Identification of three autophagy-related genes as diagnostic biomarkers and analysis of immune cell infiltration in ankylosing spondylitis patients

Mingguang Yan
the first People's Hospital of Shang Qiu

Xiao Fang
the First Affiliated Hospital of Shangqiu

Jianjun Guo
the first People's Hospital of Shang Qiu

Weibing Yin (✉️ yinweibing002@163.com)
the first People's Hospital of Shang Qiu

Research

Keywords: Ankylosing spondylitis, bioinformatics, autophagy, immune cell infiltration, GSEA

Posted Date: November 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2210447/v1

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

Background

Ankylosing spondylitis (AS) is a rheumatic immune disease that predominantly affects the sacroiliac joints and spinal joints, but the etiopathogenesis of AS remains unclear. The present research aimed to identify novel therapeutic targets and explore the molecular mechanism of AS via a bioinformatics approach.

Methods

Two microarray datasets (GSE25101 and GSE18781) were downloaded, and Gene Set Enrichment Analysis (GSEA) was used to analyze autophagy-related pathways. Autophagy-related genes (ARGs) were collected from the Human Autophagy-dedicated Database. The differentially expressed genes (DEGs) were screened with the limma package. Differentially expressed autophagy-related genes (DEARGs) were identified by intersecting the DEGs with the ARGs. Besides, GO-BP and KEGG enrichment analyses of DEARGs associated with AS were conducted by clusterProfiler package. Furthermore, hub genes among DEARGs were screened by Receiver operating characteristic (ROC) curve analysis. Finally, the expression of hub DEARGs were validated by GSE73754 dataset.

Results

GSEA results indicated that selective autophagy, programmed cell death, and endocytosis were involved in the occurrence and development of AS. A total of 10 DEARGs shared in the two datasets were identified. Besides, functional enrichment analysis results indicated these DEARGs were mainly enriched in mitophagy and autophagy. Three core DEARGs (PTEN, GABARAPL2, and PRKCQ) with AUC > 0.7 were confirmed to have the diagnostic value in AS. Immune cell infiltration analysis identified CD8 + T cells, NK cells, neutrophils, Tgd cells, Th1 cells, and Th2 cells as major participants in the AS development.

Conclusions

Overall, PTEN, GABARAPL2, and PRKCQ may be used as diagnostic biomarkers for AS. Besides, their relationships with immune cell infiltration will contribute to the development of immunotherapy in AS patients.

Introduction

Ankylosing spondylitis (AS) is a chronic inflammatory bone disease, characterized by functional impairments and inflammation of sacroiliac joints and spine, which results in progressive joint ankylosis and seriously affects patients' quality of life (1). Thus, the earlier an AS patient is diagnosed, the less
burden it is on the sufferer. Previous reports have revealed that genetic, immune, environmental, and metabolic factors are involved in the etiopathogenesis of AS (2, 3). However, the specific pathogenesis and etiology of AS remain unclear. Besides, there is a lack of effective drugs to treat AS patients, and they cannot be completely cured. Additionally, the terms disease prognosis and progression are difficult to define in AS. Thus, identifying novel and therapeutic biomarkers is very important for the early diagnosis and treatment of AS.

Autophagy is an important player in cellular metabolism involving the degradation of organelles and cytoplasmic proteins into fatty acids and amino acids for energy recycling (4). Previous reports have indicated that autophagy is associated with the development of various human diseases, including central nervous system diseases, neurodegenerative diseases, immune diseases, and cancer (5). Recent reports have revealed that mitophagy and autophagy might play an important role in bone metabolic diseases (6, 7). Furthermore, autophagy dysfunction participates in the etiopathogenesis of AS (8). Defective autophagy activity was found in AS patients and it may contribute to the development of spinal injury in AS (9). A previous study revealed that autophagy-related genes (ATG12 and ATG5) expression levels in AS sufferers were correlated with the activity and severity of the AS (10). However, the diagnostic efficiency of autophagy-related genes in AS has not been fully expounded.

In recent years, some progress has been made in bioinformatics analyses and high-throughput techniques involving the identification of novel targets and pathways to understanding the potential pathogenic process of diseases, including bone metabolic diseases (11–13). Therefore, this approach could help us systematically investigate the role of autophagy in the etiology of AS.

In the present study, the gene expression profiles were downloaded from the GEO database, and the bioinformatics approach was utilized to identify autophagy-related pathways and targets. These findings could help us to screen novel diagnostic biomarkers for AS treatment and deepen the understanding of the pathologic mechanisms of AS.

**Material And Methods**

**GEO datasets and autophagy-related genes acquisition**

The autophagy-related genes (ARGs) were collected from the Human Autophagy-dedicated Database (HADb) (14). Raw and series matrix files of three microarray datasets (GSE25101, GSE18781, and GSE73754) were downloaded from the GEO database. Before data analysis, the raw expression matrix was extracted and normalized using the R package “affy”. The probe expression matrix was converted into the gene expression matrix using a platform annotation file.

**Differentially expressed analysis of ARGs**

We used the Robust Multiarray Average method to pre-process the GEO data. The differentially expressed ARGs (DEARGs) were identified by using the limma package of R software. |log2FC| ≥ 0.2 and p-value <
0.05 was defined as criteria for DEGs identification. The “ggplot2” package was used to draw the volcano map of DEGs. The heatmap package was used to draw the heatmap of DEARGs.

Construction of protein-protein interaction network (PPI) and correlation analysis of DEARGs

In the present study, the STRING database was used to construct the PPI network of DEARGs. Cytoscape (version 3.8.2) was used to visualize the result of the PPI analysis. Besides, we performed Pearson correlation analyses of DEARGs by using the “ggplot2” package of R software.

Gene Set Enrichment Analysis (GSEA)

GSEA was performed via the clusterProfiler package of R software to screen global gene enrichment differences between control and AS groups. The expression profile of all genes in the GSE25101 and GSE18781 datasets were uploaded, and p.adjust < 0.05 and false discovery rate (FDR) < 0.25 were considered statistically significant. The c2.cp.v7.2.symbol.gmt was downloaded from the MSIGDB database (http://software.broadinstitute.org/gsea/msigdb) for GSEA.

Functional enrichment analysis of DEARGs

In the present study, we performed KEGG and GO-BP enrichment analysis to investigate the biological functions of common DEARGs in the GSE25101 and GSE18781 datasets. The Cluster Profiler of R software was used to perform enrichment analysis, and \( P < 0.05 \) was considered statistically significant.

Receiver operating characteristic (ROC) curve analysis of DEARGs

ROC curve analysis was performed to predict the diagnostic value of DEARGs. We used the pROC package of R software to perform ROC curve analysis. The area under the curve (AUC) > 0.7 was considered good diagnostic effectiveness.

**Results**

GSEA shows that autophagy participates in the progress of AS

We performed GSEA to investigate the biological function of DEGs. As shown in Fig. 1, our results revealed that most of the enriched gene sets were involved in selective autophagy (Fig. 1A, normalized enrichment scores (NES) = 1.372, p.adj = 0.047, FDR = 0.032), programmed cell death (Fig. 1B, NES = 1.481, p.adj = 0.015, FDR = 0.011), neurotrophin signaling pathway (Fig. 1C, NES = 1.406, p.adj = 0.015, FDR = 0.011), oxidative phosphorylation (Fig. 1D, NES = 1.569, p.adj = 0.015, FDR = 0.011), endocytosis (Fig. 1E, NES = 1.375, p.adj = 0.019, FDR = 0.013), and apoptosis (Fig. 1F, NES = 1.427, p.adj = 0.024, FDR = 0.016). These findings indicated autophagy may play an important role in the pathological process of AS.

Identification of DEARGs in GSE25101 and GSE18781 datasets
First, a total of 1592 DEGs were identified via the analysis of the GSE25101 dataset with R software. All genes were presented as volcano plots, of which black green dots represent 761 down-regulated DEGs, red dots represent 831 up-regulated DEGs, and grey dots indicate no significant difference genes (Fig. 2A). Next, a total of 43 DEARGs were obtained between ARGs and DEGs via a Venn tool (Fig. 2B), and the mRNA expression levels of 43 DEARGs in the GSE25101 dataset were presented as a heatmap (Fig. 2C).

We also identified 1311 DEGs via the analysis of the GSE18781 dataset with R software. All genes were presented as volcano plots, of which black green dots represent 417 down-regulated DEGs, red dots represent 894 up-regulated DEGs, and grey dots indicate no significant difference genes (Fig. 2D). Next, a total of 19 DEARGs were obtained between ARGs and DEGs via a Venn tool (Fig. 2E), and the mRNA expression levels of 19 DEARGs in the GSE18781 dataset were presented as a heatmap (Fig. 2F).

PPI network analysis and correlation analysis of DEARGs

First, we identified 10 intersection DEARGs from GSE25101 and GSE18781 datasets by using the Venn tool (Fig. 3A). Then, we used the Cytoscape to visualize the PPI network of the 10 intersection genes. The Fig. 3B exhibited the 10 intersection genes, including BNIP3L, ATG3, PRKCQ, GABARAPL2, TP53, PINK1, NRG1, NPC1, HSP90AB1, and PTEN, these 10 genes were identified as key genes involved in the development and progression of AS. The results also demonstrated that most of these ARGs interacted with each other.

Functional enrichment of DEARGs

As shown in Fig. 4A-B, KEGG enrichment analysis indicated that the DEARGs were mainly associated with mitophagy, autophagy, amyotrophic lateral sclerosis, etc. GO-BP enrichment analysis revealed that the DEARGs were significantly enriched in the process utilizing autophagic mechanism, negative regulation of the catabolic process, macroautophagy, autophagy, etc (Fig. 4C-D).

Assessment of diagnostic effectiveness of potential biomarkers

In the present study, we performed the ROC analysis of 10 intersection DEARGs to assess the diagnostic value of the biomarkers for AS. As shown in Fig. 5A-B, the AUC values of NPC1, PTEN, GABARAPL2, BNIP3L, PRKCQ, TP53, NRG1, HSP90AB1, ATG3, and PINK1 were 0.781, 0.797, 0.744, 0.715, 0.793, 0.758, 0.762, 0.805, 0.809, and 0.713, respectively. As shown in Fig. 5C-D, the AUC values of NPC1, PTEN, GABARAPL2, BNIP3L, PRKCQ, TP53, NRG1, HSP90AB1, ATG3, and PINK1 were 0.648, 0.723, 0.725, 0.675, 0.717, 0.597, 0.525, 0.683, 0.501, and 0.643, respectively. Based on the above results, the AUC of PTEN, GABARAPL2, and PRKCQ is more than 0.7, indicating those genes have the diagnostic ability to distinguish between AS and normal samples.

Validation of PTEN, GABARAPL2, and PRKCQ expression in external independent dataset
As shown in Figs. 6A-B, the expression levels of PTEN and GABARAPL2 genes in the control group were significantly lower than that of the AS group (p < 0.05), while the expression level of PRKCQ gene in the control group was significantly higher than that of the AS group (p < 0.01). Furthermore, one independent microarray datasets (GSE73754) was used for validation. As shown in Fig. 6C, the expression levels of PTEN and GABARAPL2 genes in the control group were significantly lower than that of the AS group (p < 0.01), while the expression level of PRKCQ gene in the control group was significantly higher than that of the AS group (p < 0.001).

**Immune cell infiltration analysis**

xCell was applied to assess the differences in immune cell infiltrates between the control and AS groups. As presented in Fig. 7A, the score of basophils, CD8 + T cells, CD8 + Tcm, CD8 + Tem, macrophages M2, NK cells, Tgd cells, Th1 cells, and Th2 cells was significantly decreased in the AS group, while the score of neutrophils was significantly increased in the AS group. Besides, correlation analysis showed that GABARAPL2 was negatively correlated with CD4 + T cells, CD8 + T cells, NK cells, pro-B cells, Tgd cells, Th1 cells, and Th2 cells. PRKCQ was positively correlated with CD4 + T cells, CD8 + T cells, NK cells, Tgd cells, Th1 cells, and Th2 cells (Fig. 7B).

We also used MCPCounter to assess the differences in immune cell infiltration between control and AS groups. As shown in Fig. 8A, the score of cytotoxic lymphocytes, NK cells, and monocytic lineage was significantly decreased in the AS group, while the score of neutrophils was significantly increased in the AS group. GABARAPL2 was negatively correlated with T cells, CD8 + T cells, cytotoxic lymphocytes, NK cells, monocytic lineage, and myeloid dendritic cells. PRKCQ was positively correlated with T cells, CD8 + T cells, and cytotoxic lymphocytes (Fig. 8B).

**Discussion**

AS is a chronic inflammatory disease predominantly impacting the sacroiliac and spine joints, which is associated with the dysfunctions of both skeletal and immune systems (15). Although AS was first described more than 200 years ago. However, its underlying disease mechanism remains not fully understood. Recent studies have indicated that gut mucosal inflammation, immune disorders, genetic susceptibility, and environmental incentives are being implicated in the pathogenesis of AS (16–18).

Autophagy has been demonstrated to affect the immune system and is therefore associated with some rheumatic disease processes (6, 19). Increasing evidence indicated that autophagy may also be implicated in the pathophysiology of AS (10). It has been reported that AS patients have defective autophagy activity and autophagy may represent a protective mechanism in the development of spinal injury in AS (9). However, more research is needed to support the understanding of autophagy in the pathogenesis of AS. In the present study, our results of GSEA indicated that most of the enriched genes of AS were involved in selective autophagy, programmed cell death, neurotrophin signaling pathway, oxidative phosphorylation, endocytosis, and apoptosis, indicating that autophagy may play an important
role in the pathological process of AS. Therefore, exploring novel and effective autophagy-related biomarkers for the diagnosis and treatment of AS is very important.

As far as we know, there are several reports investigating autophagy-related biomarkers in cancer, and chronic obstructive pulmonary disease (20–22). However, no studies are exploring the autophagy-related genes in AS via the bioinformatics method. In the present study, 10 potential autophagy-related biomarkers in AS were first identified via bioinformatics analysis. Some of these genes of AS have been reported in previous studies. For instance, a previous study indicated that TP53 and BNIP3L may be involved in the pathogenesis of juvenile spondyloarthritis (23).

Subsequently, the results of enrichment analysis indicated that DEARGs were significantly enriched in autophagy-related pathways, such as mitophagy, autophagy, a process utilizing autophagic mechanism, macroautophagy, etc. Autophagy is a cellular lysosomal degradation process that removes dead cells and senescent organelles and controls energy balance (4). Dysfunction of autophagy might be involved in various metabolic disorders, such as osteoporosis, obesity, cancers, and autoimmune disease (24).

Several studies have demonstrated that autophagy may be involved in the development and progression of AS. A recent report indicated that emodin could effectively induce autophagy and apoptosis of fibroblasts from AS patients and the combination of autophagy inhibitor with emodin is a potential therapy for the treatment of AS patients (25). AS patients had declined autophagy levels and miRNA-199a-5p could promote autophagy-related gene expression and prevent the pathogenesis of AS (26). DEPTOR and TNFAIP3 together medicated inammasome secretion by promoting early-onset autophagy in AS monocytes (27). Finally, we assessed the diagnostic value of 10 key DEARGs by using ROC curve analysis. Our results indicated that PTEN, GABARAPL2, and PRKCQ were potential biomarkers for the diagnosis of AS. PTEN is implicated in the inflammation and activation of fibroblast-like synoviocytes in rheumatoid arthritis (28). PTEN was also indicated to promote the development of collagen-induced arthritis, and a decrease of PTEN in myeloid cells suppressed the progression of autoimmune arthritis via inhibiting the Th17 type of immune response (29). GABARAPL2 was identified as a diagnostic biomarker in rheumatoid arthritis patients (30). PRKCQ was indicated to influence nerve regeneration and degeneration in injured sciatic nerve (31).

AS is an inflammatory disease, and a recent study revealed that immune cells participated in the development of AS (32). In our study, immune cell infiltration analysis identified CD8+ T cells, NK cells, neutrophils, Tgd cells, Th1 cells, and Th2 cells as major participants in the AS development. In addition, we found that GABARAPL2 and PRKCQ were related to most of immunocytes. Previous report indicated that CD8+ T cell frequency was decreased in blood of AS patients, and CD8+ T cell plays a central role in AS development (33). Single-cell RNA sequencing revealed altered NK cells subsets in peripheral circulation of AS patients (34). Increased of neutrophils may be a primary factor in the AS progression (35). These studies revealed that significance of immune cell infiltration in the etiopathogenesis of AS. Our results and previous reports implied that GABARAPL2 and PRKCQ may impact the immune function of AS patients.
Conclusion

In conclusion, PTEN, GABARAPL2, and PRKCQ are core DEARGs in AS patients that might be served as potential biomarkers for AS. Our findings expanded our understanding of AS.

Abbreviations

Ankylosing spondylitis (AS); Gene Set Enrichment Analysis (GSEA); Autophagy-related genes (ARGs); differentially expressed genes (DEGs); Differentially expressed autophagy-related genes (DEARGs); Receiver operating characteristic (ROC); Human Autophagy-dedicated Database (HADb); protein-protein interaction network (PPI); false discovery rate (FDR).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors are consent for the publication of this work.

Availability of supporting data

All the datasets were retrieved from the GEO database. And all of other materials are available by the corresponding authors.

Competing interests

All authors declared no conflicts of interest.

Funding

This study was supported by the grants from the Science and Technology Project of Henan Province (212102310743).

Author contributions

MGY performed this study and wrote the initial draft of the manuscript. XF and JGG contributed to the literature search and design the study. WBY reviewed and edited the manuscript. All authors read and approved the manuscript.

Acknowledgements

None.
Authors’ information

Department of Clinical Laboratory, the First People’s Hospital of Shang Qiu, Shang Qiu 476000, Henan, China.

References


**Figures**
Gene Set Enrichment Analysis (GSEA). FDR < 0.25 and p.adjust < 0.05 were considered to be significant. Enriched gene sets are mainly involved in selective autophagy (Figure 1A), programmed cell death (Figure 1B), neurotrophin signaling pathway (Figure 1C), oxidative phosphorylation (Figure 1D), endocytosis (Figure 1E), and apoptosis (Figure 1F).
Figure 2

Identifying differentially expressed autophagy-related genes (DEARGs). (A) Volcano plots of DEGs from GSE25101 between the normal samples and AS samples; black green indicates down-regulated DEGs, red indicates up-regulated DEGs, and grey indicates no significant difference genes. (B) Venn diagram used to screen DEARGs in GSE25101. (C) Heatmap of 43 DEARGs from GSE25101 dataset. (D) Volcano plots of DEGs from GSE18781 between the normal samples and AS samples; black green indicates down-regulated DEGs, red indicates up-regulated DEGs, and grey indicates no significant difference genes. (E)
Venn diagram used to screen DEARGs in GSE18781. (F) Heatmap of 19 DEARGs from the GSE18781 dataset.

**Figure 3**

Identification of key DEARGs. (A) 10 common genes were identified via the Venn tool. (B) Protein-protein interaction networks of the 10 key DEARGs.
Figure 4

Functional enrichment analysis of 10 key DEARGs. (A) The results of KEGG were presented by circle charts. (B) The results of KEGG were presented by bar plot. (C) The results of GO-BP were shown by circle charts. (D) The results of GO-BP were shown by bar plot.
Figure 5

Analysis of intersection DEARGs. (A, B) The GSE25101 dataset was used to assess the diagnostic value of intersection DEARGs. (C, D) The GSE18781 dataset was used to assess the diagnostic value of intersection DEARGs.
Figure 6

Validation of hub DEARGs in AS. The expression levels of PTEN, GABARAPL2, and PRKCQ genes in the GSE25101 (A) and GSE18781 (B). A validation cohort of GSE73754 was used to confirm the expression levels of PTEN, GABARAPL2, and PRKCQ. *p < 0.05, **p < 0.01 and ***p < 0.001 were analyzed by Wilcoxon rank sum test.
Figure 7

xCell algorithm was applied to perform immune cell infiltration analysis in AS. (A) The landscape of the 35 types of immune cells between control and AS groups. (B) Correlation analysis between the hub DEARGs and immune cell infiltration levels. *p < 0.05, **p < 0.01, and ***p < 0.001.

Figure 8
MCPCounter algorithm was applied to perform immune cell infiltration analysis in AS. (A) The landscape of the 8 types of immune cells between control and AS groups. (B) Correlation analysis between the hub DEARGs and immune cell infiltration levels. *p < 0.05, and **p < 0.01.