



Research article

Screening and validation of differentially expressed genes in polymyositis

Linmang Qin^{a,1}, Haobo Lin^{a,1}, Guangfeng Zhang^a, Jieying Wang^a, Tianxiao Feng^a, Yunxia Lei^a, Yuesheng Xie^a, Ting Xu^a, Xiao Zhang^{a,b,*}

^a Department of Rheumatology and Immunology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China

^b The Eighth Affiliated Hospital of Sun Yat-sen University, Shenzhen, China

ARTICLE INFO

Keywords:

Polymyositis
IFN- γ
DEGs
HLA
Bioinformatics

ABSTRACT

Background: Polymyositis (PM), a prevalent inflammatory myopathy, currently lacks defined pathogenic mechanism. To illuminate its pathogenesis, we integrated bioinformatics and clinical specimens to examine potential aberrant gene expression patterns and their localization.

Methods: We obtained GSE128470 and GSE3112 dataset from the Gene Expression Omnibus, performed Gene Set Enrichment Analysis (GSEA) and immune infiltration analysis using CiberSort, identified differentially expressed genes with Limma, conducted functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis, constructed a Protein-Protein Interaction network, and identified hub genes using Cytoscape. ROC analysis evaluated hub gene diagnostic accuracy for PM, validated their expression levels with clinical specimens.

Results: DEG analysis revealed 51 upregulated and 779 downregulated genes. Gene Ontology (GO) analysis implicated Type I interferon (IFN- γ) signaling, while KEGG pointed to cell adhesion molecule activation and oxidative phosphorylation inhibition. Protein-Protein Interaction (PPI) analysis identified 8 diagnostic hub genes. Clinical samples confirmed their upregulation in PM, especially IRF1 and IRF9 between muscle fibers. Different immune cell infiltrations were observed in PM patients versus controls.

Conclusions: Our study explores potential pathogenic factors, diagnostic markers, and immune cells in PM, with a focus on verifying IRF1 and IRF9 upregulation in the IFN- γ signaling pathway. These findings bear significance for PM diagnosis and treatment.

1. Introduction

Polymyositis (PM) represents a prevalent form of idiopathic inflammatory myopathy primarily affecting skeletal muscles. Its hallmark features include muscle inflammation, resulting in notable symmetrical proximal extremity weakness and tenderness. Epidemiological investigations indicate that adult PM constitutes approximately 30–45 % of idiopathic inflammatory myopathy cases.

* Corresponding author. Department of Rheumatology and Immunology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, 510080, China.

E-mail address: 13922255387@163.com (X. Zhang).

¹ These authors contributed equally to this study.

<https://doi.org/10.1016/j.heliyon.2024.e24537>

Received 9 May 2023; Received in revised form 10 January 2024; Accepted 10 January 2024

Available online 21 January 2024

2405-8440/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Notably, around 50 % of PM patients are susceptible to complications involving the heart and lungs, consequently contributing to diminished survival rates. However, despite extensive study, the precise etiology and pathogenesis of this condition remain elusive, with associations commonly acknowledged with genetic factors and viral infections.. Substantial research suggests a heightened susceptibility to idiopathic inflammatory myositis inheritance among individuals with specific human leukocyte antigen profiles. Presently, clinical management options for PM are constrained to corticosteroids, immunosuppressive agents, and biological therapies, lacking tailored pharmaceutical interventions. Consequently, patients often necessitate prolonged medication regimens, predisposing them to numerous adverse drug reactions and potentially deleterious side effects.

In recent years, notable strides have been made in second-generation sequencing, gene chip technology, and the establishment of comprehensive multi-omics databases. These advancements have notably augmented the accessibility of high-throughput sequencing data related to human diseases. As a result, they have enabled systematic inquiries into distinct gene subsets intricately linked to particular diseases. Despite a wealth of studies examining the putative pathogenesis of PM, a conspicuous gap exists in research efforts aimed at substantiating the potential pathogenic genes associated with PM. Specifically, there is a dearth of endeavors integrating bioinformatics analyses with experimental investigations conducted on clinical samples for validation purposes.

In this study, an exhaustive analysis was performed on two datasets to discern critical genes linked with PM while simultaneously evaluating immune cell infiltration and formulating a diagnostic model for this condition. Following this, clinical samples from PM patients were obtained to validate both the expression levels and subcellular localization of these hub genes. This investigation serves as a foundational framework, paving the way for further advancements in the diagnosis and therapeutic approaches for PM.

2. Materials and methods

2.1. General features

This study has been approved by the Medical Research Ethics Committee of Guangdong Provincial People's Hospital, Grant number: 2019353H (R1). Between January 2021 and December 2021, 4 PM patients aged over 18 years old took part in this study.

2.2. Inclusion criteria

Patients had to be able to communicate independently and complete questionnaires, and the disease classification criteria met the criteria proposed by Bohan & Peter in 1975 for the classification of PM and dermatomyositis.

2.3. Exclusion criteria

- Other idiopathic inflammatory myopathies, such as dermatomyositis, antisynthetase syndrome, immune-mediated necrotizing myopathy, sporadic inclusion body myositis.
- Excluding other causes of muscle damage, such as drug-related, metabolic, sports injury, endocrine related.
- patients with other autoimmune diseases, such as rheumatoid arthritis and systemic lupus erythematosus.
- Patients with any other malignant tumor were excluded.

2.4. Data acquisition

Utilizing "polymyositis" as the designated query term within the Gene Expression Omnibus (GEO) database, our investigation was confined to human-specific datasets, leading to the identification of two pertinent datasets: GSE128470 and GSE3112. Subsequently, these datasets were acquired via the GEOquery package integrated into the R software platform. Comprising muscle tissue biopsies from 13 individuals diagnosed with PM and 23 healthy controls, these datasets will serve as the foundation for subsequent analyses conducted within this study.

2.5. GSEA

GSEA stands as a methodology that assesses the collective impact of all differentially expressed genes at a comprehensive level, aiming beyond sole reliance on p-values for prioritization. GSEA encompasses genes that might lack statistically significant p-values but could wield substantial functional relevance. In this particular study, we established the threshold criteria with a False Discovery Rate <0.25 and p. adjust <0.05.

2.6. Immune infiltration analysis

The CIBERSORT database was utilized to compute the distribution of 22 distinct immune cell types within muscle samples obtained from both the disease and control groups (<https://cibersortx.stanford.edu/>). Subsequently, the outcomes were visualized to delineate the discrepancies in immune cell compositions between these two cohorts.

2.7. Identification of differentially expressed genes (DEGs)

The initial processing of the two datasets involved the elimination of duplicate probes associated with multiple molecules. In instances where a probe corresponded to the same molecule, preference was given to the probe exhibiting the highest signal value. Subsequently, the amalgamated data underwent batch-correction through the application of the ComBat function within the SVA software package. Calibration of the resulting data was assessed via principal component analysis. A normalized gene expression matrix file, encompassing 36 samples, was generated for subsequent identification of DEGs utilizing the Limma package. The threshold criteria for DEGs were established at $|\log_2 \text{fold change}| > 1$ and corrected p-value < 0.05 .

2.8. Functional enrichment analysis

Differential gene functional enrichment analysis was conducted employing the ClusterProfiler package, with subsequent visualization of the outcomes facilitated by the ggplot2 package. Genes were meticulously annotated and categorized according to their involvement across three distinct domains: biological processes, molecular functions, and cellular components, utilizing GO analysis. Furthermore, KEGG pathway enrichment analysis was executed to delineate significant signaling pathways enriched within the DEGs and to identify pivotal hub pathways.

2.9. Construction of PPI networks

The PPI network was assembled utilizing Cytoscape software, employing a combined score threshold of 0.3 to signify substantial data support for protein interactions. Isolated nodes were systematically eliminated from the network to refine its structure. The identification of hub genes was accomplished through the convergence of five topological algorithms within the network.

2.10. Diagnostic model construction

The "pROC" package in R was employed for evaluating diagnostic performance. ROC curve analysis serves to gauge the sensitivity and specificity of a test or biomarker in discriminating between two distinct groups. Within this study, ROC curves were utilized to appraise the hub genes' capacity to differentiate between PM patients and healthy controls, utilizing the Area Under the Curve (AUC) as a comprehensive measure of diagnostic performance.

2.11. Muscle tissue samples were collected

The muscle tissue specimens from 4 patients with PM were obtained from the outpatient operating room of Guangdong Provincial People's Hospital. The control group consisted of muscle tissue samples from 3 non-inflammatory myopathy patients who underwent hip joint replacement surgery or emergency trauma surgery at the Department of Orthopedics, Guangdong Provincial People's Hospital.

2.12. H&E staining

The muscle tissue underwent fixation in 4% paraformaldehyde, subsequent embedding in paraffin, and sectioning into 4 μm -thick slices utilizing a microtome, following standard staining protocols.

2.13. Immunohistochemistry (IHC)

The muscle tissue underwent fixation, embedding, slicing, dewaxing, rehydration, and subsequent incubation with primary antibodies, followed by secondary antibody labeling involving horseradish peroxidase. The visualization of DAB color reaction was carried out, and samples were examined and photographed using a microscope.

2.14. qPCR

RNA extraction was performed on muscle tissue samples obtained from each group, followed by reverse transcription into complementary DNA (cDNA). The mRNA expression levels of the target gene within each group were quantified using the 2- $\Delta\Delta\text{ct}$ method, employing the reference gene (GAPDH), and an AG fluorescent quantitative PCR kit. *HLA-B* Forward: CGGAGTATTGGGACCGGAAC, Reverse: CATGCTCTGGAGGGTGTGAG; *HLA-C* Forward: CCTAGCTGTCTTGGAGCTG, Reverse: CCTGCATCTCAGTCCCACAC; *HLA-F* Forward: AGATCCTCCAAAGGCACACG, Reverse: CTGTGTCCTGGGTCTGTTCC; *HLA-DQB1* Forward: ACCTTCGGGTAGCAACTGTC, Reverse: GTCCCGTTGGTGAAGTAGCA; *HLA-DRA* Forward: AGCCCTGTGGAAGTGAAGAGA, Reverse: GAGGGCAGGAAGGGGAGATA; *HLA-DPA1* Forward: GTTCTTCCCACAGTGCTCA, Reverse: CTGAGGGCACAAAGGTCAGG; *IRF1* Forward: GACCCTGGCTAGAGATGCGAG, Reverse: TCCTTGTGATGTCCCAGCC; *IRF9* Forward: CCAGCCAGGGACTCAGAAAG, Reverse: CAGAGGGACTGAGTGTGCGAG; GAPDH Forward: GGGAGCCAAAAGGGTCATCA, Reverse: TAAGCAGTTGGTGGTGCAGG.

2.15. Western blot (WB)

Muscle tissue underwent homogenization followed by centrifugation, and the protein concentration was quantified utilizing the BCA method. The protein was denatured, subjected to electrophoresis, and subsequently transferred onto a PVDF membrane. This membrane underwent blocking with 5 % skim milk powder, followed by overnight incubation with primary antibodies and subsequent treatment with secondary antibodies. Post-washing, a luminescent substrate solution was applied, and images were captured. The grayscale intensity of protein bands was analyzed using Image J software.

2.16. Statistical analysis

The measurement data were expressed as Mean \pm SD. Statistical analysis and graphical representation were conducted using Graph Prism 8. For comparing two groups of measurement data, a *t*-test was employed, while One-Way ANOVA was utilized for comparing multiple groups of measurement data. A significance level of $P < 0.05$ was considered statistically significant.

3. Results

3.1. Analysis flow chart

The analytical workflow is depicted in Fig. 1. Within this process, batch correction was implemented to harmonize data sets, subsequently amalgamated. GSEA enrichment analysis was conducted, succeeded by GO, KEGG, and immune infiltration analyses. Additionally, PPI networks were formulated. Notably, HLA class I and II genes, along with genes related to the interferon signaling pathway, emerged as pivotal hub genes in PM. To validate the expression levels and spatial distribution of these hub genes, clinical specimens were gathered and subjected to various experimental methodologies.

3.2. Identification of DEGs

Batch correction was performed on data sourced from GSE128470 and GSE31122 to mitigate potential batch effects, as illustrated in Fig. 2 A and B. Upon integration of these datasets and application of a threshold of $|\log_2FC| > 1$ and adjusted $P < 0.05$, a volcano plot was generated, revealing 830 DEGs (Fig. 2C). Among these, 51 genes exhibited downregulation, while 779 genes showed upregulation in PM patients in comparison to the healthy control group. A heatmap representing the top 20 DEGs is depicted in Fig. 2D.

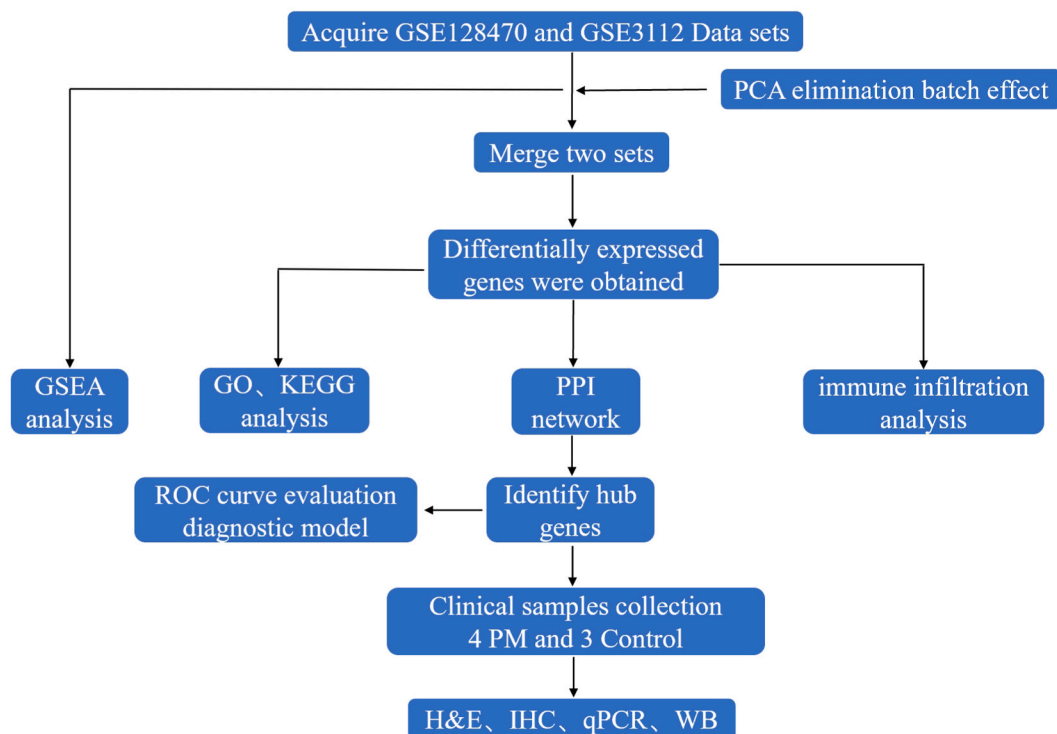


Fig. 1. Analysis flow chart.

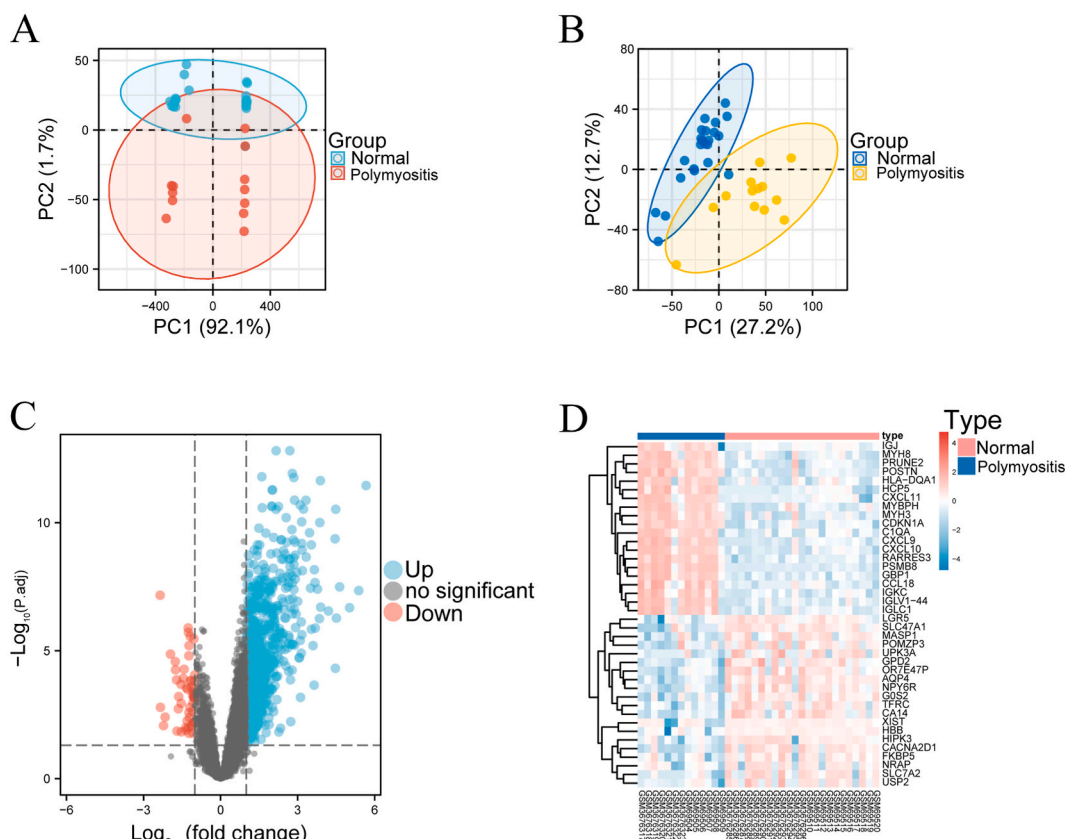


Fig. 2. Sample quality control and differential expression analysis. (A–B) Principal component analysis of two data sets before and after correction of the batch; (C–D) Volcanic and heat maps of expression patterns of DEGs.

3.3. GO and KEGG analysis

The GO functional enrichment analysis uncovered significant associations between the DEGs and various membrane regions, notably the outer segment of the plasma membrane, membrane raft, membrane microregion, secretory granule cavity, and membrane interval (Fig. 3A). Regarding molecular functions, these DEGs primarily exhibited involvement in actin binding, chemokine receptor binding, immunoreceptor activation, amide binding, and peptide binding activities (Fig. 3B). Furthermore, the biological processes enriched within these DEGs encompassed the IFN-I signaling pathways, cellular response to IFN-I, neutrophil-mediated immunity, neutrophil activation involved in immune response, response to viruses, T cell activation, and neutrophil degranulation (Fig. 3C).

Furthermore, KEGG analysis provided additional insights into the pathways implicated by these DEGs. The KEGG signal pathway enrichment analysis delineated two distinct categories: activation signal pathways and inhibition signal pathways. Bubble size and color in the visualization represented the quantity of enriched genes and their respective P-values. This analysis highlighted the prevalent enrichment of DEGs within pathways such as chemokine signaling, interactions between cytokines and viral proteins, viral protein interactions with cytokine receptors, Epstein-Barr virus infection, interactions between cytokines and their receptors, TLRs receptor signaling, and *Staphylococcus aureus* infection (Fig. 3D). Remarkably, the KEGG signaling pathways segregated into upregulated and downregulated expressions. Upregulated pathways encompassed cell adhesion molecules, Nod-like receptor signaling, and chemokine signaling, while downregulated pathways were associated with oxidative phosphorylation reduction (Fig. 3E). These findings underscore the close association between the pathogenesis of PM and the biological processes inherent in the IFN-I signaling pathway.

3.4. PPI network and identification of hub genes

The DEGs underwent analysis via Cytoscape to establish a PPI network, visually depicted by clusters represented as circles interconnected by lines denoting protein interactions between genes (Fig. 4A). Furthermore, an intricate network diagram and identification of the top 20 genes within the core module of the PPI network were provided (Fig. 4B and C). Intersection calculations utilizing five distinct topological algorithms resulted in the identification of 8 hub genes (Fig. 4D).

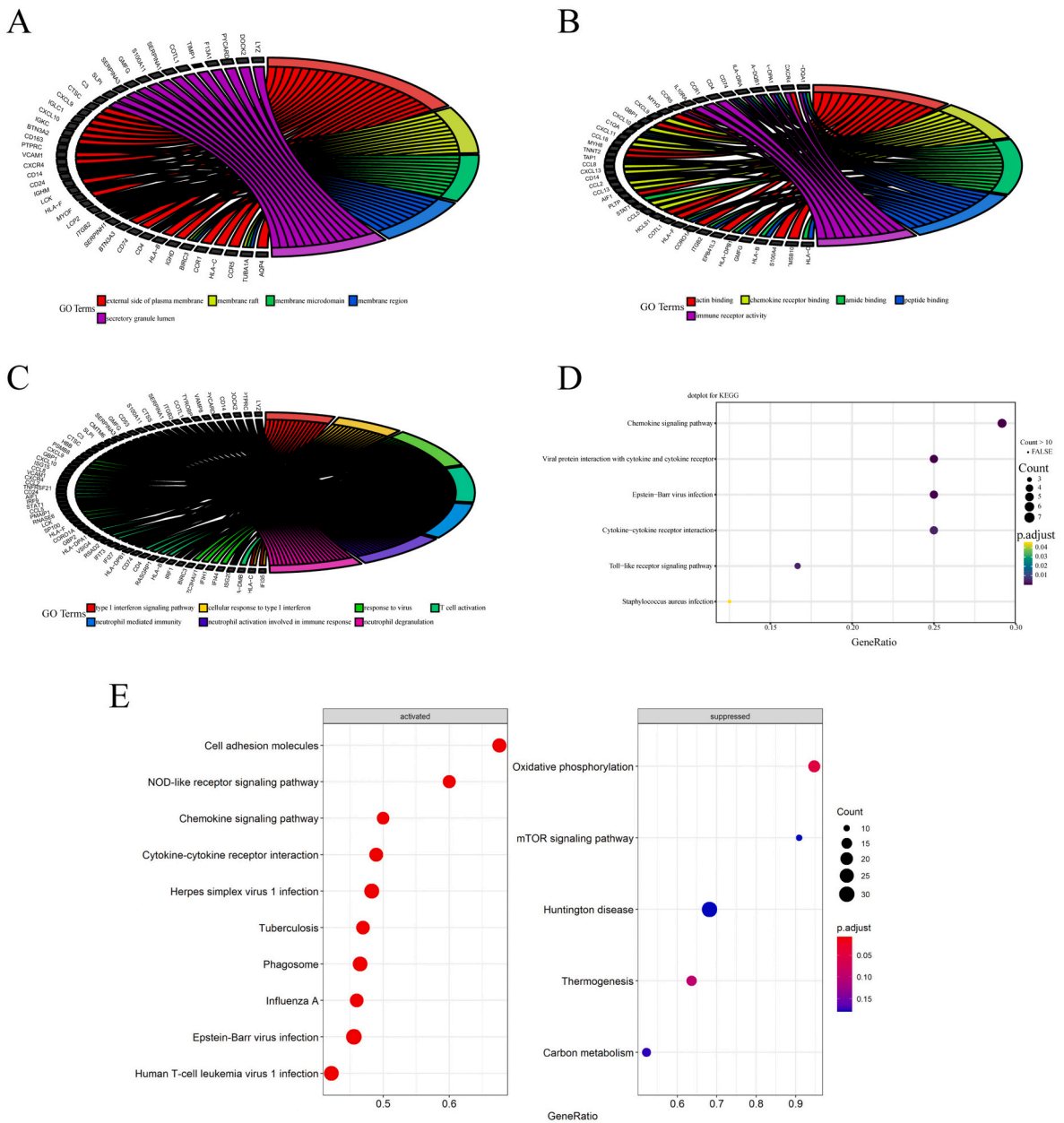


Fig. 3. GO Functional enrichment and KEGG analysis. (A) cellular components; (B) molecular functions; (C) biological processes. (D–E) Enrichment analysis of KEGG signaling pathway.

3.5. Expression pattern of hub gene and diagnostic model

The heatmap depicting expression patterns of the 8 hub genes in PM illustrates consistent upregulation across all 8 genes (Fig. 5A). Subsequently, an 8-hub gene diagnostic model was formulated and assessed through ROC curve analysis. The results showcased outstanding discriminatory capacity, with all hub genes exhibiting AUC values surpassing 0.8, indicating excellent performance in distinguishing between the disease and control groups (Fig. 5B–D).

3.6. Immune infiltration analysis and GSEA

Utilizing the CIBERSORT algorithm, we computed the relative proportions (Fig. 6A), correlations (Fig. 6C), and expression levels of 22 distinct immune cell types across 36 muscle tissue samples. These encompassed primary B cells, memory B cells, plasma cells, CD8⁺T cells, CD4⁺naive T cells, CD4⁺memory resting T cells, CD4⁺memory activated T cells, follicular helper T cells, regulatory T cells

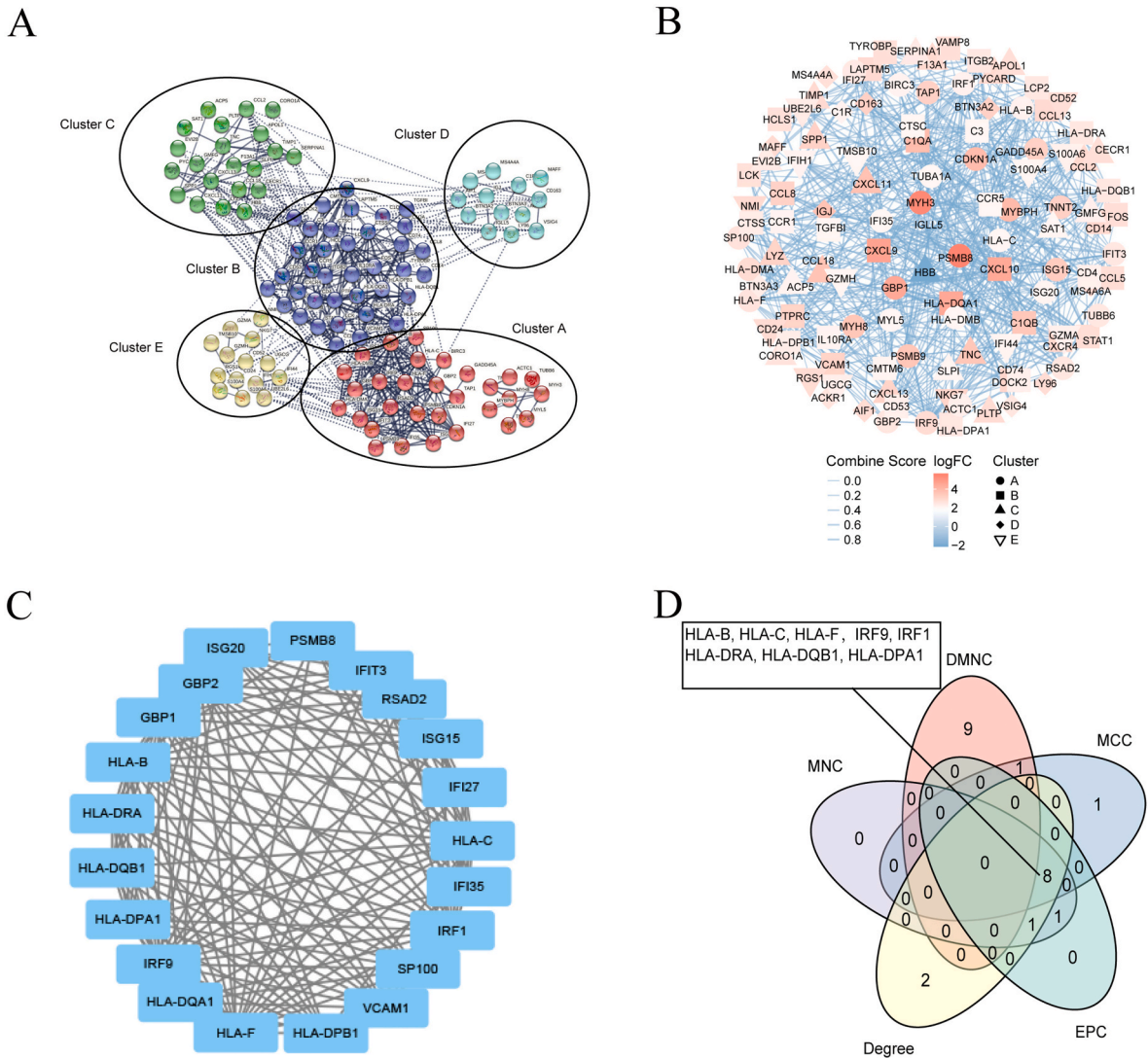


Fig. 4. PPI network diagram and identified hub genes. (A–B) PPI network graph; (C) 20 nodes; (D) Venn diagram of Hub genes.

(Tregs), gamma-delta T cells, resting natural killer (NK) cells, activated NK cells, monocytes, M0, M1, and M2 macrophages, as well as resting and activated dendritic cells, static and activated mast cells, eosinophils, and neutrophils (Fig. 6B). Comparative analysis between PM patients and the control group revealed a noteworthy increase in naive B cells, plasma cells, CD8⁺T cells, activated CD4⁺memory T cells, follicular helper T cells, and M2-type macrophages among PM patients. Conversely, PM patients exhibited a significant reduction in memory B cells, Tregs, resting NK cells, M0 macrophages, dendritic cells, resting mast cells, activated mast cells, and neutrophils. Subsequent Gene Set Enrichment Analysis unveiled enrichment of signaling pathways associated with Epstein-Barr virus infection, herpes simplex I infection, and human T-cell leukemia virus infection (Fig. 6D). The ordinate represents the enrichment items, while the horizontal axis signifies the corresponding P-value for each item.

3.7. Experimental verification

In the preceding phase, we identified hub genes characteristic of PM and subsequently conducted the collection of clinical specimens for validation purposes. Histological examination via H&E staining unveiled a notable infiltration of inflammatory cells within the muscle tissue of PM patients. These inflammatory cells were observed surrounding and infiltrating non-necrotic muscle fibers (Fig. 7A). qPCR validation of the identified hub genes exhibited a significant upregulation in mRNA levels across these 8 hub genes, notably including two pivotal transcription factors, IRF1 and IRF9 (Fig. 7F). Consequently, our focus shifted towards verifying their protein expression and localization within the muscle tissue. Immunohistochemical analyses demonstrated distinct staining patterns of IRF1 and IRF9 within PM patient muscle tissue, primarily localized between muscle fibers, with negligible expression observed in the

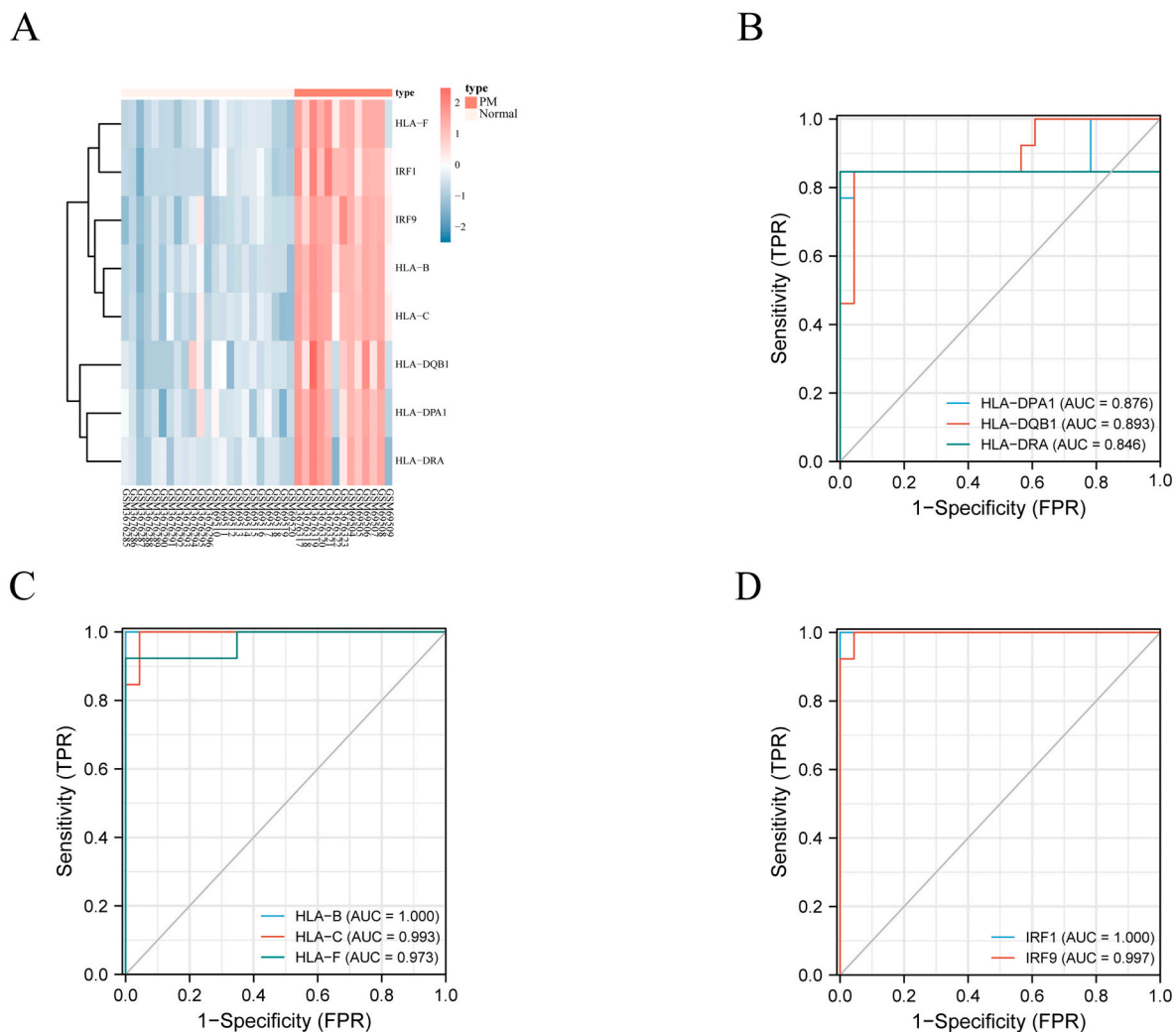


Fig. 5. Hub gene expression heat map and diagnostic model. (A) Heat maps of Hub gene expression patterns; (B–D) disease diagnostic models of Hub genes.

control group (Fig. 7B). Likewise, parallel Western blot analyses corroborated a significant elevation in IRF1 and IRF9 protein expression levels (Fig. 7C–E).

4. Discussion

In the context of this study, we conducted a comprehensive analysis of gene expression profiles and the immune cell infiltrate dynamics associated with PM, utilizing two distinct datasets. We deliberated upon the plausible signaling pathways and immune phenotypes characterizing this ailment, concurrently devising a diagnostic framework hinging on hub genes. Furthermore, we systematically procured clinical specimens to substantiate the relevance of these pivotal genes. Based on the analysis, we propose a hypothesis: the aberrant expression of specific HLA genes may disrupt immune system regulation, thereby facilitating auto-tissue aggression. Moreover, the dysregulated expression of interferon-associated genes might perturb the regulation of the interferon signaling pathway and exacerbate inflammatory responses, consequently advancing the trajectory of PM. To the best of our knowledge, this investigation represents a pioneering endeavor encompassing bioinformatic scrutiny spanning the transcriptional and proteomic strata, including insights into protein subcellular localization, to dissect the intricacies of PM.

PM constitutes a prevalent autoimmune inflammatory myopathy, and extant research posits it as a consequence of the intricate interplay between genetic predisposition and environmental factors. However, the precise pathogenesis of this condition remains enigmatic. As far back as 1977, investigations revealed the pivotal role of genetic variations within the HLA genes in the etiology of PM [1]. The HLA molecules are ubiquitously expressed on cell surfaces and represent one of the most influential susceptibility factors. Despite longstanding recognition of the significance of HLA gene polymorphisms in PM, scant efforts have been dedicated to their

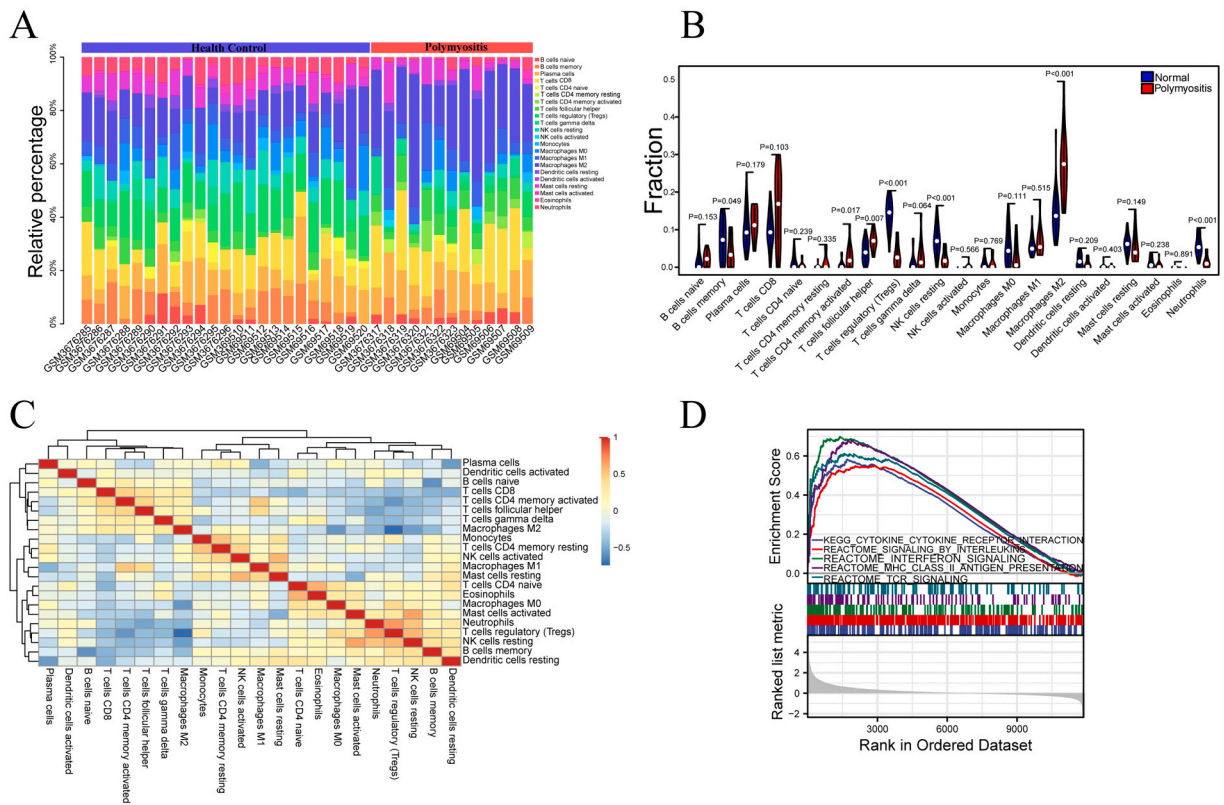


Fig. 6. Immune infiltration and GSEA analysis. (A) The relative proportion of 22 kinds of immune cells; (B) differentially expressed immune cells; (C) Correlation between immune cells; (D) GSEA based on 2 merge datasets.

detailed transcriptional-level examination. Subsequent investigations have underscored the association between genetic polymorphisms in HLA genes and diverse autoimmune disorders, including PM, exemplified by the diagnostic specificity of HLA-B27 in ankylosing spondylitis [2].

The histopathological characteristics of PM primarily encompass the infiltration of CD8⁺ cytotoxic T lymphocytes into non-necrotic muscle fibers [3–5]. The human leukocyte antigens HLA-A, HLA-B, and HLA-C constitute prominent components of the major histocompatibility complex (MHC) class I molecules embedded in the cell membrane. They serve the essential function of presenting peptides to cytotoxic CD8⁺ T lymphocytes, thus contributing significantly to immune system regulation [6]. Our study yielded analogous findings, revealing a substantial elevation in the transcriptional expression levels of HLA-B, HLA-C, and HLA-F within the muscle tissue of patients afflicted with PM. Notably, HLA-B exhibited the most pronounced upregulation. This observation aligns with a 2017 study conducted in China, encompassing 14 p.m. patients and 24 healthy controls, wherein HLA-B and HLA-DQB1 were identified as potential novel genetic variations in patients with IIM [7]. Intriguingly, another investigation suggested that individuals harboring the HLA-B allele may exhibit reduced susceptibility to pulmonary complications [8]. This observation may provide insight into the relatively lower incidence of ILD among PM patients in comparison to those with dermatomyositis.

Xiao Y’s investigation revealed that HLA-II genes serve as susceptibility factors for PM within the Han Chinese population, playing a pivotal role in the infiltration of inflammatory cells and resultant muscle damage in PM pathogenesis [9]. Our study corroborates these findings by demonstrating elevated expression levels of HLA-DQB1, HLA-DPA1, and HLA-DRA in PM patients. PM is primarily orchestrated by cellular immunity, with the HLA-DP protein, a heterodimer composed of the α chain encoded by the HLA-DPA1 gene and the β chain encoded by the HLA-DPB1 gene, assuming the crucial role of antigen presentation to CD4⁺T cells [10]. Similarly, HLA-DQB1 and HLA-DR molecules also partake in the presentation of antigens to CD4⁺T cells. HLA-DRA encodes the β chain of the HLA-II protein [11], and disruptions in the binding and recognition of cell surface peptides presented to T cells in this process lead to extensive cellular infiltration of muscle tissue, resulting in inflammation. These findings suggest that signal transmission disturbances involving HLA-DQB1, HLA-DPA1, and HLA-DRA may underlie the pathogenesis of PM.

The IFN signaling pathway is recognized to hold a pivotal role in dermatomyositis [12]. IFN- γ , in particular, has been implicated in exacerbating inflammatory assaults on compromised muscle fibers and in instigating the deposition of amyloid-like proteins within these fibers [13]. Conversely, despite limb skeletal muscle constituting the primary target tissue in PM patients, there is currently a paucity of research scrutinizing the IFN pathway in PM.

IRF1 and IRF9, members of the interferon regulatory factor family, function as transcription factors upstream of the IFN signaling cascade. IRF1 and IRF9 are pivotal transcription factors implicated in diverse cellular functions, notably in antiviral responses, cell

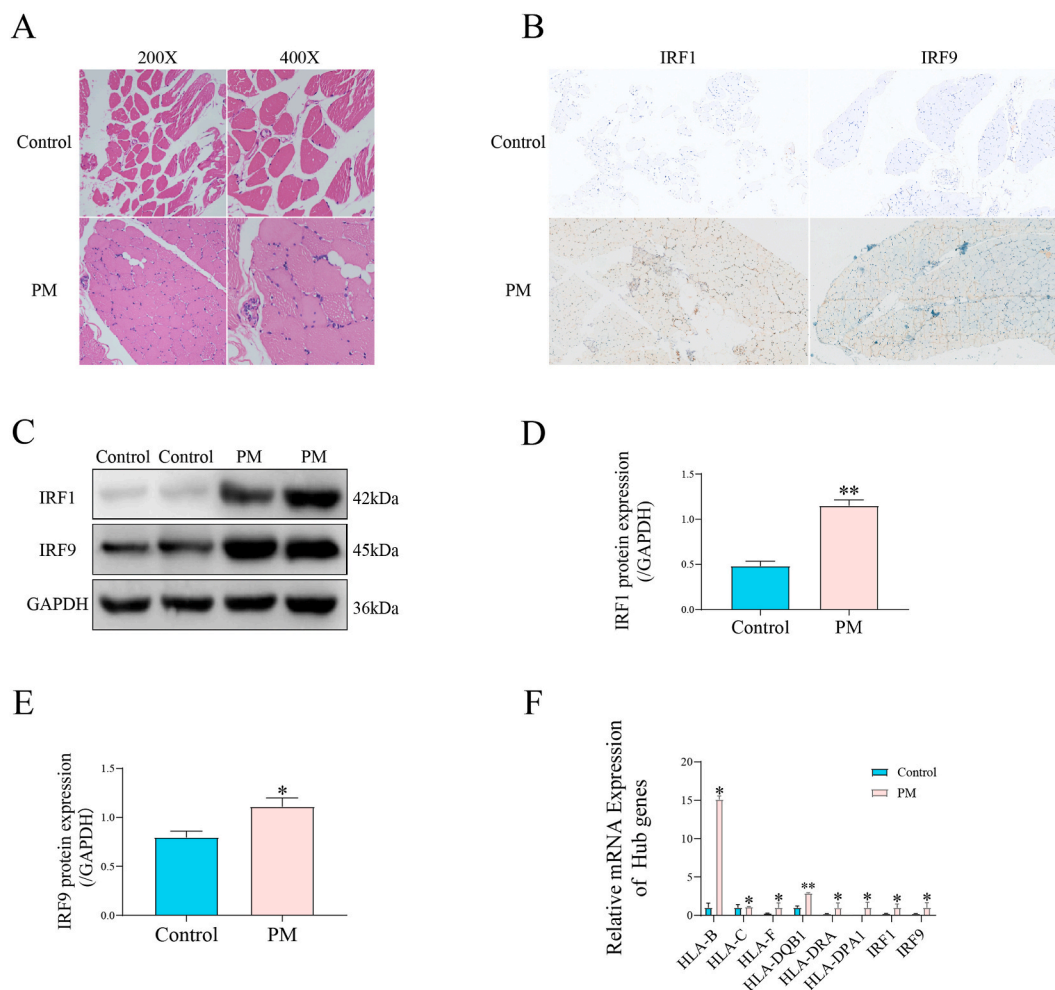


Fig. 7. Validation of Hub Genes. (A) H&E staining results of muscle tissue; (B) Immunohistochemical staining results of IRF1 and IRF9 in muscle tissue (C–E) IRF1 and IRF9 Protein expression levels in muscle tissue (F) 8 hub gene mRNA expression levels in muscle tissue. Mean \pm SD. * $P < 0.05$, ** $P < 0.01$.

proliferation, apoptosis, and immune modulation. Given their established roles within the immune system, it is plausible to infer that upon viral infection, IRF1 and IRF9 are stimulated by IFN- α , subsequently initiating the upregulation of IFN-induced genes. These genes encompass a spectrum of factors, including proinflammatory cytokines such as IL-10, and other mediators involved in apoptosis and early immune responses. This sequential activation sets forth a cascade of signal amplification within cellular pathways, culminating in PM muscle inflammation. Consequently, our focus pivoted towards scrutinizing alterations in the expression profiles of IRF1 and IRF9 in PM. Our quantitative polymerase chain reaction and Western blot analyses revealed upregulated expression of IRF1 and IRF9 in PM. Furthermore, GO and GSEA underscored significant enrichment of the Type I interferon signaling pathway in PM. Existing research elucidates their dual roles, as they can both positively and negatively modulate the transcription of interferon-stimulated genes in response to downstream interferon signaling, contingent on the presence or absence of IFN. This duality in function extends to their participation in antiviral immunity and the regulation of cellular proliferation and immune activity [14]. Our findings are in alignment with Si Chen et al.'s analysis, which also identified IRF9 and hub genes within the IFN pathway, such as TRIM22, IFI6, IFITM1, and IFI35, as noteworthy players in PM pathogenesis [15]. Additionally, Tournadre A and colleagues observed that IFN-I secretion led to an upsurge in HLA-I gene expression [16]. In concurrence, our analysis corroborated the upregulation of essential transcription factors IRF1 and IRF9 within the IFN signaling pathway, as well as heightened HLA-I gene expression. This observation may elucidate the underlying rationale for the augmented HLA-I gene expression. In summation, our investigation unveils the transcriptional and protein-level upregulation of IRF1 and IRF9 in PM, with these factors predominantly localized between muscle fibers.

In 2022, EULAR published a high-quality article recommending the inclusion of IFN-I within the scope of clinical assessments, albeit without specifying the specific molecule within the IFN-I pathway [17]. Additionally, within the US clinical trial database, clinical studies evaluating the therapeutic efficacy of interferon-alpha monoclonal antibody (MEDI-545) in PM patients have been documented. Our research underscores IRF1 and IRF9 as promising hub genes in PM. These findings collectively suggest the potential

of IRF1 and IRF9 to serve as biomarkers for PM diagnosis and treatment.

The pathogenesis of PM is predominantly characterized by cellular immunity dysregulation. Our investigation encompassed an immune infiltration analysis, revealing notable shifts in immune cell populations in PM patients. Specifically, we observed a substantial increase in the proportion of M2 macrophages, concomitant with significant reductions in regulatory T cells, resting NK cells, and neutrophils. This observed phenomenon may be attributed to the early inflammatory stage in PM, characterized by a self-protective mechanism, leading to the marked elevation of M2 macrophages. Studies have elucidated the localization of early-activated macrophages primarily within the endomysium of muscle tissue in PM patients [18]. Additionally, these activated macrophages have been linked to adverse prognostic outcomes and increased mortality rates in PM patients. In the context of sustained inflammation, regulatory T cells assume a negative regulatory role in immune modulation. The inhibitory cytokines secreted by regulatory T cells cease to induce immune tolerance to self-antigens in PM, consequently leading to their downregulation. Resting NK cells necessitate antigen presentation for effective activation [19], and our immune infiltration analysis corroborates this by demonstrating a decreased proportion of dendritic cells, the most potent antigen-presenting cell type, in PM. Consequently, resting NK cell levels exhibit a decrement in PM. Furthermore, research has highlighted widespread neutrophil infiltration in the muscle tissue of IFN- γ -deficient mice afflicted with PM [20]. In parallel, our results underscore the significant upregulation of interferon pathway genes and the notable enrichment of the interferon signaling pathway in immune cells infiltrating PM. Remarkably, we also observe a significant decrease in neutrophil levels. These collective findings emphasize the potential significance of M2 macrophages, regulatory T cells, resting NK cells, and neutrophils as promising avenues for further exploration in the context of PM pathogenesis.

5. Conclusion

Employing bioinformatics methodologies, we discerned pivotal genes central to the pathogenesis of PM and undertook a comprehensive analysis of the underlying molecular mechanisms driving this condition. To corroborate our findings, we conducted experimental validation, incorporating the collection of clinical samples from PM patients. Additionally, we scrutinized disparities in immune infiltrating cell populations between PM subjects and healthy controls. Hence, this study anticipates offering a foundational framework that holds promise for the prospective development of PM biomarkers, as well as enhanced diagnostic and therapeutic strategies.

5.1. Limitation of the study

1. Despite combining both data sets, our sample size was still relatively small and our patient population lacked diversity, which may limit generalizability to other populations.

Therefore, the identified hub genes and their regulatory mechanisms need to be further verified in a larger and independent cohort of PM patients.

2. Potential confounding factors, such as medication use and comorbidities, could affect the gene expression patterns observed in the study.
3. There is the possibility of false positive or false negative in the bioinformatics analysis, which may affect the accuracy of the identified hub genes.
4. Further functional studies are needed to elucidate the role of the identified hub genes in the pathogenesis of PM and to identify potential therapeutic targets.

Data availability statement

This study use of genome data from Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>), Gene Set Enrichment Analysis (<https://www.gsea-msigdb.org/gsea/index.jsp>), immune infiltration Analysis tool (<https://cibersortx.stanford.edu/>), Functional enrichment analysis (<http://geneontology.org/>).

Ethics declaration

This study was reviewed and approved by the Medical Research Ethics Committee of Guangdong Provincial People's Hospital (approval number:2019353H (R1)).

Additional information

No additional information is available for this paper.

CRedit authorship contribution statement

Linmang Qin: Writing – original draft, Visualization, Validation, Methodology. **Haobo Lin:** Writing – review & editing. **Guangfeng Zhang:** Resources. **Jieying Wang:** Formal analysis. **Tianxiao Feng:** Formal analysis. **Yunxia Lei:** Resources. **Yuesheng**

Xie: Resources. **Ting Xu:** Resources. **Xiao Zhang:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Natural Science Foundation of Guangdong Province (2019A1515010047 & 2019A1515010927); the National Natural Science Foundation of China (82271822); the Science and Technology Program of Guangzhou (202102021181 & 202201011346); and the Medical Scientific Research Foundation of Guangdong Province (A2019139).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2024.e24537>.

References

- [1] W.J. Cumming, et al., HLA and serum complement in polymyositis, *Lancet* 2 (8045) (1977) 978–979.
- [2] J.C. Woodrow, C.J. Eastmond, HLA B27 and the genetics of ankylosing spondylitis, *Ann. Rheum. Dis.* 37 (6) (1978) 504–509.
- [3] M.C. Dalakas, R. Hohlfield, Polymyositis and dermatomyositis, *Lancet* 362 (9388) (2003) 971–982.
- [4] A.R. Findlay, N.A. Goyal, T. Mozaffar, An overview of polymyositis and dermatomyositis, *Muscle Nerve* 51 (5) (2015) 638–656.
- [5] J.C. Millisenda, A. Selva-O'Callaghan, J.M. Grau, The diagnosis and classification of polymyositis, *J. Autoimmun.* 48–49 (2014) 118–121.
- [6] E. Olson, J. Geng, M. Raghavan, Polymorphisms of HLA-B: influences on assembly and immunity, *Curr. Opin. Immunol.* 64 (2020) 137–145.
- [7] Q.L. Peng, et al., Targeted capture sequencing identifies novel genetic variations in Chinese patients with idiopathic inflammatory myopathies, *Int J Rheum Dis* 21 (8) (2018) 1619–1626.
- [8] X. Gao, et al., HLA class II alleles may influence susceptibility to adult dermatomyositis and polymyositis in a Han Chinese population, *BMC Dermatol.* 14 (2014) 9.
- [9] Y. Xiao, et al., Global Analysis of Protein Expression in Muscle Tissues of Dermatomyositis/polymyositis Patients Demonstrated an Association between Dysferlin and Human Leucocyte Antigen A, *Rheumatology*, Oxford, 2019.
- [10] M. Anczurowski, N. Hirano, Mechanisms of HLA-DP antigen processing and presentation Revisited, *Trends Immunol.* 39 (12) (2018) 960–964.
- [11] A. Didonna, et al., A splice acceptor variant in HLA-DRA affects the conformation and cellular localization of the class II DR alpha-chain, *Immunology* 162 (2) (2021) 194–207.
- [12] L. Bolko, et al., The role of interferons type I, II and III in myositis: a review, *Brain Pathol.* 31 (3) (2021) e12955.
- [13] B. De Paepe, K.K. Creus, J.L. De Bleecker, Role of cytokines and chemokines in idiopathic inflammatory myopathies, *Curr. Opin. Rheumatol.* 21 (6) (2009) 610–616.
- [14] A. Michalska, et al., A positive feedback amplifier circuit that regulates interferon (IFN)-stimulated gene expression and controls type I and type II IFN responses, *Front. Immunol.* 9 (2018) 1135.
- [15] S. Chen, et al., Identification of hub biomarkers and immune cell infiltration in polymyositis and dermatomyositis, *Aging (Albany NY)* 14 (10) (2022) 4530–4555.
- [16] A. Tourmadre, et al., Immature muscle precursors are a source of interferon- β in myositis: role of Toll-like receptor 3 activation and contribution to HLA class I up-regulation, *Arthritis Rheum.* 64 (2) (2012) 533–541.
- [17] J. Rodríguez-Carrio, et al., EULAR points to consider for the measurement, reporting and application of IFN-I pathway activation assays in clinical research and practice, *Ann. Rheum. Dis.* 82 (6) (2022) 754–762, 2023.
- [18] K.M. Rostasy, et al., Monocyte/macrophage differentiation in dermatomyositis and polymyositis, *Muscle Nerve* 30 (2) (2004) 225–230.
- [19] R. Barreira da Silva, C. Münz, Natural killer cell activation by dendritic cells: balancing inhibitory and activating signals, *Cell. Mol. Life Sci.* 68 (21) (2011) 3505–3518.
- [20] Y. Yoshihashi-Nakazato, et al., Interferon- γ , but not interleukin-4, suppresses experimental polymyositis, *Arthritis Rheumatol.* 68 (6) (2016) 1505–1510.