Identification of key biomarkers and signaling pathways and analysis of their association with immune cells in immunoglobulin A nephropathy

GUOXIN ZHANG, LANFEN XUE, SHANSHAN ZHANG, NA LIU, XING YAO, JIEQIONG FU, LIMIN NIE

Department of Nephrology, Shijiazhuang People's Hospital, Hebei Province, China

Abstract

Introduction: Immunoglobulin A nephropathy (IgAN) is the most common glomerular disease worldwide, with a poor prognosis. The aim of our study was to identify key biomarkers and their associations with immune cells to aid in the study of IgAN pathology and immunotherapy.

Material and methods: The data of IgAN were downloaded from a public database. The metaMA package and limma package were used to identify differentially expressed mRNAs (DEmRNAs) and differentially expressed miRNAs (DEmiRNAs), respectively. Biological functions of the DEmRNAs were analyzed. Machine learning was used to screen the mRNA biomarkers of IgAN. Pearson's correlation coefficient was used to analyze the correlation between mRNA biomarkers, immune cells and signaling pathways. Moreover, we constructed a miRNAs-mRNAs targeted regulatory network. Finally, we performed in vitro validation of the identified miRNAs and mRNAs.

Results: 1205 DEmRNAs and 125 DEmiRNAs were identified. In gene set enrichment analysis (GSEA), tumor necrosis factor α (TNF- α) signaling via nuclear factor κB (NF- κB), apoptosis and MTORC-1 signaling were inhibited in IgAN. 8 mRNA biomarkers were screened by machine learning. In addition, the distribution of 8 immune cell types was found to be significantly different between normal controls and IgAN by difference analysis. Pearson correlation coefficient analysis demonstrated that AKAP8L was significantly negatively correlated with CD4⁺ memory T-cells. AKAP8L was also significantly negatively correlated with TNF- α signaling via NF- κB , apoptosis, and MTORC-1 signaling. Subsequently, 5 mRNA biomarkers predicted corresponding negative regulatory miRNAs.

Conclusions: The identification of 8 important biomarkers and their correlation with immune cells and biological signaling pathways provides new ideas for further study of IgAN.

Key words: immunoglobulin A nephropathy, AKAP8L, CD4⁺ memory T-cells, biomarkers, machine learning.

(Cent Eur J Immunol 2022; 47 (3): 189-205)

Introduction

Immunoglobulin A nephropathy (IgAN) is the most common form of primary glomerulonephritis and a major cause of renal failure, and its main feature is the deposition of IgA in the glomerular mesangium [1, 2]. IgAN is clinically associated with a variety of inflammatory and infectious diseases, liver diseases and cancers [3-5]. Its common clinical manifestations are gross hematuria, microscopic hematuria, proteinuria, and hypertension [6, 7]. In addition, IgAN is associated with significant mortality [8]. The current diagnostic criterion for IgAN remains the presence of dominant mesangial immunoglobulin A deposition in renal biopsy [9]. So far, there is no specific treatment for this disease. Therefore, the identification of new biomarkers for the disease is important for diagnosis, treatment and detection of disease progression.

The pathogenesis and progression of IgAN are related to the immune mechanism by single-cell transcriptomes [10]. Compared with the control group, IgAN patients had abnormal CD4⁺ and CD25⁺ cell levels [11]. Differentiation of CD4⁺ T cell subsets plays a vital role in the artemisinin and hydroxychloroquine combination therapy of IgAN [12]. The CD208⁺ dendritic cell level in tonsils is signifi-

Correspondence: Limin Nie, Department of Nephrology, Shijiazhuang People's Hospital, No. 365 Nan Street, Jianhua, Yuhua District, Shijiazhuang Ville 050000, China, phone: +86-17603119268, e-mail: 814320413@qq.com Received: 10.05.2022, Accepted: 10.08.2022

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0). License (http://creativecommons.org/licenses/by-nc-sa/4.0/)

cantly higher in IgAN patients, and is positively correlated with the proportion of crescentic glomeruli and urinary protein levels [13]. Compared with healthy subjects, IgAN patients had a higher percentage of activated and effector memory CD4⁺ T lymphocytes [14]. These findings suggest that exploring the distribution of immune cells will contribute to the study of IgAN pathogenesis and immunotherapy.

Although the specific pathogenesis of IgAN is still unclear, it has been found that a variety of molecules are involved in the regulation of its progression, which lays the foundation for in-depth understanding of the pathogenesis of IgAN. The expression level of core 1 synthase, glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1 (C1GALT1) is significantly down-regulated in IgAN patients and negatively correlated with the increased level of galactose-deficient IgA1 (Gd-IgA1) in peripheral B lymphocyte [15]. The expression level of toll-like receptor 4 (TLR-4) is increased in circulating monocytes of IgAN patients [16]. MiR-133a/133b inhibits regulatory T cell differentiation in IgAN through targeting forkhead box P3 (FOXP3) [17]. MiR-98-5p in pediatric IgAN patients may affect IgA1 glycosylation and thus regulate disease progression by targeting C1GALT1 [18]. In addition, differentially expressed genes participate in different biological signaling pathways in the progression of IgAN [19]. For example, sterile 20/SPS1-related proline/ alanine-rich kinase (SPAK) plays a pathogenic role in IgAN by activating the nuclear factor κ B/mitogen-activated protein kinase (NF- κ B/MAPK) signaling pathway [20]. Therefore, exploration of the IgAN molecular mechanism is helpful for the development of new management methods.

In this study, the data of IgAN were downloaded from the Gene Expression Omnibus (GEO) database. 8 mRNA biomarkers were screened by difference analysis and machine learning. The XCell tool was used to analyze the distribution of immune cells in IgAN and normal con-



Fig. 1. Flowchart of the study

GEO ID	Samples (normal control : IgAN)	Tissue type	Platform	Year	Author	Туре
GSE73953	2:15	PBMC	GPL4133	2016	Okuzaki D	mRNA
GSE58539	9:8	PBMC	GPL10558	2015	Cox SN	mRNA
GSE14795	8:12	PBMC	GPL96	2010	Cox SN	mRNA
GSE25590	7:7	Blood	GPL7731	2012	Serino G. Keller	miRNA

Table 1. Dataset retrieved from the GEO database

IgAN – immunoglobulin A nephropathy, PBMC – peripheral blood mononuclear cells

trols. Subsequently, correlations between mRNA biomarkers, immune cells and signaling pathways were analyzed. In addition, we analyzed miRNAs and constructed a miRNAmRNA targeted regulatory network (Fig. 1).

Material and methods

Data

The microarray datasets of IgAN were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/). The datasets were further chosen based on the inclusion criteria: 1) the dataset must be genome-wide mRNA/miRNA transcriptome expression data; 2) these data were obtained from blood samples of IgAN patients and normal control blood samples; 3) standardized or original datasets were considered in this study. Subsequently, four datasets were included our study, including three datasets of mRNA (GSE73953, GSE58539 and GSE14795) and 1 dataset of miRNA (GSE25590). The details of these datasets are shown in Table 1. The raw data of GSE73953, GSE58539 and GSE14795 were downloaded from the GEO database, and 12028 mRNAs shared by 3 sets of mRNA datasets were scale standardized (scale function in R, default argument). After combining the three datasets, the Combat package [21] was used to eliminate heterogeneity among the gene expression data.

Identification of DEmRNAs and DEmiRNAs

The metaMA package [22] was utilized to recognize differentially expressed mRNAs (DEmRNAs) between the IgAN patients and normal controls. The screening condition for the DEmRNAs was p value (p) < 0.05. The limma package [23] was applied to define the differentially expressed miRNAs (DEmiRNAs) between the IgAN patients and normal controls, and the selection criterion was p < 0.05. In addition, a heat map of DEmRNAs and DEmiRNAs was produced with the pheatmap package in R language (http://cran.r-project.org/web/packages/pheatmap/index.html).

Functional analysis

To uncover the potential biological function of DEmRNAs in IgAN, gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was performed using David 6.8 (https://david.ncifcrf.gov/). The selection criteria of GO and KEGG for significant enrichment were false discovery rate (FDR) < 0.05 and p < 0.05, respectively. Moreover, gene set enrichment analysis (GSEA) and gene set variation analysis (GSVA) were used to further disclose the pathways of significant enrichment. According to the reference gene set, h.all. v7.2.symbols.gmt, we used the GSEA 4.1.0 software tool (https://www.gsea-msigdb.org/gsea/index.jsp) to discover the pathways. P < 0.05 was regarded as statistically significant. Using the reference gene set described above, we used the GSVA package [24] to measure the signaling pathway variation score for each sample in IgAN. Significantly differential enriched gene sets were filtered at a cut-off criterion of p < 0.05.

Identification of optimal diagnostic mRNA biomarkers

Firstly, the least absolute shrinkage and selection operator (LASSO) algorithm in the glmnet package [25] was used to perform feature selection (a dimensionality reduction process to select important mRNAs). Then, the random forest and caret packages in R language [26] were used to find the optimal parameter, and the importance of these mRNAs was ranked from high to low according to the mean decrease in accuracy value. Subsequently, one mRNA was added from top to bottom according to the order of sorting results. The random forest algorithm was used to classify, and the 10-fold cross-validation process was used to calculate accuracy. In addition, the random forest package (https://cran.r-project.org/ web/packages/randomForest/) was employed to establish the random forest model. The rpart package (https:// cran.r-project.org/web/packages/rpart/) was used to build the decision tree model. The e1071 package (https:// cran.r-project.org/web/packages/e1071/index.html) was applied to establish the support vector machine (SVM) model. Finally, receiver operating characteristic (ROC) analysis was performed using the pROC package (http:// web.expasy.org/pROC/) in R to evaluate the diagnostic value of mRNA biomarkers.

Immune cell infiltration analysis

The level of immune cell infiltration in the immune microenvironment of each sample was analyzed using

the XCell tool (https://xcell.ucsf.edu/) according to the gene expression matrix. The cutoff value for significance was p < 0.05. To further explore the possible mechanism of mRNA biomarkers, Pearson correlation coefficient analysis was employed to calculate the correlation between immune cell types, mRNA biomarkers and signaling pathways.

Construction of the DEmiRNA-DEmRNA network

The interactions between the DEmiRNA and mRNA biomarkers were predicted using miRWalk (http://mirwalk.umm.uni-heidelberg.de/interactions/). Screening of DEmiRNAs that negatively regulate mRNA biomarkers was performed. Cytoscape software (http://www.cytoscape. org) was utilized to produce the miRNA-mRNA network.

Real-time polymerase chain reaction (real time-PCR) validation

Based on the above bioinformatics integration analysis results, 5 DEmRNAs (CD1C, TGFBRAP1, DHX32, AKAP8L and NFKBIA) and 2 DEmiRNAs (hsa-miR-922 and hsa-miR-509-5p) were selected for *in vitro* validation. Blood samples were collected from 11 IgAN patients and 19 healthy individuals. Total RNA was isolated using the RNAliquid ultra speed whole blood (liquid sample) total RNA extraction kit (Beijing HT-biotech Co., Ltd., China). The FastKing cDNA first strand synthesis kit (TIANGEN,

Table 2. Top 10 up/down-regulated DEmRNAs

China) and miRNA first strand cDNA synthesis kit (tailing method) (Sangon Biotech, China) were reverse transcribed to synthesize mRNA and miRNA, respectively. SuperReal PreMix Plus (SYBR Green) (TIANGEN, China) and the miRNA quantitative PCR kit (dye method) (Sangon Biotech, Ltd., China) were used to perform real time-PCR validation of mRNA and miRNA, respectively. Human ACTB, GAPDH and hsa-U6 were used as an internal reference. The $2^{-\Delta\Delta Ct}$ method was used to calculate the relative gene expression level [27].

Results

DEmRNAs and DEmiRNAs

According to the screen condition of p < 0.05, 1205 DEmRNAs were identified. Among them, 942 were up-regulated and 263 were down-regulated. The top 10 upregulated DEmRNAs were C19orf21, ARFRP1, PIK3R2, DHX32, PADI2, OLFM1, PAX8, SSTR2, TGFBRAP1 and EDC4 (Table 2). The top 10 down-regulated DEmRNAs were HUWE1, PDIA3, CDC25B, CDC42, CTSB, PPP1CA, SAT1, ANAPC1, ZNF710 and DPM1 (Table 2). The heat map of the top 100 DEmRNAs is shown in Fig. 2A. 125 DEmiRNAs were identified. Among them, 92 were up-regulated and 33 were down-regulated DEmiRNAs. The top 10 up-regulated DEmiRNAs were hsa-let-7i, hsa-miR-1224-5p, hsa-let-7d, hsa-let-7b, hsa-miR-760, hsa-let-7a, hsa-miR-501-5p, hsa-miR-671-

ID	Symbol	Combined ES	P-value	False discovery rate (FDR)	Up/down-regulation
126353	C19orf21	2.78642	1.64E-09	2.71E-07	Up
10139	ARFRP1	2.744126	1.22E-08	1.80E-06	Up
5296	PIK3R2	2.143624	2.87E-08	3.98E-06	Up
55760	DHX32	2.309248	2.88E-08	3.98E-06	Up
11240	PADI2	2.653745	6.76E-08	8.94E-06	Up
10439	OLFM1	2.314624	1.02E-07	1.30E-05	Up
7849	PAX8	2.699716	1.04E-07	1.30E-05	Up
6752	SSTR2	2.215677	1.16E-07	1.42E-05	Up
9392	TGFBRAP1	2.114437	1.26E-07	1.53E-05	Up
23644	EDC4	1.956832	1.44E-07	1.71E-05	Up
10075	HUWE1	-13.6214	0	0	Down
2923	PDIA3	-5.82461	2.26E-14	4.47E-12	Down
994	CDC25B	-2.68281	3.48E-10	5.89E-08	Down
998	CDC42	-3.33109	8.08E-09	1.23E-06	Down
1508	CTSB	-3.75885	9.98E-09	1.50E-06	Down
5499	PPP1CA	-5.44517	1.23E-08	1.80E-06	Down
6303	SAT1	-7.6069	1.74E-08	2.49E-06	Down
64682	ANAPC1	-2.14652	2.83E-08	3.98E-06	Down
374655	ZNF710	-4.4165	4.87E-08	6.60E-06	Down
8813	DPM1	-2.58688	4.89E-08	6.60E-06	Down



Fig. 2. Heat map of DEmRNAs and DEmiRNAs. A) Heat map of top 50 DEmRNAs; B) heat map of DEmiRNAs. Rows and columns represent genes and samples, respectively. Orange indicates above the reference channel (high expression genes). Purple indicates below the reference channel (low expression genes)

Table 3.	Top 10	up/down-regulated DEmiRNAs
----------	--------	----------------------------

miRNA_ID	Log fold change	P-value	adj. P-value	Up/down-regulation
hsa-let-7i	0.686845	0.000176	0.034484	Up
hsa-miR-1224-5p	0.82917	0.000209	0.034484	Up
hsa-let-7d	0.635086	0.000229	0.034484	Up
hsa-let-7b	0.590625	0.000337	0.03466	Up
hsa-miR-760	0.65443	0.000344	0.03466	Up
hsa-let-7a	0.695632	0.000569	0.050946	Up
hsa-miR-501-5p	0.749797	0.000748	0.058856	Up
hsa-miR-671-5p	0.736763	0.00083	0.058856	Up
kshv-miR-K12-7	0.663458	0.000881	0.058856	Up
hsa-let-7c	0.560578	0.000974	0.058856	Up
hsa-miR-30c	-1.57904	3.73E-06	0.003004	Down
hsa-miR-30a	-1.11649	8.18E-05	0.032951	Down
hsa-miR-222	-0.69527	0.000257	0.034484	Down
hsa-miR-26a	-0.53664	0.001541	0.07306	Down
hsa-miR-32	-0.80569	0.002712	0.087446	Down
hsa-miR-484	-0.55433	0.003229	0.096401	Down
hsa-miR-212	-0.63994	0.003354	0.096547	Down
hsa-miR-17*	-0.65484	0.004586	0.105904	Down
hsa-miR-140-3p	-0.45109	0.004754	0.105904	Down
hsa-miR-181c	-0.56521	0.006475	0.127292	Down



Fig. 3. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analysis of DEmRNA. **A**) GO functional enrichment of DEmRNA. Circle, triangle and square represent biological process (BP), cellular component (CC) and molecular function (MF), respectively; **B**) KEGG functional enrichment of DEmRNA

5p, kshv-miR-K12-7 and hsa-let-7c (Table 3). The top 10 down-regulated DEmiRNAs were hsa-miR-30c, hsa-miR-30a, hsa-miR-222, hsa-miR-26a, hsa-miR-32, hsa-

miR-484, hsa-miR-212, hsa-miR-17*, hsa-miR-140-3p and hsa-miR-181c (Table 3). The heat map of the top 50 DEmiRNAs is shown in Fig. 2B.



Functional analysis

Gene ontology and KEGG functional enrichment analyses were performed using David 6.8 (Fig. 3). GO enrichment analysis includes biological process (BP) terms, cell composition (CC) terms and molecular function (MF) terms. In BP terms, DEmRNAs were mainly involved in regulation of transcription from RNA polymerase II promoter and endosomal vesicle fusion. In CC terms, DEmRNAs were mainly distributed in the nucleoplasm and cytosol. In MF terms, DEmRNAs were mainly associated with protein binding and poly (A) RNA binding. According to KEGG analysis, Epstein-Barr virus infection, endocytosis, and spliceosome were significantly enriched signaling pathways. In addition, GSEA and GSVA analysis was performed on the expression matrix of all genes. In GSEA analysis, TNF- α signaling via NF- κ B (p = 0.032), apoptosis (p = 0.002) and MTORC-1 signaling (p = 0.025) were inhibited in IgAN (Fig. 4A-C). Moreover, TNF-α signaling via NF-κB, apoptosis and MTORC-1 signaling were also significantly different in GSVA analysis (p < 0.05) (Fig. 4D).

Identification of optimal diagnostic mRNA biomarkers

After feature selection of 1205 DEmRNAs, 17 important DEmRNAs were obtained (Table 4). The importance of these DEmRNAs was ranked from high to low according to the mean decrease in accuracy value (Fig. 5A). Subsequently, one mRNA was added from top to bottom according to the order of sorting results. The random forest algorithm was used to classify, and the 10-fold cross-validation process was used to calculate accuracy. The results showed that the accuracy reached the maximum value for the first time when the mRNA number reached 8 (Fig. 5B). Therefore, the first 8 DEmRNAs (AKAP8L, ANAPC1, TGFBRAP1, NFKBIA, CD1C, MRPS15, CDC25B and DHX32) were selected as the optimal diagnostic mRNA

ID	Symbol	Combined ES	<i>P</i> -value	False discovery rate (FDR)	Up/down-regulation
994	CDC25B	-2.68281	3.48E-10	5.89E-08	Down
1508	CTSB	-3.75885	9.98E-09	1.50E-06	Down
64682	ANAPC1	-2.14652	2.83E-08	3.98E-06	Down
5296	PIK3R2	2.143624	2.87E-08	3.98E-06	Up
55760	DHX32	2.309248	2.88E-08	3.98E-06	Up
9392	TGFBRAP1	2.114437	1.26E-07	1.53E-05	Up
2271	FH	2.395432	1.51E-07	1.78E-05	Up
55168	MRPS18A	2.363443	1.97E-07	2.23E-05	Up
26993	AKAP8L	1.911244	2.81E-07	3.07E-05	Up
64960	MRPS15	2.039854	7.89E-07	7.65E-05	Up
51663	ZFR	1.789206	2.77E-06	0.000205	Up
911	CD1C	1.822641	2.83E-06	0.000209	Up
3344	FOXN2	1.981824	1.03E-05	0.000536	Up
85359	DGCR6L	1.56635	2.77E-05	0.001046	Up
4792	NFKBIA	-1.36759	0.000129	0.002845	Down
8675	STX16	-1.52484	0.000164	0.003331	Down
9204	ZMYM6	-1.30245	0.000253	0.004524	Down

Table 4. 17 important DEmRNAs after feature selection

biomarkers. Heat maps of the 8 mRNA biomarkers are shown in Fig. 5C. Random forest, SVM and decision tree classification models were constructed using 8 mRNA biomarkers. The accuracy of the random forest model is up to 0.9259 (Table 5). In addition, the area under the curve (AUC) in the ROC curve of decision tree, random forest and SVM were 0.844, 0.975, and 0.992, respectively (Fig. 6A-C). In addition, ROC analysis was performed for 8 mRNA biomarkers to evaluate their diagnostic value. The AUC values of 8 mRNA biomarkers were all greater than 0.8 (Fig. 6D-K). The results showed that these 8 mRNAs may be considered as potential diagnostic gene biomarkers in IgAN.

Immune correlation analysis

XCell scores were collated into the immune cell infiltrating matrix, and the distribution of 8 immune cell types was shown to be significantly different between normal controls and IgAN by difference analysis (Fig. 7A). Pearson correlation coefficient analysis was employed to calculate the correlation between immune cell types and 8 mRNA biomarkers (Fig. 7B). The results demonstrated that AKAP8L was significantly negatively correlated with CD4+ memory T-cells. Then, Pearson correlation coefficient analysis was employed to calculate the correlation between 8 mRNA biomarkers and TNF-α signaling via NF-κB, apoptosis, MTORC-1 signaling (Fig. 7C). The results showed that AKAP8L was significantly negatively correlated with TNF-α signaling via NF-κB, apoptosis, and MTORC-1 signaling. In addition, the correlations between immune cell types and TNF- α signaling via NF- κ B, apoptosis, and MTORC-1 signaling were revealed using Pearson correlation coefficient analysis (Fig. 7D). The results showed that CD4⁺ memory T-cells were significantly positively correlated with TNF- α signaling via NF- κ B, apoptosis, and MTORC-1 signaling.

DEmiRNA-DEmRNA network

MiRWalk was used to predict DEmiRNA with targeted regulation of 8 mRNA biomarkers. After screening, 17 negative regulatory targeting pairs were obtained (Fig. 8A). Only 5 mRNA biomarkers predicted corresponding negative regulatory miRNAs (Fig. 8A).

In vitro validation

Based on the GEO integration analysis results, AKAP8L, TGFBRAP1, NFKBIA, CD1C, DHX32, hsamiR-922 and hsa-miR-509-5p were selected for real time-PCR validation (Fig. 8B). All primers are shown in Table 6. CD1C, TGFBRAP1, DHX32 and AKAP8L were up-regulated, and NFKBIA and hsa-miR-922 were up-regulated in IgAN. Among them, CD1C, TGFBRAP1 and AKAP8L showed a significant difference. In addition, expression trends of these genes were consistent with the results of bioinformatics analysis. However, we found that the expression trend of hsa-miR-509-5p was contrary to the results of bioinformatics analysis. The differences between bioinformatics and real time-PCR results may be due to the small sample size, heterogeneity between pa-





Classifier	Accuracy	Sensitivity	Specificity	Area under curve (AUC)
Decision tree mode	0.8519	0.8857	0.7895	0.8436
Random forest model	0.9259	0.9714	0.8421	0.9752
Support vector machine (SVM) model	0.963	0.9714	0.9474	0.9925

Table 5. 10-fold cross-validation results of each model

tients, and RNA degradation during extraction or preservation. Further research is needed.

Discussion

Immunoglobulin A nephropathy is the most common glomerular disease worldwide, with a poor prognosis [28]. In this study, 8 important mRNA biomarkers (AKAP8L, ANAPC1, TGFBRAP1, NFKBIA, CD1C, MRPS15, CDC-25B and DHX32) were screened based on machine learning. Then, the correlation between 8 mRNA biomarkers, immune cell types and signaling pathways were analyzed using Pearson's correlation coefficient. Subsequently, Pearson correlation analysis demonstrated that AKAP8L was significantly negatively correlated with CD4⁺ memory T-cells, TNF- α signaling via NF- κ B, apoptosis, and MTORC-1 signaling. In addition, CD4⁺ memory T-cells were significantly correlated with TNF- α signaling via NF- κ B, apoptosis, and MTORC-1 signaling. Finally, a DEmRNA-DEmiRNA network was constructed.

So far, we have not found any correlation study of A-kinase anchoring protein 8 like (AKAP8L) in nephropathy. Knockdown of AKAP8L suppressed the commitment of human hematopoietic stem cells to the erythroid lineage and cell proliferation and delayed differentiation of colony-forming unit-erythroid to the proerythroblast stage [29]. AKAP8L deletion significantly reduced human embryonic kidney 293A (HEK293A) cell size and proliferation [30]. It can also be involved in the regulation of tumor growth and the progression of diabetic liver disease [31, 32]. Moreover, high expression of AKAP8L in esophageal squamous cell carcinoma is associated with poor prognosis [33]. In this study, AKAP8L was found to be an important mRNA biomarker and was significantly negatively correlated with CD4+ memory T-cells. CD4+ memory T-cells can promote the effector response of other immune cells in the immune system and play an important regulatory role in a variety of diseases [34, 35]. Dysregulation of CD4+ memory T-cells played a role in the immune defects of chronic kidney disease patients [36]. In addition, it has been reported that the number of CD4+ memory T-cells in IgAN patients is abnormal. The IgAN patients' peripheral blood has a special immune pattern [14]. Therefore, it is speculated that the up-regulation of AKAP8L may mediate down-regulation of the CD4+ memory T-cell infiltration level, thus affecting the immune regulatory mechanism of IgAN. In addition, AKAP8L was also significantly negatively correlated with TNF-α signaling via NF-κB, apopto-

sis, and MTORC-1 signaling. TNF-α signaling via NF-κB induces epithelial-mesenchymal transformation of renal cell carcinoma cells [37]. In diabetic nephropathy, podocyte injury is associated with abnormality of the NF-KB/ TNF- α signaling pathway in macrophages [38]. Cytotoxin-associated antigen A (CagA) promotes proliferation and secretion of extracellular matrix by inhibiting the apoptosis signaling pathway of rat mesangial cells, thus inducing injury of mesangial cells [39]. MTORC-1 signaling plays a vital regulatory role in normal kidney fibrogenesis [40]. mTOR is an important mediator of CD4+ T-cells. Moreover, inhibition of mTORC1 in myeloid dendritic cells increased proinflammatory factors [41]. Therefore, we speculated that AKAP8L may play a vital mediating role in disease progression by mediating TNF- α signaling via NF-kB, apoptosis and MTORC-1 signaling. In addition, identification of AKAP8L and related signaling pathways and immune cells will contribute to the study of IgNA pathogenesis and immunotherapy.

Although we have not found relevant studies on anaphase promoting complex subunit 1 (ANAPC1) in nephropathy, previous studies have found that ANAPC1 plays an important regulatory role in colitis-associated colon cancer [42], gastric cancer [43], Rothmund-Thomson syndrome [44] and other diseases. In this study, we found that the AUC value of ANAPC1 was 0.880. Therefore, we speculated that ANAPC1 act as a potential diagnostic biomarker for IgAN. Hsa-miR-509-5p was found to negatively regulate ANAPC1 in the DEmRNA-DEmiRNA network. In renal cell carcinoma, the abnormal expression of hsa-miR-509-5p affects cell proliferation, migration and apoptosis [45]. Hsa-miR-509-5p can also influence inflammatory responses by targeting relevant genes [46]. Therefore, we hypothesized that the role of ANAPC1 in IgAN was regulated by hsa-miR-509-5p. In addition, hsamiR-509-5p negatively regulates the cell division cycle 25B (CDC25B) in the DEmRNA-DEmiRNA network. Abnormal expression of CDC25B affects the proliferation and migration of renal cell carcinoma cells [47]. Up-regulation of CDC25B can protect lens epithelial cells from oxidative stress [48]. Moreover, in this study, the AUC value of CDC25B was 0.832. This indicates that CDC25B plays a vital role in IgAN.

We found that hsa-miR-922 negatively regulates transforming growth factor β receptor associated protein 1 (TGFBRAP1) in the DEmRNA-DEmiRNA network. So far, we have not found a correlation study between hsa-miR-922 and TGFBRAP1 in nephropathy. However, TGFBRAP1







Fig. 7. Immune correlation analysis. **A**) The distribution of 8 immune cell types was significantly different between normal controls and IgAN according to difference analysis; **B**) Pearson correlation coefficient analysis of the correlation between immune cell types and 8 mRNA biomarkers; **C**) Pearson correlation coefficient analysis of the correlation between 8 mRNA biomarkers and TNF-α signaling via NF-κB, apoptosis, and MTORC-1 signaling; **D**) Pearson correlation coefficient analysis of the correlation between immune cell types and TNF-α signaling via NF-κB, apoptosis, and MTORC-1 signaling. Blue and red represent positive and negative correlations, respectively. The larger the absolute value of the numerical value, the stronger the correlation. *p < 0.05, **p < 0.01, ***p < 0.001

was found to be associated with diabetes and liver damage [49, 50]. Hsa-miR-922 also was found to be associated with liver damage [51]. Therefore, we hypothesized that TGF-BRAP1 may play a regulatory role in IgAN progression, and this regulation may be regulated by hsa-miR-922.

MiR-381-3p affects the cellular immune response by targeting the CD1c molecule (CD1C) to regulate the antigen presentation capacity of dendritic cells [52]. Hsa-miR-140-3p can affect the growth of mesangial cells by regulating dachshund family transcription factor 1 (DACH1)



and then participate in the regulation of IgAN development [53]. Hsa-miR-140-3p expression is down-regulated in chronic kidney disease and participates in the regulation of multiple biological pathways through targeted mRNAs [54]. In this study, hsa-miR-140-3p was found to negatively regulate CD1C in the DEmRNA-DEmiRNA network. Therefore, we speculated that CD1C may play a vital role in IgAN progression through the targeted regulation of hsa-miR-140-3p.

As a mitochondrial ribosomal protein gene, mitochondrial ribosomal protein S15 (MRPS15) plays a vital role in early embryonic development and cancer [55, 56]. HsamiR-296-5p is involved in the regulation of cervical cancer [57], liver cancer [58] and thoracic aortic aneurysm [59]. In addition, hsa-miR-296-5p can participate in host antiviral defense and play an important role in virus infection [60, 61]. So far, we have not found a correlation study between hsa-miR-296-5p and MRPS15 in nephropathy. This article may present the first discovery that hsa-miR-296-5p and MRPS15 are abnormally expressed in IgAN. Moreover, in this study, hsa-miR-296-5p negatively regulated MRPS15 in the DEmRNA-DEmiRNA network. The discoveries of hsa-miR-296-5p and MRPS15 provide new ideas for future studies of IgAN research.

In this study, we found that NFKB inhibitor alpha (NFKBIA) was abnormally expressed in IgAN. NFKBIA is an important inflammatory response gene, and its polymorphism can be used as a biomarker to predict the risk of acute kidney injury in children [62]. Dominant-negative NFKBIA mutation leads to hepatic disease with severe immune deficiency [63]. DEAH-box helicase 32 (DHX32) can induce vascular endothelial growth factor A (VEGFA) expression by enhancing the β -catenin signaling pathway to promote angiogenesis [64]. Previous studies have found that DHX32 may be involved in the differentiation of normal lymphocytes [65]. DHX32 may also be involved in regulating the response of T cells to certain apoptotic stimuli [66]. In this study, DHX32 was also significantly correlated with CD4+ memory T-cells in IgAN. In addition, the AUC values of NFKBIA and DHX32 were both greater than 0.8. Therefore, we speculated that NFKBIA and DHX32 play an important regulatory role in IgAN. Identification of NFKBIA and DHX32 provides potential directions for further research on the molecular mechanism of IgAN.

Although the above findings have opened up new directions for the molecular research of IgAN, this research still has certain limitations. Firstly, there is a certain degree of error between the real time-PCR results and the bioinformatics analysis results, so it is necessary to collect samples for further research. Secondly, the specific molecular mechanism of the genes identified in IgAN is still unclear.

Conclusions

In this study, GSEA analysis revealed that TNF- α signaling via NF-kB, apoptosis and MTORC-1 signaling were inhibited in IgAN. Then, 8 important mRNA biomarkers (AKAP8L, ANAPC1, TGFBRAP1, NFKBIA, CD1C, MRPS15, CDC25B and DHX32) were screened based on machine learning. Subsequently, the distribution of 8 immune cell types was shown to be significantly different between normal controls and IgAN by difference analysis. Pearson correlation analysis demonstrated that AKA-P8L was significantly negatively correlated with CD4+ memory T-cells, TNF- α signaling via NF- κ B, apoptosis, and MTORC-1 signaling. Finally, 5 mRNA biomarkers (ANAPC1, TGFBRAP1, CD1C, MRPS15 and CDC25B) predicted corresponding negative regulatory miRNAs. Therefore, we hypothesized that AKAP8L, ANAPC1, TGFBRAP1, CD1C, MRPS15 and CDC25B play important roles in the progression of IgAN, and the specific molecular mechanism deserves further study. These results provide new ideas for understanding IgAN and potential directions for further research on the molecular mechanism of IgAN.

Ethical statement

All participants were informed as to the purpose of this study, and that this study complied with the Declaration of Helsinki. This study was approved by the ethics committee of Shijiazhuang People's Hospital (2019045).

Table 6. Primer sequences in RT-PCR

Primer name	Primer sequence (5' to 3')				
GAPDH-F (Internal reference)	5-CTGGGCTACACTGAGCACC-3				
GAPDH-R (Internal reference)	5-AAGTGGTCGTTGAGGGCAATG-3				
ACTB-F (Internal reference)	5-GATCAAGATCATTGCTCCTCCT-3				
ACTB-R (Internal reference)	5-TACTCCTGCTTGCTGATCCA-3				
CD1C-F	5-CCAGAATACAACATGGGTGCC-3				
CD1C-R	5-TCTGTGACGCCTTCATACTGA-3				
TGFBRAP1-F	5-TGCACTGGTCGGAGAATGTG-3				
TGFBRAP1-R	5-GGCAGCGTCTGCTTCTGTT-3				
DHX32-F	5-GTGGTAAGAGCGCTCAGGTT-3				
DHX32-R	5-CGTAGCCAACCTCATGACCA-3				
AKAP8L-F	5-CAGTGCCGATTCCGTTTTATCC-3				
AKAP8L-R	5-CCTCCTTGCATCATGTCTGTCT-3				
NFKBIA-F	5-ACCTGGTGTCACTCCTGTTGA-3				
NFKBIA-R	5-CTGCTGCTGTATCCGGGTG-3				
hsa-U6 (Internal reference)					
hsa-miR-922	5-GCAGCAGAGAATAGGACTACGTC-3				
hsa-miR-509-5p	5-TACTGCAGACAGTGGCAATCA-3				

Data availability statement

All data generated or analyzed during this study are included in this published article. The data sets (GSE73953, GSE58539, GSE14795 and GSE25590) analyzed during the current study are available in the Gene Expression Omnibus (GEO) database; the accessible web link to the database is https://www.ncbi.nlm.nih.gov/geo/.

The authors declare no conflict of interest.

References

- Goodman S, Reid-Adam J (2019): Immunoglobulin A nephropathy. Pediatr Rev 40: 439-441.
- Lafayette RA (2014): Immunoglobulin A nephropathy: insights and progress. Transl Res 163: 3-7.
- Lafayette RA, Kelepouris E (2018): Immunoglobulin A nephropathy: advances in understanding of pathogenesis and treatment. Am J Nephrol 47 Suppl 1: 43-52.
- Alghamdi SA, Saadah OI, Almatury N, Al-Maghrabi J (2012): Hepatic-associated immunoglobulin-A nephropathy in a child with liver cirrhosis and portal hypertension. Saudi J Gastroenterol 18: 214-216.

- Ryu J, Ryu H, Kim S (2019): Comparison of cancer prevalence between patients with glomerulonephritis and the general population at the time of kidney biopsy. PLoS One 14: e0224024.
- Wyatt RJ, Julian BA (2013): IgA nephropathy. N Engl J Med 368: 2402-2414.
- Suzuki K, Honda K, Tanabe K, et al. (2003): Incidence of latent mesangial IgA deposition in renal allograft donors in Japan. Kidney Int 63: 2286-2294.
- Rajasekaran A, Julian BA, Rizk DV (2021): IgA nephropathy: an interesting autoimmune kidney disease. Am J Med Sci 361: 176-194.
- Yeo SC, Goh SM, Barratt J (2019): Is immunoglobulin A nephropathy different in different ethnic populations? Nephrology (Carlton, Vic) 24: 885-895.
- Zheng Y, Lu P, Deng Y, et al. (2020): Single-cell transcriptomics reveal immune mechanisms of the onset and progression of IgA nephropathy. Cell Rep 33: 108525.
- Campos A, Rivera F, Egido J, Parera M (1992): T-cell alterations in immunoglobulin A nephropathy. Int J Clin Pharmacol Res 12: 191-195.
- 12. Bai L, Li H, Li J, et al. (2019): Immunosuppressive effect of artemisinin and hydroxychloroquine combination therapy on IgA nephropathy via regulating the differentiation of CD4+ T cell subsets in rats. Int Immunopharmacol 70: 313-323.
- Takechi H, Oda T, Hotta O, et al. (2013): Clinical and immunological implications of increase in CD208+ dendritic cells in tonsils of patients with immunoglobulin A nephropathy. Nephrol Dial Transplant 28: 3004-3013.
- Esteve Cols C, Graterol Torres FA, Quirant Sánchez B, et al. (2020): Immunological pattern in IgA nephropathy. Int J Mol Sci 21: 1389.
- 15. Xing Y, Li L, Zhang Y, et al. (2020): C1GALT1 expression is associated with galactosylation of IgA1 in peripheral B lymphocyte in immunoglobulin a nephropathy. BMC Nephrol 21: 18.
- Coppo R, Camilla R, Amore A, et al. (2010): Toll-like receptor 4 expression is increased in circulating mononuclear cells of patients with immunoglobulin A nephropathy. Clin Exp Immunol 159: 73-81.
- Jin LW, Ye HY, Xu XY, et al. (2018): MiR-133a/133b inhibits Treg differentiation in IgA nephropathy through targeting FOXP3. Biomed Pharmacother 101: 195-200.
- Liu C, Li X, Shuai L, et al. (2021): Astragaloside IV inhibits galactose-deficient IgA1 secretion via miR-98-5p in pediatric IgA nephropathy. Front Pharmacol 12: 658236.
- Chen X, Sun M (2020): Identification of key genes, pathways and potential therapeutic agents for IgA nephropathy using an integrated bioinformatics analysis. J Renin Angiotensin Aldosterone Syst 21: 1470320320919635.
- 20. Lin TJ, Yang SS, Hua KF, et al. (2016): SPAK plays a pathogenic role in IgA nephropathy through the activation of NFκB/MAPKs signaling pathway. Free Radic Biol Med 99: 214-224.
- Wang M, Huang J, Liu Y, et al. (2017): COMBAT: a combined association test for genes using summary statistics. Genetics 207: 883-891.
- Marot G, Foulley JL, Mayer CD, Jaffrézic F (2009): Moderated effect size and P-value combinations for microarray meta-analyses. Bioinformatics (Oxford, England) 25: 2692-2699.
- 23. Ritchie ME, Phipson B, Wu D, et al. (2015): limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43: e47.

- Hänzelmann S, Castelo R, Guinney J (2013): GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 14: 7.
- Engebretsen S, Bohlin J (2019): Statistical predictions with glmnet. Clin Epigenetics 11: 123.
- 26. Kuhn M (2008): Building predictive models in R using the caret package. J Statist Software 28: 1-26.
- Livak KJ, Schmittgen TD (2001): Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods (San Diego, Calif) 25: 402-408.
- Suzuki H (2019): Biomarkers for IgA nephropathy on the basis of multi-hit pathogenesis. Clin Exp Nephrol 23: 26-31.
- 29. Huang P, Zhao Y, Zhong J, et al. (2020): Putative regulators for the continuum of erythroid differentiation revealed by single-cell transcriptome of human BM and UCB cells. Proc Natl Acad Sci U S A 117: 12868-12876.
- Melick CH, Meng D (2020): A-kinase anchoring protein 8L interacts with mTORC1 and promotes cell growth. J Biol Chem 295: 8096-8105.
- Gao JL, Lv GY, He BC, et al. (2013): Ginseng saponin metabolite 20(S)-protopanaxadiol inhibits tumor growth by targeting multiple cancer signaling pathways. Oncol Rep 30: 292-298.
- 32. Chen JY, Chou HC, Chen YH, Chan HL (2013): High glucose-induced proteome alterations in hepatocytes and its possible relevance to diabetic liver disease. J Nutr Biochem 24: 1889-1910.
- Luo QY, Di T, Qiu MZ, et al. (2022): High AKAP8L expression predicts poor prognosis in esophageal squamous cell carcinoma. Cancer Cell Int 22: 90.
- 34. Stockinger B, Bourgeois C, Kassiotis G (2006): CD4+ memory T cells: functional differentiation and homeostasis. Immunol Rev 211: 39-48.
- Schreiner D, King CG (2018): CD4+ memory T cells at home in the tissue: mechanisms for health and disease. Front Immunol 9: 2394.
- 36. Xu G, Gong Y (2013): Deregulation from CD4+ memory T cells to regulatory cells in patients with chronic renal failure: a pilot study. J Clin Lab Anal 27: 423-426.
- 37. Wu ST, Sun GH, Hsu CY, et al. (2011): Tumor necrosis factor-α induces epithelial-mesenchymal transition of renal cell carcinoma cells via a nuclear factor kappa B-independent mechanism. Exp Biol Med (Maywood) 236: 1022-1029.
- Yang H, Xie T, Li D, et al. (2019): Tim-3 aggravates podocyte injury in diabetic nephropathy by promoting macrophage activation via the NF-κB/TNF-α pathway. Mol Metab 23: 24-36.
- 39. Wang L, Tan RZ, Chen Y, et al. (2016): CagA promotes proliferation and secretion of extracellular matrix by inhibiting signaling pathway of apoptosis in rat glomerular mesangial cells. Ren Fail 38: 458-464.
- 40. Rozen-Zvi B, Hayashida T, Hubchak SC, et al. (2013): TGF-β/Smad3 activates mammalian target of rapamycin complex-1 to promote collagen production by increasing HIF-1α expression. Am J Physiol Renal Physiol 305: F485-494.
- Haidinger M, Poglitsch M, Geyeregger R, et al. (2010): A versatile role of mammalian target of rapamycin in human dendritic cell function and differentiation. J Immunol 185: 3919-3931.
- 42. Sharp SP, Malizia RA, Walrath T, et al. (2018): DNA damage response genes mark the early transition from colitis to neoplasia in colitis-associated colon cancer. Gene 677: 299-307.

- 43. Lima EM, Leal MF, Burbano RR, et al. (2008): Methylation status of ANAPC1, CDKN2A and TP53 promoter genes in individuals with gastric cancer. Braz J Med Biol Res 41: 539-543.
- 44. Zhang Y, Qin W, Wang H, et al. (2021): Novel pathogenic variants in the RECQL4 gene causing Rothmund-Thomson syndrome in three Chinese patients. J Dermatol 48: 1511-1517.
- 45. Zhang WB, Pan ZQ, Yang QS, Zheng XM (2013): Tumor suppressive miR-509-5p contributes to cell migration, proliferation and antiapoptosis in renal cell carcinoma. Ir J Med Sci 182: 621-627.
- 46. Zhu M, Cao S, Zheng W, et al. (2021): miR-509-5p anti-infection response for mycoplasma pneumonia in sheep by targeting NF-κB pathway. Vet Immunol Immunopathol 238: 110275.
- 47. Yu XY, Zhang Z, Zhang GJ, et al. (2012): Knockdown of Cdc25B in renal cell carcinoma is associated with decreased malignant features. Asian Pac J Cancer Prev 13: 931-935.
- 48. Xu J, Li D, Lu Y, Zheng TY (2021): Aβ monomers protect lens epithelial cells against oxidative stress by upregulating CDC25B. Free Radic Biol Med 175: 161-170.
- 49. Yang S, Chen X, Yang M, et al. (2019): The variant at TGF-BRAP1 is significantly associated with type 2 diabetes mellitus and affects diabetes-related miRNA expression. J Cell Mol Med 23: 83-92.
- 50. Zhang J, Zhao Z, Bai H, et al. (2019): The variant at TGF-BRAP1 but not TGFBR2 is associated with antituberculosis drug-induced liver injury. Evid Based Complement Alternat Med 2019: 1685128.
- Liu Y, Chen SH, Jin X, Li YM (2013): Analysis of differentially expressed genes and microRNAs in alcoholic liver disease. Int J Mol Med 31: 547-554.
- 52. Wen Q, Zhou C, Xiong W, Su J (2016): MiR-381-3p regulates the antigen-presenting capability of dendritic cells and represses antituberculosis cellular immune responses by targeting CD1c. J Immunol 197: 580-589.
- 53. Zhou X, Lu Y, Guo P, Zhou C (2021): Upregulation of microRNA-140-3p mediates dachshund family transcription factor 1 expression in immunoglobulin A nephropathy through cell cycle-dependent mechanisms. Mol Med Rep 23: 134.
- Rudnicki M, Perco P, Haene BD, et al. (2016): Renal micro-RNA- and RNA-profiles in progressive chronic kidney discase. Eur J Clin Invest 46: 213-226.
- 55. Cheong A, Lingutla R, Mager J (2020): Expression analysis of mammalian mitochondrial ribosomal protein genes. Gene Expr Patterns 38: 119147.
- 56. Sotgia F, Whitaker-Menezes D, Martinez-Outschoorn UE, et al. (2012): Mitochondria "fuel" breast cancer metabolism: fifteen markers of mitochondrial biogenesis label epithelial cancer cells, but are excluded from adjacent stromal cells. Cell Cycle (Georgetown) 11: 4390-4401.
- 57. Xie J, Chen Q, Zhou P, Fan W (2021): Circular RNA hsa_ circ_0000511 improves epithelial mesenchymal transition of cervical cancer by regulating hsa-mir-296-5p/HMGA1. J Immunol Res 2021: 9964538.
- Luo X, Shen S, Yi S, et al. (2020): Screening of differentially expressed miRNAs in tensile strain-treated HepG2 cells by miRNA microarray analysis. Mol Med Rep 21: 2415-2426.
- 59. Magouliotis DE, Fergadi MP (2021): In-depth bioinformatic study of the cadherin 5 interactome in patients with thoracic aortic aneurysm unveils 8 novel biomarkers. Eur J Cardiothorac Surg 61: 11-18.

- 60. Zheng Z, Ke X, Wang M, et al. (2013): Human microRNA hsa-miR-296-5p suppresses enterovirus 71 replication by targeting the viral genome. J Virol 87: 5645-5656.
- 61. Gao J, Gao L, Li R, et al. (2019): Integrated analysis of microRNA-mRNA expression in A549 cells infected with influenza A viruses (IAVs) from different host species. Virus Res 263: 34-46.
- 62. He J, Xie G, Wu H, et al. (2018): Association between inflammatory-response gene polymorphisms and risk of acute kidney injury in children. Biosci Rep 38: BSR20180537.
- 63. Tan EE, Hopkins RA, Lim CK, et al. (2020): Dominant-negative NFKBIA mutation promotes IL-1β production causing hepatic disease with severe immunodeficiency. J Clin Invest 130: 5817-5832.
- 64. Lin H, Fang Z, Su Y, et al. (2017): DHX32 promotes angiogenesis in colorectal cancer through augmenting β-catenin signaling to induce expression of VEGFA. EBioMedicine 18: 62-72.
- Abdelhaleem M, Sun TH, Ho M (2005): DHX32 expression suggests a role in lymphocyte differentiation. Anticancer Res 25: 2645-2648.
- 66. Alli Z, Chen Y, Abdul Wajid S, et al. (2007): A role for DHX32 in regulating T-cell apoptosis. Anticancer Res 27: 373-377.