Toll like receptor signaling pathway     | 0.561| 1.830 | 0.000 | 0.186 |
| Alzheimer's disease                      | 0.565| 1.811 | 0.004 | 0.114 |
| apoptosis                                | 0.472| 1.763 | 0.004 | 0.123 |
| Nod like receptor signaling pathway      | 0.659| 1.757 | 0.004 | 0.098 |
| Type I diabetes mellitus                 | 0.617| 1.637 | 0.010 | 0.233 |
| Graft versus host disease                | 0.657| 1.616 | 0.008 | 0.230 |
| Natural killer cell mediated cytotoxicity| 0.434| 1.598 | 0.028 | 0.229 |
| Amyotrophic lateral sclerosis (ALS)      | 0.463| 1.594 | 0.053 | 0.207 |
| Leishmania infection                     | 0.522| 1.589 | 0.019 | 0.190 |
| Viral myocarditis                        | 0.545| 1.575 | 0.041 | 0.192 |
| Alograft rejection                       | 0.564| 1.539 | 0.036 | 0.227 |
| Chemokine signaling pathway              | 0.465| 1.523 | 0.018 | 0.232 |
remained robust in two validation groups. Moreover, the AUC values of the training groups and validation groups were larger than 0.7, so the model was reliable.

Gene set enrichment analysis was performed to have a deeper understanding of the underlying molecular mechanisms of the occurrence and development of DLBCL. GSEA enrichment indicated that these pathways (chemokine signaling pathway, allograft rejection, viral myocarditis, leishmania infection, natural killer cell-mediated cytotoxicity, type I diabetes mellitus, graft versus host disease, nod-like receptor signaling pathway, apoptosis, Alzheimer’s disease and toll-like receptor signaling pathway,) were significantly related to the development of DLBCL. The abnormal expression of some chemotaxis, such as CCL3 and CCL4, were associated with bad prognosis in DLBCL (Takahashi et al., 2015). Viral myocarditis pathway involved autoimmune diseases (Zheng et al., 2016). The high expression of possible poor prognostic biomarker GJB2 caused bad prognostic through natural killer cell-mediated cytotoxicity pathway (Tang et al., 2020). Intrinsic anti-apoptosis was related to drug resistance and eventual fatal outcome in DLBCL patients to some extent (Muris et al., 2007; Cillessen et al., 2010; Liu et al., 2021).
Surprisingly, many studies had proved that there were gender differences in the occurrence and development of many diseases (Haitao et al., 2020; Strope et al., 2020). However, in this study, it was found that the OS rate and immune cell infiltration of DLBCL was not significantly related to gender. The differences in immune cells of gender in the same risk group were studied,

FIGURE 7 | Box-plots showed the difference in the scores of ESTIMATE, immune and stromal between the low- and high-risk groups.

FIGURE 8 | The proportion of immune cell infiltration between the low-risk group (red) and high-risk group (blue) (significant differences between high and low risk groups).
**FIGURE 9** | The results of differential expression of 14 IRGs in two GEO data sets were shown.

**TABLE 3** | The information of 13 IRGs related diseases and functions.

| IRGs  | Disease type                                                                 | Functions                                                                                                                                 |
|-------|-----------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| AQP9  | Astrocytoma (Lv et al., 2018), hepatocellular carcinoma (Liao et al., 2020), breast cancer (Zhu et al., 2019), renal cell carcinoma (Yamada et al., 2019), melanoma (Gao et al., 2011), colorectal cancer and gastric cancer (Huang et al., 2017; Thapa et al., 2018) | Promote the invasion and motility inhibit cell apoptosis. Overexpression of AQP9 were correlated to have better OS |
| LMBR1L| in mice (Choi et al., 2019)                                                  | Mutant LMBR1L impaired the development of lymphoid lineages.                                                                             |
| FGF20 | Human hepatocellular carcinoma cells (Wang et al., 2017)                     | Stimulate the proliferation and migration of cancer cells.                                                                                   |
| TANK  | BLBCL (Cao et al., 2012; Wang et al., 2016), breast cancer, non-metastatic clear cell renal cell carcinoma, non-small cell lung cancer and colorectal cancer risk (Nimptsch et al., 2015; Hu et al., 2016; Akamine et al., 2018; Preet et al., 2018a b) | Antiviral innate immunity and negatively regulates NF-NF-κ B signaling.                                                                      |
| CRP   |                                                                            | Inflammatory marker                                                                                                                                 |
| ORM1  | Rat model of septic stroke pathology                                         | Neuroinflammation (Astrup et al., 2018; Higuchi et al., 2020; Moguckin et al., 2020; Sharma et al., 2020)                                        |
| JAK1  | Lung cancer (Viel et al., 2019), epithelial ovarian cancer, colon cancer, prostate cancer and colorectal cancer (Davudian et al., 2016; Shajari et al., 2017; Zhu et al., 2018; Han et al., 2019) | Immune deficiency in natural killer cells (Witalisz et al., 2019), the mutation of JAK1 was related to immune escape in many cancers (Xie et al., 2009; Albacker et al., 2017) apoptosis and growth in several cancers (Siavash et al., 2004) optimal fitness of activated B cell (Zhu et al., 2017) |
| BACH1 | Lung cancer (Viel et al., 2019), epithelial ovarian cancer, colon cancer, prostate cancer and colorectal cancer (Davudian et al., 2016; Shajari et al., 2017; Zhu et al., 2018; Han et al., 2019) | Stimulates glycolysis dependent lung cancer metastasis required for metastatic AsPC-1 cells (Sato et al., 2020) stimulate cancer cell metastasis |
| IFITM1| Lung cancer, colorectal cancer, inflammatory breast cancer, cervical squamous cell carcinoma and lung cancer (FangYu et al., 2015; Ogony et al., 2016; Jin et al., 2017; Zheng et al., 2017; Yan et al., 2019) | Silenced IFITM1 caused inhibited migration and invasion.                                                                                     |
| TNFSF10| Amyloid-related disorders (Cantarella et al., 2015)                           | Stimulated proliferation and inflammation, and inhibited apoptosis (Huang et al., 2019) improvement and restrained immune/inflammatory response |
| FGF12 | Esophageal squamous cell carcinoma (Bhushan et al., 2017a)                   | Silencing FGF12 inhibited apoptosis of radiation-induced cell and the cell migration and proliferation (Fumia et al., 2008; Bhushan et al., 2017b) |
| RXF5  | Hepatocellular carcinoma                                                    | Promoted the progression of the cell cycle (Chen et al., 2020)                                                                             |
| LAP3  | Glioma cells (He et al., 2015), esophageal squamous cell (Zhang et al., 2014), ovarian carcinoma (Suganuma et al., 2005; Wang et al., 2015), breast cancer (Wang et al., 2020) and hepatocellular carcinoma (Tian et al., 2014) | Regulate cell proliferation, invasion and/or angiogenesis expressed in several malignant and affects tumor angiogenesis |
but no significant differences were found. Male and female DLBCL patients may have the same immune pattern.

Our analysis result from ESTIMATE showed that high-risk group had higher immune cell infiltration scores. These research results were similar to the study of immune-related LncRNA in breast cancer (Shen et al., 2020). Naïve B cells and M0 macrophages cells had higher expression in low-risk score group, the expression of CD8 T cells, activated memory CD4 T cells and M1 macrophages cells were higher in the high-risk group. The infiltrating levels of CD8 T cells were associated with high-risk level and low OS (Jeong et al., 2017). The present and previous findings suggested that the infiltration of specific immune cells could cause bad prognosis.

It had some limitations in this work. First, the prognostic value of 14 IRGs was validated by two external data, but no in vivo or in vitro experimental study was carried out. Second, six IRGs (BACH2, FGF20, IFITM1, LMBR1L, MTCP1, and RFX5) part of risk model was all correlated to OS, but no difference was detected between low- and high-risk groups. The reason and rule of these IRGs were not explored.

In this work, we not only explored the role of the immune system in DLBCL development, but we provided an advanced treatment way.

**CONCLUSION AND OUTLOOK**

In this work, a risk score formula was established based on the 14 IRGs signature in DLBCL. According to these signatures, our study might present valuable clinical applications in personalized and precise treatment. The result of GSEA enrichment analysis indicated that the deterioration of DLBCL involved natural killer cell-mediated cytotoxicity and other 11 pathways. These 12 pathways were valuable to further analyze in cell and animal testing. Two strong correlation ($R > 0.8$) were found between TNSF10 and TIM-3 (HAVCR2), LAP3 and TIM-3. The expression level of TNSF10 and LAP3 could provide basis and guidance for immunotherapy. The infiltration and potential immune checkpoint blockade immunotherapy could be predicted. Furthermore, 79 small molecular were screened as potential drugs for DLBCL. Nevertheless, our conclusion should be tested by other public data and verified in future research.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

**ETHICS STATEMENT**

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

**AUTHOR CONTRIBUTIONS**

PF and GL were the major contributors in writing of this manuscript. GL was the responding author. YH, HL, and JP provided a few studies, ideas, and revised opinion. All authors have read and agreed to the published version of the manuscript.

**ACKNOWLEDGMENTS**

The authors would like to thank the Department of Basic Medicine at Changzhri Medical College.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.625414/full#supplementary-material

**SUPPLEMENTARY MATERIAL**

**Supplementary Figure 1** | A flowchart for methods was shown.

**Supplementary Figure 2** | The proportion of immune cell infiltration between the low-risk group (red) and high-risk group (blue) (no significant differences between high and low risk groups).

**REFERENCES**

Akamine, T., Takada, K., Toyokawa, G., Kinoshita, F., Matsubara, T., et al. (2018). Association of preoperative serum CRP with PD-L1 expression in 508 patients with non-small cell lung cancer: A comprehensive analysis of systemic inflammatory markers. Surgical Oncol. Oxford 27, 88–94. doi: 10.1016/j.suronc.2018.01.002

Albacker, L. A., Wu, J., Smith, P., Warmuth, M., and Chmielecki, J. (2017). Loss of function JAK1 mutations occur at high frequency in cancers with microsatellite instability and are suggestive of immune evasion. PLoS One 12:e176181. doi: 10.1371/journal.pone.0176181

Astrup, L. B., Skovgaard, K., Rasmussen, R. S., Iburg, T. M., Agerholm, J. S., et al. (2019). Staphylococcal aureus infected embolic stroke upregulates Orm1 and Cxcl2 in a rat model of septic stroke pathology. Neurol. Res. 5, 399–412. doi: 10.1080/01616412.2019.1573455

Bhushan, A., Singh, A., Kapur, S., Borthakar, B. B., Sharma, J., et al. (2017a). Identification and validation of fibroblast growth factor 12 gene as a novel potential biomarker in esophageal cancer using cancer genomic datasets. Omics J. Integrat. Biol. 21, 616–631. doi: 10.1089/omi.2017.0116

Bhushan, A., Singh, A., Kapur, S., Borthakar, B. B., Sharma, J., et al. (2017b). Identification and validation of fibroblast growth factor 12 gene as a novel potential biomarker in esophageal cancer using cancer genomic datasets. Omics J. Integrat. Biol. 21, 616–631. doi: 10.1089/omi.2017.0116

Borghaei, H., Paz-Ares, L., Horn, L., Spigel, D. R., Steins, M., et al. (2015). Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. N. Engl. J. Med. 2, 123–135. doi: 10.1056/NEJMoa1507643

Calado, D. P., Zhang, B., Srinivasan, L., Sasaki, Y., Seagal, J., et al. (2010). Constitutive canonical NF-κB activation cooperates with disruption of BLIMP1 in the pathogenesis of activated b cell-like diffuse large cell lymphoma. Cancer Cell 18, 580–589. doi: 10.1016/j.ccr.2010.11.024
Hu, H., Xia, W. K., Wu, X., Yu, T. H., Wu, Y., et al. (2016). Prognostic significance of lymphocyte-to-monocyte ratio and CRP in patients with nonmetastatic clear cell renal cell carcinoma: A retrospective multicenter analysis. Oncotargets Ther. 9, 2759–2767. doi: 10.2147/OTT.S101458

Huang, B., Yu, H., Li, Y., Zhang, W., and Liu, X. (2019). Upregulation of long noncoding TNSFSF10 contributes to osteoarthritis progression through the miR-376-3p/FGFR1 axis. J. Cell. Biochem. 120, 19610–19620. doi: 10.1002/jcb.29267

Huang, D., Feng, X., Liu, Y., Deng, Y., Chen, H., et al. (2017). AQP9-induced cell cycle arrest is associated with RAS activation and improves chemotherapy treatment efficacy in colorectal cancer. Cell Death Dis. 8:e2894. doi: 10.1038/cddis.2017.282

Jeong, J., Oh, E. J., Yang, W. I., Kim, S. J., and Yoon, S. O. (2017). Implications of infiltrating immune cells within bone marrow of patients with diffuse large B-cell lymphoma. Hum. Pathol. 64, 222–231. doi: 10.1016/j.humpath.2017.04.012

Jin, B., Jin, H., and Wang, J. (2017). Silencing of Interferon-Induced transmembrane protein 1 (IFITM1) inhibits proliferation, migration, and invasion in lung cancer cells. Oncol. Res. doi: 10.3727/096504017X1484360974116 [Epub ahead of print].

Kadkhodazadeh, M., Amid, R., Ebadian, A. R., Shams, E., and Tamizi, M. (2012). TRAF family member-associated NF-KB activator (TANK) gene polymorphism in chronic periodontitis and peri-implantitis patients. J. Long Term Eff Med. Implants 22, 127–136. doi: 10.1615/longtermmedimplants.v22.i2.30

Kim, Y. R., Min, Y. H., Hyunyoon, D., Shin, H., Mun, Y., et al. (2012). Prognostic factors in primary diffuse large B-cell lymphoma of adrenal gland treated with rituximab-CHOP chemotherapy from the Consortium for Improving Survival of Lymphoma (CISL). J. Hematol. Oncol. 5:49. doi: 10.1186/1756-7722-5-49

Liao, S., Chen, H., Liu, M., Gan, L., Li, C., et al. (2020). AQP9 inhibits growth and metastasis of hepatocellular carcinoma cells via Wnt/b-catenin pathway. Aging 12, 1527–1544. doi: 10.18632/aging.102698

Lin, Z., Chen, X., Li, Z., Luo, Y., Fang, Z., et al. (2016). PD-1 antibody monotherapy for malignant melanoma: A systematic review and Meta-Analysis. PLoS One 11:e0160485. doi: 10.1371/journal.pone.0160485

Liu, K., Song, J., Yan, Y., Zou, K., Che, Y., et al. (2021). Melatonin increases the chemosensitivity of diffuse large B-cell lymphoma cells to epirubicin by inhibiting P-glycoprotein expression via the NF-kB pathway. Transl. Oncol. 14:100876. doi: 10.1016/j.tranon.2020.100876

Liu, X., Wu, S., Yang, Y., Zhao, M., Zhu, G., et al. (2017). The prognostic landscape of tumor-infiltrating immune cell and immunomodulators in lung cancer. Biomed. Pharmacother. 95, 55–61. doi: 10.1016/j.biopha.2017.08.003

Llover, J. M., Robert, M., Daniela, S., and Finn, R. S. (2018). Molecular therapies and precision medicine for hepatocellular carcinoma. Nat. Rev. Clin. Oncol. 15, 599–616. doi: 10.1038/s41571-018-0073-4

Lv, Y., Huang, Q., Dai, W., Jie, Y., Yu, G., et al. (2018). AQP9 promotes astrocytoma cell invasion and motility via the AKT pathway. Oncol. Lett. 16, 6059–6064. doi: 10.3892/ol.2018.9361

Mecguidin, M. C., Giessy, S. L., Davis, A. N., Aybeyta, M. A., and Boisclair, Y. R. (2020). The acute phase protein orosomucoid 1 is upregulated in early lactation but does not trigger appetite-suppressing STAT3 signaling via the leptin receptor. J. Dairy Ence 103, 4765–4776. doi: 10.3168/jdr.2019-18094

Motzer, R., Rini, B. I., Mcdermott, D. F., Redman, B. G., Kuzel, T. M., et al. (2014). Nivolumab for metastatic renal cell carcinoma: Results of a randomized phase II trial. J. Clin. Oncol. 33, 1430–1437. doi: 10.1200/JCO.2014.59.0703

Muris, J. J. F., Ybstra, B., Cillessen, A. G. M., Osenkoppke, G. J., Kluin-Nelemans, J. C., et al. (2007). Profiling of apoptosis genes allows for clinical stratification of primary nodal diffuse large B-cell lymphomas. Br. J. Haematol. 136, 38–47. doi: 10.1111/j.1365-2141.2006.06375.x

Mushatq, M. U., Papadis, A., Pagenkopf, A., Fleitner, E., Morrow, Z., et al. (2018). Tumor matrix remodeling and novel immunotherapies: The promise of matrix-derived immune biomarkers. J. Immunother. Cancer 6:65. doi: 10.1186/s40425-018-0376-0

Nimptsch, K., Aleksandрова, K., Boeing, H., Janke, J., Lee, Y., et al. (2015). Association of CRP genetic variants with blood concentrations of C-reactive protein and colorectal cancer risk. Int. J. Cancer 136, 1181–1192. doi: 10.1002/ijc.29086

Ogny, J., Choi, H. J., Lui, A., Cristofanilli, M., and Lewis-Wambi, J. (2016). IFITM1 overexpression enhances the aggressive phenotype of SUM149 inflammatory breast cancer cells in a signal transducer and activator of transcription 2 (STAT2)-dependent manner. Breast Cancer Res. 18:25. doi: 10.1186/s13058-016-0683-7
Opinto, G., Veglionate, M. C., Negri, A., Skrypets, T., and Ciavarella, S. (2020). The tumor microenvironment of DLBCL in the computational era. Front. Oncol. 10:351. doi: 10.3389/fonc.2020.00351

Pitt, J. M., Vézérou, M., Daillére, R., Roberti, M. P., Yamazaki, T., et al. (2016). Resistance mechanisms to Immune-Checkpoint blockade in cancer: Tumor- Intrinsic and -Extrinsic factors. Immunity 44, 1255–1269. doi: 10.1016/j.immuni.2016.06.001

Popovic, A., Jaffee, E. M., and Zaidi, N. (2018). Emerging strategies for combination checkpoint modulators in cancer immunotherapy. J. Clin. Investig. 128, 3209–3218. doi: 10.1172/JCI127075

Preet, K. R., Rubal, Singh, B. R. P., Rajesh, V., Monisha, D., et al. (2018a). Association of elevated levels of C-reactive protein with breast cancer, breast cancer subtypes and poor outcome. Curr. Problems Cancer 2, 123–129. doi: 10.1016/j.curprobncancer.2018.05.003

Preet, K. R., Rubal, Singh, B. R. P., Rajesh, V., Monisha, D., et al. (2018b). Association of elevated levels of C-reactive protein with breast cancer, breast cancer subtypes and poor outcome. Curr. Problems Cancer 223:S1897824381. doi: 10.1016/j.curprobncancer.2018.05.003

Reck, M., Rodriguez-Abreu, D., Robinson, A. G., Hui, R., Cs, Szi, T., et al. (2017). Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. N. Engl. J. Med. 375:1823. doi: 10.1056/NEJMoa1606774

Rittmeyer, A., Barlesi, F., Waterkamp, D., Park, K., and Gandara, D. D. (2016). Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): A multicentre, open-label, phase 2 randomised controlled trial. Lancet 2016, 1837–1846. doi: 10.1016/S0140-6736(16)00587-0

Rowshanravan, B., Halliday, N., and Sansom, D. M. (2017). CTLA-4: A moving target in immunotherapy. Blood 2017;741033. doi: 10.1182/blood-2017-06-741033

Sato, M., Matsumoto, M., Saiki, Y., Alam, M., and Igarashi, K. (2020). BACH1 promotes pancreatic cancer metastasis by repressing epithelial genes and enhancing Epithelial-Mesenchymal transition. Cancer Res. 80, 4018–4099. doi: 10.1158/0008-5472.CAN-18-4099

Schumacher, T. N., and Schreiber, R. D. (2015). Neoantigens in cancer immunotherapy. Science 348:69. doi: 10.1126/science.aaa4971

Scott, D. W., and Gascoyne, R. D. (2014). The tumour microenvironment in B cell lymphomas. Nat. Rev. Cancer 14, 517–534. doi: 10.1038/nrc3774

Sebastian, K., Stanislav, P., Felicitas, R., and Johannes, V. B. (2018). Rationale for combining bispecific t cell activating antibodies with checkpoint blockade for cancer therapy. Front. Oncol. 8:285. doi: 10.3389/fonc.2018.00285

Shajari, N., Davudian, S., Kazemi, T., Mansoori, B., Salehi, S., et al. (2017). Silencing of leucine aminopeptidase 3 (LAP3) correlates with prognosis and malignant development of human hepatocellular carcinoma (HCC). Int. J. Clin. Exp. Pathol. 7, 3752–3762. doi: 10.1016/B978-0-12-800992-2.0013-7

Velcheti, V., Schalper, K. A., Carvajal, D. E., Anagnostou, V. K., Syrigos, K. N., et al. (2014). Programmed death ligand-1 expression in non-small cell lung cancer. Lab. Invest. J. Tech. Methods Pathol. 94, 107–116. doi: 10.1038/labinvest.2013.130

Wallin, J. J., Bendell, J. C., Funke, R., Szol, M., Korski, K., et al. (2016). Atezolizumab in combination with bevacizumab enhances antigen-specific T-cell migration in metastatic renal cell carcinoma. Nat. Commun. 7:12624. doi: 10.1038/ncomms12624

Wang, J., Zhou, M., Wang, X., Ju, J., Chen, B., et al. (2016). Pretreatment C-reactive protein was an independent prognostic factor for patients with diffuse large B-cell lymphoma treated with RCHOP. Clin. Chim. Acta 459, 150–154. doi: 10.1016/j.cca.2016.05.033

Wang, S., Lin, H., Zhao, T., Huang, S., Fennig, D., et al. (2017). Expression and purification of an FGF9 fusion protein in E. Coli, and the effects of the FGF9 subfamily on human hepatocellular carcinoma cell proliferation and migration. Appl. Microbiol. Biotechnol. 101, 7823–7835. doi: 10.1007/s00253-017-8468-1

Wang, X., Li, S., Ma, Y., Xing, Z., Zhou, Y., et al. (2020). Vimentin plays an important role in the promotion of breast cancer cell migration and invasion by leucine aminopeptidase 3. Cytotechnology 72, 639–647. doi: 10.1007/s10616-020-00402-x

Wang, X., Shi, L., Deng, Y., Qu, M., Mao, S., et al. (2015). Inhibition of leucine aminopeptidase 3 suppresses invasion of ovarian cancer cells through down-regulation of fascin and MMP-2/9. Eur. J. Pharmacol. 768, 116–122. doi: 10.1016/j.ejphar.2015.10.039

Wiel, C., Le Gal, K., Ibrahim, M. X., Jahangir, C. A., Kashiif, M., et al. (2019). RAC1 stabilization by antioxidants stimulates lung cancer metastasis. Cell 178, 330–345. doi: 10.1016/j.cell.2019.06.005

Witalisz, A., Klein, K., Prinz, D., Leidenfrost, N., Schabbauer, G., et al. (2019). Loss of JAK1 drives innate immune deficiency. Front. Immunol. 9:3108. doi: 10.3389/fimmu.2018.03108

Xie, H., Bae, H., Noh, J., Eun, J., Kim, J., et al. (2009). Mutational analysis of JAK1 drives innate immune deficiency. J. Exp. Med. 200, 01429–y. doi: 10.1186/s12935-020-1104-7

Xia, C., Dai, D., Wu, J., Liu, L., and Zhang, M. (2019). Inhibiting of proliferation, migration, and invasion in lung cancer induced by silencing Interferon-Induced transmembrane protein 1 (IFITM1). BioMed. Res. Int. 2019, 1–9. doi: 10.1155/2019/8085435

Zhang, J., Zhang, J., Yuan, C., Luo, Y., Li, Y., et al. (2020). High GJB2 mRNA expression and its prognostic significance in lung adenocarcinoma: A study based on the TCGA database. Med. 99:e19054. doi: 10.1007/MD.100000000019054

Zhao, S., Chen, S. H., Shao, B. F., Cai, H. Y., and Xu, A. B. (2014). Expression of leucine aminopeptidase 3 (LAP3) correlates with prognosis and malignant development of human hepatocellular carcinoma (HCC). Front. Clin. Exp. Pathol. 7, 3752–3762. doi: 10.1016/B978-0-12-800992-2.0013-7

Zheng, Z., Chen, Z., Zuo, W., Zhu, X., Li, Y., et al. (2016). Omacetaxine mepesuccinate induces apoptosis and cell cycle arrest, promotes cell differentiation, and reduces telomerase activity in diffuse large B-cell lymphoma cells. Mol. Med. Rep. 13, 3092–3100. doi: 10.3892/mmr.2016.4899
Zhang, S., Yang, X., Shi, H., Li, M., Xue, Q., et al. (2014). Overexpression of leucine aminopeptidase 3 contributes to malignant development of human esophageal squamous cell carcinoma. J. Mole. Histol. 45, 283–292. doi: 10.1007/s10735-014-9566-3
Zheng, W. Y., Zheng, W. X., and Hua, L. (2016). Detecting shared pathways linked to rheumatoid arthritis with other autoimmune diseases in a in silico analysis. Mole. Biol. 50, 462–469. doi: 10.1134/S0026893316030146
Zheng, W., Zhao, Z., Yi, X., Zuo, Q., Li, H., et al. (2017). Down-regulation of IFITM1 and its growth inhibitory role in cervical squamous cell carcinoma. Cancer Cell Int. 17:88. doi: 10.1186/s12935-017-0456-0
Zhou, Y., Li, G., and Xu, G. (2020). Kaempferol protects cell damage in in vitro ischemia reperfusion model in rat neuronal PC12 cells. BioMed. Res. Int. 2020, 2461079. doi: 10.1155/2020/2461079
Zhu, F., Hwang, B., Miyamoto, S., and Rui, L. (2017). Nuclear import of JAK1 is mediated by a classical NLS and is required for survival of diffuse large b-cell lymphoma. Mole. Cancer Res. 15, 348–357. doi: 10.1158/1541-7786.MCR-16-0344
Zhu, G. D., Liu, F., Ouyang, S., Zhou, R., Jiang, F. N., et al. (2018). BACH1 promotes the progression of human colorectal cancer through BACH1/CXCR4 pathway. Biochem. Biophys. Res. Commun. 499, 120–127. doi: 10.1016/j.bbrc.2018.02.178
Zhu, L., Ma, N., Wang, B., Wang, L., Zhou, C., et al. (2019). Significant prognostic values of aquaporin mRNA expression in breast cancer. Cancer Manag. Res. 11, 1503–1515. doi: 10.2147/CMAR.S193396

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 © 2021 Feng, Li, Pei, Huang and Li. 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.