

## Research Article

# The Diagnostic and Predictive Significance of Immune-Related Genes and Immune Characteristics in the Occurrence and Progression of IgA Nephropathy

Jian-Bo Qing <sup>1</sup>, Wen-Zhu Song <sup>2</sup>, Chang-Qun Li <sup>1</sup>, and Ya-Feng Li <sup>3,4,5,6</sup>

<sup>1</sup>The Fifth Clinical Medical College of Shanxi Medical University, Taiyuan, Shanxi 030001, China

<sup>2</sup>School of Public Health, Shanxi Medical University, Taiyuan, Shanxi 030001, China

<sup>3</sup>Department of Nephrology, Shanxi Provincial People's Hospital (Fifth Hospital) of Shanxi Medical University, Taiyuan, Shanxi 030012, China

<sup>4</sup>Core Laboratory, Shanxi Provincial People's Hospital (Fifth Hospital) of Shanxi Medical University, Taiyuan, Shanxi 030012, China

<sup>5</sup>Shanxi Provincial Key Laboratory of Kidney Disease, Shanxi Provincial People's Hospital (Fifth Hospital) of Shanxi Medical University, Taiyuan, Shanxi 030012, China

<sup>6</sup>Academy of Microbial Ecology, Shanxi Medical University, Taiyuan, Shanxi 030000, China

Correspondence should be addressed to Ya-Feng Li; [dr.yafengli@gmail.com](mailto:dr.yafengli@gmail.com)

Received 27 February 2022; Revised 25 March 2022; Accepted 4 April 2022; Published 28 April 2022

Academic Editor: Dawei Cui

Copyright © 2022 Jian-Bo Qing et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Objective.** To investigate the potential diagnostic and predictive significance of immune-related genes in IgA nephropathy (IgAN) and discover the abnormal glomerular inflammation in IgAN. **Methods.** GSE116626 was used as a training set to identify different immune-related genes (DIRGs) and establish machine learning models for the diagnosis of IgAN; then, a nomogram model was generated based on GSE116626, and GSE115857 was used as a test set to evaluate its clinical value. Short Time-Series Expression Miner (STEM) analysis was also performed to explore the changing trend of DIRGs with the progression of IgAN lesions. GSE141344 was used with DIRGs to establish the ceRNA network associated with IgAN progression. Finally, ssGSEA analysis was performed on the GSE141295 dataset to discover the abnormal inflammation in IgAN. **Results.** Machine learning (ML) performed excellently in diagnosing IgAN using six DIRGs. A nomogram model was constructed to predict IgAN based on the six DIRGs. Three trends related to IgAN lesions were identified using STEM analysis. A ceRNA network associated with IgAN progression which contained 8 miRNAs, 14 lncRNAs, and 3 mRNAs was established. A higher macrophage ratio and lower CD4+ T cell ratio in IgAN compared to controls were observed, and the correlation between macrophages and monocytes in the glomeruli of IgAN patients was inverse compared to controls. **Conclusion.** This study reveals the diagnostic and predictive significance of DIRGs in IgAN and finds that the imbalance between macrophages and CD4+ immune cells may be an important pathomechanism of IgAN. These results provide potential directions for the treatment and prevention of IgAN.

## 1. Introduction

IgA nephropathy (IgAN) is inflammatory nephropathy characterized by IgA deposition in the mesangial area of the glomeruli [1]. It represents the most common primary glomerular disease globally [2], and its prevalence varies geographically, more frequent in Asian populations (45 mil-

lion people/year in Japan) than in Caucasians (31 million people/year in France) [3]. Also, IgAN owns a higher incidence in young adults [4], with 20-40% of patients subject to end-stage renal disease within 10-20 years [5]. Worse still, IgAN is a lifelong disease with severe signs and symptoms closely associated with poor prognosis, posing a heavy mental and financial burden to patients [6].

TABLE 1: The summary of those four GEO datasets.

GSE number	Samples	Organization	Purpose
GSE116626	52 IgAN patients, 22 non-IgAN GN, and 7 controls	Kidney	Training set
GSE115857	55 IgAN patients, 7 controls	Kidney	Test set
GSE141344	6 IgANp patients and 6 IgANnp patients	Kidney	ceRNA network
GSE141295	14 IgAN patients and 10 controls	Glomeruli	ssGSEA

Abbreviation: IgAN: IgA nephropathy; non-IgAN GN: non-IgA nephropathy glomerulonephritis; IgANp: IgA nephropathy with progression; IgANnp: IgA nephropathy with no progression.

TABLE 2: The characteristics of samples in the GSE116626 dataset.

Group	Number	MEST-C classification
Minimal lesion IgAN	22	(M0,1; E0; S0,1; T0; C0)
Active lesion IgAN	8	(M0,1; S0,1; T0; E1 and/or C,1,2)
Chronic lesion IgAN	12	(M0,1; E0; S0,1; T1,2; C0)
Mixed lesion IgAN	10	(M0,1; E0,1; S0,1; T1,2; C0,1,2)
Non-IgAN GN	7	—
Healthy control	22	—

Abbreviation: IgAN: IgA nephropathy; non-IgAN GN: non-IgA nephropathy glomerulonephritis.

It is crucial to make an early diagnosis of IgAN since delayed diagnosis would contribute to a poor prognosis [7]. Nowadays, pathological biopsy represents the golden standard in IgAN diagnosis, but the results are changeable with the stage of the disease [8]. Besides, the pathological results are sometimes uncertain, making IgAN diagnosis and evaluation tricky. As such, an effective and reliable diagnostic method for IgAN is all the more important. Currently, it has been documented that the occurrence and development of IgA are closely related to genetic [9, 10]. As the microarray technique plus bioinformatics advances, it is a good approach to utilize genes to make a diagnosis and risk assessment for IgAN. Marker genes can not only help us diagnose, but also help us explore the molecular mechanism, signaling pathway, and pathological progress of IgAN.

This study is aimed at exploring the role of genes in the occurrence and progression of IgAN, using integrated gene expression profiling data downloaded from the Gene Expression Omnibus (GEO) database, and at further identifying immune-related genes as diagnostic biomarkers for IgAN patients, which may contribute to the diagnosis and treatment of IgAN. Additionally, abnormal immune infiltration in the glomerulus was studied in patients with IgAN. These results will contribute to better diagnosis, prevention, and treatment of IgAN. In this study, we first identified genes for constructing models and then further explored the gene network related to IgAN progression. Finally, we analyzed the immune infiltration of IgAN glomerulus.

## 2. Materials and Methods

### 2.1. Data Collection

**2.1.1. GEO Dataset Download and Process.** We used the keyword “IgA nephropathy” to search IgAN gene expression profiles in the GEO database. The following steps to obtain

dataset: first, screen datasets for constructing a diagnostic model, whose organization used for sequencing must be the kidney and have cases and controls. Second, search the noncoding RNA profiling dataset used for exploring the ceRNA network of IgAN. Third, as IgAN features glomerular disease and the proportion of differentially expressed genes detected by RNA-SEQ was higher than that by CHIP, the RNA-SEQ dataset of glomerular tissue was selected. Fourth, the datasets must be published within the last five years. Finally, the GEO datasets numbered GSE116626, GSE115857, GSE141344, and GSE141295 were selected. The summary of those four GEO datasets is shown in Table 1.

With more evenly distributed data and more detailed records of the IgAN lesions, GSE116626 was employed for the training set and used for differential expression analysis, and the characteristics of samples are shown in Table 2. Moreover, GSE115857 was utilized as the test set. We then performed log<sub>2</sub> transformed for gene expression profiling and gene symbol conversion; in addition, quality control was performed on each dataset to improve the efficiency of subsequent analysis.

**2.1.2. Immune-Related Gene Download.** We downloaded immune-related genes (IRGs) data from the import database (<https://www.immport.org/shared/>) [11]. After eliminating duplicate genes, we finally obtained 1793 immune-related genes (Supplementary Table 1). The schematic of the research is shown in Figure 1.

**2.2. Differential Expression Analysis and Enrichment Analysis.** Differentially expressed genes (DEGs) between IgAN and healthy control in GSE116626 were achieved by the package “limma” [12] in R software with a criterion of  $P < 0.05$  and  $|\log_2 \text{fold change}| > 1$ . Afterwards, a comparison was made between 245 DEGs and 1793 immune-related genes (IRGs) to obtain different immune-related genes (DIRGs). Furthermore, for deep insight into the gene functions, the package “clusterProfiler” [13] in R was employed for the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis. Finally, we visualized the results by bubble map in R software.

**2.3. Construction and Assessment of RF, GBM, and Treebag Model.** There are 83 samples in our test set, which contains 52 IgAN patients, 22 non-IgAN patients (non-IgAN GN), and 7 healthy people; although artificial neural networks (ANN) or deep neural networks (DNN) are effective learning models in the prediction process, they are poorly

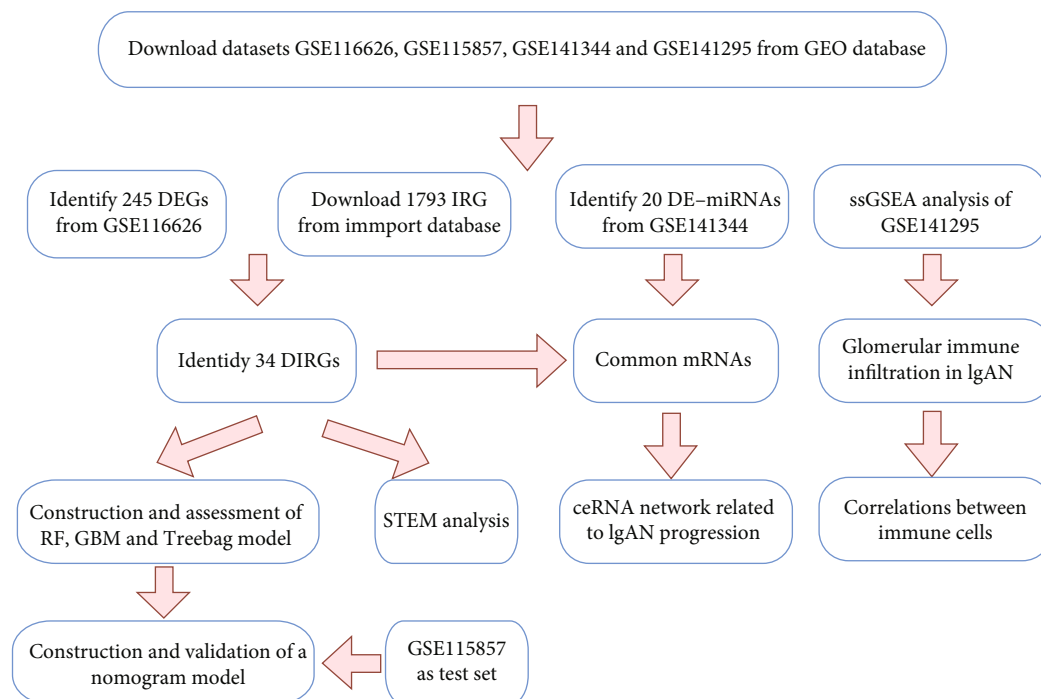


FIGURE 1: Workflow of the research. Abbreviations are defined as follows: Gene Expression Omnibus database (GEO), differentially expressed gene (DEG), immune-related gene (IRG), different immune-related gene (DIRG), differentially expressed miRNA (DE-miRNA), random forest (RF), Gradient Boosting Machine (GBM), Short Time-Series Expression Miner (STEM).

interpretable and require more samples [14]; thus, machine learning (ML) will be more advantageous in this study. Firstly, as elastic net [15] has less prediction error than lasso and ridge regression in our test, it was employed for gene screening in 34 DIRGs using the “glmnet” package in R; the  $\alpha = 1$  and  $\lambda = 6$  (CV.Glmnet function automatically produces the most appropriate value). Besides, we divided the samples of GSE116626 into three types, random forest (RF), Gradient Boosting Machine (GBM), and treebag model as they could achieve the optimal effect in the ML of the three classification samples. After which, we used GSE116626, with IgAN as the response variable, and DIRGs as explanatory variables, to establish RF, GBM, and treebag model. What is more, we employed the package “DALEX” in R software to analyze the above three models and plot the residual distribution.

**2.4. Construction and Validation of a Nomogram Model for IgAN Diagnosis.** A nomogram model was established to predict the occurrence of IgAN using the “rms” package after three ML models verified the accuracy of the DIRGs. “Points” indicates the score of the corresponding factors below, and “total points” indicates the sum of all the scores of factors above used to estimate the risk of IgAN. In addition, a calibration curve was used to evaluate the predictive performance of the nomogram model, while decision curve analysis and clinical impact curve were further used to evaluate the clinical value of the model.

**2.5. Short Time-Series Expression Miner (STEM) Analysis.** We used STEM software [16] to cluster the 34 DIRGs in 4

different lesions of IgAN patients and healthy controls in the GSE116626 datasets;  $\alpha < 0.05$  was considered statistically significant clustering. This method can help us to simulate and estimate the spatiotemporal variation of DIRGs in IgAN. The significantly clustered genes showed a gradual up- or downregulation trend with the change of IgAN lesion.

**2.6. Identification of miRNA Related to IgAN Progression.** The package “DESeq2” [17] in R was used to seek for differentially expressed miRNAs (DE-miRNAs) between IgAN progression (IgAp) and IgAN nonprogression (IgAnp) with  $P < 0.05$  and  $|\log_2 \text{fold change}| > 1$  criterion in GSE141344. Samples were classified as IgAnp if the serum creatinine had changed by  $< 10\%$  over the 10 years since diagnosis, while samples whose serum creatinine had at least doubled or the patient had developed end-stage renal disease (ESRD) in the same time period were classified as IgAnp. Afterwards, mRNAs targeted by miRNAs with experimental support were identified using miRDB [18], miRWalk [19], and miRanda [20] databases. To make our results more convincing, we first incorporated the common genes from the three databases and then compared them with the 34 DIRGs for the final target mRNAs. Also, starBase v 3.0 [21] was used to predict the lncRNAs of miRNAs with final target RNAs, with search parameters set at “CLIP-Data  $\geq 5$ ”. Finally, we built the ceRNA network based on the above data, and the ggalluvial (version 0.9.1) package [22] was used to visualize the network.

**2.7. Immune Signatures of the Glomeruli in IgAN.** The ssGSEA was applied to explore the different infiltrations of

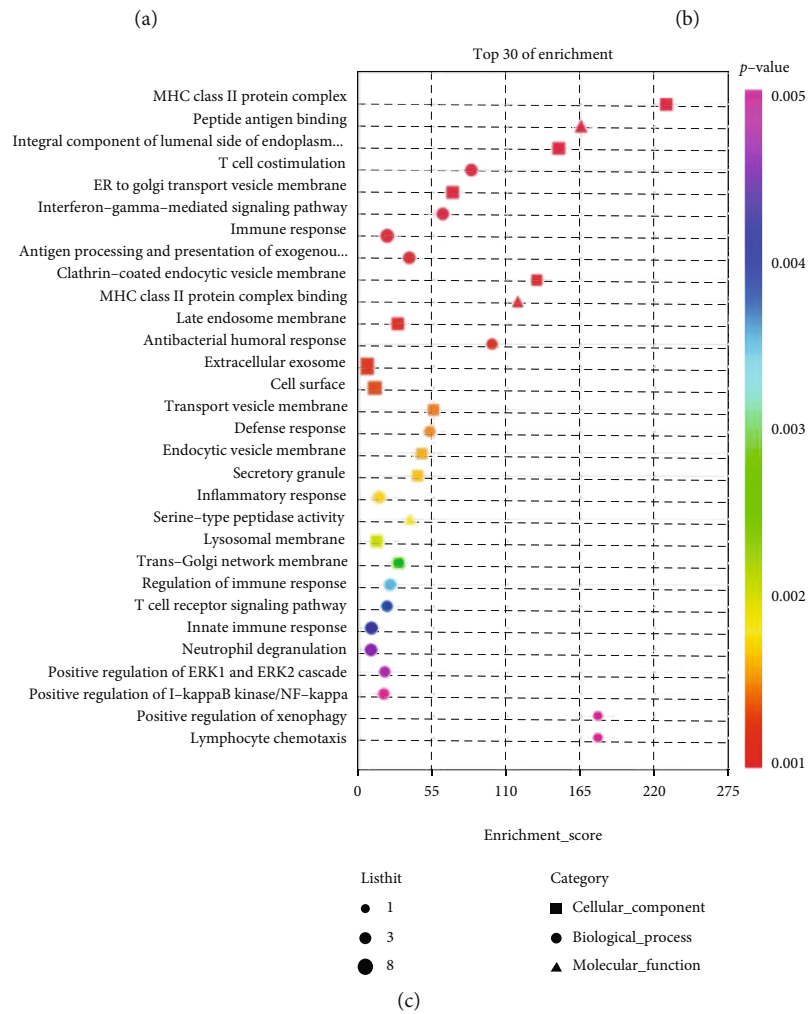
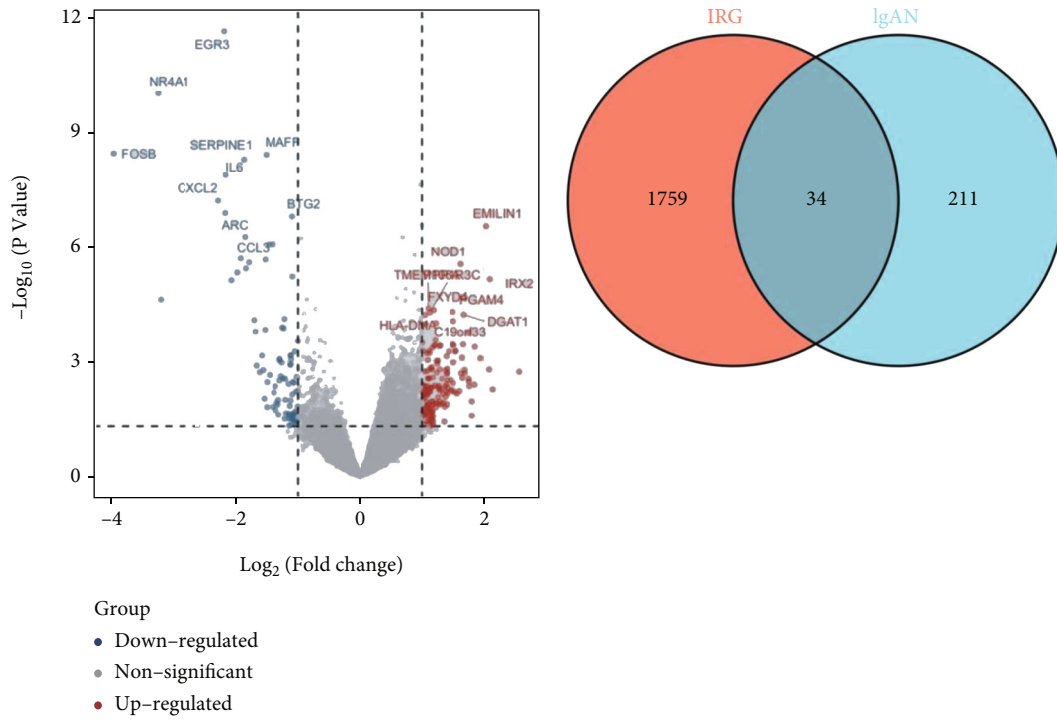


FIGURE 2: Continued.

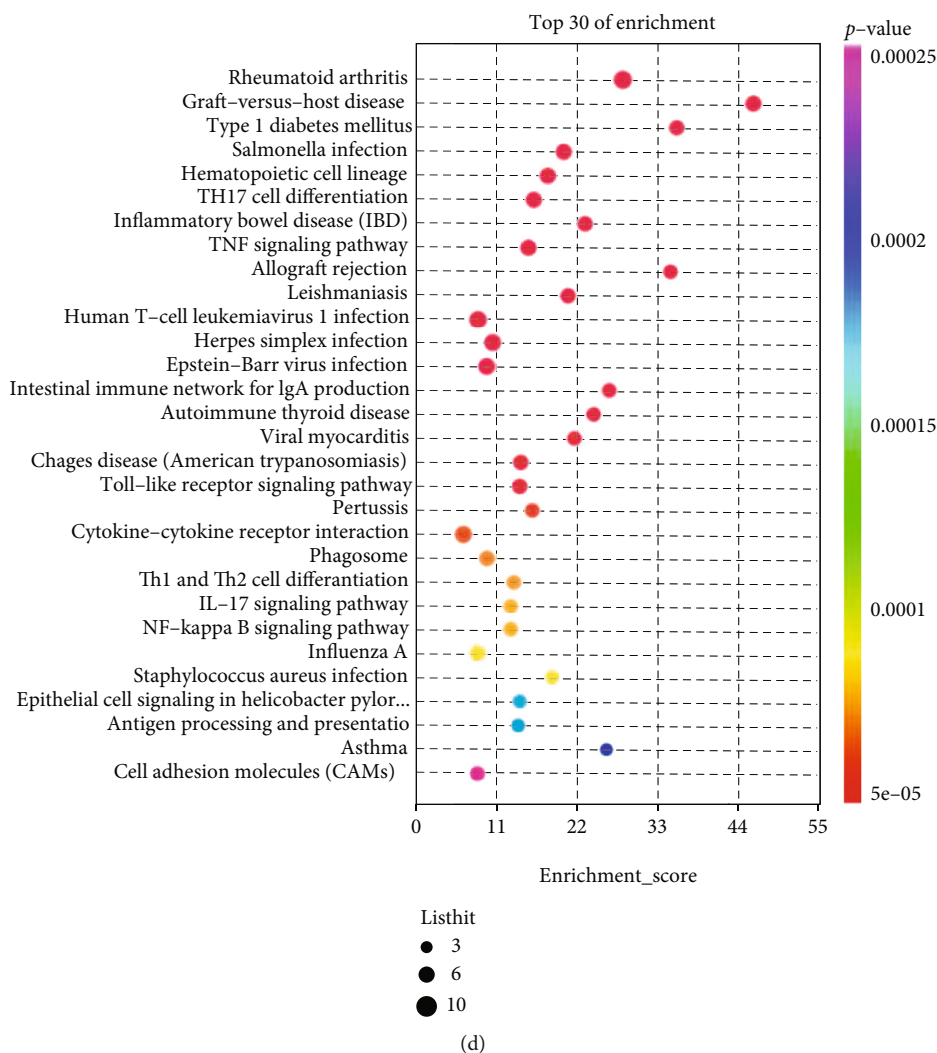


FIGURE 2: Differential expression analysis and enrichment analysis in GSE116626. (a) Volcano map of all DEGs of GSE116626: 245 DEGs were identified with  $P < 0.05$  and  $|\log_2 \text{fold change}| > 1$ , including 160 upregulated and 85 downregulated genes. (b) Venn diagram of DEGs and IRGs: 34 DIRGs were identified. (c) Top 30 GO functional enrichment of 34 DIRGs. Biological process (BP, circle), cellular component (CC, square), and molecular function (MF, triangle) analysis results of 34 DIRGs. (d) Top 30 KEGG functional enrichment of 34 DIRGs. The size of the graph represents the number of genes, and the  $x$ -axis represents the enrichment score.

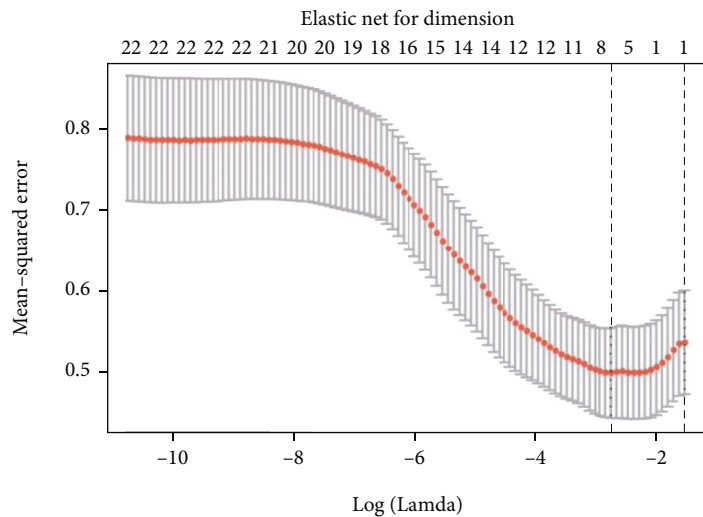
immune cells in GSE141295 using the R package “GSVA” [23]. Afterwards, we compared the ratio of 24 immune cells between the IgAN and control groups and the correlation change of 10 main immune cells between the two groups.

### 3. Results

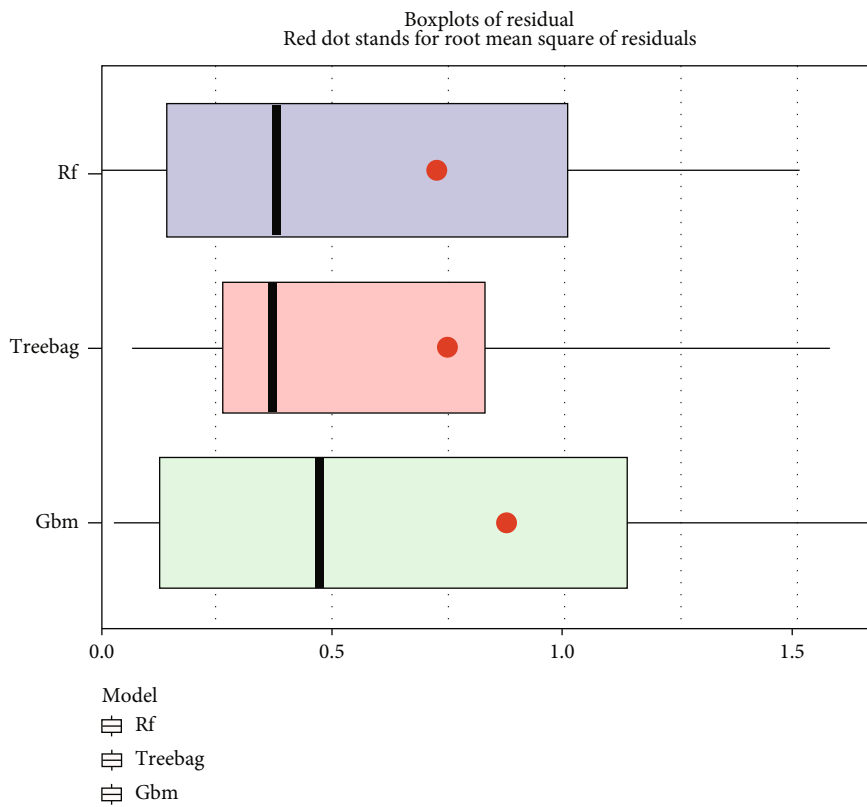
**3.1. Differential Expression Analysis and Enrichment Analysis.** 245 DEGs between IgAN and control groups in GSE116626 were identified with  $P < 0.05$  and  $|\log_2 \text{fold change}| > 1$ , including 160 upregulated and 85 downregulated genes (Figure 2(a)). Also, 34 DIRGs were obtained from 245 DEGs and 1793 IRGs (Figure 2(b)), whose enriched pathways and functions were further understood by GO and KEGG enrichment analysis (Supplementary Table 2 and 3). GO analysis showed that the 34 DIRGs were mainly associated with MHC class II protein complex, peptide antigen binding, and interferon-gamma-

associated pathway, interferon-gamma-mediated signaling, and pathway positive regulation of I-kappaB kinase/NF-kappaB signaling (Figure 2(c)). Meanwhile, KEGG analysis showed that these 34 DIRGs were closely related to autoimmune diseases, differentiation of T cell helper, and intestinal immune network for IgA production. It suggested that these 34 DIRGs were bound up with autoimmune diseases (Figure 2(d)).

**3.2. Construction and Assessment of RF, GBM, and Treebag Model.** The elastic net revealed six genes (PPIA, CCL3L3, CXCL2, TFRC, IL6, and LIF) (Figure 3(a)), based on  $\alpha = 1$  and  $\lambda = 6$ , which were taken into the ML, namely, RF, GBM, and treebag models. There were 100 trees in the RF models; besides, the final values used for the GBM model were  $n.\text{trees} = 50$ ,  $\text{interaction.depth} = 2$ ,  $\text{shrinkage} = 0.1$ , and  $n.\text{minobsinnode} = 10$ ; what is more, the bagging regression trees of the treebag model have 25 bootstrap replications.

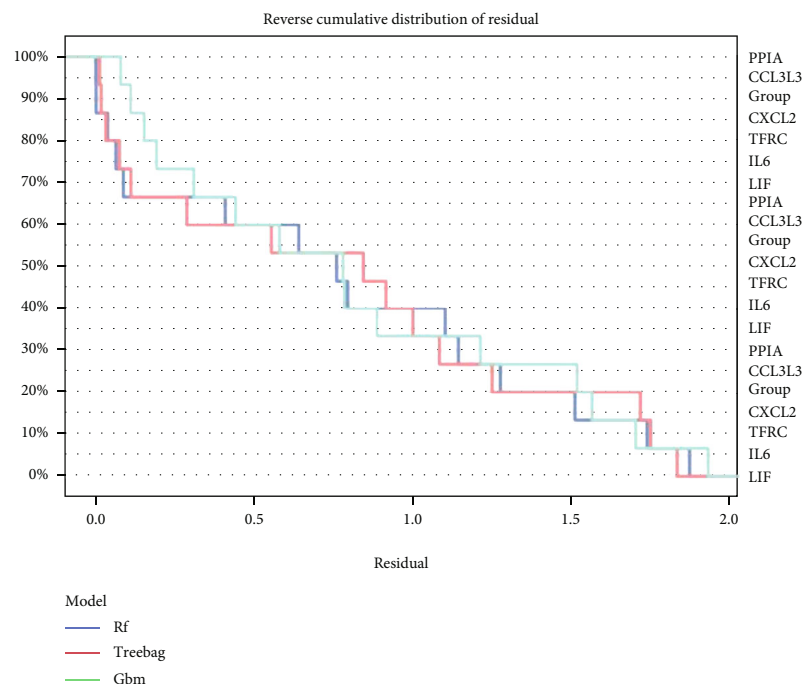


(a)

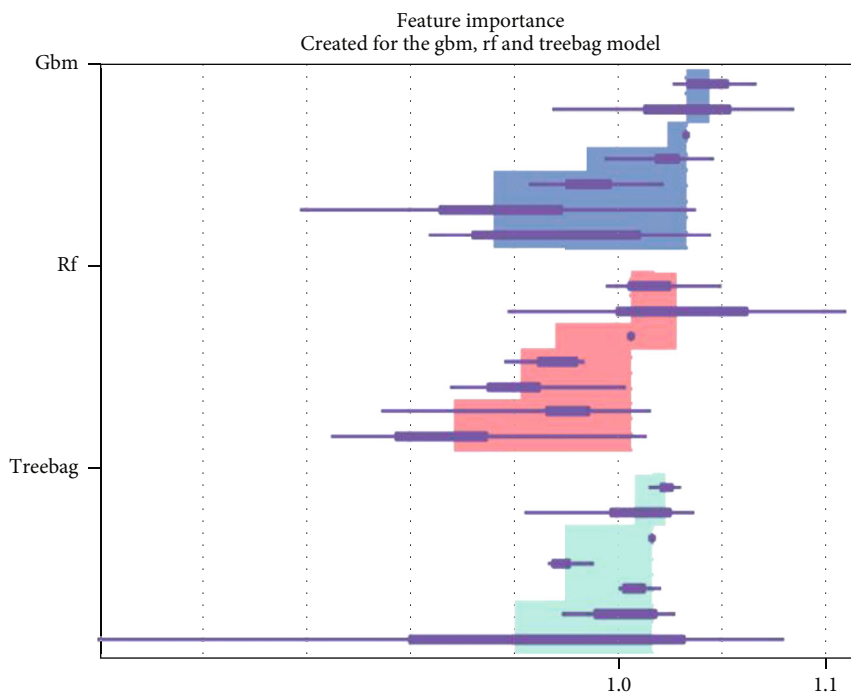


(b)

FIGURE 3: Continued.



(c)

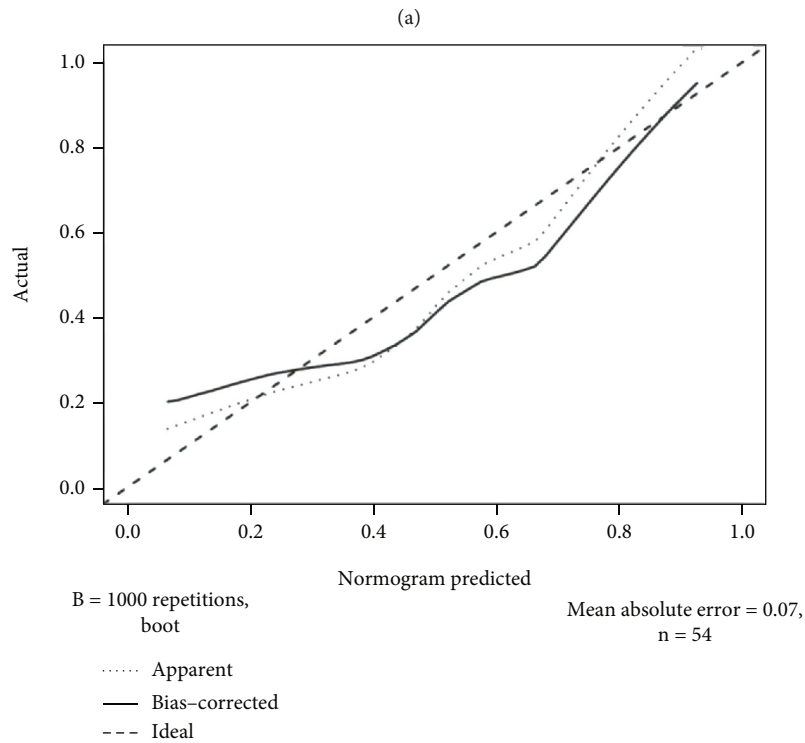
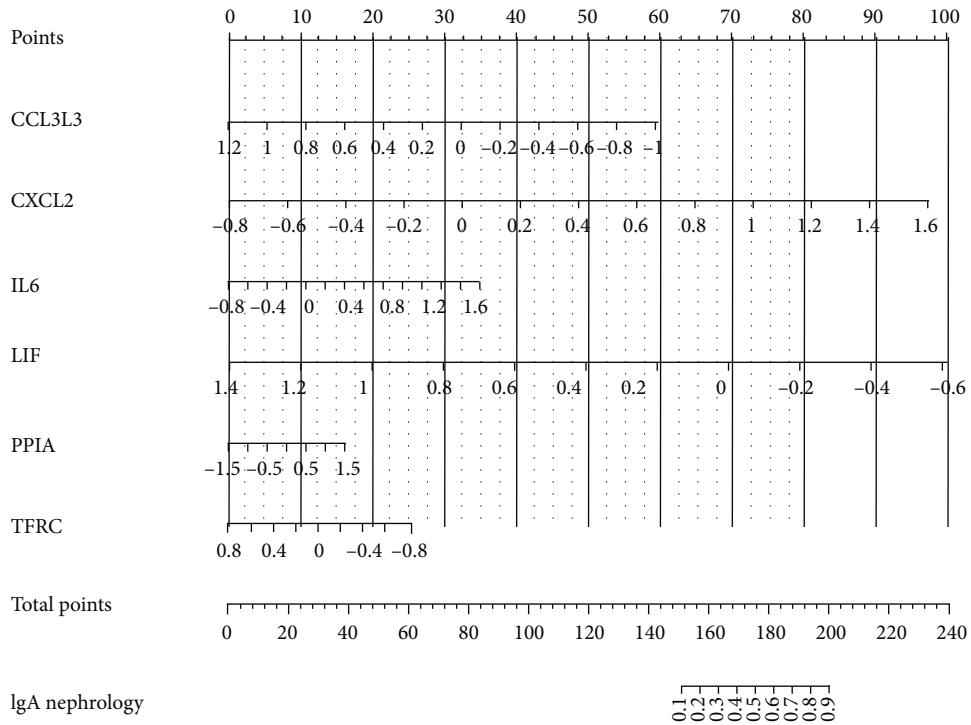


(d)

FIGURE 3: Construction and assessment of RF, GBM, and treebag model. (a) Elastic net of 34 DIRGs in GSE116626. The main parameters are  $\alpha = 1$  and  $\lambda = 6$  (CV.Glmnet function automatically produces the most appropriate value.). (b) Cumulative residual distribution map of the sample. (c) Boxplots of the residuals of the sample. Red dot stands for root mean square of residuals. The residual distributions were very close in the three models. (d) Importance of the variables in RF, GBM, and treebag model. The six DIRGs have different importance in three models.

The three models were then analyzed using the explanatory features of the “DALEX” package in R, and the residual distributions were plotted to obtain the optimal model at the bottom of the test set (Figures 3(b) and 3(c)); additionally,

Figure 3(d) shows us the importance of six genes. The results showed that the three machine learning models enjoy a good performance; ML was successfully used to identify six genes that could be applied to clinical diagnosis and prediction.



(b)

FIGURE 4: Continued.



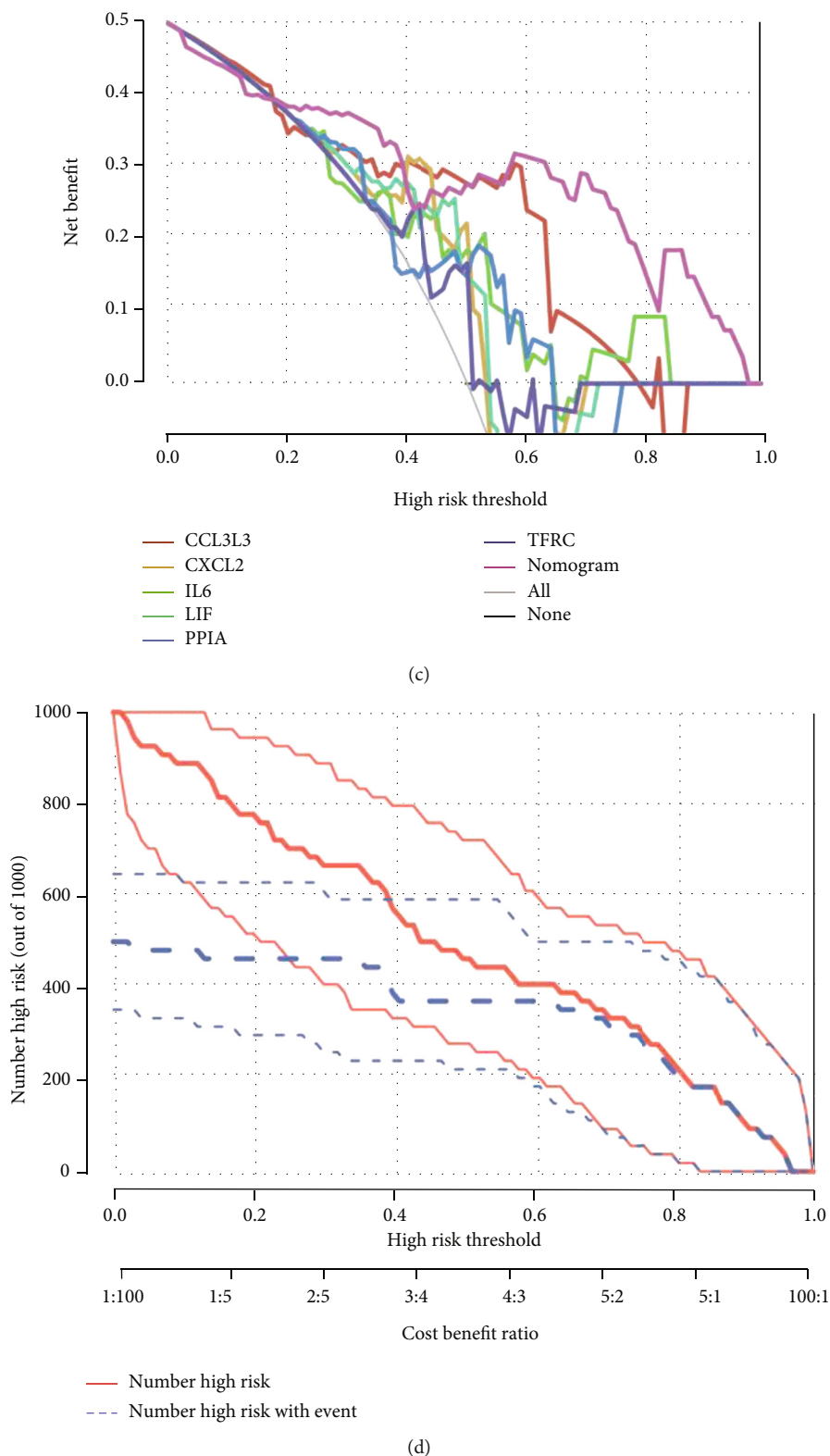
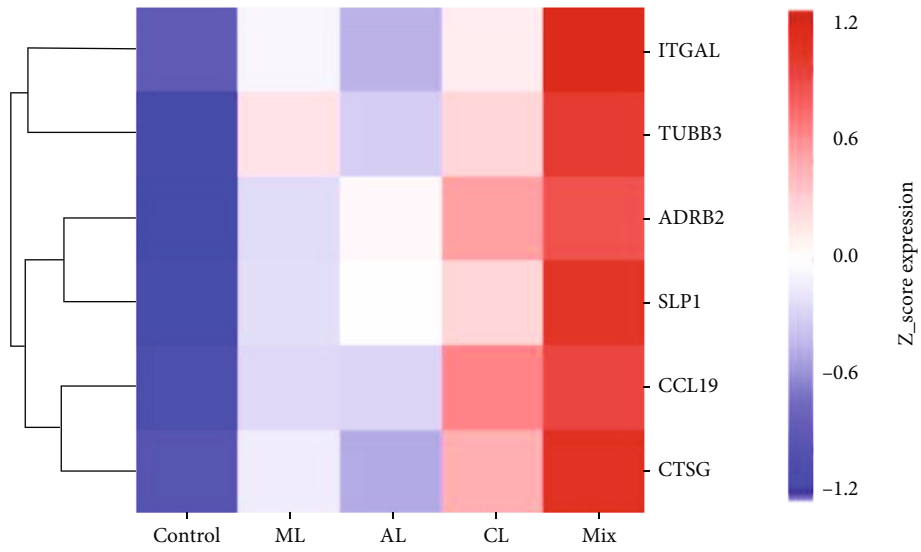
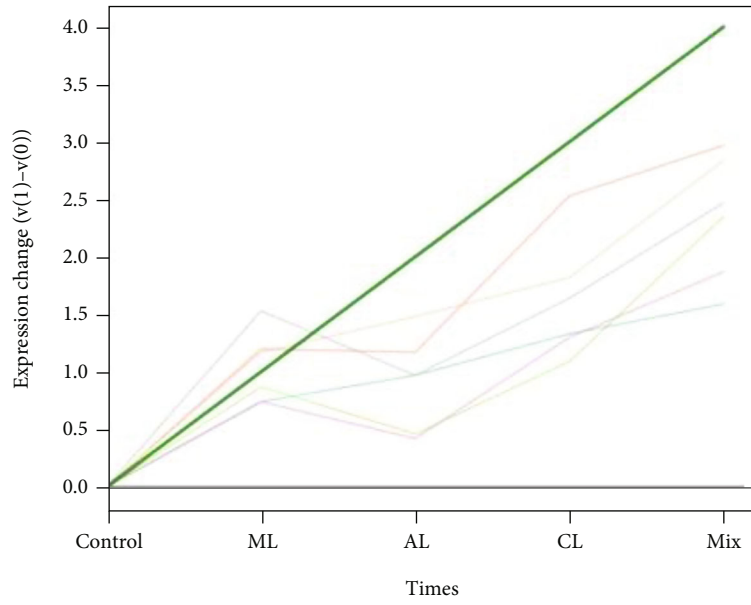


FIGURE 4: Construction and validation of a nomogram model for IgAN diagnosis. (a) Nomogram model for IgAN diagnosis, based on the 6 DIRGs (PPIA, CCL3L3, CXCL2, TFRC, IL6, and LIF). (b) Calibration curve to evaluate the nomogram model. The actual IgAN risk and the predicted risk are very close. (c) DCA curve to assess the nomogram model. Different colors represent different combinations of variables, and the nomogram model has better benefits than other models in most risk thresholds. (d) The clinical impact curve based on the DCA curve to evaluate the nomogram model, and the nomogram model has a better clinical effect when the risk threshold exceeded 0.6.



(a)

Profile model: (0.0,1.0,2.0,3.0,4.0)  
6 Genes assigned: 0 genes expected: p-value (FDR) = 0.00017



(b)

FIGURE 5: Continued.

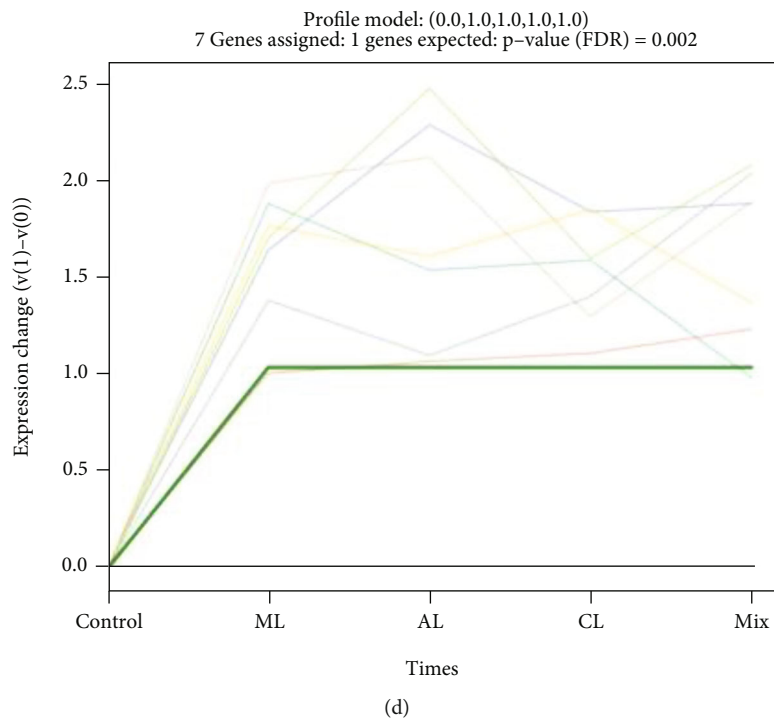
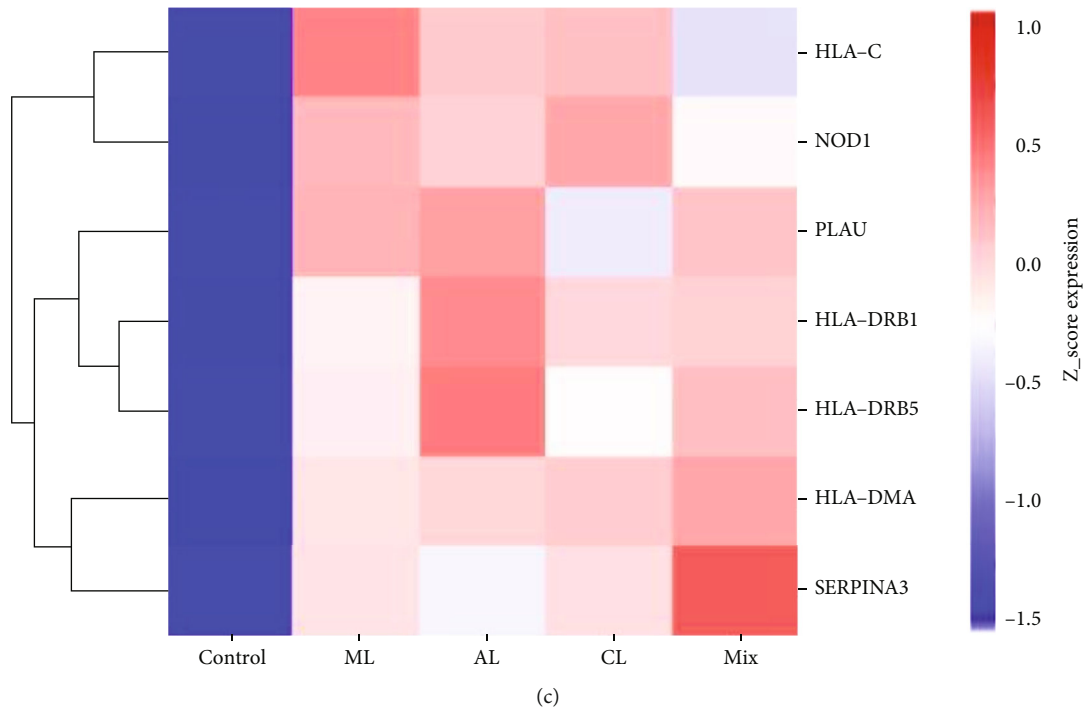
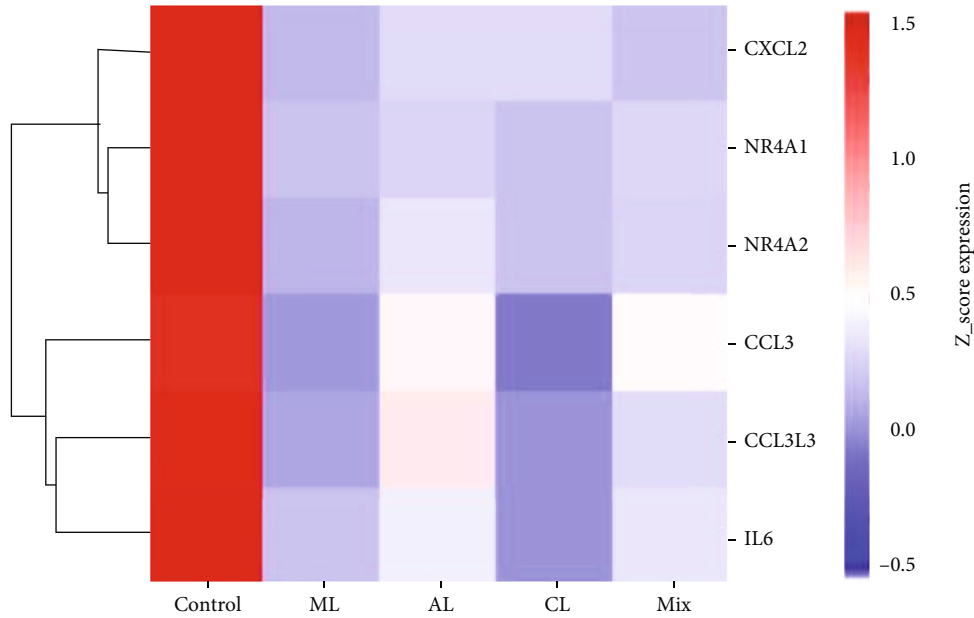
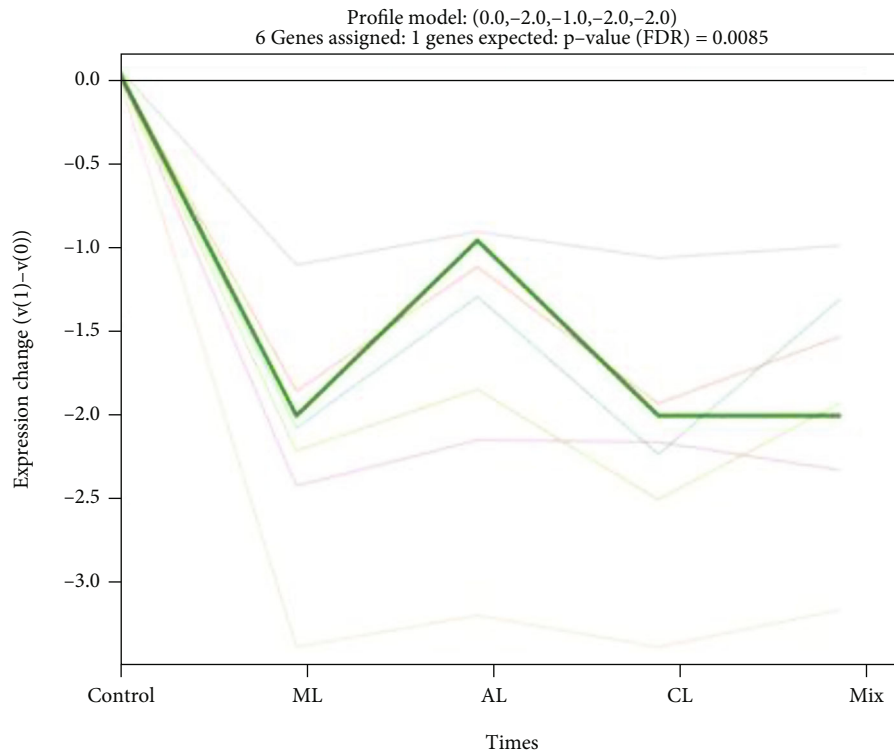


FIGURE 5: Continued.



(e)



(f)

FIGURE 5: Short Time-Series Expression Miner (STEM) analysis. (a) Heatmap of six gene expressions in trend 1, including 6 DIRGs, ITGAL, TUBB3, ADRB2, SLP1, CCL19, and CTSG. (b) Line chart of six gene expressions in trend 1, FDR = 0.00017. (c) Heatmap of seven gene expressions in trend 2, including 7 DIRGs, HLA-C, NOD1, PLAU, HLA-DRB1, HLA-DRB5, HLA-DMA, and SERPINA3. (d) Line chart of seven gene expressions in trend 2, FDR = 0.002. (e) Heatmap of six gene expressions in trend 3, including 6 DIRGs, CXCL2, NR4A1, NR4A2, CCL3, CCL3L3, and IL6. (f) Line chart of six gene expressions in trend 3, FDR = 0.0085.

3.3. Construction and Assessment of a Nomogram Model for IgAN Diagnosis. The “rms” package was used to establish the nomogram model for IgAN diagnosis based on the 6 DIRGs (PPIA, CCL3L3, CXCL2, TFRC, IL6, and LIF) (Figure 4(a)).

The calibration curve which was used to evaluate the predictive power of the nomogram model indicated that the actual IgAN risk and the predicted risk are very close, suggesting that the nomogram model owns high accuracy to predict

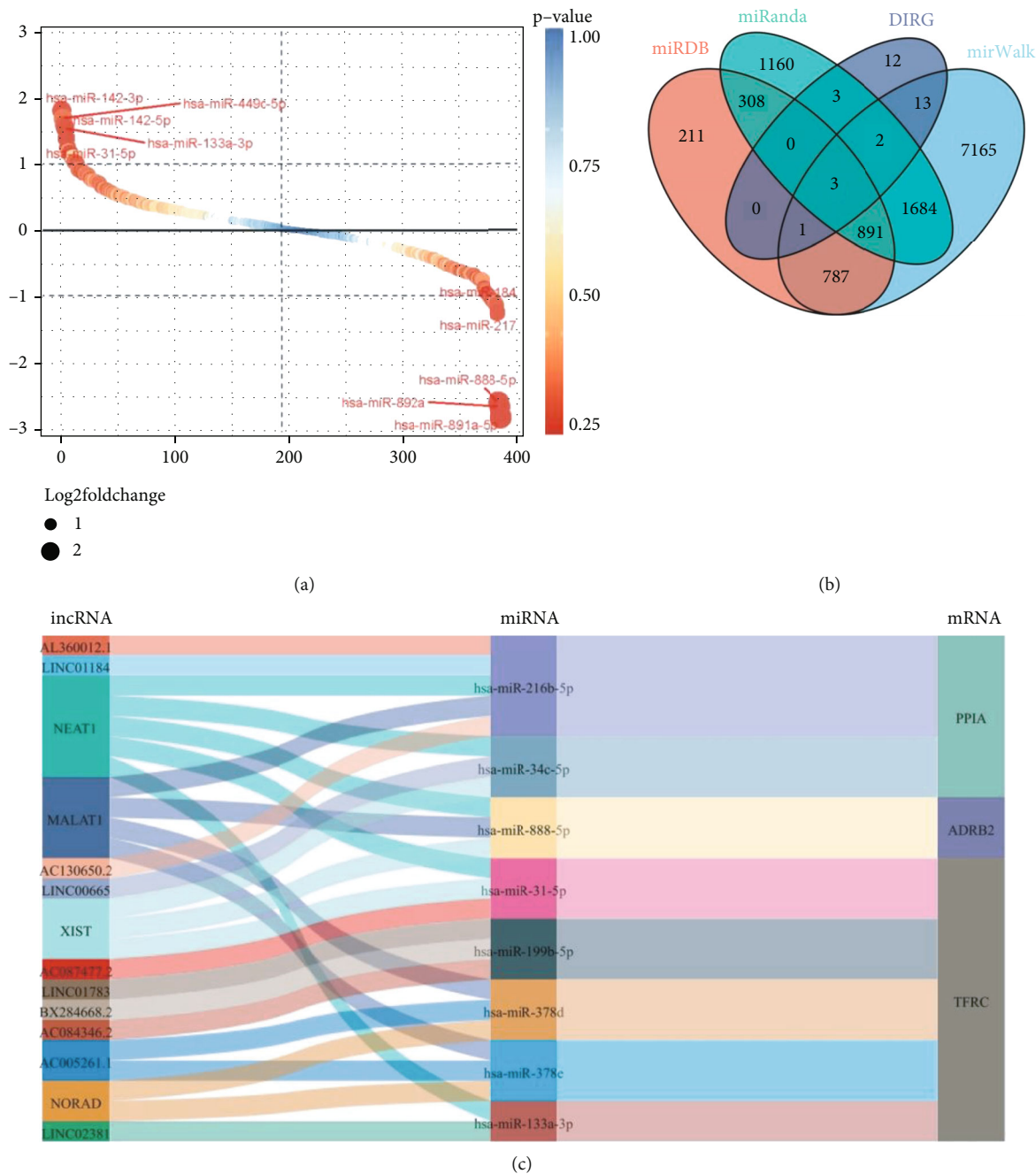
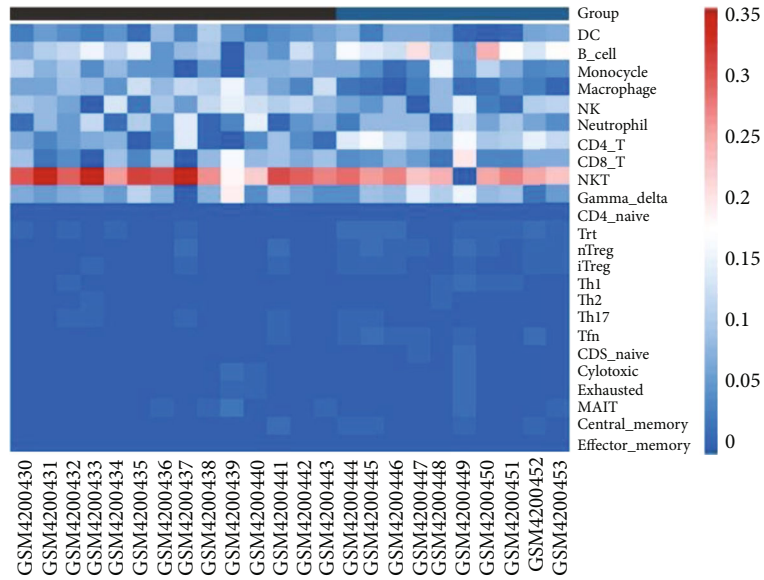


FIGURE 6: Identification of ceRNA related to IgAN progression. (a) Gene rank of DE-miRNAs in GSE141344. 10 upregulated and 10 downregulated DE-miRNAs were identified using the package “DESq2” at  $P < 0.05$  and  $|\log_2 \text{fold change}| > 1$ . (b) Venn diagram of 34 DIRGs and the data from miRDB, miWALK, and miRanda databases. There were three common targeted DIRGs (PPIA, ADRB2, and TFRC) in these three databases. (c) ceRNA network related to IgAN progression, including 8 miRNAs, 14 lncRNAs, and 3 mRNAs.

IgAN (Figure 4(b)). Furthermore, decision curve analysis (DCA) indicated that the “nomogram” curve had the highest benefits in most risk threshold ranges, especially if the risk threshold exceeds 0.6, indicating that the nomogram model boosts good predictions (Figure 4(c)), and the clinical impact curve on the ground of the DCA curve was conducted to evaluate the clinical effects of the nomogram model more visually (Figure 4(d)), which indicated that the prediction of the nomogram model is very close to the actual

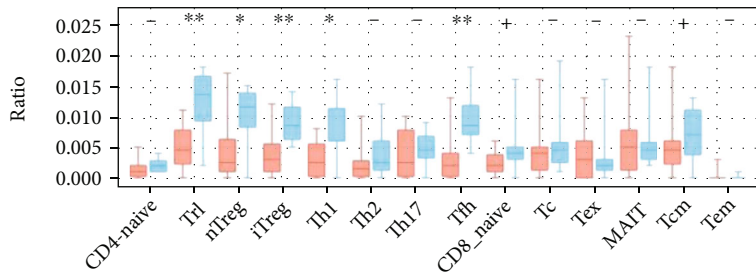
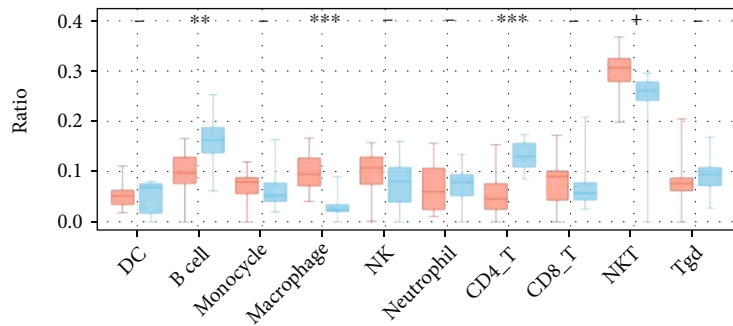
event when the risk threshold exceeds 0.6. The results indicated that the nomogram model owns great potential in predicting the risk of IgAN.

**3.4. Short Time-Series Expression Miner (STEM) Analysis.** STEM software was used to determine the alteration in 34 DIRGs during the progression from healthy controls to different IgAN lesions. These genes were divided into three significant clusters (Figures 5(a)–5(f)). Among them, ITGAL,



Group  
 ■ IgAN  
 ■ Control

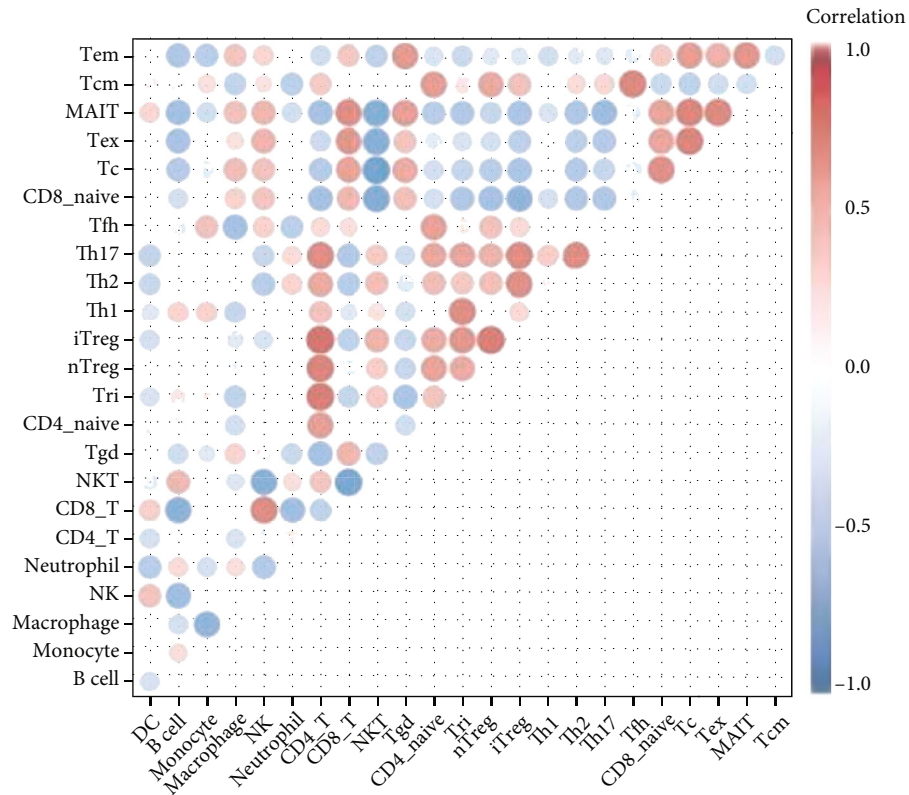
(a)



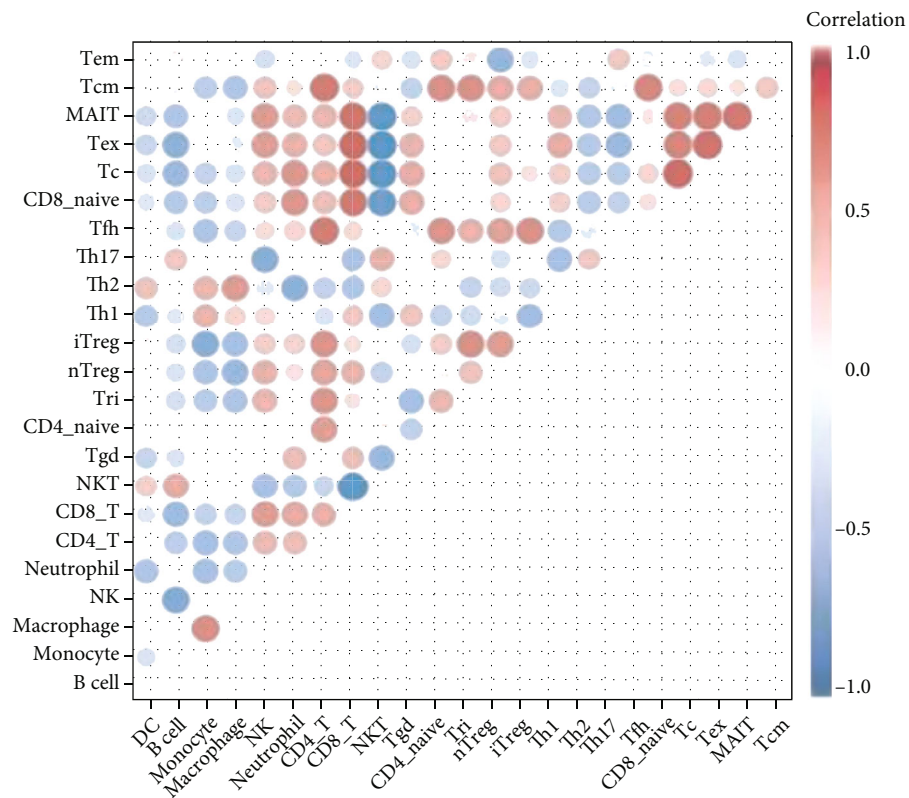
Group  
 ● IgAN  
 ● Control

(b)

FIGURE 7: Continued.



(c)



(d)

FIGURE 7: Continued.

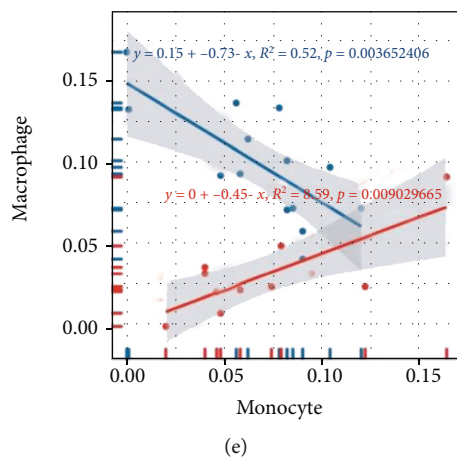


FIGURE 7: Immune signatures of the glomeruli in IgAN. (a) Heatmap of 24 immune cells calculated with ssGSEA in IgAN and healthy controls. (b) Boxplot of 24 immune cells in IgAN and healthy controls. Macrophages and NKT cells were significantly increased, while B cells, CD4+ T cells, Tr1, Treg, and Th1 cells were significantly decreased in IgAN. (c) Correlation heatmap of 10 main immune cells in IgAN. (d) Correlation heatmap of 10 main immune cells in healthy controls. (e) Line regression of monocytes and macrophages in IgAN and healthy controls. The proportion of monocytes and macrophages was negatively correlated in IgAN ( $R^2 = 0.52$ ,  $P = 0.003$ ) but negatively correlated in the control group ( $R^2 = 0.59$ ,  $P = 0.009$ ).

TUBB3, ADRB2, SLP1, CCL9, and CTSG showed a trend of gradual increase with different IgAN stages, with a peak at the mixed lesion, suggesting that these six genes may be associated with different pathological types of IgAN alterations. These genes in the three significant clusters may represent biomarkers of IgAN lesions.

**3.5. Identification of ceRNA Related to IgAN Progression.** We first detected 10 upregulated and 10 downregulated DE-miRNAs in IgAp using the package “DESeq2” at  $P < 0.05$  and  $|\log_2 \text{fold change}| > 1$  (Figure 6(a)), after which miRDB, miRWalk, and miRanda databases were employed for target predictions of these 20 DE-miRNAs (Supplementary Tables 4–6). Surprisingly, PPIA, ADRB2, and TFRC were obtained in comparison to 34 DIRGs of IgAN in the GSE116626 (Figure 6(b)). Also, lncRNAs, which target the DE-miRNA related to PPIA, ADRB2, and TFRC, were explored by the starBase v 3.0 database (Supplementary Table 7). Finally, we established a ceRNA network with 8 miRNAs, 14 lncRNAs, and 3 mRNAs (Figure 6(c)).

**3.6. Immune Signatures of the Glomeruli in IgAN.** ssGSEA analysis was first applied to explore the abnormal glomerular immune infiltration in IgAN. Afterwards, 24 types of immune cells in IgAN patients and controls were generated and compared, as shown in Figure 7(a). Clearly, IgAN patients showed a significant increase in macrophages and NKT cells, and a significant decrease in B cells, CD4+ T cells, Tr1, Treg, and Th1 cells in the glomeruli (Figures 7(b) and 7(c)). Besides, by comparing correlations between 10 main immune cells, correlation changes were detected between IgAN patients (Figure 7(d)) and controls (Figure 7(e)), among which changes between macrophages and monocytes are the most obvious (Figure 7(f)). As such, the occurrence of IgAN is closely related to abnormal immune cell infiltration, and the correlation change between immune cells may be implicated in the occurrence and progression of IgAN.

## 4. Discussion

With the increase in kidney biopsy, the underestimated prevalence of IgAN will be revealed [24]. Many studies have shown that IgAN involves not only young adults but also the elderly [25]. The existing diagnosis owns some limitations for early diagnosis of IgAN, and the pathogenesis of IgAN is extremely complicated, involving biochemical, immunological, and genetic [26], which brings challenges to the prevention and therapy of IgAN. As such, new biomarkers and diagnostic models are needed urgently. Recent genetic findings have confirmed the strong role of genetics and suggested geographical and ethnic differences in IgAN susceptibility, among which Asians are more inclined to genetic risks [10].

**4.1. The Function of the DIRGs.** The enrichment analysis indicated that 34 DIRGs not only represent an important player in IgAN but also serve as a bridge between IgAN and other immune diseases. GO analysis showed that the interferon- $\gamma$ -mediated signaling pathway is closely associated with IgAN, and the production of IFN- $\gamma$  in IgAN was positively correlated with total IgA levels [27]. Meanwhile, altered ERK1/2 expression, which has been shown to highly influence mesangial cell proliferation [28], positively regulates the ERK1 and ERK2 cascades that may be involved in IgAN development. Positive regulation of the I-kappaB kinase/NF-kappaB signaling pathway would promote IgAN [29], and many studies have shown that drugs could treat IgAN by inhibiting this pathway [30]. Regulation of related gene expression may have a positive effect on IgAN.

What is more, KEGG enrichment analysis showed that T cell helper (Th) differentiation, TNF signaling pathway, and IgA-generating intestinal immune network are involved in the pathogenesis of IgAN. IgAN is regulated by Th [31]. Interestingly, it has been shown that the use of antitumor necrosis factor- (TNF-) alpha agents in rheumatoid arthritis (RA) could lead to the development of IgAN and diabetes



[32, 33], which may be interrelated because of similar abnormal gene expression and thus have some signaling pathways running through many diseases, such as inflammatory bowel disease (IBD) [34]. Notably, intestinal immunity is a major player in IgAN [35], and it has become an important foothold for IgAN treatment and exploration, but the mechanisms involved are complex [36]. The discovery of DIRGs may help us to further uncover their underlying pathogenesis. The results of our enrichment analysis for 34 DIRGs are parallel to the current study, and exploring the pathogenesis from a genetic perspective may provide new ideas for treatment of IgAN.

#### 4.2. The Models Used for Diagnosis and Prediction of IgAN.

We first employed elastic net to screen six genes (PPIA, CCL3L3, CXCL2, TFRC, IL6, and LIF) from 34 DIRGs, which were used for model construction, including RF, tree-bag, and GBM models. It turned out that three models performed excellently, the residuals of the sample in three models are small and close, which indicated that the six DIRGs were used to diagnose and predict IgAN with credible precision and accuracy. We managed to narrow down the number of IRGs to six, which is of great significance for clinical practice. Afterwards, the six genes were utilized for the construction of the nomogram model, which owns great prediction performance from the nomogram model at a high-risk threshold from 0 to 1; the calibration curve, DCA curve, and clinical impact curve indicated that the nomogram model boosts good predictions. With growing renal biopsies in patients with kidney disease, patients can benefit from the nomogram model.

#### 4.3. STEM Analysis and ceRNA Network for Exploring the Progression of IgAN.

The possible space-time variations of the 34 DIRGs in different IgAN lesions were discovered through STEM analysis, which could help find potential biomarkers related to disease progression. One of the most interesting changes was in ITGAL TUBB3, ADRB2, SLP1, and CCL19 CTSG, with slightly lower expression of the minimal lesion and active lesion and significantly higher than the chronic and mixed lesion. ITGAL is involved in various immune phenomena including leukocyte-endothelial cell interactions, cytotoxic T cell-mediated killing, and antibody-dependent killing by granulocytes and monocytes [37]. CCL19 plays an important role in the trafficking of T cells in the thymus and T cell and B cell migration to secondary lymphoid organs [38]; CTSG regulates inflammation, activating matrix metalloproteinases (MMPs) and coagulation [39]. Besides, TUBB3 [40], ADRB2 [41], and SLP1 [42] are also involved in the immune response, such as cytokine network and antibacterial response. Furthermore, HLA-DRB1, HLA-DRB5, and HLA-DMA are broadly expressed in different lesions of IgAN, with the strongest correlation between HLA and IgAN [43]. Studies have shown that HLA-DB1 may be associated with IgAN progression [44]. These genes may serve as biomarkers for predicting IgAN lesion and progression as IgA progression is closely related to changes in gene expression.

Moreover, we further analyzed miRNAs associated with IgAN progression and generated a ceRNA network of 8 miRNAs, 14 lncRNAs, and 3 mRNAs based on miRDB miWALK and miRNADA and starBase v 3.0. miRNAs have been associated with immune and pathological changes in the kidney through regulating gene expression [45]. The TFRC gene involved in ferroptosis and some other important pathogenesis has been linked to a variety of diseases [46] and may be linked to IgAN susceptibility; meanwhile, NEAT1, MALAT1, and XIST which have the most degree with miRNA in the ceRNA network were identified; NEAT1 may affect IgAN by regulating the TLR2/NF- $\kappa$ B signaling pathway [47]. Besides, research once showed that IgAN mesangial cells displayed increased expression of MALAT1 [48], and XIST could mediate inflammatory response via NF- $\kappa$ B/NLRP3 inflammasome pathway, which may be the potential pathomechanism by which it affects the IgAN progression.

#### 4.4. The Changes of Immune Landscape in the Glomeruli of IgAN.

IgAN is the most common primary glomerulonephritis worldwide and is characterized by abnormal immune cell infiltration [49]. ssGAEA analysis was applied to explore the immunological signatures of IgAN in glomerular biopsy tissue. Deposits of immune complexes containing galactosylated IgA1 activate mesangial cells, leading to an overproduction of local cytokines, chemokines, and complement [50]. Macrophages, which are significantly higher in IgAN, could amplify the effects of inflammatory factors and participate in and contribute to the process of mesangial hyperplasia and glomerulosclerosis [51]. The prolonged macrophage infiltration may promote mesangial cell proliferation or the development of extracapillary lesions that eventually lead to the progression of IgAN [52]. Besides, activation of NKT cells is associated with the production of mucosal IgA [53] and an increased ratio of NKT is involved in IgAN immunopathology [54].

Additionally, the ratio of CD4+ T cells is significantly reduced in IgAN patients. Its reduction is mainly manifested in regulatory T cell (Treg) and helper T cell (Th). Since the reduction in Treg is associated with abnormal immune function in patients with IgAN [55], it is of great significance for the prognosis of IgAN. Thus, increasing the ratio of Treg may protect the renal function of patients [56, 57]. Meanwhile, the reduction of Th1 and Tfh in glomeruli is an integral part of IgAN and they play an active role in maintaining a normal immune response to IgAN and stopping its progression [58, 59]. Although B cell infiltration also increased in IgAN, it owns an insignificant effect. Finally, by analyzing the correlation between immune cells, we found that the correlation between macrophages and monocytes in the glomeruli of IgAN patients was inverse. The results may indicate a significant increase in differentiation of monocytes into macrophages, which would emerge as a potential mechanism and therapeutic target for IgAN.

Although lots of work have been done, there are some limitations in our study. Firstly, the clinical information of the samples is not rich enough, such as renal function, urine protein, and blood pressure; more risk factors make more

accurate model. Secondly, the expression level of DIRGs and the immune landscape of IgAN may need further verification. We will enrich and verify our conclusions on real clinical cases in the following study.

## 5. Conclusion

The diagnostic model of IgAN is constructed using IRGs in our study, and we find that the imbalance between macrophages and CD4+ immune cells may be an important pathomechanism of IgAN. These results provide potential directions for the treatment and prevention of IgAN. Our study draws on a wide range of data and methods to reach rich conclusions that will be translated into clinical practice in the future, and we will explore new directions for IgAN treatment based on our findings in this study.

## Data Availability

The raw data supporting the conclusions of this article will be made available in the GEO database. All data generated or analyzed during this study are included in this article and its supplementary information files.

## Consent

Consent is not applicable.

## Conflicts of Interest

The authors declare that they have no competing interests.

## Authors' Contributions

Jianbo Qing collected and processed the data and drafted the paper. Yafeng Li and Changqun Li gave constructive suggestions for the manuscript. Wenzhu Song helped polish the manuscript. All authors approved the final version.

## Acknowledgments

We express our sincere gratitude to those who helped us a lot during our writing process. This work was supported by the National Nature Science Foundation of China (8217032147), the Key Laboratory Construction Plan Project of Shanxi Provincial Health Commission (2020SYS01), the Key Project of Shanxi Provincial Health Commission (2020XM21), and the Research Project Supported by Shanxi Scholarship Council of China (2020-183).

## Supplementary Materials

Supplementary Table 1: the immune-related genes downloaded from the import database (<https://www.immport.org/shared/>). Supplementary Table 2: all data of GO enrichment analysis. Supplementary Table 3: all data of KEGG enrichment analysis. Supplementary Table 4: the target predictions of these 20 DE-miRNAs in the miRDB database. Supplementary Table 5: the target predictions of these 20 DE-miRNAs in the miRWalk database. Supplementary

Table 6: the target predictions of these 20 DE-miRNAs in the miRanda database. Supplementary Table 7: the lncRNAs which target the DE-miRNA related to PPIA, ADRB2, and TFRC in the starBase v 3.0 database. (*Supplementary Materials*)

## References

- [1] J. Hassler, "IgA nephropathy: a brief review," *Seminars in diagnostic pathology*, vol. 37, no. 3, pp. 143–147, 2020.
- [2] E. Gutiérrez, F. Carvaca-Fontán, L. Luzardo, E. Morales, M. Alonso, and M. Praga, "A personalized update on IgA nephropathy: a new vision and new future challenges," *Nephron*, vol. 144, no. 11, pp. 555–571, 2020.
- [3] F. Schena and I. Nistor, "Epidemiology of IgA nephropathy: a global perspective," *Seminars in nephrology*, vol. 38, no. 5, pp. 435–442, 2018.
- [4] R. Wyatt and B. Julian, "IgA nephropathy," *The New England journal of medicine*, vol. 368, no. 25, pp. 2402–2414, 2013.
- [5] S. Reid, P. Cawthon, J. Craig et al., "Non-immunosuppressive treatment for IgA nephropathy," *The Cochrane database of systematic reviews*, no. 3, p. CD003962, 2011.
- [6] E. Pillebout and J. Vérine, "IgA nephropathy," *Nephrologie & thérapeutique*, vol. 12, no. 4, pp. 238–254, 2016.
- [7] H. Suzuki, "Biomarkers for IgA nephropathy on the basis of multi-hit pathogenesis," *Clinical and experimental nephrology*, vol. 23, no. 1, pp. 26–31, 2019.
- [8] G. D'Amico, "Natural history of idiopathic IgA nephropathy and factors predictive of disease outcome," *Seminars in nephrology*, vol. 24, no. 3, pp. 179–196, 2004.
- [9] Y. Neugut and K. Kiryluk, "Genetic determinants of IgA nephropathy: western perspective," *Seminars in nephrology*, vol. 38, no. 5, pp. 443–454, 2018.
- [10] R. Magistroni, V. D'Agati, G. Appel, and K. Kiryluk, "New developments in the genetics, pathogenesis, and therapy of IgA nephropathy," *Kidney international*, vol. 88, no. 5, pp. 974–989, 2015.
- [11] S. Bhattacharya, P. Dunn, C. G. Thomas et al., "ImmPort, toward repurposing of open access immunological assay data for translational and clinical research," *Scientific data*, vol. 5, no. 1, article 180015, 2018.
- [12] M. Ritchie, B. Phipson, D. Wu et al., "Limma powers differential expression analyses for RNA-seq and microarray studies," *Nucleic acids research*, vol. 43, no. 7, article e47, 2015.
- [13] G. Yu, L. Wang, Y. Han, and Q. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *OmicS: a journal of integrative biology*, vol. 16, no. 5, pp. 284–287, 2012.
- [14] S. Chung and L. Abbott, "Neural population geometry: an approach for understanding biological and artificial neural networks," *Current Opinion in Neurobiology*, vol. 70, pp. 137–144, 2021.
- [15] Y. Sanjougand, "Elastic network models: theoretical and empirical foundations," *Methods in molecular biology (Clifton, NJ)*, vol. 924, pp. 601–616, 2013.
- [16] J. Ernst and Z. Bar-Joseph, "STEM: a tool for the analysis of short time series gene expression data," *BMC Bioinformatics*, vol. 7, no. 1, 2006.
- [17] M. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2," *Genome biology*, vol. 15, no. 12, 2014.

- [18] Y. Chen and X. Wang, "miRDB: an online database for prediction of functional microRNA targets," *Nucleic acids research.*, vol. 48, no. D1, pp. D127–D131, 2020.
- [19] H. Dweep, N. Gretz, and C. Sticht, "miRWalk database for miRNA-target interactions," *Methods in molecular biology (Clifton, NJ)*, vol. 1182, pp. 289–305, 2014.
- [20] E. Gu, W. Song, A. Liu, and H. Wang, "SCDb: an integrated database of stomach cancer," *BMC Cancer*, vol. 20, no. 1, p. 490, 2020.
- [21] Y. Gao, H. Ma, and D. Hou, "Sevoflurane represses proliferation and migration of glioma cells by regulating the ANRIL/let-7b-5p axis," *Cancer biotherapy & radiopharmaceuticals*, 2020.
- [22] S. Jin, R. Zhou, X. Guan, J. Zhou, and J. Liu, "Identification of novel key lncRNAs involved in periodontitis by weighted gene co-expression network analysis," *Journal of Periodontal Research*, vol. 55, no. 1, pp. 96–106, 2020.
- [23] S. Hänzelmann, R. Castelo, and J. Guinney, "GSVA: gene set variation analysis for microarray and RNA-seq data," *BMC Bioinformatics*, vol. 14, no. 1, 2013.
- [24] M. Perše and Ž. Večerić-Haler, "The role of IgA in the pathogenesis of IgA nephropathy," *International journal of molecular sciences.*, vol. 20, no. 24, p. 6199, 2019.
- [25] W. Cheungpasitporn, S. Nasr, C. Thongprayoon, M. Mao, and Q. Qian, "Primary IgA nephropathy in elderly patients," *Nephrology (Carlton, Vic.)*, vol. 20, no. 6, pp. 419–425, 2015.
- [26] H. Suzuki, K. Kiryluk, J. Novak et al., "The pathophysiology of IgA nephropathy," *Journal of the American Society of Nephrology: JASN*, vol. 22, no. 10, pp. 1795–1803, 2011.
- [27] S. Fujieda, S. Suzuki, H. Sunaga, H. Yamamoto, M. Seki, and H. Sugimoto, "Production of interferon- $\gamma$  by tonsillar mononuclear cells in IgA nephropathy patients," *Acta oto-laryngologica.*, vol. 120, no. 5, pp. 649–654, 2000.
- [28] Y. Gorin, J. Ricono, B. Wagner, N. Kim, B. Bhandari, and G. Choudhury, "Angiotensin II-induced ERK1/ERK2 activation and protein synthesis are redox-dependent in glomerular mesangial cells," *The Biochemical journal.*, vol. 381, no. 1, pp. 231–239, 2004.
- [29] H. Zhang and S. Sun, "NF- $\kappa$ B in inflammation and renal diseases," *Cell & Bioscience*, vol. 5, no. 1, 2015.
- [30] C. Wu, K. Hua, W. Hsu et al., "IgA nephropathy benefits from compound K treatment by inhibiting NF- $\kappa$ B/NLRP3 inflammasome and enhancing autophagy and SIRT1," *Journal of immunology (Baltimore, Md: 1950)*, vol. 205, no. 1, pp. 202–212, 2020.
- [31] S. Chintalacheruvu, M. Yamashita, N. Bagheri et al., "T cell cytokine polarity as a determinant of immunoglobulin A (IgA) glycosylation and the severity of experimental IgA nephropathy," *Clinical and experimental immunology.*, vol. 153, no. 3, pp. 456–462, 2008.
- [32] V. Di Lernia, "IgA nephropathy during treatment with TNF-alpha blockers: could it be predicted?," *Medical hypotheses.*, vol. 107, pp. 12–13, 2017.
- [33] V. Premužić, I. Padjen, M. Cerovec, M. Ćorić, B. Jelaković, and B. Anić, "The association of TNF-alpha inhibitors and development of IgA nephropathy in patients with rheumatoid arthritis and diabetes," *Case reports in nephrology.*, vol. 2020, pp. 1–8, 2020.
- [34] J. Ambruzs and C. Larsen, "Renal manifestations of inflammatory bowel disease," *Rheumatic diseases clinics of North America.*, vol. 44, no. 4, pp. 699–714, 2018.
- [35] R. Coppo, "The gut-renal connection in IgA nephropathy," *Seminars in nephrology.*, vol. 38, no. 5, pp. 504–512, 2018.
- [36] N. Zhou, Y. Shen, L. Fan et al., "The characteristics of intestinal-barrier damage in rats with IgA nephropathy," *The American journal of the medical sciences.*, vol. 359, no. 3, pp. 168–176, 2020.
- [37] T. Browne, R. Dearman, and A. Poles, "Human neutrophil antigens: nature, clinical significance and detection," *International Journal of Immunogenetics*, vol. 48, no. 2, pp. 145–156, 2021.
- [38] R. Förster, A. Davalos-Misslitz, and A. Rot, "CCR7 and its ligands: balancing immunity and tolerance," *Nature reviews Immunology.*, vol. 8, no. 5, pp. 362–371, 2008.
- [39] L. Pérez-Is, M. Ocaña, A. Montes et al., "The N125S polymorphism in the cathepsin G gene (rs45567233) is associated with susceptibility to osteomyelitis in a Spanish population," *PLoS One*, vol. 14, no. 10, article e0220022, 2019.
- [40] K. Sobierajska, K. Wiczorek, W. Ciszewski et al., " $\beta$ -III tubulin modulates the behavior of snail overexpressed during the epithelial-to-mesenchymal transition in colon cancer cells," *Biochimica et Biophysica Acta*, vol. 1863, no. 9, pp. 2221–2233, 2016.
- [41] F. Wu, T. Fang, L. Yu et al., "ADRB2 signaling promotes HCC progression and sorafenib resistance by inhibiting autophagic degradation of HIF1 $\alpha$ ," *Journal of hepatology.*, vol. 65, no. 2, pp. 314–324, 2016.
- [42] Y. Wada, K. Kitamoto, T. Kanbe, K. Tanaka, and Y. Anraku, "The SLP1 gene of *Saccharomyces cerevisiae* is essential for vacuolar morphogenesis and function," *Molecular and cellular biology.*, vol. 10, no. 5, pp. 2214–2223, 1990.
- [43] J. Feehally, M. Farrall, A. Boland et al., "HLA has strongest association with IgA nephropathy in genome-wide analysis," *Journal of the American Society of Nephrology: JASN.*, vol. 21, no. 10, pp. 1791–1797, 2010.
- [44] J. In, K. Jung, S. Shin, K. Park, H. Lee, and E. Song, "Association of HLA-DRB1 and -DQB1 alleles with susceptibility to IgA nephropathy in Korean patients," *Annals of laboratory medicine.*, vol. 42, no. 1, pp. 54–62, 2022.
- [45] C. Szeto and P. Li, "MicroRNAs in IgA nephropathy," *Nature reviews Nephrology.*, vol. 10, no. 5, pp. 249–256, 2014.
- [46] H. Jabara, S. Boyden, J. Chou et al., "A missense mutation in *TFR1*, encoding transferrin receptor 1, causes combined immunodeficiency," *Nature genetics.*, vol. 48, no. 1, pp. 74–78, 2016.
- [47] S. Wang, G. Liu, H. Xian, J. Si, S. Qi, and Y. Yu, "LncRNA NEAT1 alleviates sepsis-induced myocardial injury by regulating the TLR2/NF- $\kappa$ B signaling pathway," *European review for medical and pharmacological sciences.*, vol. 23, no. 11, pp. 4898–4907, 2019.
- [48] R. Tang, T. Meng, W. Lin et al., "A partial picture of the single-cell transcriptomics of human IgA nephropathy," *Frontiers in immunology.*, vol. 12, article 645988, 2021.
- [49] H. Selvaskandan, C. Cheung, M. Muto, and J. Barratt, "New strategies and perspectives on managing IgA nephropathy," *Clinical and experimental nephrology.*, vol. 23, no. 5, pp. 577–588, 2019.
- [50] K. Lai, S. Tang, F. Schena et al., "IgA nephropathy," *Nature reviews Disease primers.*, vol. 2, no. 1, 2016.
- [51] L. Lv, Y. Feng, Y. Wen et al., "Exosomal CCL2 from tubular epithelial cells is critical for albumin-induced tubulointerstitial inflammation," *Journal of the American Society of Nephrology: JASN.*, vol. 29, no. 3, pp. 919–935, 2018.

- [52] M. Nagata, Y. Akioka, Y. Tsunoda et al., "Macrophages in childhood IgA nephropathy," *Kidney international.*, vol. 48, no. 2, pp. 527–535, 1995.
- [53] C. Wu, Z. Li, X. Fu, S. Yu, S. Lao, and B. Yang, "Antigen-specific human NKT cells from tuberculosis patients produce IL-21 to help B cells for the production of immunoglobulins," *Oncotarget*, vol. 6, no. 30, pp. 28633–28645, 2015.
- [54] H. Kaminski, L. Couzi, and M. Eberl, "Unconventional T cells and kidney disease," *Nephrology*, vol. 17, no. 12, pp. 795–813, 2021.
- [55] S. Yang, B. Chen, J. Shi, F. Chen, J. Zhang, and Z. Sun, "Analysis of regulatory T cell subsets in the peripheral blood of immunoglobulin A nephropathy (IgAN) patients," *Genetics and molecular research: GMR.*, vol. 14, no. 4, pp. 14088–14092, 2015.
- [56] F. Chen, Y. Ma, H. Ding, and B. Chen, "Effects of Tripterygium wilfordii glycosides on regulatory T cells and Th17 in an IgA nephropathy rat model," *Genetics and molecular research: GMR.*, vol. 14, no. 4, pp. 14900–14907, 2015.
- [57] G. Vlachopoulos, A. Georgalis, P. Korkolopoulou, E. Patsouris, and H. Gakiopoulou, "Immunohistochemical expression of FOXP3+ regulatory T cells in proteinuric primary glomerulopathies," *International journal of nephrology.*, vol. 2021, article 9961713, 2021.
- [58] J. Ruskowski, K. Lisowska, M. Pindel, Z. Heleniak, A. Dębska-Ślizień, and J. Witkowski, "T cells in IgA nephropathy: role in pathogenesis, clinical significance and potential therapeutic target," *Clinical and experimental nephrology.*, vol. 23, no. 3, pp. 291–303, 2019.
- [59] L. Zhang, Y. Wang, X. Shi, H. Zou, and Y. Jiang, "A higher frequency of CD4<sup>+</sup>CXCR5<sup>+</sup> T follicular helper cells in patients with newly diagnosed IgA nephropathy," *Immunology letters.*, vol. 158, no. 1-2, pp. 101–108, 2014.