Identification of GBP2 and TMSB10 as Immune-associated Genes in Hypertension Nephropathy by Integrated Bioinformatics Analysis and Machine Learning

Xiaolin Liao  
Hunan University of Chinese Medicine

Huaguan Lu  
Hunan University of Chinese Medicine

Jianjun Liu  
Hunan University of Chinese Medicine

Yuhong Wang  (✉ wyh107@hnucm.edu.cn)  
Hunan University of Chinese Medicine

Article

Keywords: Hypertension nephropathy (HN), Immune infiltration, Differentially expressed genes (DEGs), GBP2, TMSB10

Posted Date: April 7th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2733974/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Clinical and experimental proof suggests that hypertension nephropathy (HN) is a chronic inflammatory disease. Our study aims to disclose the role of immune-related genes in the progression of HN. Using the Gene Expression Omnibus (GEO) database, two human HN gene expression datasets (GSE37455 and GSE37460; n = 35) along with the relevant controls (n = 43) could be as the discovery metadata to analyze for differentially expressed genes (DEGs) in HN. Three different machine-learning algorithms were integrated to screen immune-related genes in HN. Receiver-operating characteristic (ROC) curves were generated to estimate diagnostic efficacy. The diagnostic value and expression levels of these candidate genes were validated in the GSE104954 dataset (20 HN patients and 3 controls). Single sample gene set enrichment analysis (ssGSEA) was used to evaluate immune cell infiltrations, and immune checkpoints were quantified. The expression of the potential genes was confirmed in vivo. As a result, a total of 220 DEGs were identified between HN and control samples in these datasets, of which 52 were immune differential genes. The magenta module in WGCNA was the highest correlation. Two immune-associated genes GBP2 (guanylate binding protein 2) and TMSB10 (Thymosin β10) for HN were obtained after the intersection of genes screened by machine learning. The expression levels of GBP2 and TMSB10 were validated using discovery and validation cohort data sets. Following the ssGSEA analysis, we identified potential immune cell types in HN patients, as well as revealed the correlation between immune-related genes (GBP2 and TMSB10) and immune cells. Furthermore, the mRNA and protein levels of GBP2 and TMSB10 in vivo were consistent with the bioinformatics analysis which confirms the accuracy of our analysis. Our results demonstrated that GBP2 and TMSB10 are promising immune-related genes for the diagnosis of HN, which may help in the development of more precisely tailored HN immunotherapy.

1. Introduction

Hypertension is a disease that usually impairs target organs, especially kidneys. Hypertensive nephropathy (HN) refers to damage to the kidneys caused by chronic high blood pressure. The problem of HN has become an important issue in public health due to its significant morbidity and mortality. The incidence of HN has been rising recently, and 27.5% of dialysis patients are hypertensive before starting treatment in America. According to the China Kidney Disease Network (CK-NET2016), chronic kidney disease (CKD) in China accounts for 20.8% of patients with HN, which is the second cause after diabetic nephropathy. However, it has been found that lowering blood pressure to the recommended level only slows, not stops, the progression of HN. It is therefore crucial to understand the pathogenesis of HN in addition to reducing blood pressure.

Microarray technology and integrated bioinformatics analysis have lately led to the discovery of novel genes associated with some diseases that may be used as prognostic and diagnostic biomarkers. In the realm of HN, for example, several DEGs, including AGTR1, CYP4A11, and CYP3A4 are reported to be involved in the development of HN. In addition, three ferroptosis-related differentially expressed genes, including ALB, NNMT, and ATF3, may be useful for the prevention and control of HN disease. Immune-associated genes are gradually recognized to play a significant role in the progression of various diseases. Therefore, it is of great significance to evaluate the immune-related genes in HN and find the relative changes of immune cell types as potential markers for further elucidating the molecular mechanism of HN and developing new
immunotherapy targets. However, as far as we know, there has been limited research on immune-associated diagnostic candidates for HN.

This study examined immune-associated genes and identified relevant biomarkers in HN patients. Firstly, the Gene Expression Omnibus (GEO) database was utilized to download three HN datasets. A limma package was adopted to identify DEGs, and the important module genes were selected. Then, based on machine learning algorithms and validation in a separate cohort, possible diagnostic HN biomarkers were identified. A ssGSEA analysis of immune cell gene expression in HN and control samples was successful in identifying and quantifying a wide range of immune infiltrated cells. In addition, the relationship between identified HN genes and immune cell infiltration was also characterized. Lastly, RT-qPCR and western blot were used to further verify the differential expression of the most critical genes.

2 Results

2.1 Differential expression analysis

Following batch effects removal with R package “SVA”, two expression array data sets from the GEO database (GSE37460 and GSE37455) were merged into a discovery metadata dataset. 220 differentially expressed genes were identified, of which 114 were up-regulated while 106 were downregulated (Fig. 1A). Figure 1B shows the expression of the first 20 up-regulated DEGs as well as the first 20 down-regulated DEGs. The intersected DEGs were enriched by GO and KEGG pathways to investigate their potential biological function. Supplementary Tables S1 and S2 showed 454 BP pathways, 67 CC pathways, 40 MF pathways, and 47 KEGG pathways significantly enriched. Figure 1C shows that these DEGs were mainly enriched in antigens via MHC class II processes related to BPs. In addition, tuberculosis, phagosome and cytokine-cytokine receptor interaction showed to have a close relationship with HN by GO analysis (Fig. 1D).

2.2 Building the WGCNA Network and Selecting Modules Related to Traits

To construct a scale-free network, we chose a soft threshold $\beta = 5$ (Fig. 2A). As shown in Fig. 2B, based on merged dynamic tree cuts, different colors were used to identify gene co-expression modules. Biological significance is often found in modules closely related to clinical features. A heat map illustrates the correlation between modules and traits in Fig. 2C. Compared with other modules, the magenta module (1,029 genes) was the most positively correlated with the HN groups [cor = 0.62, p = 1e - 09] and was regarded as the pivotal module for subsequent analysis. A high correlation [cor = 0.64, p = 1.2e-119] between GS and MM in the magenta module suggested that the magenta module was largely responsible for the above result (Fig. 2D). Therefore, the magenta module was considered to be an important module associated with HN disease and was considered pivotal for further analysis.

2.3 Identification of candidate immune-related characteristic genes via Machine Learning
The profile of immune cell infiltration in HN and control samples was explored by ssGSEA. In the ImmPort database, 1793 immune-related genes were obtained. Subsequently, by intersection with 220 differentially expressed genes, 52 immune-related genes were obtained. By overlapping these immune-related genes with the MEmagenta module genes, 16 candidate genes were able to obtain (Fig. 2E). To identify gene signatures, these genes were further screened for diagnostic markers using three bioinformatic methods, including LASSO, RF and SVM-RFE. Firstly, the differentially expressed genes were filtered using the LASSO regression algorithm. As a result, three gene signatures were obtained, including GBP2, PSMB8 and TMSB10 (Figs. 3A, B). In the meantime, we identified 6 gene signatures by RF method when the important score was set greater than 2.0 (Fig. 3C, D). Using the SVM-RFE algorithm, 11 genes were identified as immune-related characteristic genes. It was found that the SVM-RFE algorithm was the most accurate when there were 11 eigengenes genes, up to 0.85 (Fig. 3E). As a result of overlapping genes from LASSO, RF and SVM-RFE, we obtained two robust gene signatures for HN ultimately (Fig. 3F), namely GBP2 and TMSB10.

2.4 Prediction and verification of HN progression using a characteristic gene nomogram

To predict HN progression, a nomogram was constructed by incorporating characteristic genes. Nomograms assign scores to characteristic genes, and scores for all characteristic genes were added together to determine the total score (Fig. 4A). It was used to estimate the risks of HN and the calibration curve demonstrated their accuracy in predicting HN progression (Fig. 4B). As seen in Fig. 4C, the nomogram could provide clinical benefits for patients diagnosed with HN, according to the decision curve analysis. To further check the diagnostic efficacy of GBP2 and TMSB10, we validated them using a metadata (combined GSE37460 and GSE37455) and the GSE10495 dataset, respectively. A ROC curve analysis was used to assess the diagnostic ability of each characteristic gene to predict HN progression in the metadata (Fig. 4D-F) and GSE10495 dataset (Fig. 4G-I). The AUCs of GBP2 and TMSB10 in the metadata cohort were 0.902 and 0.849, respectively. In addition, the AUCs of GBP2 and TMSB10 in the GSE10495 dataset were 0.950 and 0.917, respectively, indicating that both GBP2 and TMSB10 have high diagnostic values.

2.5 The immunological analysis of HN

The expression of infiltration of immune cells and immune checkpoints were evaluated according to immunological features. The occurrence and development of HN is closely related to immune infiltrating cells, according to an increasing body of research\(^{15,16}\). Twenty-eight subpopulations of immune infiltrated cells in HN and control samples were identified and presented in the heatmap (Fig. 5A). Compared with the control group, more innate and adaptive immune cells were infiltrated in the HN group (Fig. 5B). In addition, the expression of immune checkpoint-related genes, such as B2M, CD40, CD86, PTPRC and TNFRSF4 were remarkably up-regulated in the HN (Fig. 5C). Combined with the above results, we considered HN might benefit from immune therapy.

2.6 Correlation analysis between GBP2, TMSB10 and infiltrating immune cells
The correlation of two characteristic genes (GBP2 and TMSB10) with 28 types of immune cell infiltration in HN was established (Fig. 5D). Our result showed that characteristic genes were positively associated with the majority of immune infiltrated cells across HN. GBP2 was significantly positively correlated with Type 1 T-helper cell ($P < 0.001$), T follicular helper cell ($P < 0.001$), regulatory T-cell ($P < 0.001$), natural killer T-cell ($P < 0.001$), natural killer cell ($P < 0.001$), MDSC ($P < 0.001$), mast cell ($P < 0.001$) and activated dendritic cell ($P < 0.001$). TMSB10 was significantly positively correlated with Type 1 T-helper cell ($P < 0.001$), regulatory T-cell ($P < 0.001$), plasmacytoid dendritic cell ($P < 0.001$), natural killer T-cell ($P < 0.001$), natural killer cell ($P < 0.001$), MDSC ($P < 0.001$), mast cell ($P < 0.001$) and activated dendritic cell ($P < 0.001$). As a result, the characteristic genes GBP2 and TMSB10 might modulate these infiltrating immune cells during HN.

### 2.7 In vivo Validation of the Key Genes

Both WKY and SHR rats were conducted in animal experiments for *vivo* validation. In Supplementary Figure S1, the blood pressures of WKY and SHR rats are shown. Kidney tissues were examined with hematoxylin-eosin and Masson staining to determine pathological effects. There were significant glomerular abnormalities in SHR rats, such as mesangial hyperplasia, disordered glomerular clusters, and expansion of mesangial matrix (Fig. 6A). As chronic kidney diseases progress, renal fibrosis often plays a significant role\(^\text{17}\). Our results showed that SHR model rats displayed significant renal fibrosis, as evidenced by Masson staining (Figs. 6B) and the expression of fibrosis maker α-SMA (Figs. 6D). Renal function parameters of rats in the SHR group changed significantly compared with the WKY group (Supplementary Table S3). Based on these results, an animal model of HN was successfully constructed.

RT-qPCR detection showed that the expression levels of GBP2 and TMSB10 in the HN groups were significantly up-regulated compared with the control ($P < 0.05$ or $P < 0.01$) (Fig. 6C). The same results were also verified in the Western blotting experiment (Fig. 6D). Therefore, based on these results we conclude that GBP2 and TMSB10 are overexpressed in HN rats and may be involved in the development of HN.

### 3. Discussion

Increasing prevalence and epidemic levels of HN have been observed worldwide. Hospitalization and mortality rates for HN continue to be high, despite significant progress in medications and interventions. Currently, there are no markers that can be used to diagnose HN and the pathogenesis of the disease is heterogeneous, many patients haven’t achieved satisfactory clinical efficacy. In recent literatures, immune-related genes appear to play an important role in HN occurrence and development\(^\text{18,19}\). Finding some novel and effective diagnostic genes for HN among immune cell components is a hopeful approach to research that may lead to early intervention in HN, improving clinical outcomes. Therefore, we are searching for diagnostic genes for HN and investigating how immune-related genes influence HN progression.

The present study provided an in-depth examination of GEO datasets to identify likely diagnostic biomarkers of HN patients among DEGs and immune-related genes.

We created a metadata cohort by merging two GEO datasets and then identifying DEGs between HN and control samples. As a result, 114 genes were remarkably up-regulated, while 106 genes were down-regulated. Subsequent enrichment analysis indicated that these DEGs were mainly linked to leukocyte-mediated immunity, humoral immune response, MHC class II presentation of exogenous peptide antigens, along with other
biological processes. Our findings were consistent with previous reports that DEGs between the HN and control groups are primarily involved in MHC II antigens processes\textsuperscript{20,21}.

Machine learning algorithms are integrated into screening biomarkers to increase their accuracy. LASSO logistic regression determines variables based on the exploration of the data $\lambda$. SVM-RFE is an efficient function selection technique\textsuperscript{22}. RF algorithm can handle data with unbalanced values and missing values. Using these three machine learning algorithms, feature variables were screened and the best classification model was established. Meanwhile, as a system biology method, WGCNA was used to identify gene clusters or modules to investigate the characteristics of disease-gene cluster relationships\textsuperscript{23}. In our study, by combining WGCNA and machine learning algorithms, GBP2 and TMSB10 were chosen as immune-related genes for HN identification lastly.

Guanylate-binding protein (GBP) is a class of large GTPase proteins that can be induced by interferon (IFN) and can mediate active cellular immunity and participate in a wide range of biological responses against pathogenic microorganisms\textsuperscript{24,25}. As a member of the GTPase family, guanylate binding protein 2 (GBP2) plays an important role in host immunity. Wang found that the expression of GBP2 was up-regulated in the immune class, which was strongly associated with the infiltration of CD8 + T cells. Meanwhile, GBP2 knockout further reduced CD8 + T cell infiltration\textsuperscript{26}. This result was in agreement with our correlation analysis of immune cell infiltrations with GBP2 (Fig. 6D). There have been many reports that GBP2 can be used as a predictive molecule for tumor disease detection and diagnosis\textsuperscript{27--29}. To the best of our knowledge, we first discovered that GBP2 may serve as an immune-related gene for the diagnosis of HN.

Thymosin β10 (TMSB10), a thymosin family member, plays a role in cytoskeleton assembly and formation in cells\textsuperscript{30}. It has been shown to promote macrophage M2 conversion and act as a biomarker for malignant tumors\textsuperscript{31,32}. Some researchers reported that the high expression of TMSB10 in clear cell renal cell carcinoma (ccRCC) may serve as a biomarker for diagnosis and prognosis, as well as a potential therapeutic target\textsuperscript{30,33}. In addition, it was found that TMSB10 was strongly associated with immune cells and immune-related pathways in Duchenne Muscular Dystrophy\textsuperscript{34}. Tang et al found that the expression of TMSB10 was increased in mesangial cells of HN patients compared with control samples\textsuperscript{35}. However, few studies investigating TMSB10 as an immune-related gene associated with HN have been performed. Our results suggest that TMSB10 may be valuable for the early diagnosis and immunotherapy of HN.

Increasing evidence suggests that immune cell infiltration affects HN\textsuperscript{18,36,37}. As a result of the ssGSEA evaluation of 28 immune cell types in HN, significant differences were found between HN and the healthy control groups. Especially, the abundance of activated CD4 + T cells, activated dendritic cells, activated CD8 + T cells, γδT, MDSC, natural killer T cells, and natural killer cells et al were significantly up-regulated ($P<0.001$) in HN tissues, while eosinophil cells showed lower levels of expression ($P<0.001$). These findings indicated that they may play a crucial role in the etiology of HN. It was found that patients with hypertensive nephrosclerosis had an increased number of CD4 + and CD8 + cells in the tubulointerstitial space\textsuperscript{38}. These findings are consistent with our results. In addition, some immune checkpoint-related genes, such as B2M, CD40 and CD86 were significantly increased in HN. As shown in Fig. 7D, our experimental results also confirmed a significant increase in CD86 protein expression in HN rats compared with normal rats, which indicates that our findings are
accurate and further consolidate the key role of immune response in HN. It is therefore possible to develop novel treatments for HN by precisely controlling various immune cells.

However, several limitations need to be taken into account. Firstly, the batch effect cannot be completely eliminated due to the heterogeneity of different annotation platforms and clinical covariates. Secondly, these results may be affected by different pathological stages of HN. Thirdly, there may also be a problem with the number of samples. Although this study was finally conducted in vivo studies, due to the small sample size of the data, there is still a distance from the clinical application of these characteristic genes. Therefore, the biological functions of these two immune-related genes need a further larger prospective verification.

4. Conclusion

In summary, we identified GBP2 and TMSB10 may serve as immune-related biomarkers associated with HN by comprehensive bioinformatics analysis. Activated CD4+ T cells, activated CD8+, activated dendritic cells, γδT cells, MDSC, natural killer T cells, natural killer cells and so on are related to the occurrence of HN. These immune cells may play an important role in the development of HN. Further exploration of the interaction between immune-related genes and immune cell infiltration will help to determine the goal of immunotherapy for HN and improve immune regulation therapy in patients with HN.

5. Materials And Methods

5.1 Data acquisition and variance analysis

In the current study, three raw datasets GSE37455, GSE37460, and GSE104954 including gene expression data for HN patients and controls were downloaded from the GEO (https://www.ncbi.nlm.nih.gov/geo/) database. For further analysis, a metadata cohort combining the GSE37455 dataset and the GSE37460 dataset was created through the ‘rbind’, “normalizeBetweenArrays” and “combat” in the R package SVA, and the GSE104954 dataset was used for verification after normalization. In these datasets, gene expression profiles were collected from patients with HN and healthy controls. Detailed descriptions of clinical features can be found in related literature (Table 1). The difference analysis was adopted by the “limma” R package and the absolute value of |log₂ FC| > 0.585 and adjusted \( p < 0.05 \) were used as screening conditions to obtain DEGs

The pheatmap and “ggplot2” packages were used for heat maps and volcano map drawings. The control and HN samples were analyzed together following hierarchical clustering to identify and eliminate outliers. The research process is shown in Fig. 7.

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Platforms</th>
<th>Organism</th>
<th>HN</th>
<th>Control</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE37460</td>
<td>GPL14663</td>
<td>Homo sapiens</td>
<td>15</td>
<td>22</td>
<td>Berthier et al., 2012;</td>
</tr>
<tr>
<td>GSE37455</td>
<td>GPL11670</td>
<td>Homo sapiens</td>
<td>0</td>
<td>18</td>
<td>Berthier et al., 2012;</td>
</tr>
<tr>
<td></td>
<td>GPL14663</td>
<td>Homo sapiens</td>
<td>20</td>
<td>3</td>
<td>Ju et al., 2015</td>
</tr>
<tr>
<td>GSE104954</td>
<td>GPL24120</td>
<td>Homo sapiens</td>
<td>20</td>
<td>3</td>
<td>Grayson et al., 2018</td>
</tr>
</tbody>
</table>
5.2 Weighted gene co-expression network analysis construction (WGCNA)

The co-expression network was analyzed by the “WGCNA” package\(^{42}\). First, optimal weighted parameters of adjacent functions were obtained by using the pick Soft Threshold function, which served as a soft threshold for ensuing network construction. Then, following the construction of a weighted adjacency matrix, related gene modules were established on the basis of the hierarchical clustering of topological overlap measures. Several module characteristic genes (MEs) were investigated for their possible correlation with the clinical traits, and the expression patterns of genes in each module were summarized. Lastly, the average gene significance (GS) and the module significance (MS) within the module were computed to measure the correlation between the module expression profile related to the sample traits. Using WGCNA analysis, MS modules of importance were identified.

5.3 Acquisition of immune-related genes

Genes associated with immunity were obtained from the ImmPort database (http://www.immport.org/shared/home), and then intersected with DEGs and the most significant modules to obtain key immune-related genes.

5.4 Analysis of DEG in the KEGG and GO enrichment pathways

To further investigate the biological significance of the overlapping immune-related differential genes, BioManager and ClusterProfiler were used to perform GO function and KEGG pathway enrichment analyses on the DEGs\(^{34}\), and the cutoff value was set to \(P < 0.05\).

5.5 Screening immune-related genes via the comprehensive strategy

The LASSO regression analysis was performed using the R package “glmnet” to identify genes associated with immunity\(^{43}\). The SVM-RFE method was used to select characteristic genes and the RF method can help predict continuous variables by reducing the effects of variable conditions while improving accuracy, sensitivity, and specificity\(^{44}\). Overlapping genes among the three algorithms were considered as candidate immune-related biomarkers in HN. To assess the associated efficiency of immune-related genes, the area under the curve (AUC) of the receiver-operating characteristic (ROC) curve was measured by an R package of pROC\(^{45}\). For clinical HN diagnosis, nomogram construction was valuable, so the R package “rms” was used to construct a nomogram scoring system based on candidate genes. “Points” represents the score of candidate genes, while “Total Points” represents the total score of all candidate genes.

5.6 Evaluation of immune cell infiltration and analysis of its relationship with characteristic genes

The R software was used to link the ssGSEA deconvolution algorithm to simulate the transcriptional characteristics of immune cells. In each sample, 28 types of immune cells were identified. The differences in immune cell infiltration between HN and control samples were analyzed by the Wilcoxon rank sum test. An in-
depth analysis of correlations between characteristic genes and infiltrating immune cells was possible using Pearson correlation analysis. The results of the analysis were visualized using “ggplot2” from the R package.

5.7 The Establishment of HN Model

The spontaneously hypertensive rats (SHRs) have been widely used as an essential hypertension animal model, and multiple structural and functional changes in the kidneys mark hypertensive nephropathy. Gan et al reported that 20-week-old SHR rats exhibited significant renal injury characteristics compared to normal rats.

A total of six male Wistar Kyoto normotensive rats (WKY) and six male spontaneously hypertensive rats (SHRs) at the age of 10 weeks were purchased from Beijing Vital River Laboratory Animal Technology (Beijing, China).

Six male Wistar Kyoto normotensive rats (WKY) and six spontaneously hypertensive rats (SHRs) with 10-week-old were purchased from Beijing Vital River Laboratory Animal Technology Co., Ltd. WKY rats served as healthy controls. Food and water were freely available to all rats in a temperature-controlled room. Kidney tissue samples were collected after feeding 10 weeks for histology and biochemical detection. Animal welfare regulations and guidelines established by the Animal Experimental Ethical Inspection Committee of Hunan University of Chinese Medicine were strictly followed in the rearing and treatment of rats. The use of rat approved by the Experimental Animal Ethics Committee of Hunan University of Chinese Medicine in accordance with NIH guidelines. This study was carried out in compliance with the ARRIVE guideline.

5.8 HE and Masson staining

At the end of the experiments, rats were anesthetized with isoflurane and then sacrificed. Kidney tissues were collected. Paraffin-embedded kidney tissues were cut into slices for HE and Masson staining, which were performed as previously described.

5.9 PCR experiment

Total RNA was isolated using a TRIZOL reagent according to the manufacturer’s instructions (Invitrogen). SuperScript First Strand cDNA System was used to reverse-transcribe the RNA. On the Exicycler 96 system, a real-time PCR assay was performed. Results were calculated using the $2^{-\Delta\Delta CT}$ method. β-actin was a reference gene. The primer sequences are as follows. GBP2: F, CTGTGGTGTTGCTGGAGCTGTTG; R, TGAGACTTCCTGTGGAGCTAGAG; TMSB10: F, CGATAAGGCAAGCTGAGGATTG; R, CACTCCGCTTCTCCTGCTCAATG. β-actin: F, AGGCCCTCTGAACCCTAAG; R, CCAGAGGCATACACG.

GGGACAAAC.

5.10 Western Blot Assay

Protein concentrations were measured with the Bioss BCA kit (Bioss) after the total proteins were extracted from kidney tissues. By separation with SDS-PAGE, the protein was then transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore). Non-fat milk was used to block the membrane, and then GBP2 (Affinity, 1:1000), TMSB10 (Affinity), α-SMA (Affinity, 1:1000), CD86 (Santa cruz, 1:500) and β-actin (Servicebio, 1:2000) antibodies were incubated respectively. After overnight incubation at 4°C, TBST was used to wash the membrane, followed by secondary antibody incubation and ECL (Biosharp) visualization.
5.11 Statistical Analysis

Data analysis was performed using GraphPad Prism 8 (GraphPad Software, La Jolla, CA). Student’s t-test or Wilcoxon rank sum test was used to analyze all data. Values are expressed as Mean ± SEM. P-value < 0.05 was considered statistically significant.

Declarations

Conflict of Interest

The authors declare that there are no commercial or financial relationships that could construe the research as having a conflict of interest.

Author Contributions

LXL and LHG analyzed and interpreted the data. LXL, LHG wrote the manuscript. LJJ and WYH designed and edited the manuscript. All authors contributed to the article and approved the manuscript.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (81874464; 82174357), the Education Department Scientific Research Foundation of Hunan Province of China (18A228) and the Hunan University of Chinese Medicine postgraduate innovation project in 2022 (2022CX95).

Competing interests

The authors declare no competing interests.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

References


(2013).


Figures
Figure 1

Analyses of transcriptome profiles of HNs and controls. (A) Significant DEGs between HN and Control samples on a volcano plot. (B) Heatmap showing the top 20 DEGs that are significantly regulated. (C) Top 10 enriched GO terms for DEGs in each category. (D) The bubble plot shows the KEGG pathways significantly enriched in DEGs.
Figure 2

WGCNA Analysis. (A) Analyzing the scale-free fit index and mean connectivity for various soft-threshold powers to determine the WGCNA soft-threshold power. (B) A cluster dendrogram derived from independent data sets created 12 gene coexpression modules based on branches of the cluster dendrogram. (C) The relationship between consensus modules and clinical trait. (D) Scatter plot of gene significance for HN related to module membership in the magenta module. (E) The DEGs and MEmagenta modules have 16 overlapping candidate genes.
Figure 3

Screening candidate immune-related genes for HN with machine learning. (A) Candidate gene profiles were calculated using LASSO coefficients. (B) Analyzing LASSO regression using cross-validation to select the optimal tuning parameter log (Lambda). (C, D) A random forest algorithm is used to find the error in HN and control groups; Sort according to the importance score of genes. (E) By using SVM-RFE analyses, 11 gene signatures are identified with an accuracy of 0.85. (F) The Venn diagram of two overlapping candidate genes shared by LASSO, RF and SVM-RFE algorithms.
Figure 4

Constructing a nomogram and evaluating its diagnostic value. (A) The visible nomogram for diagnosing HN. (B) Calibration curve evaluates the prediction efficacy of the nomogram. (C) The nomogram's clinical benefit is demonstrated by a decision curve analysis. (D-F) The ROC curve of candidate genes (GBP2, TMSB10) in the combined GSE37455 and GSE37460 datasets. (G-I) The ROC curve of candidate genes (GBP2, TMSB10) in the GSE104954 dataset.
Figure 5

Immune cell infiltration analysis between HN and control. (A) The proportion of 28 kinds of immune cells in different samples was visualized from the barplot. (B) Box plots depicting the infiltration levels of immune cells in Control and HN. (C) Box plots showing the gene expression of immune checkpoints in Control and HN. (D) Correlation analysis of immune cell infiltrations with characteristic genes.
Figure 6

The mRNA and protein expression of GBP2 and TMSB10 were upregulated in SHR rats relative to WKY rats. (A) Representative microscopy images of HE staining and Masson staining of kidney sections from the Control group and HN group (×200; scale bar = 100 µm). (C) The expression of GBP2 and TMSB10 mRNA in HN group relative to Control group by RT-qPCR. (D) Representative Western blots and quantification of α-SMA, CD86, GBP2 and TMSB10 protein levels in the Control group and HN group (**P < 0.01, *P < 0.05, n = 6).
Figure 7
The study flow chart.

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigureS1thebloodpressureofWKYandSHRrats.jpg
- SupplementaryTablesS1summaryof454BPpathways67CCpathways40MFpathwayssignificantlyenriched.txt
- SupplementaryTablesS2summaryof47KEGGpathwayssignificantlyenriched.txt
- SupplementaryTablesS3RenalfunctionparametersofratsinWKYgroupsandSHRgroups..doc