Immune-related genes probably increase the occurrence of Interstitial Lung Disease in Dermatomyositis

Changjian Liu  
China-Japan Friendship Hospital

Wei Jiang  
China-Japan Friendship Hospital

Yongpeng Ge  
gyp2016@163.com  
China-Japan Friendship Hospital

Research Article

Keywords: dermatomyositis, interstitial lung disease, differentially expressed genes, immune cells

Posted Date: March 23rd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2657185/v1

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

Background Interstitial lung disease (ILD) is one of the significant complications of dermatomyositis (DM), but the mechanisms by which it occurs remain incompletely elucidated. This study aimed to explore further the possible genetic mechanisms by which this complication occurs.

Methods Gene expression profiles for dermatomyositis (GSE39454, GSE46239, GSE143323) and interstitial lung disease (GSE32537, GSE110147, GSE150910) were downloaded from the Gene Expression Omnibus (GEO) database. After identifying common differentially expressed genes (DEGs) to dermatomyositis and interstitial lung disease using the "limma" R package and the "VennDiagram" R package, functional annotation, relationship to immune cell infiltration, identification of transcription factors (TFs). We also collected clinical cases of dermatomyositis-associated interstitial lung disease (DM-ILD), including 3 cases of rapidly progressive interstitial lung diseases and 3 cases of none-rapidly progressive interstitial lung diseases, and explored whether there were differences in serum lymphocyte subpopulations.

Results A total of 4 common DEGs (SLAMF7, SPP1, TDO2, and VCAM1) were screened and GO enrichment analysis showed that these genes were mainly enriched in T cell activation, regulation of lymphocyte activation, lymphocyte differentiation, leukocyte proliferation and regulation of T cell activation. In terms of KEGG pathways, the three significantly enriched pathways were the PI3K-Akt signaling pathway, MAPK signaling pathway, and Cytokine-cytokine receptor interaction. In lung and muscle tissues, 21 and 3 TFs may regulate the expression of these genes, respectively. Finally, by analysing the serum lymphocyte subpopulations, we also found a decrease in the absolute number of CD8+ T cells and an increase in the CD4+/CD8+ T cell ratio in dermatomyositis combined with rapidly progressive interstitial lung disease.

Conclusion These common pathways and key genes may provide new ideas for further research into DM-ILD.

Key Points
- We explored the common DEGs in DM and ILD.
- CIBERSORT analysis showed that DM and ILD are associated with immune cell infiltration.
- TFs that may regulate common DEGs in muscle and lung tissue.
- The CD4+/CD8+ T cell ratio was increased in patients with DM-associated rapidly progressive ILD.

Introduction

Dermatomyositis (DM) is a group of systemic autoimmune diseases that can affect skin, muscle and lung tissue[1–3]. The most common extra-muscular target organ is the lung tissue. It is estimated that approximately 30–40% of DM can be complicated by interstitial lung disease (ILD), with limited efficacy
of immunosuppressive combined with glucocorticoid therapy and often associated with a poor prognosis\textsuperscript{[4–7]}. In particular, DM patients with rapidly progressive interstitial lung disease (RP-ILD) have been identified as a major cause of death\textsuperscript{[8, 9]}. Consequently, it is crucial to explore for pathogenetic processes shared by DM and ILD, which will give novel diagnostic and therapeutic strategies for the disease.

Even though ILD is a the common complication of DM, its exact mechanism remains unclear. Thus, no specific targeted therapies are available to treat DM while controlling the progression of ILD\textsuperscript{[10]}. It is currently believed that quantitatively and qualitatively, immune cells are abnormal in the tissues and circulation of DM, including T cells, B cells, NK cells, neutrophils and macrophages\textsuperscript{[11, 12]}. Adaptive immunity is mediated by T and B cells, which recognize specific antigens and produce specific cell-mediated and antibody-mediated responses\textsuperscript{[13]}. NK cells are mainly involved in innate immunity and act through the release of immunomodulatory cytokines or cytotoxic particles upon activation\textsuperscript{[14]}. Other immune cells also exert their corresponding immune responses in the body. Ultimately, these aberrantly activated immune cells may attack lung tissue by changing the target\textsuperscript{[15, 16]}. For example, CD163 can be shed from activated macrophage membranes and released into peripheral blood\textsuperscript{[17]}, and immunohistochemical analysis also showed a significant alveolar infiltration of CD163\textsuperscript{+} macrophages in the lungs of patients with DM-associated ILD compared to normal lungs\textsuperscript{[18]}. Furthermore, levels of neutrophil-secreted α-defensin, which increases collagen production and causes lung fibrosis, are significantly elevated in the plasma and alveolar lavage fluid of DM-ILD patients and rise as the disease progresses\textsuperscript{[19]}

It is important to note that most of the current mechanistic research on DM-ILD is focused on serological aspects, and very little research has been done on the transcriptomic characteristics of the tissues involved in either or both diseases. This study aimed to search for commonly related genes in the pathogenesis of DM and ILD. We analysed six gene expression datasets downloaded from the GEO database (GSE39454, GSE46239, GSE143323, and GSE32537, GSE110147, GSE150910). A bioinformatic approach was used to identify common differentially expressed genes (DEGs) and their biological roles and to analyze the association between common DEGs and immune cell infiltration. In addition, we analysed the transcription factors (TFs) that may regulate these genes. Finally, we investigated whether the serum lymphocyte subsets of DM-RP-ILD were different from those of DM-NRP-ILD. The above analysis is expected to provide a more detailed understanding of the possible mechanisms of DM-ILD.

**Materials And Methods**

**Data sources**

The datasets of DM (GSE39454, GSE46239, GSE143323) and ILD (GSE32537, GSE110147, GSE150910) generated and/or analysed during the current study are available in the GEO database (http://www.ncbi.nlm.nih.gov/geo/). The GSE39454 dataset contains 31 DM and 5 normal samples. The GSE46239 dataset contains 48 DM and 4 normal samples. The GSE143323 dataset contains 39 DM and
20 normal samples. The GSE32537 consists of 167 ILD and 50 normal samples. The GSE110147 consists of 37 ILD and 11 normal samples. The GSE150910 consists of 103 ILD and 103 normal samples. Of these, GSE143323 and GSE150910 were used as validation sets. The current study follows GEO’s data access policy and publication guidelines.

In addition, we retrospectively analysed information on lymphocyte subpopulations in six patients with DM diagnosed during their stay in the rheumatology department of the China-Japan Friendship Hospital from May 2021 to October 2022. Inclusion criteria: (i) serological features of anti-MDA5 antibody; (ii) initial diagnosis at our hospital; (iii) no systemic treatment before admission. The collected cases were grouped and patients were considered dermatomyositis-associated rapidly progressive interstitial lung disease (DM-RP-ILD) when they had an acute exacerbation within 3 months of diagnosis, otherwise dermatomyositis-associated none-rapidly progressive interstitial lung disease (DM-NRP-ILD). This study protocol complied with the ethics requirements of the China-Japan Friendship Hospital (reference number: 2019-25-K19).

**Obtaining the common differentially expressed genes (DEGs)**

The “limma” R package was used to analyze the DEGs between disease and normal tissues, including two datasets of DM (GSE39454, GSE46239) and two datasets of ILD (GSE32537, GSE110147). DEGs were identified when P-value < 0.05 and |LogFC| ≥ 1. Volcanoes were mapped using the “ggplot2” R package. We then obtained the common DEGs and visualised the results, venn diagram using the “VennDiagram” R package and a heatmap using the “pheatmap” R package. The receiver operator characteristic curve (ROC) was drawn using the “pROC” R package to assess the sensitivity and specificity of the common DEGs for disease diagnosis.

**Functional analysis of the common DEGs**

A Spearman rank correlation analysis was performed in order to obtain the associated genes for the common DEGs in the four datasets discussed above (Cor > 0.3 and P-value < 0.05). We then obtained the shared associated genes for each common DEGs in each of the four datasets and combined them to visualize the number of shared associated genes using the “UpSetR” R package. Gene ontology (GO) and Kyoto Encyclopedia of Genomes (KEGG) analyses were performed on the associated genes of the above combined common DEGs using the "clusterProfiler" R package, and the P values were adjusted using the B-H method. The relationship between expression levels of the common DEGs and 22 immune cell infiltration was assessed using the CIBERSORT algorithm.

**Validation of the common DEGs in other datasets**

The common DEGs were validated for expression levels in GSE143323 and GSE150910. As described above, the GSE143323 dataset consists of 39 DM and 20 normal samples. GSE150910 consists of 103 ILD and 103 normal samples. The Wilcox rank sum test was used to compare the two data groups. P-value < 0.05 was considered significant.
Prediction of TFs for common DEGs

Possible TFs of the common DEGs in muscle and lung tissues were obtained from the hTFtarget (http://bioinfo.life.hust.edu.cn/hTFtarget/)\(^\text{[20]}\) database and the regulatory network was visualised using Cytoscape software.

Lymphocyte subpopulations in DM-RP-ILD and DM-NRP-ILD sera

We collected lymphocyte subpopulations from 6 patients with anti-MDA5 type DM (including 3 cases of DM-RP-ILD and 3 cases of DM-NRP-ILD), and Student's t-test was used to analyse whether there were differences in lymphocyte subpopulations. We also analysed the chest's high-resolution CT (HRCT) features in 6 patients.

Statistical analysis

All statistical analyses were performed using R software (version 4.0.0) or GraphPad Prism (version 8.0.0). DEGs were analysed using the 'limma' package with thresholds set at \(|\log\text{FC}| \geq 1\) and \(P\)-value < 0.05. Two groups were compared using the Wilcoxon rank sum test or Student's t-test, and correlations were analysed using the Spearman method. \(P\)-value < 0.05 was considered statistically significant and all \(P\)-values were two-tailed.

Results

Four common DEGs were obtained

The flow chart of this study is shown in Fig. 1. Identification of DEGs using the "limma" R package (544 in GSE39454, 281 in GSE46239, 298 in GSE32537, and 3464 in GSE110147) (Fig. 2A, B, C, D). A total of 4 co-upregulated DEGs (SLAMF7, SPP1, TDO2, VCAM1) and no co-downregulated DEGs were obtained after taking the intersection using the Venn diagram (Fig. 2E, F). The heatmap showed the expression of SLAMF7, SPP1, TDO2, and VCAM1 in the above four datasets (Figure 2G). The ROC showed that SLAMF7, SPP1, TDO2, and VCAM1 were of high diagnostic value (AUC>0.7) for both DM and ILD (Fig. S1).

Obtaining common DEGs-associated genes and enrichment analysis

Firstly, the associated genes of SLAMF7, SPP1, TDO2, and VCAM1 in GSE39454, GSE46239, GSE32537, and GSE110147 were obtained separately, and the intersection of each gene in the above dataset was taken separately (499 for SLAMF7, 132 for SPP1, 366 for TDO2, and 380 for VCAM1) (Fig. S2 A, B, C, D) and visualized the network using Cytoscape (Fig. S2E). The intersecting genes were combined and analysed for GO and KEGG pathway enrichment. It was determined from the GO analysis that these
genes were mainly associated with T cell activation, regulation of lymphocyte activation, lymphocyte differentiation, leukocyte proliferation, and regulation of T cell activation (Fig. 3A). In terms of KEGG pathways, the three most significantly enriched pathways were found to be PI3K-Akt signaling pathway, the MAPK signaling pathway, and the Cytokine-Cytokine Receptor Interaction (Fig. 3B). It is evident from these findings that immune regulation plays a critical role in the development of these two diseases in conjunction with each other.

**Immune infiltration analysis**

We also assessed the relationship between the expression levels of SLAMF7, SPP1, TDO2, VCAM1 and 22 immune cell infiltrates based on CIBERSORT. High expression of SLAMF7, TDO2, and VCAM1 in DM was found to be positively correlated with the degree of infiltration of activated CD4 memory T cells, γδ T cells. In addition, high expression of SLAMF7 was positively correlated with the degree of infiltration of M1 macrophages and negatively correlated with the degree of infiltration of regulatory T cells (Fig. S3). In the ILD, high expression of SPP1 positively correlated with the degree of plasma cell infiltration; high expression of TDO2 positively correlated with activated CD4 memory T cells; and high expression of VCAM1 positively correlated with activated CD4 memory T cells and M1 macrophage infiltration (Fig. S4). Further evidence suggests that SLAMF7, SPP1, TDO2, and VCAM1 could be involved in immune cell infiltration during disease progression.

**Validation of common DEGs and Prediction of TFs**

To validate the reliability of the above common DEGs levels. We selected two other datasets containing DM and ILD and analysed the expression levels of SLAMF7, SPP1, TDO2, and VCAM1. The results showed that these genes were significantly upregulated in the DM compared to normal tissue (Fig. 4A). Similarly, the expression of all genes was higher in the ILD than in normal tissue (Fig. 4B), which further demonstrates the stability of differential expression of SLAMF7, SPP1, TDO2, and VCAM1 in the two diseases. In addition, based on the hTFtarget database, we identified 3 and 21 TFs in muscle tissue and lung tissue, respectively, that may regulate the expression of these genes (Fig. 5).

**CD4+/CD8+ T cells elevation in the serum of DM-RP-ILD patients**

The analysis revealed that the percentage and absolute number of CD8+ T cells were reduced and CD4+/CD8+ T cells were elevated in DM-RP-ILD sera relative to DM-NRP-ILD (Fig. 6), which may indicate severe abnormal activation of the immune system in DM-RP-ILD patients, attacking their own organs and tissues but with a resistance to bacteria, viruses or fungi diminished. In addition, chest HRCT shows that the lung lesions in DM-RP-ILD are more severe and often involve the upper and middle lobes of the lung (Fig. 7). Whether the above-mentioned common DEGs accelerate the development of DM-ILD by causing
abnormalities in CD8+ T cells and CD4+/CD8+ T cells need to be further analysed in vitro and vivo experiments. In addition, the reliability of the results needs further validation due to the small sample size.

Discussion

Some DM patients are insidious and frequently present to the respiratory department with pulmonary symptoms, making early diagnosis of DM-ILD difficult based on clinical symptoms, signs, and laboratory markers alone, which significantly impacts the prognosis and outcome of the disease [2]. Therefore, an active search for new markers of DM-ILD is urgent.

To the best of our knowledge, this is the first study to examine gene expression profiles common to DM and ILD at the transcriptome level. We found that the expression of SLAMF7, SPP1, TDO2, and VCAM1 was higher in tissue in both DM and ILD patients than in the corresponding normal tissue, and had some diagnostic value for the disease. Subsequently, GO and KEGG pathway enrichment analysis showed that these genes might affect the body's immune response through the regulation of signalling pathways such as the PI3K-Akt signaling pathway or MAPK signaling pathway, leading to the development of the disease. Through further validation, we likewise found that the expression of SLAMF7, SPP1, TDO2, and VCAM1 was associated with the infiltration of activated CD4 memory T cells and M1 macrophages in DM and ILD. In addition, we found that 21 and 3 TFs may regulate the expression of these genes in lung and muscle tissues, respectively. Moreover, CTCT and SPI1, which are both TFs, have been shown to be involved in the regulation of the expression of common DEGs in the lung and muscles. Taking into account the diverse results and complexity of each of the 4 common DEGs, we will focus on the main analytical findings, potential clinical relevance, and current findings of each.

Signaling lymphocytic activation molecule (SLAM) family member 7 (SLAMF7) gene is located on chromosome 1, has 8 exons, and is an I type transmembrane protein. It expresses a subpopulation of myeloma cells and immune cells, capable of participating in the body's immune response[21]. It was found that in a mouse model of multiple sclerosis, the immune cell receptor SLAMF7 was expressed on various subsets of immune cells in the central nervous system (CNS) and that its lack of expression could activate specific subsets of B cells and T memory cells, thereby regulating the susceptibility of the CNS to autoimmunity leading to the development of disease[22]. Also, an analysis of RNA-seq data from the synovial macrophages of people with rheumatoid arthritis showed a close link between SLAMF7 and macrophage activation. This link was also seen in the intestinal macrophages of people with Crohn's disease and in the lung macrophages of people with severe COVID-19[23], consistent with our findings in DM and ILD that common DEGs may contribute to disease development by inducing infiltration of M1 macrophages. Although there are no results from studies correlating SLAMF7 with DM and ILD, this common potential biomarker may be a cue for immune cell infiltration, i.e., SLAMF7 expression may lead to the development of DM-ILD by inducing an immune response.

Secreted phosphoprotein 1 (SPP1) is involved in osteoclast attachment to mineralized bone matrix and is a cytokine that upregulates interferon-γ and interleukin-12 expression[24]. Interferon-γ signaling was
upregulated in SPP1\textsuperscript{hi} macrophages, cytotoxic T cells, and natural killer cells in idiopathic pulmonary fibrosis, whereas type I interferon signaling and production were upregulated in the corresponding systemic sclerosis-associated interstitial lung disease population. In addition, gene expression changes in normal macrophages during the transition to SPP1 macrophages may be one of the reasons for systemic sclerosis-associated interstitial lung disease\cite{25}. In addition, the expression of SPP1 in bronchoalveolar lavage fluid of ILD patients was significantly higher than that of controls, and it regulated the occurrence and progression of ILD by affecting the expression of COL1A1\cite{26}. Although SPP1 has not been reported in DM, to some extent, it is difficult to conclude the correlation between the abnormal expression of SPP1 in the two diseases. However, the above findings of systemic sclerosis-associated interstitial lung disease and our analysis may offer new insights into the pathogenesis of DM-ILD.

Tryptophan 2,3-dioxygenase (TDO2) encodes a haemoglobinase that plays a crucial role in tryptophan metabolism\cite{27}. In vitro and in vivo assays found that TDO2 increased intracellular tryptophan metabolism levels in the kynurenine (Kyn) pathway and that the increase in the tryptophan metabolite Kyn led to sustained proliferation of glioma cells through AhR/AKT pro-survival signalling and immunosuppressive effects\cite{28}. Similar to SLAMF7, no studies have reported a role for TDO2 in DM or ILD. However, in other diseases, such as hepatocellular carcinoma, it has been found that TDO2 expression contributes to the secretion of interleukin-6 (IL-6), which promotes tumour cell proliferation through STAT3 and NF-kB/TIM4 signalling\cite{29}. In osteoarthritis patients, TDO2 levels were significantly and positively correlated with IL-1\textsubscript{\beta} and TNF-\alpha levels, suggesting that high levels of TDO2 in the synovium may correlate with pro-inflammatory cytokines and the severity of osteoarthritis\cite{30}. Among these, IL-6 plays a vital role in the inflammatory process, and studies have found higher levels of IL-6 in DM patients than in healthy populations\cite{31}. Gono et al. also demonstrated higher serum IL-6 levels in DM patients, particularly DM-ILD. The IG superfamily gene vascular cell adhesion molecule 1 (VCAM1) has been reported to play an essential role in the development of systemic sclerosis-associated interstitial lung disease\cite{32}. Although systemic sclerosis and DM are distinct diseases, they share similar pathogenesis. Furthermore, serum VCAM1 levels were significantly higher in DM-ILD patients compare to patients without ILD, suggesting that VCAM1 might be used as a biomarker to determine the severity of DM-related lung disease\cite{33}. Our study found that high expression of VCAM1 was associated with infiltration of activated CD4 memory T cells in both DM and ILD, suggesting an essential role for activated CD4 memory T cells in DM-ILD. However, this remains a speculative hypothesis that requires further validation.

In conclusion, our bioinformatics and clinical data analysis revealed 4 potential biomarkers associated with immune regulation in the interconnection between DM and ILD. We found that common DEGs were associated with immune cell infiltration in both diseases, which supports the possibility that common DEGs may contribute to DM-ILD through modulation of the body's immune system. In addition, we found a decrease in the absolute number of CD8 + T cells and an increase in CD4+/CD8 + T cells in DM-RP-ILD serum, suggesting that the body is in a state of immune dysfunction, which should be further analysed for association with common DEGs and needs to be verified in a large sample. The inferences that have
been made from the data are obviously hypotheses that need to be tested further. In addition, the specific mechanisms of common DEGs in the development of both diseases need to be further elucidated and are currently thought to be related to the PI3K-Akt signaling pathway or MAPK signaling pathway. In addition, it is unclear which gene is most relevant in the onset of the process between DM and ILD. A complex interplay between all these mechanisms appears to be the cause of the overlap between the two diseases. We have also explored the TFs that may regulate common DEGs in DM and ILD. This integrated bioinformatics approach has been shown to be reliable in various diseases. The present study will provide potential directions for the molecular mechanisms of DM-ILD.

Our study also has some limitations: (i) the main problem is the lack of validation of the findings by a large clinical sample. In addition, the DM in this study had a small number of healthy controls in the corresponding dataset, which needs to be considered when interpreting the results. (ii) In the case of DM-RP-ILD, we have limited the study to immune cells in serum, the sample size is very small, and further results in tissue are needed to explain this phenomenon.

**Declarations**

**Ethics approval and consent to participate**

Our study was approved by the ethical committee of the China-Japan Friendship Hospital (reference number: 2019-25-K19). Written informed consent was obtained from individual or guardian participants.

**Consent for publication**

Not applicable.

**Availability of data and materials**

The raw datasets used in this work can be downloaded in the GEO database (https://www.ncbi.nlm.nih.gov/geo/). All the databases above are open access and can be login in directly without relevant accession numbers.

**Competing interests**

The authors have no conflict of interest.

**Funding**

This work was supported by the National High Level Hospital Clinical Research Funding [2022-NHLHCRF-YS-02].

**Authors’ contributions**
Conceptualization, data curation, formal analysis, resources: Changjian Liu and Wei Jiang; writing—original draft: Changjian Liu; writing—review and editing: Yongpeng Ge.

Acknowledgements

We acknowledge the free use of GEO databases.

References


Figures
Figure 1

Flow chart of the study design.

Figure 2

DEGs. (A, B, C, D) Volcano plots of GSE39454, GSE46239, GSE32537, GSE110147. (E, F) 4 datasets show 4 common upregulated DEGs and no common downregulated DEGs. (G) Heatmap showing the expression of 4 common DEGs in the above 4 datasets, and relative to normal tissue, common DEGs are highly expressed in disease.

Figure 3

GO Enrichment

KEGG Enrichment
Results of common DEGs-associated genes GO (A) and KEGG (B) pathway enrichment analysis. The size of the circle indicates the count of genes involved and the horizontal coordinates indicate the proportion of genes involved to the total number of genes.

Figure 4
Validation of the expression levels of the common DEGs in GSE143323 (A) and GSE150910 (B), respectively. The stability of differential expression of all common DEGs in the two diseases.

**Figure 5**

Transcription factors regulatory network in muscle and lung tissues, TFs labeled in yellow, common DEGs labeled in blue, and tissues labeled in green.
Figure 6

Characteristics of lymphocyte subpopulations in serum from patients with DM-RP-ILD and DM-NRP-ILD. (A-B) The percentage and absolute number of CD3+ T cells and the percentage and absolute number of CD4+ T cells did not differ in the two groups. (C) The percentage and absolute number of CD8+ T cells were reduced in DM-RP-ILD sera. (D) CD4+/CD8+ T cells were elevated in DM-RP-ILD sera. Student's t-test compares the differences between the two groups.
Figure 7

HRCT characteristics of patients with DM-RP-ILD (A) and DM-NRP-ILD (B). Compared to DM-NRP-ILD, DM-RP-ILD lung lesions are often more severe and involve the lung's upper and middle lobes.

Supplementary Files

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

- SupplementaryInformation.docx